Integrative Structural Analysis of the Type III Secretion System from Gram-negative Bacteria

DISSERTATION

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

Department of Chemistry
Universität Hamburg

Lara Kristin Flacht

born in Berlin

Hamburg, 2022

Reviewers

1. Prof. Dr. Michael Kolbe

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

Department of Structural Infection Biology, Helmholtz Centre for Infection Research (HZI), Braunschweig, Germany

Centre for Structural Systems Biology (CSSB), Hamburg, Germany

2. Dr. Christian Löw

European Molecular Biology Laboratory (EMBL), Hamburg Unit, Hamburg, Germany

Centre for Structural Systems Biology (CSSB), Hamburg, Germany

Examination Commission

1. Chair: Prof. Dr. Michael Kolbe

2. Vice Chair: Prof. Dr. Kay Grünewald

3. Member: Prof. Dr. Jens Bosse

Thesis submission: 07.11.2022

Oral defense: 13.01.2023

Approval for publication: 13.01.2023

This dissertation was prepared from July 2016 to October 2022 under the supervision of Prof. Dr. Michael Kolbe in the working group 'Structural Infection Biology' at the Centre for Structural Systems Biology in Hamburg and of Prof. Dr. Charlotte Uetrecht in the working group 'Dynamics of Viral Structures' at the Leibniz Institute of Virology in Hamburg, Germany.

Parts of this work are published or are in preparation for publication:

Flacht L, Lunelli M, Kaszuba K, Chen ZA, O'Reilly FJ, Rappsilber J, Kosinski J, Kolbe M (2022) Integrative structural analysis of the type III secretion system needle complex from *Shigella flexneri*. Submitted.

Bernal I, Römermann J, Flacht L, Lunelli M, Utrecht C, Kolbe M (2019) Structural analysis of ligand-bound states of the *Salmonella* type III secretion system ATPase InvC. *Protein Science* 28:1888-1901.

Table of Contents

Abbreviations

AHT	Anhydrotetracycline
ATP	Adenosine triphosphate
BS3	Bis(sulfosuccinimidyl)suberate
C	Cyclic (symmetry)
C-terminus	Carboxyl(-COOH)-terminus
CAS RN	Chemical abstracts service registry number
CMC	Critical micelle concentration
Cryo-EM	Cryo-electron microscopy
Cryo-ET	Cryo-electron tomography
DDM	N-Dodecyl-β-D-maltoside
DMSO	Dimethyl sulfoxide
DSS	Disuccinimidyl suberate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EPEC	Enteropathogenic *Escherichia coli*
ETEC	Enterotoxigenic *Escherichia coli*
FDR	False discovery rate
FWHM	Full width half maximum
GHS	Globally Harmonized System of Classification and Labelling Chemicals
HRP	Horseradish peroxidase
IM	Inner membrane
LB	Lysogeny broth
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LMNG	Lauryl maltose neopentyl glycol
Lol	Localization of lipoproteins
M cell	Microfold cell
m/z	Mass-to-charge ratio
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
N-terminus	Amino(-NH$_2$)-terminus
NMR	Nuclear magnetic resonance spectroscopy
OD$_{600nm}$	Optical density (measured at a wavelength of) 600 nm
OM	Outer membrane
PBS	Phosphate-buffered saline
PDB ID	Protein data bank identifier
PNMs	Polymorphonuclear leukocytes
PVDF	Polyvinylidene difluoride
RBM	Ring-building motive

RMSD	Root-mean-square deviation
rpm	Revolutions per minute
RT	Room temperature
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SEC-MALS	Size-exclusion chromatography multi-angle light scattering
SPI-1	*Salmonella* pathogenicity island 1
Spp.	Species pluralis
T2SS	Type II secretion system
T3SS	Type III secretion system
T4PS	Type IV pilus system
TBS-T	Tris-buffered saline with Tween 20
TEM	Transmission electron microscopy
TSB	Tryptic soy broth
v/v	Volume per volume
w/v	Weight per volume
x *g*	Times gravitational force
Δ	Deletion
::	Insertion

Zusammenfassung

Das Typ-III-Sekretionssystem (T3SS) ist ein Megadalton Proteinkomplex der von vielen Gram-negativen Bakterien während der Infektion genutzt wird. Es besteht aus dem membrandurchspannenden Nadelkomplex und weiteren zytosolischen Komponenten, wie die Sortierungsplattform, Regulatoren und die zu transportierenden Effektorproteine sowie deren zugehörigen Chaperonen. Während der Infektion werden Effektorproteine mithilfe des T3SS direkt aus dem bakteriellen Zytosol in die Wirtszelle transportiert, um die Funktion der Wirtszelle zu Gunsten des Erregers zu modulieren. Für die künftige Entwicklung von Medikamenten, welche auf dieses System abzielen, ist es von entscheidender Bedeutung, die Struktur und die zugrundeliegenden molekulare Mechanismen des T3SS zu verstehen. Jedoch stammt das meiste strukturelle Wissen über den T3SS Nadelkomplex bisher von nur einem einzigen Bakterium aus der Gattung der *Salmonellen*. Daher konzentriert sich diese Arbeit auf die Strukturanalyse des T3SS Nadelkomplexes von einem Bakterium aus einer weiteren Gattung: *Shigella flexneri*.

Zunächst wurde das bestehende Verfahren zur Isolierung des T3SS Nadelkomplexe für die darauffolgenden Versuche angepasst. Dazu gehört der Austausch des Puffers (zu HEPES) und des Detergens (zu Triton X-100 oder Lauryl Maltose Neopentyl Glycol (LMNG)). Die Gesamtqualität der Nadelkomplexprobe war besser, wenn die *Shigellen* bei niedrigeren optischen Zelldichten (gemessen bei einer Wellenlänge von 600 nm (OD_{600nm})) geerntet wurden als denen, die bei höheren dichten geerntet wurden. Die Wahl des Kulturmediums, entweder lysogeny broth (LB, engl.) oder tryptic soy broth (TSB, engl.) wiederum, hatte keine Auswirkung auf die Qualität der Probe, soweit dies anhand negativ-kontrastierten transmissionselektronenmikroskopischen Aufnahmen beurteilt werden kann. Allerdings wuchsen die Bakterien (gemessen an der OD_{600nm}) im TSB-Medium schneller als im LB-Medium, und das Wachstum wurde durch die Plasmidinduktion negativ beeinflusst. Darüber hinaus wird ein schnelles, nicht-denaturierendes Dot-Blot-Verfahren beschrieben, mit dem die Nadelkomplex-Untereinheiten MxiG [SctD] sowie das Strep-MxiH [SctF] nachgewiesen werden können. Dieses Verfahren wird die Überwachung künftiger Isolierungen vereinfachen.

Die Analyse der Proteinstruktur vom Sekretin (MxiD [SctC]), welche durch Einzelpartikel-Kryo-Elektronenmikroskopie (Kryo-EM) des isolierten Nadelkomplexes gewonnen wurde, zeigte, dass der Sekretinkomplex ein großes, pentadekameres, doppelwandiges β-Fass bildet. Obwohl dieser Sekretinkomplex seinem *Salmonella*-Homolog sehr ähnlich ist, nimmt die C-

terminale S-Domäne, eine einzigartige Konformation an. Um Regionen zu untersuchen, welche nicht durch Kryo-EM bestimmt werden konnten, wurde ein Quervernetzungsverfahren (engl. cross-linking) angewandt. Dieses Verfahren musste für den isolierten Nadelkomplex, sowie für *Shigella*-Zellen *in vivo* zunächst etabliert und anschließend mithilfe von Massenspektrometrie (MS) analysiert werden. Die resultierenden Quervernetzungen aus dem isolierten Nadelkomplex bestätigten zunächst das Vorhandensein des Pilotins (MxiM [SctG]). Des Weiteren konnte durch eine integrative Modellierungsmethode die Lokalisierung des Pilotins weiter eingegrenzt werden, welche jedoch auch die hohe Flexibilität des Pilotins offenlegte. Im Gegensatz dazu, ließen *in vivo* erzeugte Querverbindungen auf eine definiertere Lokalisierung mit geringerer Flexibilität des Pilotins schließen. Zusammen mit früheren Erkenntnissen deuten diese Ergebnisse darauf hin, dass das Pilotin in *Shigella* an den vollständig assemblierten T3SS-Nadelkomplex gebunden bleibt, was wahrscheinlich durch die einzigartige Faltung der S-Domäne des Sekretins ermöglicht wird.

Native MS-Analysen der rekombinant exprimierten *Salmonella*-Sortierplattform-ATPase InvCΔ79 [SctN] zeigten, dass das Konstrukt sowohl monomere als auch dimere Zustände annehmen kann. Darüber hinaus wurde ein dichtes Netzwerk von Querverbindungen aus dem löslichen Sortierplattform-Unterkomplex SpaO/SpaO$_C$/OrgB/InvC [SctQ/SctQ$_C$/SctL/SctN] generiert, dass in einer zukünftigen integrativen Modellierungsanalyse verwendet werden kann.

Diese Arbeit trägt zu einem besseren Verständnis des T3SS bei, welches, wenn in Zukunft fortentwickelt, dazu verwendet werden könnte (z.B. bei der Medikamentenherstellung) um Infektionen mit Gram-negativen Bakterien die das T3SS benötigen, zu bekämpfen.

Abstract

The type III secretion system (T3SS) is a giant, megadalton protein complex utilized by many Gram-negative bacteria to initiate infection. It is composed of the membrane-spanning needle complex and further cytosolic components, such as the sorting platform, regulators, effector proteins, and their corresponding chaperones. During infection, effectors are transported from the bacterial cytosol directly into the host cell by the T3SS to govern host cell function in favor of the pathogen. Understanding the structure and underlying molecular mechanisms of the T3SSs is essential for future drug development targeting this system. Still, the most detailed structural knowledge of the T3SS needle complex originates from a single pathogen within the *Salmonella* genus. To expand the understanding of T3SSs, this work focused on the needle complex from a pathogen of an additional genus, *Shigella flexneri.*

First, the existing needle complex isolation procedure was adapted for downstream analysis. This included the exchange of buffering agent (to HEPES) and detergent (to Triton X-100 or lauryl maltose neopentyl glycol (LMNG)). Cell harvest at lower optical density (measured at a wavelength of 600 nm (OD_{600nm})) improved overall sample quality as far as could be determined by negative stain electron microscopy. Further, the choice of culture medium, either lysogeny broth (LB) or tryptic soy broth (TSB), did not affect sample quality. However, the choice of culture media did impact growth behavior. The utilized *Shigella* strain grew faster in TSB than in LB (as measured by OD_{600nm}), and plasmid induction negatively impacted the growth rate in all samples. In addition, a fast, non-denaturing dot blot procedure is described, which can detect the needle complex subunits MxiG [SctD] and the strep-tagged-MxiH [SctF] that will aid in monitoring future isolation procedures.

Analysis of the isolated needle complex's atomic model of the secretin (MxiD [SctC]) obtained by single particle cryo-electron microscopy (cryo-EM) demonstrated that it adopts a large, pentadecameric, doubly-walled β-barrel highly similar to its *Salmonella* homolog. The secretin C-terminal S domain, however, adopts a unique, laterally extended fold. To investigate regions that could not be structurally resolved by cryo-EM, a cross-linking procedure for the isolated needle complex and *Shigella* cells *in vivo* was established and subsequently analyzed with mass spectrometry (MS). Cross-links obtained from the isolated needle complex confirmed the models obtained by cryo-EM and disclosed the presence of an additional subunit: the pilotin (MxiM [SctG]). An interactive modeling approach could state the localization of the pilotin, which, however also revealed high flexibility of this subunit. In contrast, cross-links generated

in vivo suggested a more defined localization of the pilotin. Together with previous work, these findings indicate that the pilotin of *Shigella* remains bound to the fully assembled T3SS needle complex, likely enabled by the unique fold of the secretin's S domain.

Native MS analysis of the recombinantly expressed *Salmonella* sorting platform ATPase InvCΔ79 [SctN] demonstrated that the construct can adopt monomeric and dimeric states. Additionally, a dense network of cross-links was generated from the soluble sorting platform sub-complex SpaO/SpaOC/OrgB/InvC [SctQ/SctQ$_C$/SctL/SctN] that could be utilized in a future integrative modeling approach.

This work contributes to a better understating of the T3SS, which, if further advances in the future, could be used (e.g., in drug development) to treat infections caused by T3SS-dependent Gram-negative bacteria.

1. Introduction

1.1. The type III secretion system and its role in pathogenic Gram-negative bacteria

Gram-negative bacteria appear red to pink when stained according to the Gram-protocol due to their cell envelope composition [1,2]. While Gram-positive bacteria retain the primary dye after solvent treatment and subsequently appear purple-blue, Gram-negative do not, thus, are stained red-pink by the subsequent counter dye [1]. The cell wall of Gram-negative is composed of an inner membrane (IM) and an outer membrane (OM) with a periplasmic space in-between that includes the peptidoglycan layer [3].

In order to transport substrates across these barriers, multiple secretion systems have evolved [4]. One of these systems is the type III secretion system (T3SS). It transports effector proteins from the bacterial cytosol directly into the host cell [5,6]. While some bacteria utilize the T3SS in a more symbiotic relationship with plants they inhabit [7,8], the T3SS is mainly associated with its essential role during infection [9]. Additionally, the virulence T3SS is closely related to the flagellar apparatus, which is often referred to as the flagellar T3SS [5,10]. This work, however, focuses on the virulent T3SS and will refer to it as T3SS, while the flagellum will be mentioned as such.

Many pathogenic Gram-negative bacteria rely on the T3SS to initiate infection. This includes species of multiple genera such as *Shigella, Salmonella, pathogenic Escherichia coli* (enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC)), *Pseudomonas, Yersinia, Vibrio, Chlamydia, Bordetella and Burkholderia* [9]. These pathogenic bacteria infest a broad variety of tissues ranging from e.g., the intestine, for *S. flexneri* causing dysentery [11], the respiratory tract, for *B. pertussis* causing whooping cough [12], or opportunistic infections of *P. aeruginosa* evoking different diseases [13]. This demonstrates that the T3SS is not restricted to a host environment or clinical disease spectrum [9]. Because of its essential role during infection of various Gram-negative bacteria, the T3SS displays an attractive potential target for drug development [9,14].

This is of particular interest since the increasing antibiotic resistance is emerging to one of the biggest global health threats with a high economic burden [15–17]. In 2017, the WHO released a

list of twelve drug-resistant bacteria of particular concern to global health [18]. From these, at least three utilize the T3SS for virulence (*P. aeruginosa*, *Salmonella* spp., *Shigella* spp. [9]).

The T3SS is a large protein machinery with more than 20 proteins in multiple copies and a molecular mass around 6 MDa [19] (Fig. 1 A-B). It can be roughly divided into the membrane-spanning needle complex and cytosolic components, such as the sorting platform and the effectors with their corresponding chaperons and regulators [19,20] (Fig. 1 B). Because of its shape and function resembling that of a syringe, it is often referred to as 'the injectisome' [21]. The needle complex is composed of the basal body and a long filamentous hollow needle structure protruding from the bacterial surface [22,23]. During infection, effectors are transported from the bacterial cytosol into the host cell [5,24]. To passage the narrow channel of the needle, effectors must be at least partially unfolded [25–27]. Even though the structural components of the T3SS are relatively conserved among the different bacteria, secreted effectors are species-specific to meet the individual pathogen requirements [5,19]. Secretion follows a hierarchical order, with the filamentous needle, the needle tip, and the translocon pore being the first proteins secreted [5,19].

Apart from the needle complex, the T3SS includes an additional cytosolic complex located below the basal body, termed the sorting platform. It is highly dynamic and determines the order of effector secretion [28–30]. While the needle complex can be isolated as whole [22,23], and therefore could be subjected to analyses such as single-particle cryo-electron microscopy (cryo-EM) [31] in near-atomic resolution [32–34] the sorting platform is too fragile or dynamic and dissociates from the needle complex upon isolation. Therefore, structural information of the entire sorting platform is limited to lower resolution cryo-electron tomography (cryo-ET) analysis *in situ* [35–41].

The molecular details of T3SS energetics are not yet fully understood [20,42]. It was demonstrated that the ATPase within the sorting platform is responsible for the dissociation of the effectors from their cognate chaperons and subsequent unfolding for secretion [43,44]. The translocation of effectors through the needle complex, however, is mainly energized by proton motive force [45,46].

Even though major structural advances of the T3SS have been made (for a recent review, see [20]), many questions remain. This includes the molecular details of energetics, along with a high-resolution structure of the full-length export gate, and the molecular details of selective substrate sorting, including high-resolution structures of the sorting platform [20,47]. Additionally,

analyzing the T3SS in a species-specific manner to identify conserved and unconserved features might facilitate drug design targeting the T3SS to combat a broad range of common pathogens.

Shigella	Salmonella	Unified	Function / Location
IpaB	SipB	SctE	1st translocator
IpaC	SipC	SctB	2nd translocator
IpaD	SipD	SctA	Needle tip
MxiH	PrgI	SctF	Needle filament
MxiI	PrgJ	SctI	Inner rod
MxiD	InvG	SctC	Secretin, OM ring
MxiM	InvH	SctG	Pilotin, (OM ring)
SpaP (Spa24)	SpaP	SctR	Export apparatus core
SpaQ (Spa9)	SpaQ	SctS	Export apparatus core
SpaR (Spa29)	SpaR	SctT	Export apparatus core
SpaS (Spa40)	SpaS	SctU	Autoprotease
MxiA	InvA	SctV	Export gate
MxiJ	PrgK	SctJ	Inner IM ring
MxiG	PrgH	SctD	Outer IM ring
MxiK	OrgA	SctK	Adaptor
Spa33	SpaO	SctQ	Pod
MxiN	OrgB	SctL	Radial spoke
Spa47	InvC	SctN	ATPase
Spa13	InvI	SctO	Stalk
MxiC	InvE	SctW	Gatekeeper
Spa32	InvJ	SctP	Needle lenght regulator

Figure 1: Architecture of the *Shigella* type III secretion system. (A) Top: Scheme of a *Shigella* spp. bacterium utilizing the T3SS. Effectors (shades of green), which are in complex with their corresponding chaperons [48] (shades of orange) in the bacterial cytoplasm, are transported into the host cell via the T3SS. In order to pass through the narrow channel of the needle, the effectors must be at least partially unfolded [25–27]. **Bottom:** Table of T3SS subunits from *Shigella* spp., *Salmonella* spp. pathogenicity island 1 (SPI-1) and the unified secretion and cellular translocation (Sct) nomenclature [49]. For other Gram-negative bacteria, see [5,50,51]. (Protein naming derived from: Inv: Invasion; Ipa: Invasion plasmid antigen; Mxi: Membrane expression of Ipa; Org: Oxygen-regulated gene; Prg: *phoP*-repressed gene; Sip: *Salmonella* invasion protein; Spa: surface presentation of antigen [51])**. (B)** Scheme of the T3SS of *Shigella* spp. spanning from the cytosol through the inner (IM) and outer membrane (OM) of the bacteria and extending up to the host cell. Subunits [19,51] are labeled on the left, and substructure nomenclature [20] is indicated on the right.

1.1.1. The pathogen *Shigella spp.*

Shigella spp. are rod-shaped, non-sporulating, facultative anaerobic, pathogenic, Gram-negative bacteria restricted to humans and primates [52]. *Shigella* infection causes bacillary dysentery, also known as shigellosis. Clinical symptoms can range from mild to severe. Usually, symptoms are initiated with fever, headache, and malaise, followed by watery or bloody and mucosal diarrhea with abdominal pain and cramps [11]. 269 million cases were estimated globally for 2016, with more than 200 000 deaths, which primarily affects children under the age of 5 years [53].

Generally, the infection is self-limiting and can be overcome by supportive care e.g., counteracting the diarrhea-induced dehydration and electrolyte loss. Further, treatment with antibiotics such as ciprofloxacin is advised, particularly in severe cases (for detailed treatment guidelines, see: [11]). However, the accelerating resistance of *Shigella* to antibiotics, especially to multiple drugs simultaneously, is an increasing matter of concern [54]. Recently, *Shigella* was among the priority list of drug-resistant bacteria released by the WHO to promote development of novel therapeutics against these pathogens [18]. Up to date, no licensed vaccine against *Shigella* is available; however, multiple candidates are currently in development [55].

The genus *Shigella* comprises four species (*S. boydii*, *S. dysenteriae*, *S. flexneri*, *S. sonnei*), which are further divided into various serotypes depending on the O antigens on their lipopolysaccharides [56]. Globally *S. flexneri* and *S. sonnei* are the most common causes of shigelloses. While *S. flexneri* is the main cause of shigelloses in low- and middle-income countries, *S. sonnei* is predominantly responsible for outbreaks in high-income countries [11].

The *Shigella* infection follows the fecal-oral route or by the uptake of contaminated food or water [11]. *Shigella* requires a relatively low infection dose, with 10-500 cells being sufficient to cause disease (for comparison, in *Salmonella,* a dose of 10^5-10^9 cells was reported) [57]. After ingestion, the bacteria follow the digestion route until they reach the colon epithelium (Fig. 2). There, *Shigella* are uptaken by microfold (M) cells [52,58]. M cells are specialized cells within Peyer's patch, which sample the lumen for antigens (e.g,. bacteria) as part of the host's immunosurveillance. Thereby M cells transfer antigens across the epithelium to antigen-presenting cells, where they are phagocytosed, broken down, and further processed by the immune system [59]. *Shigella*, however, with the help of the T3SS, escapes the macrophage's phagosome, replicates, and actively induces inflammatory cell death [52,60,61]. This leads to the release of the bacteria, which can now invade the colonic epithelial cells from the basolateral

side [62]. Utilizing effectors, *Shigella* are uptaken by actin rearrangement of the cell and subsequently lyse phagosome to reach their main replication niche, the cytosol [63,64]. Within the cell, *Shigella* induces actin polymerization at one side of the pole for motility and spreads into adjacent epithelial cells [64,65]. Thereby effectors intervene with the inflammatory response to promote host cell survival [52,66]. In addition, polymorphonuclear leukocytes (PMNs) are recruited by the generated inflammatory signals and transmigrate the epithelia layer. This destabilization of the epithelial layer enables further invasion of luminal *Shigella* and initially amplifies infection and tissue damage [67,68]. However, eventually, the infection can be cleared by the host's immune response, in which PMN neutrophils play an important role [52].

Shigella's virulence is based on genes encoded on a large $\approx$ 210 kilobase virulence plasmid pWR100 [69,70], and loss of the plasmid can lead to avirulent phenotypes [71]. Genes necessary for invasion cluster within a 31 kilobase region of the plasmid termed the 'entry region' [69,72,73]. These genes encode 34 proteins comprising the T3SS apparatus, the to-be-transported effectors, and their corresponding chaperones as well as transcriptional regulators [69]. Mainly because of its importance, the virulence plasmid was sequenced as early as 20 years ago by two groups independently [69,70]. The rest of the genome followed around a decade later [74]. However, the most complete genome sequence and annotation of the laboratory reference strain *S. flexneri* serotype 5a M90T was just published recently [75]. It comprises the chromosome (GeneBank database accession number: CP037923) and the latest sequence of the virulence plasmid (CP037924).
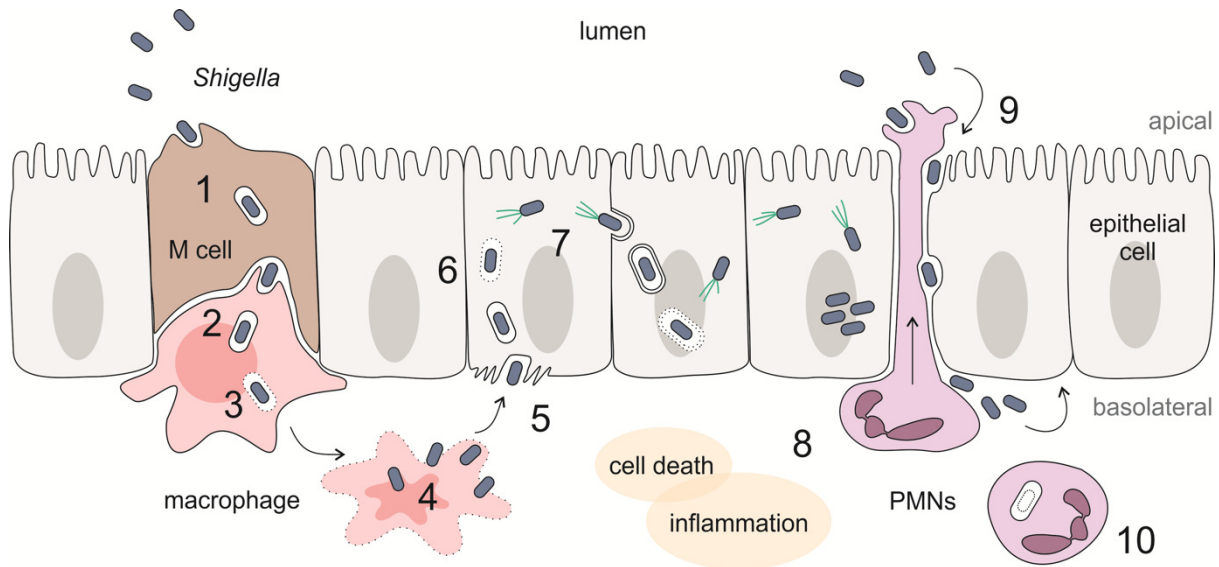
Figure 2: *Shigella* infecting the colonic epithelium. Scheme of the *Shigella* spp. infection route of the colonic epithelium. **(1)** Transfer of *Shigella* across the epithelium to the basolateral side through M cells and **(2)** phagocytoses by residential macrophages as part of the host's mucosal immune surveillance [52,58]. **(3)** *Shigella* escapes by lysing the phagosome and **(4)** actively inducing inflammatory cell death [52,60,61]. **(5)** Basolateral entry of *Shigella* in epithelial cells by induced actin rearrangement [62,64]. **(6)** Phagosomal lysis and replication of bacteria within the cytosol [52,63,64] **(7).** Intra- and inter-cellular motility by actin polymerization cytosol [52,64,65]. **(8)** Inflammatory signals recruit PMNs that transmigrate the epithelium [52,67,68]. **(9)** Transmigration reduces barrier function, which enables the passage of further bacteria, and enhances infection and tissue damage [52,67,68]. **(10)** Eventually, the infection will be cleared by the immune system, in which PMN neutrophils play an important role [52].

1.2. The architecture of the *Shigella* needle complex

The needle complex consists of the long filamentous-like needle and the membrane-spanning basal body (Fig. 1). The exterior of the basal body is composed of three proteins (in *Shigella*: MxiG, MxiJ, and MxiD) forming multiple ring-like structures [22,23,76]. Based on their native location, they are referred to as the IM ring, the OM ring, and as the connector, which is located in-between the previously mentioned structures [33].

The IM ring exhibits a 24 symmetry and is composed of the periplasmic domains of MxiJ and MxiG. While MxiJ locates within the interior of the ring, MxiG surrounds the exterior, facing the periplasm. Recently, high-resolution structures of the IM ring from *Shigella* [33] and

Salmonella [32] were solved by cryo-EM analysis of isolated needle complexes. They revealed multiple conserved ring-building motives (RBM) domains, a reoccurring motive within the T3SS [77]. Interestingly, solvent-accessible channels were found within the IM ring. They connect the periplasm with the export cage, potentially providing a pathway for proton flux for energy generation [33].

The OM ring and the connector are mainly composed of the protein MxiD [23,33]. MxiD belongs to a group of proteins termed secretins [78], which will be discussed in further the following chapter (1.2.1 The secretins). In brief, the connector is formed by the N-terminal domains, and the OM ring is formed by the C-terminal, pore-forming domain of MxiD. Intriguingly, there is a symmetry mismatch between the OM ring (cyclic (C)15) and the connector (C16) [33,79,80]. In addition to the secretin, it was suggested that its pilotin, MxiM, might be part of the OM ring [81], which will be discussed in more detail in a later chapter (1.2.2.1. The pilotin (MxiM) of *Shigella*).

The connector provides the connection between the IM and OM ring and consists of the C-termini of MxiG and the N0 domain of MxiD. The MxiG C-termini structurally vary and can be pictured as repeating groups of three [33,79]. From each of these triplets, two MxiG C-termini reach to the connector and interact with MxiD as β-sheet augmentation. This omission of every third MxiG enables interaction of the IM ring (C24) with the connector (C16), despite the different symmetries [33].

Located within these ring-like structures is the export apparatus core, which serves a gateway for substrates [27,82]. Up to date, serval export apparatus core structures have been solved: either directly within the whole, isolated needle complexes [27,79] or as heterologously expressed sub-complexes [83–85]. They are composed of multiple alpha-helical proteins arranged in a pseudo-helical manner. Most of them share a 5:4:1 (in *Shigella* SpaP:SpaQ:SpaR) stoichiometry. SpaS, an appendant subunit that is most likely lost during isolation, has only been structurally described in the flagellar homolog (FlhB) from *Vibrio* [85]. SpaS is an autoprotease which's cleavage switches the secretion hierarchy from late to middle substrates [86,87].

The complete export apparatus additionally includes the export gate, MxiA. Yet, because of its multiple transmembrane domains, high-resolution structural analysis of the interface between MxiA and the export apparatus core remain unaccomplished. However, the crystal structure of

the cytoplasmic C-terminal domain of MxiA revealed a nonameric ring assembly [88]. The same number was reported for other homologs, both isolated as well as *in situ* [89–94].

On the distal side, the export apparatus connects to the needle via the inner road protein, MxiI in *Shigella* [23,33,95]. In *Salmonella,* its orthologue PrgJ adopts a helix-turn-helix conformation similar to the needle protein [27,79]. Six of its subunits initiate the first turn of the needle structure before being continued by the needle protein itself, PrgI, in *Salmonella* [27,79,96,97]. Additionally, the inner rod protein anchors the export apparatus and needle within the basal body by a β-sheet augmentation with the secretin, InvG in *Salmonella* [27,79].

The needle is a long filamentous structure composed of a small protein, MxiH in *Shigella*, arranged in a helical fashion with ≈ 5.6 subunits per turn [23,98,99]. The needle reaches from within the basal body and protrudes several nm from the bacterial surface [76]. Thereby, the filamentous structure creates a hollow channel with a ≈ 20 Å diameter for the passage of virulent proteins [98]. Consequently, secreted proteins must be at least partially unfolded to transit through this narrow opening [26,27].

Even though the *S. flexneri* T3SS needle complex has been analyzed by single-particle cryo-EM [33], some regions remain unresolved. This comprises the OM ring, which includes the secretin and potentially the pilotin [35,81], the export apparatus, and its interface with the needle filament via the inner road protein. For further reading on the *Shigella* T3SS also see detailed review [100].

1.2.1. The secretins

Secretins are large pore-forming proteins found within the bacterial T3SS, the type II secretion system (T2SS), and the type IV pilus system (T4PS) [78,101] (Fig. 3). Additionally, secretins are present in filamentous phages where they are utilized during egress from the bacterial membranes [78,101]. Secretins assemble into homo-oligomeric, doubly-walled beta barrels within the periplasm and OM, with their pores being wide enough to allow transfer of folded proteins [102] (Fig. 3). Among the T2- and T3SS pentdadecameric assembly seems to be the most common, but other arrangements have been found as well [32,34,103–111].

Secretins can roughly be divided into the C-terminal, pore-forming secretin-domain, and multiple N-terminal periplasmic domains that can vary in number [78,101,112]. The T2-and T3SS

additionally carry a small C-terminal S domain, which acts as a binding site for a pilot protein during assembly and localization of the secretin [101,103,107,113] (Fig. 3). While the barrel-forming secretin domain is one of the most conserved regions within the protein (even among the different systems) [32,114], the S domain shares no sequence similarity among the T3SS [32,33,114].

Within the T3SS, the N0- and the N1- domain form the periplasmic 'connector', and the N3-, secretin- and S domain form the 'OM ring'. Interestingly, within the T3SS needle complex, these two regions display a symmetry mismatch [33,79,80]. While the OM ring is pentadecameric (C15), the connector displays a hexadecameric assembly (C16) which enables a symmetrical interaction between the connector and the IM ring (C24). The mismatch is probably achieved by proteolytic cleavage of a single subunit between the two structures [80]. Nevertheless, the exact mechanism causing the symmetry mismatch, as well as the time point when it occurs during assembly, remain unclear.

Up to date, serval T2SS secretin structures form multiple species were solved, including *Vibrio cholerae* (GspD) [106], *E. coli* K12 (GspD) [106], *Pseudomonas aeruginosa* (XcpQ) [105], EPEC (GspD) [104], ETEC (GspD) [107], *V. vulnificus* (EpsD) [103], *Aeromonas hydrophila* (ExeD) [103] and *Klebsiella pneumoniae* (PulD) [108]. However, within the T3SS, the only complete secretin structures solved so far originate from the same pathogen - *Salmonella* [27,34,79]. Even though my research group recently accomplished cryo-EM analysis of the *Shigella* needle complex, the generated maps just allowed *de novo* modeling of the periplasmic domains (N0, N1) of the secretin MxiD [33]. Because of the low resolution of the OM ring map, the remaining domains (N3, secretin) had to be modeled based on homology, utilizing the *Salmonella* orthologue InvG as a template [34]. Since *Salmonella* and *Shigella* share no sequence similarity within the S domain, it could not be considered during modeling; hence, its structure remained unobserved.

In *Salmonella,* the S domain adopts a helix-turn-helix conformation, mid-height on the exterior of the secretin barrel, with the helices oriented perpendicular to the membrane [32,34]. The important role of this domain was demonstrated with deletion experiments. Removing the last 20 C-terminal residues of InvG (which is within the last helix of the S domain (residues 542-562)) led to secretion-deficient phenotypes [32]. Further, truncation of the last ten residues led to the entire abortion of needle complex formation and function, and even the omission of only the last six residues already displayed impairments [80]. Most likely, this is caused by the loss of the interaction between the S domain and its pilotin, InvH, which is required for proper localization and assembly of the secretin [32,115,116]. Despite the key role the S domain

demonstrated in *Salmonella*, the sequence diversity raises the question of how this is accomplished in other species. Structural analysis of further secretins will allow a detailed comparison and help to understand the underlying mechanisms of secretin assembly and T3SS function.
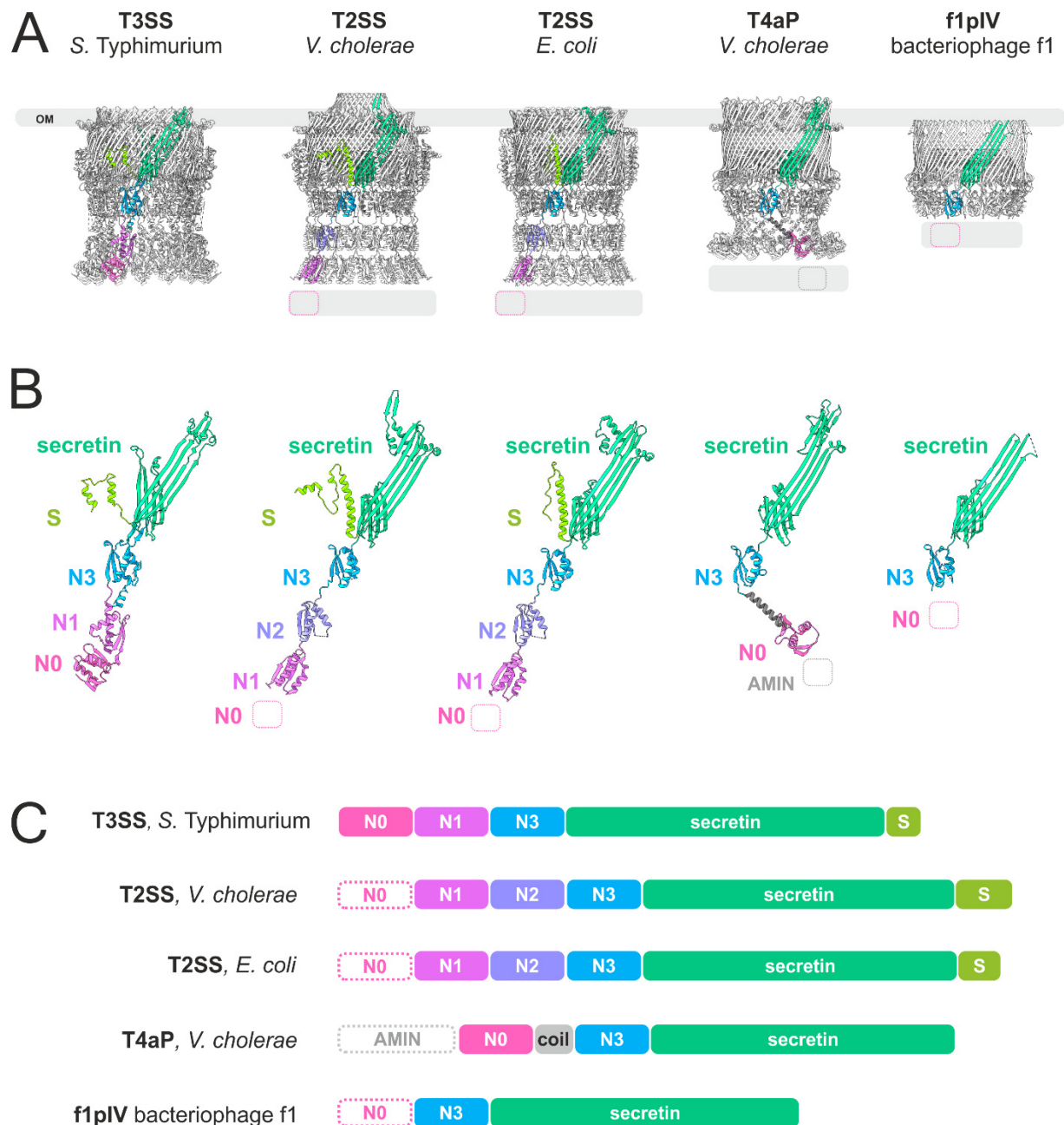
Figure 3: Structural comparison of selected secretins. (A) Atomic models in cartoon representations of the secretin pores (from left to right) from the type III secretion system (**T3SS**) of *Salmonella enterica* serovar Typhimurium (PDB ID: 6DV6 [34]), the type II secretion system (**T2SS**) of *Vibrio cholerae*

(PDB ID: 5WQ8 [106]) and *Escherichia coli* (PDB ID: 5WQ7 [106]), type IVa pilus (**T4aP**) from *V. cholerae* (PDB ID: 6W6M [117]), and pIV from the f1 filamentous bacteriophage (**f1pIV**) (PDB ID: 7OFH [118]). One subunit of each complex is highlighted in color. Domains missing in the atomic models are indicated as gray-filled boxes. The potential outer membrane (OM) area is displayed as a gray bar. **(B)** Single subunit of each secretin (order as in (A)) in cartoon representation, with each domain individually colored. N0 (pink), N1 (deep pink), N2 (lilac), N3 (blue), Secretin (green) and S domain (lime). Structurally unresolved domains are indicated as gray bar with dashed, blank boxes of indicating a single subdomain. The T4aP specific amidase N-terminal (AMIN) domain is depicted in gray. **(C)** Schematic diagram of the domain organizations of the displayed secretins. Domains are colored as in (B).

1.2.2. The pilotins

Pilotins are a diverse group of small lipoproteins which facilitate the assembly and localization of their corresponding secretins [113]. They share an N-terminal signal sequence accompanied by a conserved, lipidatable cysteine residue by which they are thought to be transported to the OM via the localization of lipoproteins (Lol) pathway [119,120]. Prior binding of the pilotin to the secretin, and subsequent co-shuttling to the OM, has been proposed as a mechanism for how pilotins aid secretin localization [120,121]. Yet, localization is not the only role of the pilotins. The absence of the pilotin can have multiple effects on secretin formation: I) either incapability of oligomerization [122] or II) inefficient/no localization to the OM [115,123] III) both [124], or IV) none of the previously mentioned effects [125].

Even though pilotins generally share the same function of aiding secretin formation of some sort, their structures vary considerably (Fig. 4). They range from super helical arrangements with several tetratricopeptide motifs [124,126], to being composed of predominantly α-helices [116,127,128], or predominately out of β-strands [129] or a mixture of both α-helices and β-strands [103,130,131].

Within the T2- and the T3SS, most pilotins interact with the last C-terminal residues (within the S domain) of their cognate secretins [101,113]. Four structures of this interaction have been solved so far. This includes AspS-GspD (T2SS of ETEC PDB ID: 5ZDH [107]), OutS-OutD (T2SS of *Dickeya dadantii* PDB ID: 4K0U [132]), InvH-InvG (T3SS of *Salmonella,* PDB ID: 6XFK [116]) and MxiM-MxiD (T3SS of *Shigella,* PDB ID: 2JW1 [121]). In all cases, the C-terminal residues of the secretin form a short, amphipathic helix across a hydrophobic area within the pilotin (colored in purple in Fig. 4). Except the first structure, all of these only include a small peptide of the secretin. In contrast, the structure of the T2SS from ETEC obtained by cryo-EM

includes the completely assembled secretin barrel [107]. It demonstrates that the secretin (GspD) and its pilotin (AspS) form a stable complex with a 15:15 stoichiometric ratio. Thereby the C-terminal helix of the secretin is oriented outward, enabling pilotin interaction. In return, the pilotin has a stabilizing effect on the secretin's S domain [107]. This is unlike the *Salmonella* (pathogenicity island 1 (SPI-1)) T3SS secretin. When fully assembled into a barrel, the C-terminal helix huddles closely against the secretin barrel [32,34], making it sterically impossible for the pilotin to remain bound in this conformation. Consistent with this are previous unsuccessful attempts to co-isolate both proteins in complex [115]. Most recently, the structure of the *Salmonella* (SPI-1) pilotin InvH was solved [116]. In the absence of its secretin, it forms α-helical homo-dimers in solution and when crystalized. In the presence of the C-terminal secretin peptide, however, the pilotin forms a complex with the secretin peptide in a 1:1 ratio (InvH in Fig. 4). For both homo-dimerization and binding of the secretin, the same hydrophobic interface is used. These findings suggest that the *Salmonella* pilotin does not persist at the secretin after assembly.

Little is known why some pilotins remain associated to their secretin (e.g., T2SS ETEC [107]), while others seemingly do not (e.g., T3SS (SPI-1) *Salmonella* [32,34,115]). The significant variation, in general, remains poorly understood. Considering the high degree of structural conservation among the secretins (Fig. 3), the structural diversity found among the pilotins is rather striking (Fig. 4). Equally versatile are the various secretin phenotypes that arise from the absence of their pilotins, as mentioned previously. Therefore, investigating pilotins in a species-specific manner is crucial in order to fully comprehend the underlying molecular mechanisms and their function beyond secretin assembly.
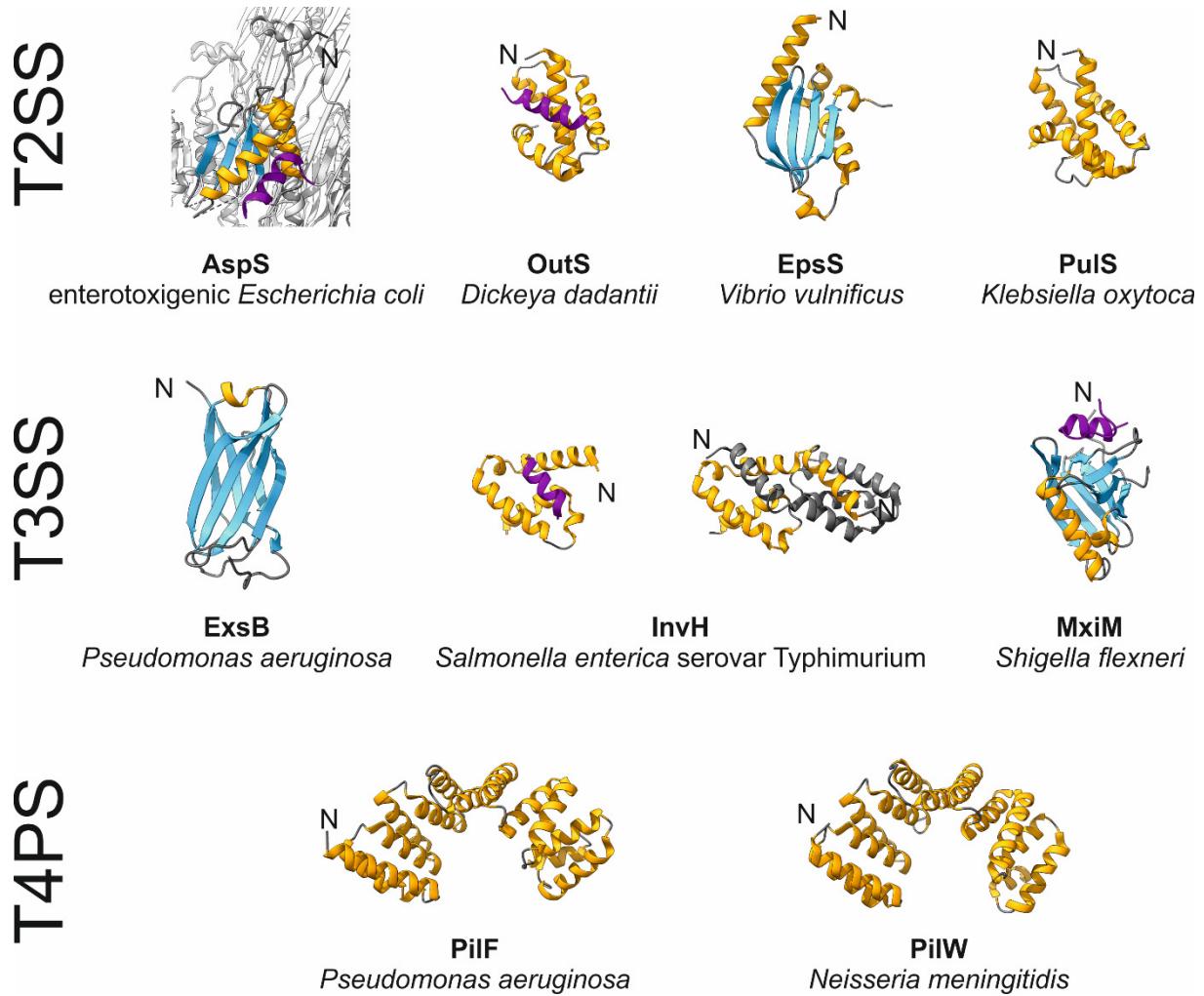
Figure 4: Structures of selected pilotins. Atomic models in cartoon representation of pilotins from the type II secretion system **(T2SS)**, type III secretion system **(T3SS)**, and the type IV pilus system **(T4PS)**. From left to right: AspS (PDB ID: 5ZDH [107]), OutS (4K0U [132]); EpsS (6I2V [103]) PulS (4A56 [127]), ExsB (2YJL [129]), InvH (6XFK & 6XFJ [116]), MxiM (2JW1), PilF (2HO1 [124]), PilW (2VQ2 [126]). Pilotin helices are colored orange, and β-sheets are colored light blue. Secretin peptides are indicated in purple. In the case of InvH, the second dimer subunit is colored in gray for clarity. N-termini of pilotins are indicated with a N.

1.2.2.1. The pilotin (MxiM) of *Shigella*

MxiM is the pilotin of *Shigella*. Its absence abolishes T3SS function and results in an inability to invade cells [133]. Therefore, MxiM is indispensably required for *Shigella* virulence and consequently displays a potential drug target. MxiM influences the stability, oligomerization, and OM association of its cognate secretin MxiD [123].

MxiM was one of the first pilotin structures to be solved (Fig. 4). Its shape was described as a 'cracked barrel' by the authors. Thereby ten β-strands form a curved β-sheet and resemble the structure of a β-barrel. Yet, this barrel is discontinued and closed on one side by a helix, giving it an overall conical shape with a deep hydrophobic cavity (PDB ID: 1Y9T) [130]. It was found that detergent and lipids could bind within the hydrophobic cavity and that a peptide from the C-terminal residues of the secretin MxiD could preclude this lipid binding [130]. Later studies showed that the MxiD peptide, which is unstructured in solution, rearranges into an amphipathic helix across the hydrophobic cavity of MxiM (Fig. 4, PDB ID: 2JW1) [121]. Thereby it obstructs lipid binding in two different ways. First, by simply blocking the access to the cavity and second, by conformational reorientation of aromatic side chains within the cavity, which would make lipid insertion unfeasible. The finding that detergent and lipids could bind within the hydrophobic cavity led to the hypothesis that MxiM could bind its own lipid moiety [121,130]. (The N-terminus, including lipidated residue, was not included in the constructs of both structural studies [121,130], most likely to enhance protein stability.) Because of the mutually exclusive fashion of binding either lipids or the secretin, it was proposed that this could be a mechanism to avoid premature transport of the pilotin to the OM before encountering it's to-be-transported secretin [121,130]. However, it cannot be ruled out that the presence of a hydrophobic molecule within the hydrophobic cavity is just a physical artifact and may not be of biological relevance.

Nevertheless, indisputable is the crucial role of MxiM, during secretin formation [123,133]. Beyond this, it was suggested that the pilotin is part of the *Shigella* needle complex since its presence could be detected by immuno-gold labeling transmission electron microscopy (TEM) of isolated particles [134]. Cryo-ET of the T3SS within *Shigella* minicells displayed an unassignable density on the exterior of the secretin right below the OM, in which the authors speculated the pilotin [35]. Additionally, Sani and collogues [81] found that MxiM co-purifies with isolated *Shigella* needle complexes. This finding, in addition to the critical role of the pilotin [133], and the similarity to the T2SS PulD-PulS complex [135], led to the assumption that spokes found around the OM ring of the isolated *Shigella* needle complexes originate from the pilotin, MxiM [81]. Generally, more research is needed to confirm the location of the pilotin and elaborate its role within the T3SS needle complex of *Shigella*

1.3. Architecture of the *Salmonella* Typhimurium (SPI-1) sorting platform

The sorting platform is a large cytoplasmic T3SS protein complex located beneath the needle complex [5]. It is highly dynamic and responsible for effector recruitment and their hierarchical translocation [28–30]. In *Salmonella* (SPI-1), it is composed of five subunits: SpaO, OrgA, OrgB, InvC and InvI [28,36]. It should be noted that through an internal translation initiation site within the *spaO* gene, two products arise from this gene: the full-length SpaO, and the shorter C-terminal variant, $SpaO_C$ [136–138].

Because it is usually lost upon needle complex isolation, the only structural information of the entire sorting platform complex originates from cryo-ET of bacterial minicells [35,36,39–41]. Thereby it assembles in a wheel-like structure with six L-shaped spokes converging at a central hub below the export gate. By adding a protein tag to individual sorting platform proteins, the emerging additional density of the tag allowed the assignment of specific densities to individual subunits [36]. It was found that SpaO is the central segment of each spoke, which is connected to the needle complex on the proximal side by OrgA, and connected to the central hub on the distal side by OrgB. The central hub is composed of the hexametric ATPase, InvC, and the stalk protein, InvI, that locates above the hub oriented towards the export gate [36].

Native mass spectrometry (MS) and small-angle X-ray scattering (SAXS) analysis of recombinantly expressed soluble sorting platform subunits gave insight into stepwise increasing sub-complexes [137]. It was suggested that one of these complexes with a $SpaO$-$2SpaO_C$-$2OrgB$-$InvC$ stoichiometry adopts an extended L-shaped conformation which most likely resembles one of the spokes obtained *in situ* by cryo-ET [36,137]. Nevertheless, it was not possible to fully reconstitute the wheal-like complex, likely due to missing interactions with other components e.g., InvI, OrgA or the needle base [137].

Together these findings revealed the overall shape and composition of the sorting platform; however, the molecular details remain concealed. High-resolution structural models are limited to small regions of SpaO and to the SpaO-OrgB interaction [139] (as well as the recently solved structure of InvC [140] from which parts of this publication will be presented here).

Because of its infeasible retrieval by either isolation or reconstitution, the sorting platform remains one of the least understood units of the T3SS, and more research is required to comprehend its molecular mechanisms.

1.3.1. The ATPase (InvC) from *Salmonella*

InvC is the ATPase of the *Salmonella* SPI-1 T3SS, and it shares sequence homology with the β-subunit of the F_0F_1 ATPases [141]. It plays a role in substrate recognition, chaperon release, and unfolding of the substrate, which requires adenosine triphosphate (ATP) hydrolysis [43,44].

The T3SS ATPases exhibit three domains: First, the N-terminal domain, that binds to the stator (OrgB in *Salmonella*) [137] and is considered critical for membrane association [142] and oligomerization [143]. Second, the core domain that displays catalytic activity and holds the Walker A motive that is a common phosphate-binding loop (GXXXXGKT/S) in ATP dependent enzymes [142,144,145]. Third, the C-terminal domain that acts as a recognition site for chaperone-effector complexes [142,146].

Even though multiple T3SS ATPase structures were solved [146–150], the *Salmonella* T3SS (SPI-1) ATPase, InvC, remained to be structurally characterized.

1.4. Aims of this study

This study aimed to better understand the infection-essential T3SS by analyzing its architecture. The focus was on the *Shigella* T3SS needle complex that can only be isolated from the pathogen itself. Additional side experiments addressed recombinantly expressed subunits from the *Salmonella* T3SS (SPI-1) sorting platform.

Single-particle cryo-EM analysis of the *Shigella* needle complex was performed by my research group [33] and could be further improved recently [151]. These newly obtained structures (in particular, the secretin, MxiD) should be analyzed and compared to existing models from *Salmonella* [34] to detect conserved and species-specific features.

To reveal the architecture of regions within the *Shigella* needle complex that could not be structurally resolved by cryo-EM, cross-linking MS should be applied. This included the isolated T3SS needle complex and *Shigella* cells *in vivo*. The generated cross-links should be compared with the newly obtained cryo-EM structure [151]. Further, by combining all available information (such as the recently generated cryo-EM models, the cross-links along with previously published findings, e.g., from X-ray crystallography or cryo-ET), additional structural insights should be obtained.

For cross-linking MS on the *Shigella* needle complex, it was required to produce adequate samples. This included adaptions of the existing isolation procedure to enable cross-linking reactions and identify the optimal protein-to-cross-linker ratio. Additionally, an *in vivo* cross-linking protocol for *Shigella* cells with subsequent needle complex isolation was developed.

Because the sorting platform cannot be isolated as a whole, its recombinantly expressed subunits are structurally analyzed individually. Colleagues from my laboratory recently determined the stoichiometry of sorting platform complexes formed by SpaO/SpaO$_C$/OrgB/InvC of the *Salmonella* (SPI-1) T3SS by native MS [137]. To obtain further details on structural arrangement, a cross-linking procedure should be established for future studies on these subunits. Further, in the scope of solving the crystal structure of the ATPase construct InvCΔ79 [140], the oligomeric states of that construct should be identified by native MS.

2. Results

2.1. Analysis of the isolated *Shigella* type III secretion system needle complex

The needle complex is one of the main components of the T3SS. Even though the *Shigella* needle complex has been previously isolated and structurally analyzed [23,26,33,76,81,134], much remains unknown. The following chapter (2.1) describes the analysis of the isolated needle complex of *S. flexneri* M90T. This includes adjustments of the isolation procedure, novel structural features within the secretin obtained by cryo-EM, and additional interactions with the pilotin acquired by cross-linking MS and an integrative modeling approach.

2.1.1. Purification adjustments for downstream analysis of the needle complex

The needle complex was purified multiple times and was further developed and applied in my laboratory [26,33,152,153]. Nevertheless, the isolation of needle complexes in satisfying quality and quantity remains to be a challenging task. Therefore, it was looked for ways to improve the purification procedure. In addition, buffer reagents of the existing protocol interfered with the planned downstream analysis. Subsequently, the isolation protocol for those experiments had to be adapted accordingly.

2.1.1.1. Dot blot for anti-strep-tag-MxiH detection

It is helpful to monitor the purification procedure when adjustments are made. This can be done by localizing the protein in the continuous purification steps by western blot. Since the T3SS needle complex consists of many proteins, monitoring only one of them is insufficient to describe the integrity of the complex. This could be overcome by utilizing two antibodies against different proteins of the needle complex. If the complex were correctly assembled, both antibodies would be detected positive for a sample.

In my laboratory, a self-made antibody against MxiG (Max-Planck Institute for Infection Biology, Protein Purification facility, Berlin) is commonly used to detect the *Shigella* needle

complex with western blotting [26,33,152]. However, when applying an anti-strep antibody against the same western blot samples, nothing could be detected. Not even in the sample that demonstrably contains the MxiH-strep protein as confirmed by negative stain TEM of the strep-affinity purified complex (data not shown). To further identify the applicability of the available strep antibody (Qiagen) for the needle complex, dot blots with different conditions were tested.

A dot blot of induced *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* cells was performed, testing different pre-treatments of the sample before loading them on the dot blot. This included extreme temperature exposure (95°C or -80 °C) and different loading buffers (sodium dodecyl sulfate (SDS), Blue Native, and Clear Native loading buffer).

As seen in figure 5, under denaturing conditions, such as heat or addition of SDS loading buffer, no signal for strep-MxiH could be detected. In contrast, untreated or only frozen at -80 °C MxiH-strep could clearly be detected. Adding loading buffers from native gel applications (blue native page or clear native page) resulted in a reduction of signal, yet could still be detected.

To exclude that the result is caused by the possibility that intact bacteria might adhere better to the membrane than denatured ones, a dot blot of the different steps during T3SS needle complex isolation was performed (Fig. 6). The dot blot shows that MxiH-strep can be detected in all the expected factions (Fig. 6, bottom). When applying the exact same protocol, which excludes heating and the addition of loading buffers to the sample, MxiG can equally be detected (Fig. 6, top).

Taken together, the dot blot protocol described here is a quick and easy method to monitor the expression and localization of the *Shigella* T3SS needle complex proteins (MxiG and MxiH-strep) during the isolation procedure.

		S. flexneri M90T ΔipaD ΔmxiH + strep-Xa-mxiH cells						Control
Temperature	95 °C	—	95 °C	—	- 80 °C	—	—	95 °C
Sample buffer	SDS	SDS	—	—	—	BN	CN	SDS
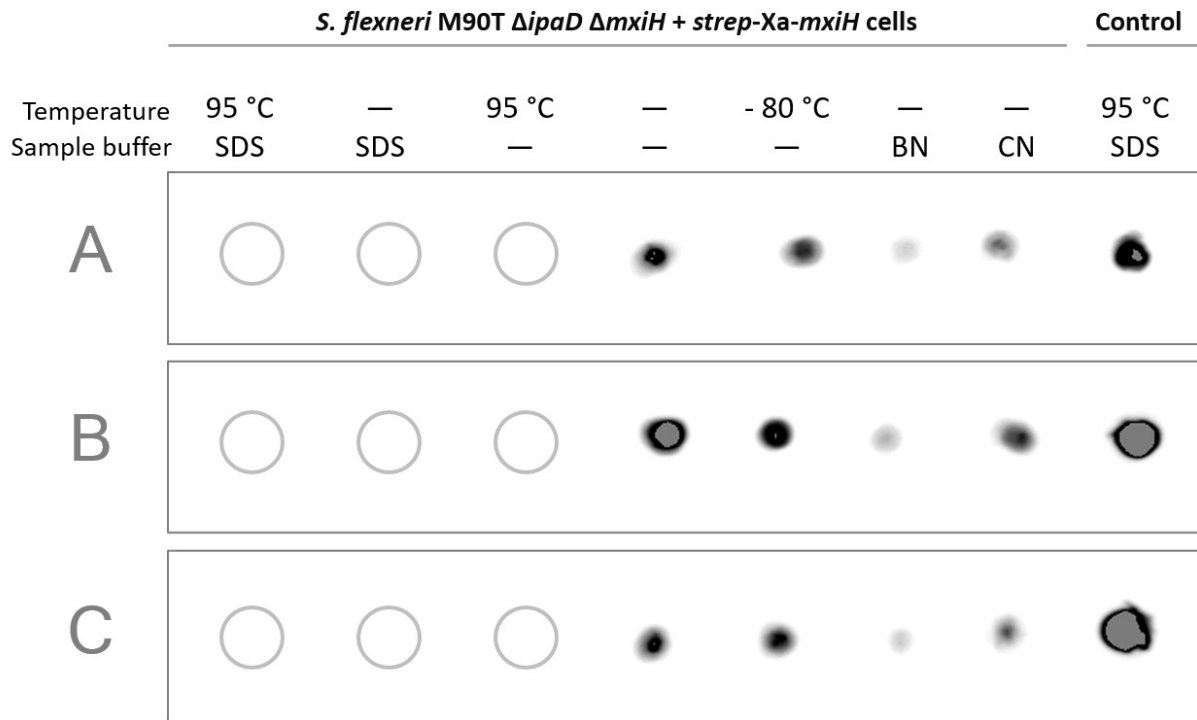
Figure 5: Anti-strep-MxiH dot blot with native and denaturing pretreatment conditions. Dot blot and subsequent anti-strep detection of *S. flexneri* M90T *ΔipaD ΔmxiH* complemented with *strep-mxiH* cells with different pretreatment conditions. This included heating **(95 °C)**, freezing **(-80 °C)**, or no extreme temperature exposure **(—)**, as well as the addition of different sample loading buffers such as **SDS**, Blue Native **(BN)**, or Clear Native **(CN)** loading buffer. Purified, strep-tagged protein served as control. Gray-outlined circles indicate the lack of signal detection. A, B and C are technical and biological replicates. The dot blots demonstrate that non-denaturing conditions enable the detection of strep-MxiH from *S. flexneri* cells.

| before induction	after induction	lysis	load on resin	flow through	elution	strep control
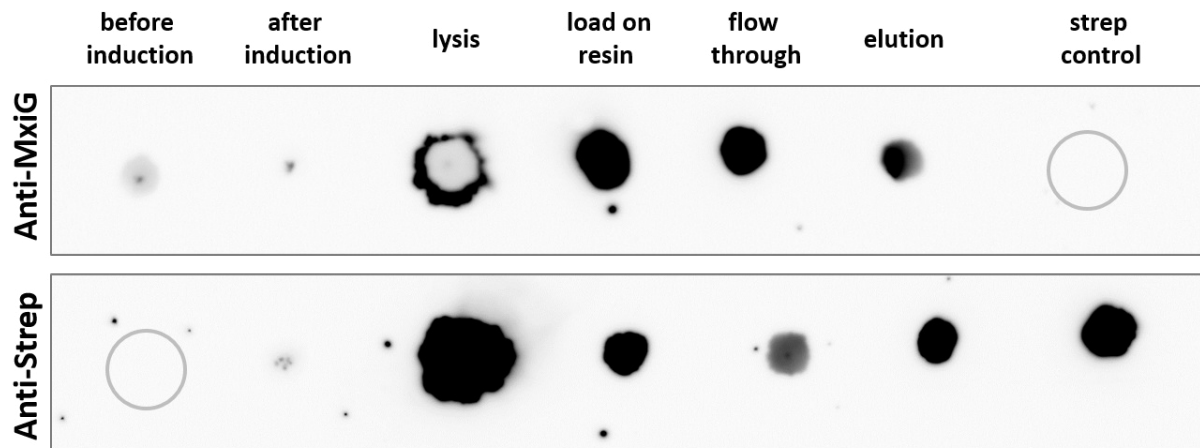
Figure 6: Anti-MxiG and anti-strep-MxiH dot plot of needle complex isolation procedure. Dot blot and subsequent anti-MxiG (top) or anti-strep-tag (bottom) detection of T3SS needle complex isolation steps from *S. flexneri* M90T *ΔipaD ΔmxiH* complemented with *strep-mxiH* with novel detergent lauryl

maltose neopentyl glycol (LMNG). Isolation steps are shown in sequential arrangement from left to right and are indicated on top. Gray-outlined circles indicate the lack of signal detection. Purified, strep-tagged protein served as control (last lane on the right). MxiG and strep-MxiH are detectable in all expected fractions, making this dot blot protocol suitable for detecting the T3SS needle complex during isolation procedure.

2.1.1.2. Cell harvest at low optical density enhances needle complex sample quality

In general, TSB is the medium of choice when culturing *Shigella* in my laboratory. However, recently, a former college could achieve a slightly higher yield of the substrate-trapped T3SS needle complexes by prior adaption of the bacteria to minimal media with increasing amounts of antibiotics over time [152]. This raised the question of whether the choice of culture medium could potentially affect the yield of isolatable needle complexes. Therefore, parallel needle complex isolations from bacteria cultured in either lysogeny broth (LB) (Luria/Miller) or tryptic soy broth (TSB) medium were compared.

While culturing *Shigella* in parallel, in either TSB or LB medium for subsequent needle complex isolation, it became apparent that the bacteria grow faster in TSB than in LB. While in LB, the OD_{600nm} was at 0.53 it was already twice the amount in TSB with OD_{600nm} at 1.07. Because of that, the cells were harvested at the same time point but displayed different OD_{600nm}. Micrographs of the resulting needle complex isolations are shown in figure 7 A, B. As can be seen in the figure, both samples contained T3SS needle complexes. However, the TSB-cultured sample (Fig. 7 B) contained significantly more background than the LB-cultured sample (Fig. 7 A). This finding raised the question of whether the high background was caused by the TSB medium or was a result of the overall higher OD_{600nm}. Therefore, the experiment was repeated, comparing both media but harvested at a similar OD_{600nm}. Consequently, for both cultures to reach a similar OD_{600nm}, the TSB-cultured *Shigella* had to be harvested earlier than the LB-grown culture (OD_{600nm} at cell harvest for TSB: ≈ 0.29 and LB: ≈ 0.38). Micrographs of the resulting isolated needle complex samples (Fig. 7 C, D) suggested similar sample quality in terms of needle complex abundance and, in particular, the amount of background. In general, this experiment shows that harvesting *Shigella* at lower OD_{600nm} affects sample quality more than the choice of the two tested media.
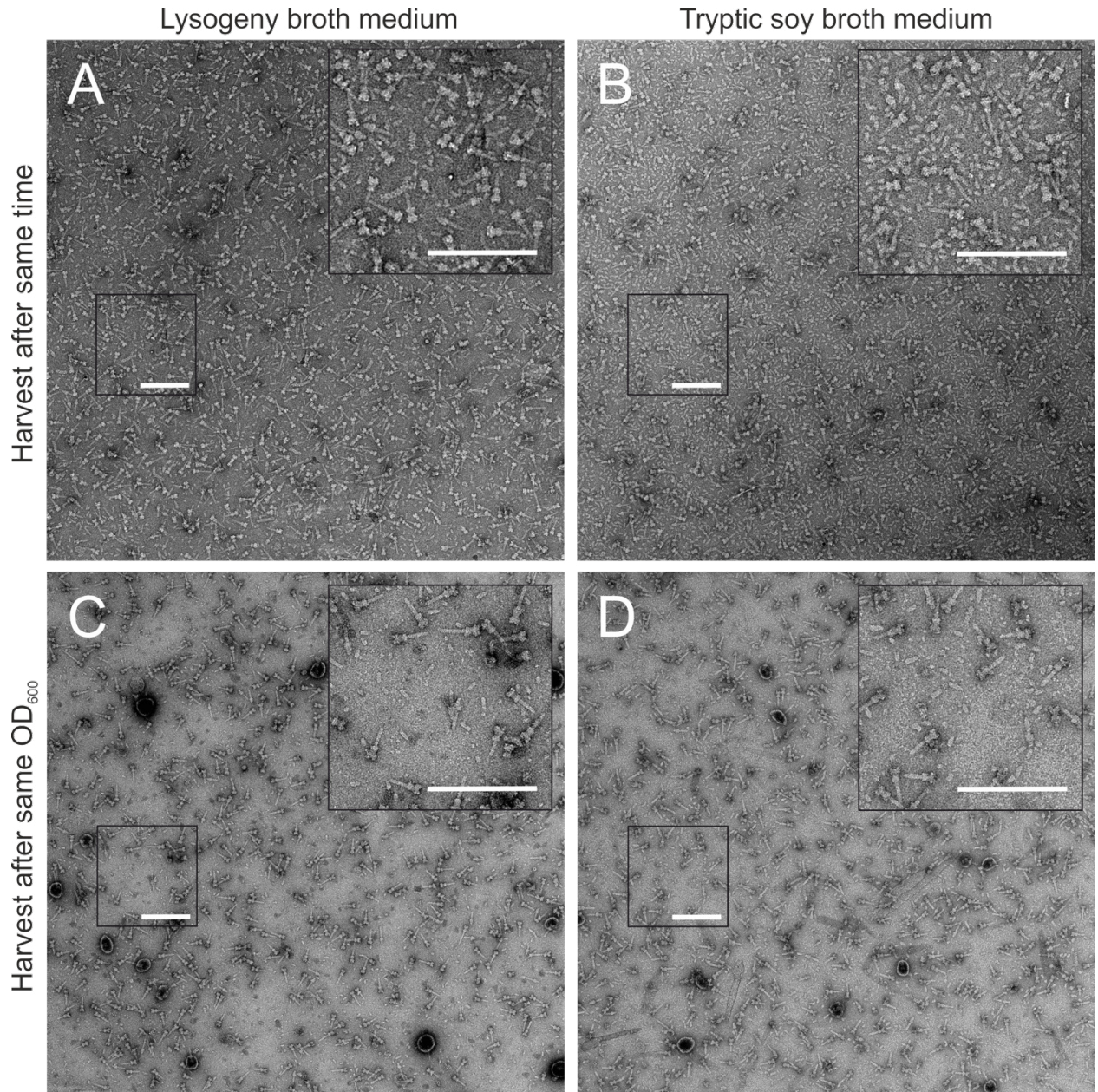
Figure 7: Isolated needle complexes from *S. flexneri* cultured in different media. 1% Uranyl acetate negatively stained TEM micrographs of T3SS needle complex isolations from S. *flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* grown either in LB medium (A, C) or TSB medium (B, D) either harvested after the same amount of time (A, B) or at the same OD_{600nm} (C, D). Isolations A + B, as well as**,** C + D were carried out in parallel. OD_{600nm} at cell harvest: A: 0.53, B: 1.07, C: 0.38 and D: 0.29. Inlet at the upper right of each micrograph displays an enlarged area of the boxed region on the left. Scale bar 500 nm. While the culture medium has little effect on sample quality, harvest at lower OD_{600nm} reduces background.

2.1.1.3. The utilized *Shigella* mutant grows faster in tryptic soy broth than in lysogeny broth medium

Experiments of the previous chapter (2.1.1.2.) should determine if the choice of growth medium could affect the sample quality of the isolated needle complexes. However, while doing these experiments, it was noticed that the *Shigella* grew faster in TSB than in LB. To further describe the growth of the utilized mutant strain *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* in detail, growth curves in two different media were determined. In addition, it was monitored whether the induction of the plasmid would affect the growth rate.

When comparing the two media, *S. flexneri* grew faster in TSB than in LB (Fig. 8, Suppl. table 1). After 5 h post-inoculation, the OD_{600nm} was around twice as high in the TSB cultured sample (1.74 ± 0.05) than the LB cultured sample (0.88 ± 0.01). In both cultures, the induction of the plasmid and subsequent production of the N-terminal strep-tagged MxiH, decelerated the growth rate considerably. Thus, after 5 h, the TSB plasmid-induced culture had an OD_{600nm} of 1.16 ± 0.02, and the LB plasmid-induced culture an OD_{600nm} of 0.71 ± 0.01.
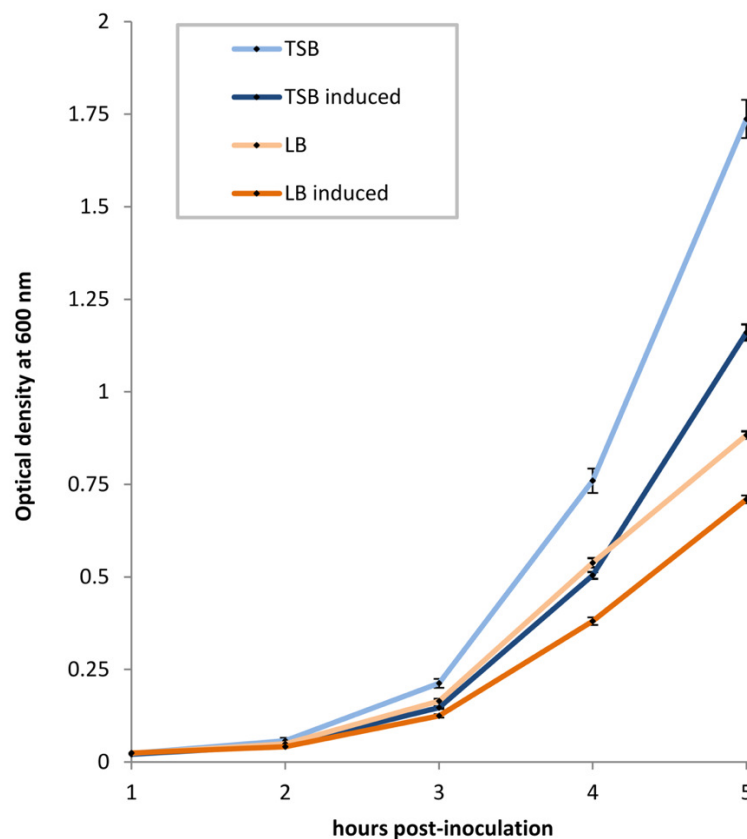
Figure 8: Growth curves of the *S. flexneri* mutant under different conditions. Hourly measured optical densities at 600 nm of *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* cultured in either

TSB or LB media under inducing or non-inducing conditions. Average values of biological quadruplicates are plotted with corresponding standard deviations indicated as error bars (also see Suppl. table 1). The bacteria multiply faster in the TSB than in the LB medium. In both media, the induction of the plasmid leads to deceleration of the growth rate.

2.1.1.4. Exchanging the detergent and buffer to Triton X-100 and HEPES

The *Shigella* T3SS needle complex isolation procedure [33] was adapted so that the buffer components meet the needs of downstream experiments.

Since preliminary attempts to measure the needle complex with native MS remained unsuccessful (data not shown), it was thought that exchanging the detergent might improve these measurements. Therefore, it was tested whether the needle complex could be isolated in the detergent Triton X-100 instead the initially utilized n-dodecyl-β-D-maltoside (DDM). Triton X-100 is usually not a detergent of choice for protein characterization, mainly because it absorbs within the UV range of proteins [154,155] and because of its heterogeneous formulation due to varying chain lengths [155,156]. Nevertheless, it can easily be dissociated from membrane proteins in the gas phase of native MS analysis[157], which is beneficial when deriving the masses of multimeric membrane proteins. Further, it is cheap, stable, well suited for lysing cells, and usually available in every molecular biology laboratory.

In addition, the buffering reagent used in the previous extraction protocol was incompatible with the downstream cross-linking procedure and therefore needed to be exchanged. The desired cross-linkers used (bis(sulfosuccinimidyl)suberate (BS3) and disuccinimidyl suberate (DSS)), predominately react on primary amines on the side chains of lysine residues and with the N-termini of the proteins [158]. The existing protocol included the buffering reagent Tris, which has primary amines and would consequently quench the cross-linking reaction. Because Tris can increase the permeability of the OM [159], the buffer reagent Tris was changed to HEPES in a later step during isolation. Increased permeability of the OM is desired in the initial step of purification during spheroplast formation. Spheroplast formation is mainly achieved by permeabilization of the OM with EDTA (ethylenediaminetetraacetic acid) so that the added lysozyme can hydrolyze the peptidoglycan layer [160]. Because spheroplasts lack their stabilizing peptidoglycan layer, they are osmotically fragile, and their appearance shifts from a rod to a spherical cell shape. This process can be monitored by light microscopy (Suppl. fig. 1).

As seen in figure 9, buffer and detergent could successfully be replaced. Further, more minor adjustments, such as culturing bacteria in LB medium and harvesting at a low optical density (see chap. 2.1.1.2-3), were also applied. For the exact protocol, see material and methods (chap. 4.1.1.1.).
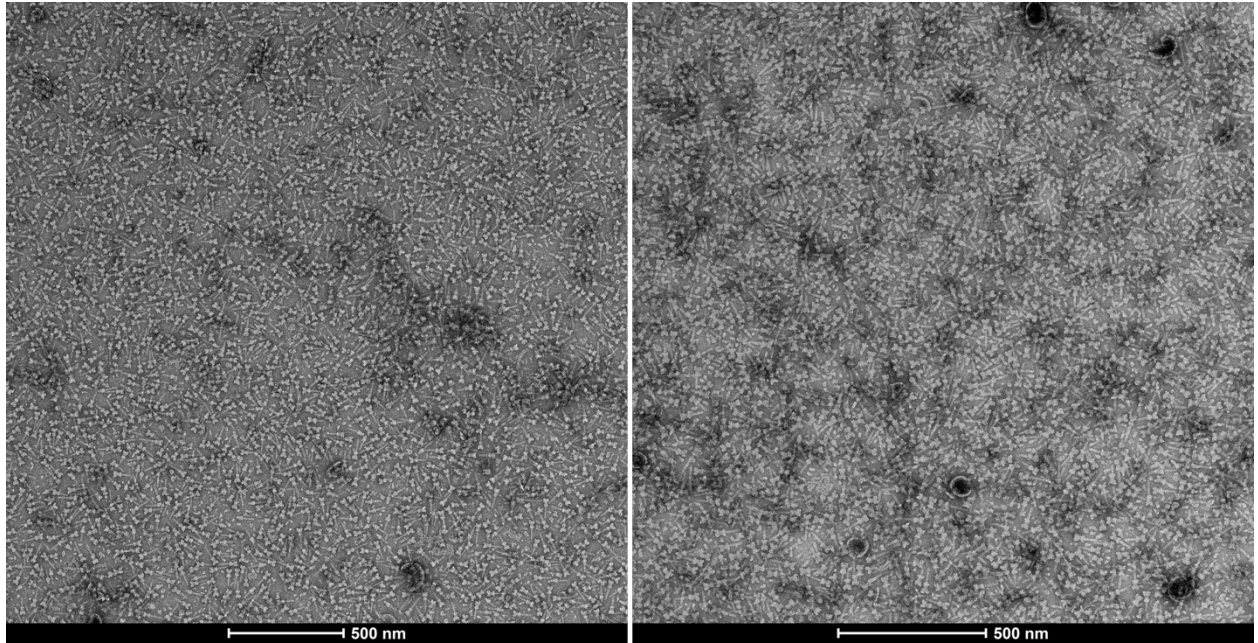

Figure 9: *S. flexneri* needle complexes obtained with adjusted isolation protocol. 1% Uranyl acetate negatively stained TEM micrographs from two individual large-scale T3SS needle complex isolations from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* in 50 mM HEPES (pH 8.0), 150 mM NaCl, 5 mM EDTA, 25 mM Biotin, 0.04% Triton X-100. The applied protocol adjustments, which included a different buffering reagent, detergent, growth medium, and a harvest at a low OD_{600nm}, successfully yielded needle complex samples suitable for downstream analysis.

2.1.1.5. Exchanging the detergent to lauryl maltose neopentyl glycol (LMNG)

The detergent LMNG (or MNG-3) consists of two hydrophilic maltose groups and two linked lipophilic chains [161]. LMNG has a low critical micelle concentration (CMC) (for comparison: LMNG: 0.01 mM or 0.001%, DDM: 0.17 mM or 0.0087% [161]) and can form rod-shaped micelles of increasing lengths in increasing concentrations [162]. Its most exciting trait, however, is its slow off-rate [163], which enables the removal of free detergent and micelles without causing immediate aggregation of the protein. An approach of removing free LMNG by gradient centrifugation was successfully developed for cryo-EM applications [164]. A different method

exploiting this trait of LMNG is mass photometry. By a 2000-fold drop dilution of a protein sample containing LMNG, into a LMNG-free buffer prior to measurements, reduced the detergent background and improved resolution [165].

Therefore, isolating the *Shigella* T3SS needle complex with detergent LMNG could be beneficial for further downstream investigations such as EM or mass analysis. For this purpose, the protocol described in the previous chapter was further adapted so that all the buffers after the ultracentrifugation step contained 2x CMC LMNG (0.02 mM) instead of Triton X-100.

Figure 10 shows a negatively stained micrograph of T3SS needle complexes isolated in LMNG, demonstrating the potential applicability of the LMNG-adjusted protocol. The corresponding dot blot of individual isolation steps can be seen in figure 6.
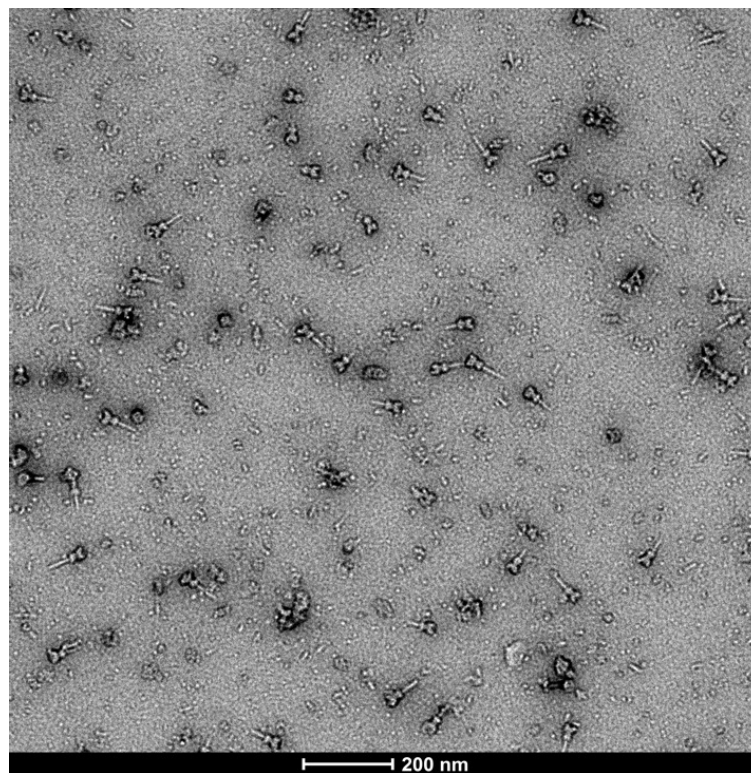
Figure 10: *S. flexneri* needle complexes isolated with LMNG. 1% Uranyl acetate negatively stained TEM micrograph of small-scale T3SS needle complex isolation from *S. flexneri* M90T *ΔipaD ΔmxiH* complemented with *strep-mxiH* in 50 mM HEPES (pH 8.0), 150 mM NaCl, 5 mM EDTA, 25 mM Biotin, 0.02 mM LMNG. As shown in the micrograph, T3SS needle complexes can be potentially isolated in the detergent LMNG with the reported protocol.

2.1.2. Mass photometry of the needle complex

Mass analysis can provide detailed information on proteins and protein complexes. It allows insights into stoichiometry, interactions with ligands, and post-translational modifications (for review on native MS, see [166]). Therefore, a procedure to derive the mass of the T3SS needle complex could be beneficial for future applications, such as e.g., identifying novel interaction partners. Since attempts to measure the needle complex with native MS remained unsuccessful (data not shown), a different, just recently available technique, termed mass photometry [167,168], was tested alternatively.

Mass photometry of the T3SS needle complex sample generated from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* in a buffer containing 2x CMC Triton X-100 displays a broad distribution of binding events within the MDa range. As expected, the detergent induced a strong background signal [165] resulting in a peak at 865 kDa (green arrow Fig. 7 and Fig. 7 inlet). Nonetheless, additional landing events could be detected in the higher mass range (Fig. 7). The most abundant species was detected at 4.6 MDa, followed by additional peaks at around 3 MDa, 6 MDa, and 8 MDa. A few landing events could be detected around 10 MDa as well.

Compared to the theoretical mass derived from the stoichiometry obtained from single-particle cryo-EM analysis (Suppl. table 2), the most abundant peak could represent the needle complex, which has a calculated theoretical mass of 4.22 MDa (Suppl. table 2). The peak around 3 MDa could display the basal body (Suppl. table 2), which lacks most parts of the needle either because it was damaged or still in the process of assembling itself [169,170]. This interpretation may also apply to all lading events between 3-5 MDa, which could display the needle complex with varying needle lengths.
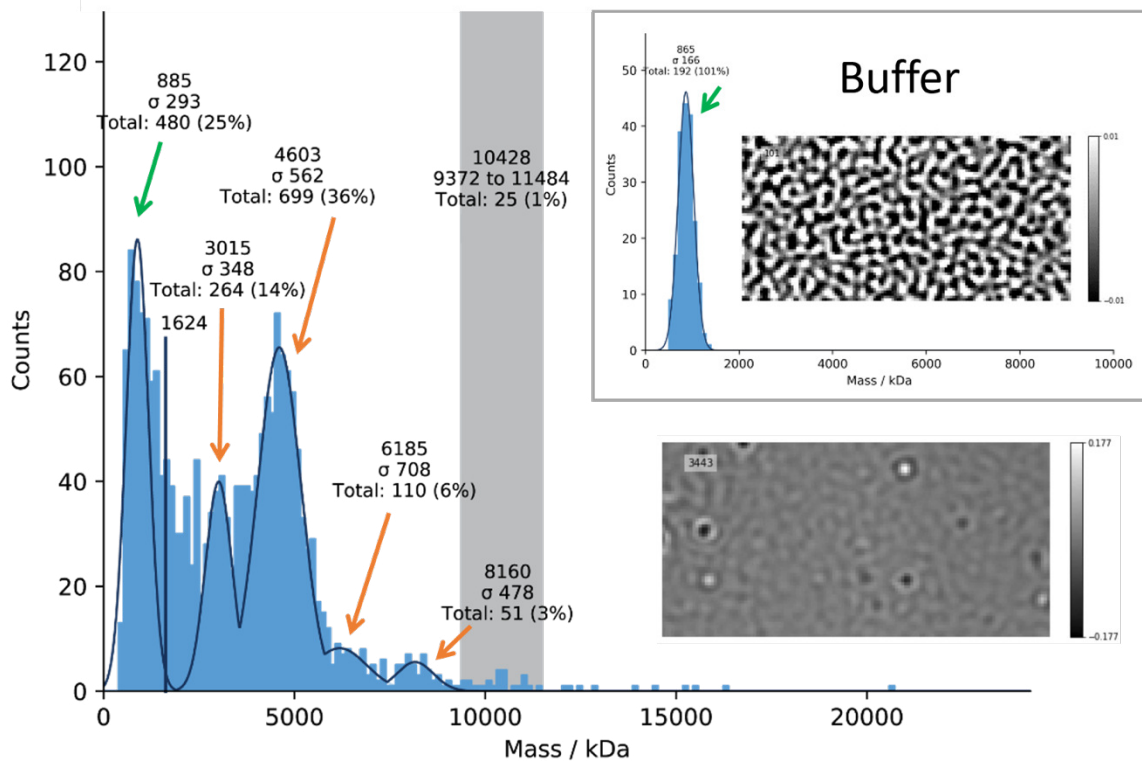
Figure 11: Mass distribution of the *S. flexneri* needle complex isolated with Triton X-100. Mass distribution histogram of the isolated T3SS needle complex sample from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* in 50 mM HEPES (pH 8.0), 150 mM NaCl, 5 mM EDTA, 25 mM Biotin, 0.04% Triton X-100. A snapshot of the corresponding mass photometry recording is displayed on the bottom right. The grayscale of the contrast was adjusted (-0.177 - 0.177) to accentuate landing events from the detergent-induced background. Inlet: mass distribution histogram and a snapshot of corresponding mass photometry recording (grayscale -0.01 - 0.01) of the buffer. Even though the detergent generates a background signal with a peak at 865 kDa (green arrow), larger molecular masses can be detected, with the 4.6 MDa species being the most abundant.

2.1.3. Analysis of the needle complex structures obtained by cryo-electron microscopy

Recently members of my laboratory solved the structure of the *Shigella* T3SS needle complex [33] by cryo-EM. However, the resolution of some areas, such as the OM ring, only allowed for homology modeling based on the *Salmonella* InvG (PDB ID: 6DV6 [34]). Now, by removing the micrograph fractions with the highest beam exposures, and subsequent reprocessing of the data, higher resolution maps could be obtained. These allow *de novo* modeling of many areas of the needle complex, including the OM ring [151].

2.1.3.1. The secretin pore complex exhibits a novel S domain

The improved *Shigella* T3SS needle complex OM ring map resolved up to 3.4 Å and allowed *de novo* atomic model building (PDB ID: 8AXL, EMD-15701 [151]). The obtained OM ring structure covers the N3 (residues K180 to H300) and the secretin (residues I301 to I520) domains, as well as the residues K521 to P548 of the previously unresolved S domain of MxiD (Fig. 12).

Apart from the OM ring, the N0 and N1 domain of MxiD (Fig. 12 E) form a second ring structure, termed the connecter, which exhibits a different symmetry (C16) than the OM ring (C15). However, this substructure and the symmetry mismatch have been characterized earlier and will not further be described here [33,79,80]. The following work will focus mainly on the novel substructure of the *Shigella* secretin, the OM ring.

Originally located within the OM, the OM ring forms a large pentadecameric (C15 symmetry) pore with a diameter of ≈ 70 Å (Fig. 12 A, B). Thereby, the secretin domain assembles into a large doubly-walled, antiparallel β-barrel with 60 strands in each wall (four stands per wall in each of the 15 subunits (Fig. 12 C)). The backbone hydrogen bonds of the β-sheets stabilize each individual wall, and the hydrophobic interactions of the side chains stabilize the two walls with one another. While the β-strands of the inner wall run parallel to the cylindrical axis, the β-strands of the outer wall adopt a conformation with a 35° angle tilt to the axis. On the distal side of the secretin, the four β-strands of every subunit in the inner wall are connected by two partially unresolved turns. Therefore, the actual inner diameter of the pore might be even narrower than the ≈ 70 Å measured from the atomic model. One of these turns (residues S392 to L401) forms a hairpin loop in *Salmonella*, which closes the secretin pore in the absence of the needle (PDB ID: 5TCQ [32]). On the outer wall, the four β-strands of each subunit create a hydrophobic L-shaped grove on the distal side of the secretin. This grove lines the exterior, distal opening of the secretin pore and would natively associate with the OM (Fig. 12 B). Therefore, it is sometimes referred to as the membrane-associated subdomain (Fig. 12 C). On the opposite side, the secretin is lined by N3 domain.

The main feature of the N3 domain is the RBM. In each subunit it is created by two α-helices packed against a three-stranded, antiparallel β-sheet (Fig. 12 A, C). This structurally conserved motive promotes ring oligomerization [77]. Further, the hydrophobicity and the large area (≈ 1040 Å^2) between the interfacing N3 domains are likely to contribute to ring stability. In addition, the N3 domain contains two inserts. The first insert (residues A196 to G209) includes a β-

hairpin, which extends distally against the inner wall of the secretin domain. Since it is not present in secretins from other systems, such as the T2SS, this insert is specific to the T3SS [32]. The second insert (residues N220 to L261) is only partially resolved and is not conserved among the T3SS [33]. Nevertheless, the resolved β-strand of this insert augments the β-sheet on the proximal side of the RBM from the neighboring subunit and adds further stability to the complex.

The C-terminal S domain (residues K521 to P548) adopts the conformation of a long loop (residues K521 to I535) followed by a small amphipathic helix (residues S536 to Q545) (Fig. 12 C). It extends laterally, mid-height on the exterior of the secretin barrel, thereby reaching over the following two subunits (Fig 12 D, E). While the loop interacts with the directly neighboring subunit (n+1) mainly via hydrogen bonds (Fig. 12 D), the amphipathic helix interacts with a hydrophobic patch (residues W306, I308, V368, L470, V512, L514) on the surface of the subunit thereafter (n+2) (Fig. 12 E). Since each S domain shares an interface area with two adjacent subunits (n+1 ≈ 880 Å², n+2 ≈ 340 Å²), this creates a network in which every single MxiD subunit is in contact with at least four neighboring subunits, which most likely contributes to the overall stability of the complex. It should be noted that the atomic model does not include the last eighteen amino acid residues of the S domain (residues S549 to Y566). These were not resolved in the EM map, which indicates flexibility of this region. In previously published nuclear magnetic resonance spectroscopy (NMR) structures, however, which include these missing residues in the form of a peptide in complex with the pilotin MxiM (PDB ID: 2JW1 [121]), they adopt a conformation of a loop and a second, small S domain helix (residues E557 to N656).

When compared to the *Salmonella* orthologue InvG (PDB ID: 6DV6 [34]), which is the only T3SS-related secretin structure solved up to date, both structures share a very similar architecture. The root-mean-square deviation of atomic positions (RMSD) of 1.2 Å for Cα atoms of the 290 aligned residues displays the high level of structural conservation between the species. As already indicated in previously published sequence alignments [32,33], the conservation is restricted to the N3 and the secretin domain of the OM ring. The S domains, however, share no sequence or structural conservation. While the *Shigella* S domain extends laterally across the secretin, the *Salmonella* S domain adopts a helix-turn-helix conformation with the helices oriented perpendicular to the membrane (Fig. 12 F). In *Salmonella*, the first helix (α1) interacts with the directly adjacent subunit (n+1), and the second helix (α2) interacts with the subunit thereafter (n+2). In *Shigella*, however, the first helix (α1) skips the first subunit

and interacts directly with the subunit after that (n+2). The C-terminal rest of the *Shigella* S domain, which contains the putative second helix (α2) [121] was not resolved in the EM map.

Taken together, the *Shigella* MxiD OM ring is a large, pentadecameric pore complex, which shares high structural similarity with its *Salmonella* orthologue. Yet, its S domain adopts a novel and unique fold.
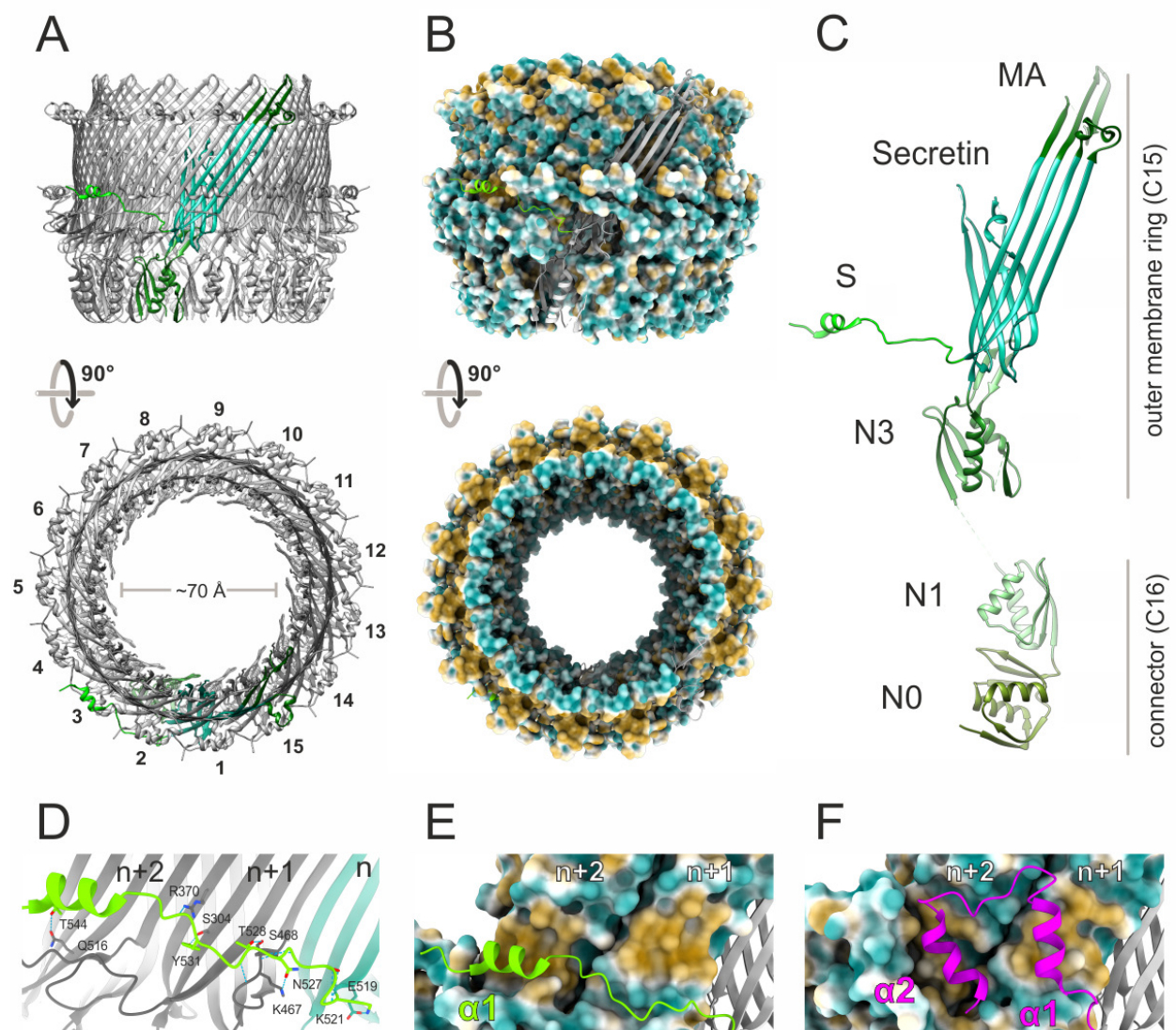
Figure 12: Structure of the secretin pore from the *S. flexneri* needle complex. (A) Side and top view of the OM ring (MxiD$_{180-548}$) atomic model (PDB ID: 8AXL, EMD-15701 [151]) in cartoon representation with a single subunit highlighted in green. Consecutively numbered subunits in the top view highlight pentadecameric (C15) assembly. **(B)** Side and top view of the OM ring (MxiD$_{180-548}$) molecular lipophilicity potential ranging from gold for hydrophobic, throughout white, to cyan for hydrophilic.

(C) A single, full-length MxiD$_{34-548}$ subunit with domains colored in different shades of green. While the domains N0 and N1 form the connecting region with a 16-fold symmetry, the N3, Secretin, membrane-associated (MA), and S domain assemble into the secretin pore complex with a 15-fold symmetry. **(D)** Hydrogen bonds between the S domain (lime green) and the two following subunits (n+1&2, gray). Dashed blue lines indicate hydrogen bonds, and involved side-chain residues are displayed as sticks. Structure of the **(E)** *Shigella* and **(F)** *Salmonella* (PDB ID: 6DV6 [34]) S domains (as a cartoon) interacting with the two following secretin subunits (n+1&2) displayed as molecular lipophilicity potential. While the first amphipathic helix (α1) interacts with the hydrophobic patch on the n+2 subunit in *Shigella*, this position, in contrast, is occupied by the second helix (α2) in *Salmonella*.

The data collection and atomic modeling of the OM ring were performed by Dr. Michele Lunelli. Parts of this chapter will be published in [151].

2.1.4. Cross-linking mass spectrometry of the isolated needle complex

Single-particle Cryo-EM has developed into one of the most powerful tools in structural biology, routinely providing near-atomic resolution structures of otherwise hard to characterize samples such as large complexes or membrane proteins [171]. Nevertheless, because the method depends on combining the information of multiple particles, structural heterogeneity within a sample can lead to lower resolution, and consequently, impede atomic modeling of those areas. Thus, dynamic interactions, transient states, and structural flexibility within protein samples can remain challenging to resolve (for review on single-particle cryo-EM challenges and opportunities, see [172]). Indeed, the cryo-EM map of the *Shigella* needle complex also exhibited some unresolved areas, such as an unidentifiable ring structure surrounding the distal side of the secretin pore complex (Suppl. fig. 2)[151]. Therefore, adding a complementary method, such as cross-linking MS, could provide additional information for these low-resolution areas. This method provides proximities of amino acid residues in proteins by chemically linking these. After the chemical reaction, protein samples are cleaved into peptides (e.g., with trypsin), then cross-linked peptides are enriched, and analyzed with MS [173]. The generated restraints can be combined with the EM data in an integrative modeling approach.

2.1.4.1. Finding the optimal protein-to-cross-linker ratio

Before MS analysis, it is required to empirically optimize the cross-linking reaction conditions for each protein sample individually. This includes determining the optimal protein-to-cross-

linker ratio to ensure sufficient, but also to avoid, over-cross-linking of the sample. Because purification of the T3SS needle complex outcome does vary, and quantity and quality of the complexes cannot be determined merely based on the protein concentration, a large amount of sample was produced by pooling to ensure consistent sample quality. Then, the protein-to-cross-linker ratio was tested on aliquots, and the desired ratio would be subsequently applied to the large-scale sample for MS analysis. An optimal ratio was determined at ≈ 0.2 mM BS3. At this concentration, both the silver-stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel (Fig. 13 A), and western blot against the needle complex IM ring protein, MxiG, (Fig. 13 B), show a shift of bands to higher molecular masses. Yet, a distribution of bands is still visible, indicating that the needle complex is not entirely cross-linked, as seen in higher BS3 concentrations. To confirm that the shift to higher molecular masses is due to desired intra-needle complex cross-links and not caused by inter-needle complex cross-link-induced aggregates, the sample was viewed by TEM. At 0.2 mM BS3, the cross-linked sample did not show enhanced clustering of needle complexes (Fig. 13 C).

Taken together, ≈ 0.2 mM BS3 was considered an optimal amount and was applied for subsequent cross-linking MS experiments of the T3SS needle complex.
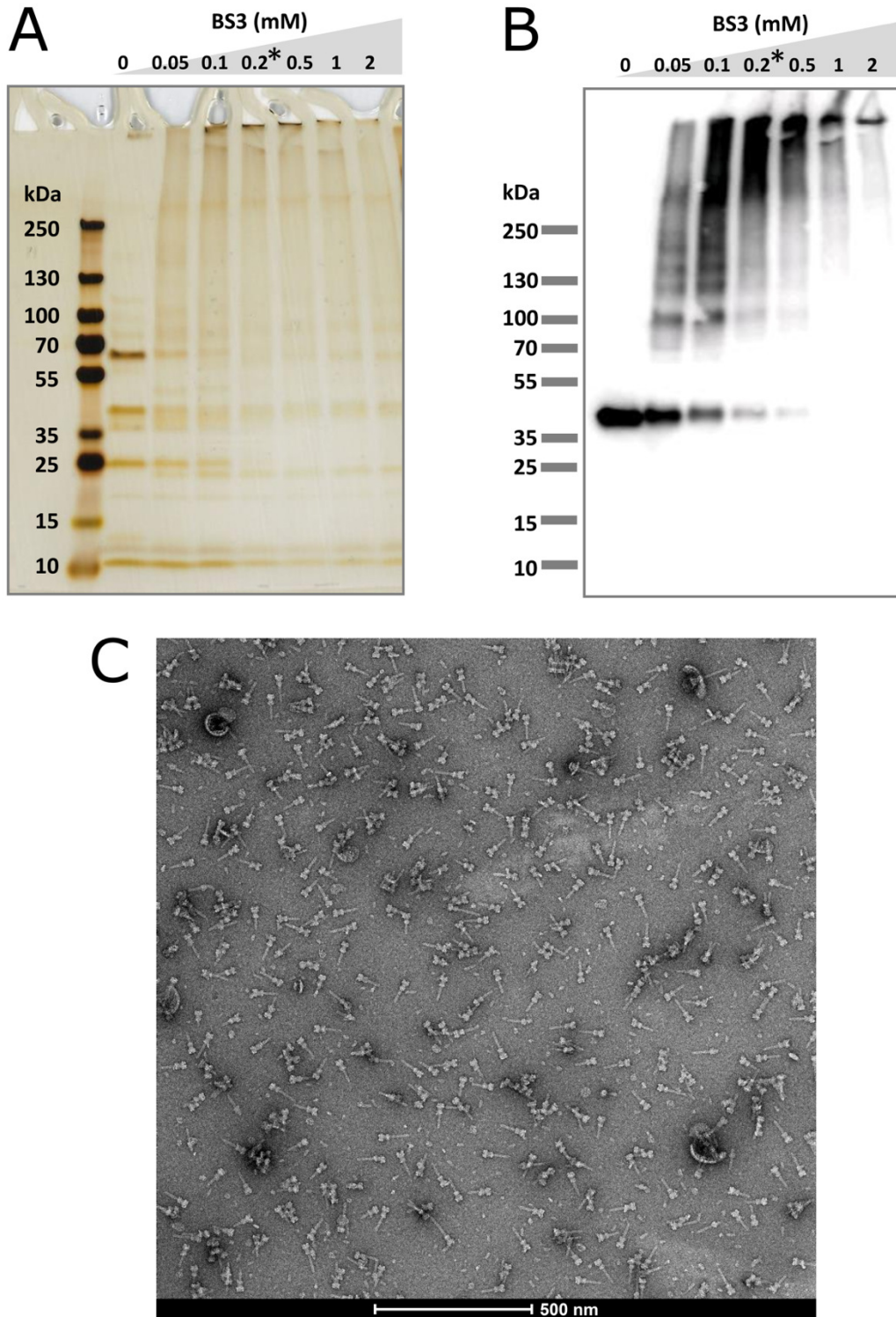
Figure 13: Sample preparation for cross-linking MS of the isolated needle complex from *S. flexneri.* Gradient (4-12%) SDS-PAGE of purified T3SS needle complexes titrated with increasing BS3 concentrations followed by **(A)** silver staining or **(B)** western blotting with anti-MxiG antibody. Upon cross-linking, individual subunit bands disappear, and higher molecular weight species become visible. Protein-to-BS3-ratio for MS analysis marked with an asterisk. Negatively stained electron micrograph of T3SS needle complexes after BS3 cross-linking procedure indicating integrity and no major aggregation of particles **(C)**.

Parts of this chapter will be published in [151].

2.1.4.2. Cross-links agree with structures obtained by cryo-electron microscopy

The needle complexes were cross-linked according to the established ratio (chap. 2.1.4.1.) and further analyzed with a cross-linking MS approach. The generated peak list was searched against 34 proteins [69], encoded on the 'entry region' of the *Shigella* virulence plasmid pWR100 (GenBank accession number: CP037924) [75]. This 31 kb region of the virulence plasmid is necessary for cell entry and includes T3SS-relevant proteins [69,72,73].

Thereby 364 cross-links at a 2% false discovery rate (FDR) at linkage level were found [174,175] (Suppl. table 3). For subsequent modeling, cross-links below the search score of 6 and cross-links originating from the protein purification tag (strep-tag) on the needle protein (MxiH) were excluded. This additional filtering provided 68 heteromeric, 24 homomultimeric, and 132 self-links. Figure 14 A displays these cross-links and reveals a dense network of heteromeric links between the needle complex subunits. This includes different regions of the needle complex, such as the IM ring proteins (MxiG and MxiJ), the OM ring (MxiD), the needle (MxiH), the inner rod protein (MxiI), and one of the export apparatus core proteins (SpaP). In addition, these proteins display multiple self-cross-links. From these, it is not possible to distinguishable whether the cross-links originate from the actual same subunit or whether they are derived from oligomers of the same protein. However, a few homomultimeric cross-links could be obtained in almost all proteins as well (Suppl. table 3). These cross-links are from the same protein, yet, because both peptides overlap, they must be from different subunits therefore indicate multimerization of these proteins within the needle complex.

From the obtained cross-links, 106 could be mapped onto the atomic model generated from cryo-EM. 90 of these cross-links were within the acceptable range of 30 Å or less (Cα- Cα) from the BS3 cross-linker, overall confirming the cryo-EM structure (Fig. 14 B). 16 cross-links exceed the accepted distance 30 Å (Fig. 14 C). These were mainly found between the connector and the OM ring of MxiD, indicating an additional, more compact conformation of this protein within the sample.

Apart from proteins present in the cryo-EM structure, the cross-links revealed additional interactions. This includes the export gate of the export apparatus (MxiA), the gate keeper (MxiC), and the pilotin (MxiM). While MxiA and MxiC only display a few links, MxiM is heavily cross-linked to the secretin MxiD and therefore is an interesting candidate for a further integrative modeling approach (see the following chap. 2.1.4.3.).

Taken together, the cross-links mostly agree with the atomic model obtained by cryo-EM, and beyond that, indicate additional, not yet resolved interactions within the T3SS needle complex.
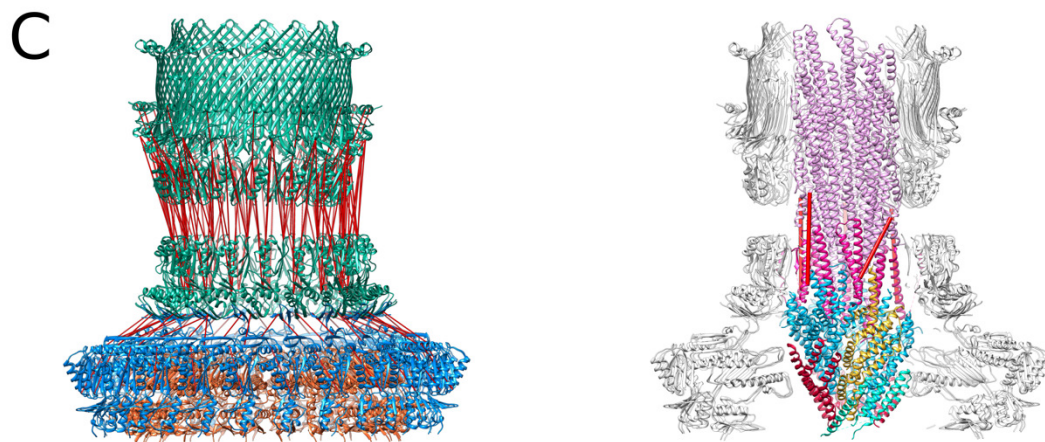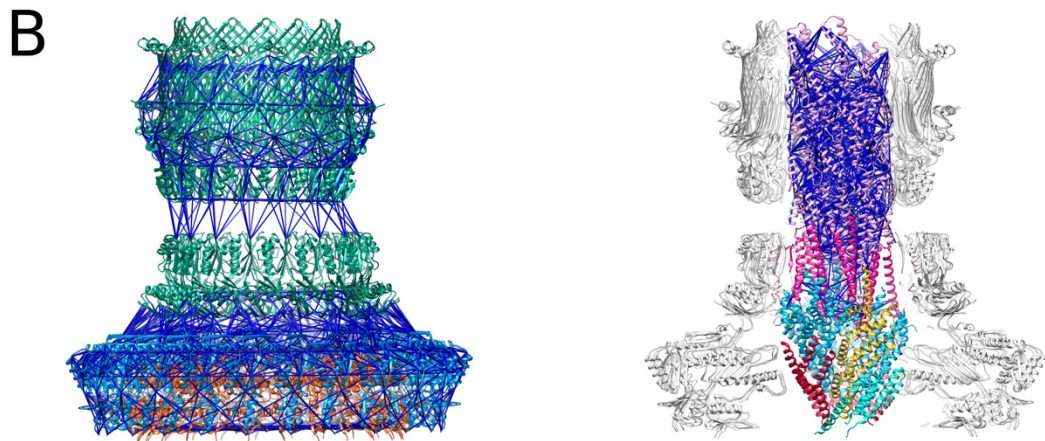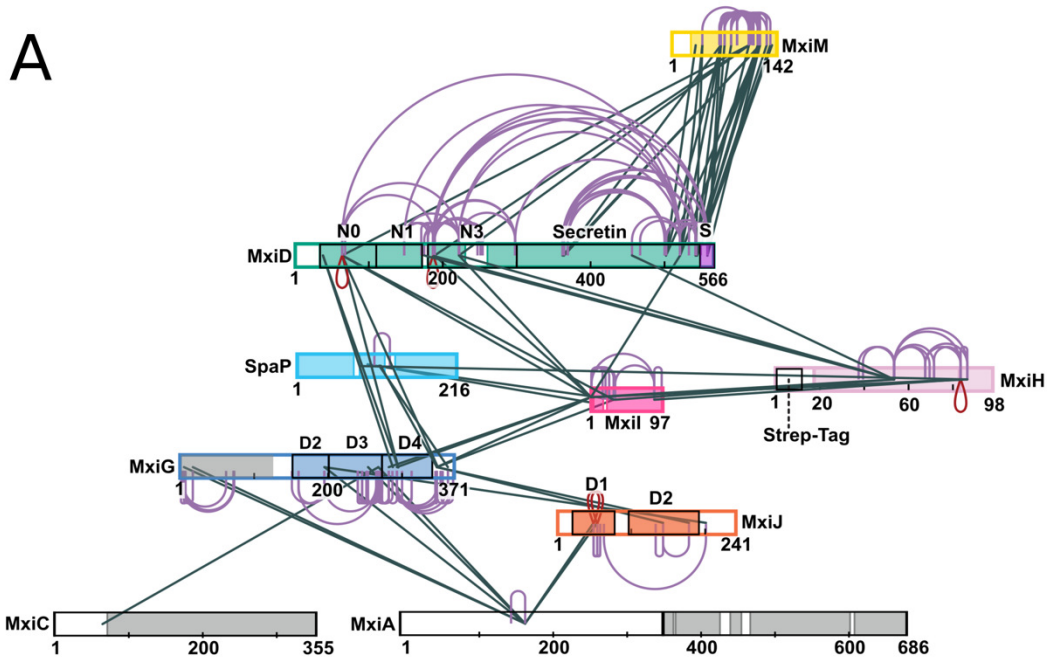
Figure 14: Cross-links of the isolated needle complex from *S. flexneri*. (A) Cross-linking network of T3SS proteins. Protein sequences are displayed as bars. Protein regions that are structurally solved are indicated by color fill. Subunits that were solved by cryo-EM of the needle complex are vividly colored (PDB ID: 8AXK [151], MxiJ orange; MxiG medium blue; MxiD green, SpaP light blue; MxiI magenta; MxiH pink). Structures available from other studies are indicated in either gray (for MxiA PDB ID: 4A5P [88], MxiC PDB ID: 2VJ5 [176], and the cytoplasmic domain of MxiG (PDB ID: 4A4Y [177])), yellow (for MxiM) or purple (S domain of MxiD) filling (PDB ID: 2JW1 [121]). Heteromeric cross-links are shown in black, self-cross-links in violet, and homomultimeric cross-links in red. Cross-links descending from the purification tag of MxiH are not shown for clarity. **(B)** Confirmatory cross-links mapped onto the basal body rings (MxiJ orange; MxiG medium blue; MxiD green) (left) and the export apparatus including the needle (SpaP light blue; SpaQ red; SpaR dark yellow; SpaS turquoise; MxiI magenta; MxiH pink) (right). Structures are shown in the cartoon representation and cross-links below the distance threshold of ≤30 Å are indicated as dark blue sticks. Cross-links exceeding the distance threshold of >30 Å are shown in **(C)** as red sticks.

The cross-linking MS measurements were performed by Dr. Zhuo Angel Chen and Dr. Francis O'Reilly. Filtering the cross-links for subsequent modeling was implemented by Dr. Karol Kaszuba. The cryo-EM model of the needle complex was provided by Dr. Michele Lunelli. Parts of this chapter will be published be published in [151].

2.1.4.3. The pilotin (MxiM) remains associated to the secretin ring (MxiD) upon isolation and displays high flexibility

Cross-linking the isolated T3SS needle complex revealed additional interactions, which are not resolved in the cryo-EM maps. One of these proteins, MxiM, was extensively cross-linked and therefore required further investigation (Fig. 14 A). MxiM belongs to the pilotins, a diverse group of lipoproteins, which stabilize and/ or localize secretins during assembly [101,113]. Consistent with that, 26 cross-links between MxiM and the *Shigella* secretin, MxiD, were found. Many pilotins bind their secretin on their C-terminal S domain. Part of this interaction has been structurally described for the *Shigella* proteins by NMR (PDB ID: 2JW1 [121]). In this structure, the C-terminal residues from MxiD (residues S549 to Y566) form a loop and a small amphipathic helix (α2, residues E557 to N565) which extends above a hydrophobic groove on the pilotin. However, in the obtained cryo-EM secretin structure (chap. 2.1.3.1.) the S domain adopts a conformation of a loop and a short helix, yet ends at P548, consequently lacking the last residues that are described in complex with the pilotin. Nevertheless, many cross-links originating from MxiM localize in a region around the S domain (Fig. 15). Interestingly, intra-

MxiD cross-links show that the last C-terminal residues (α2) react with almost the same residues as MxiM does (Fig. 15), indicating similar localization.
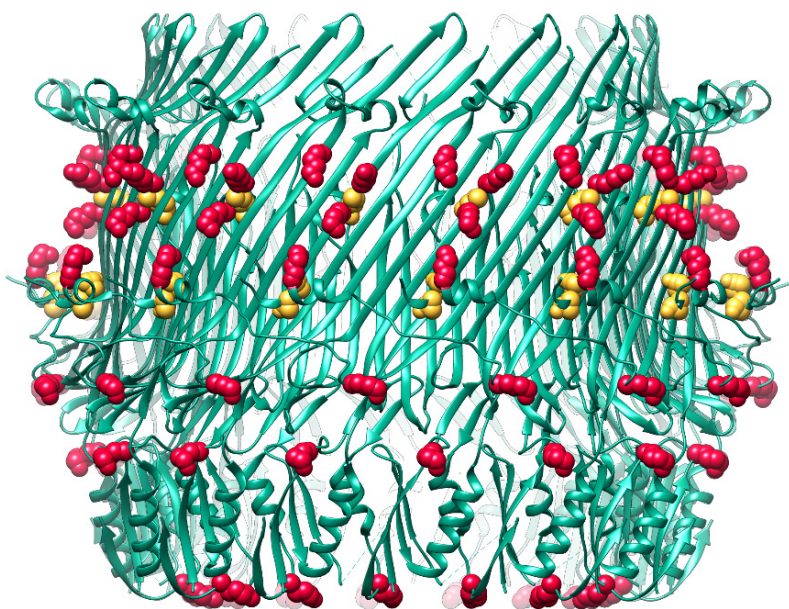
Figure 15: Distribution of cross-links of MxiM and C-terminal MxiD$_{549-566}$ across the secretin MxiD$_{180-548}$. Side view of the atomic model of MxiD$_{180-548}$ (PDB ID: 8AXL [151]) as cartoon representation (green) obtained by cryo-EM. Residues of MxiD (K189, K222, K362, K364, K502, K521, K542) cross-linking to both MxiM and the residue K558 and S549 belonging to the C-terminal MxiD (not observed in the EM model) are displayed as red spheres. Residues of MxiD (K363, T543, T544) cross-linking solely to MxiM are displayed in yellow. All residues of MxiD$_{180-548}$ that cross-link with MxiD$_{559-566}$ also cross-link with MxiM.

Parts of this chapter will be published in [151].

For further investigation, the MxiM structure, solved in complex with the C-terminal residues (S549 to Y566) (PDB ID 2JW1 [121]), was docked to the cryo-EM structure of the OM ring of the secretin (chapter 2.1.3.1.) using the Integrative Modeling Platform [178] and Assembline software [179]. (For the integrative modeling workflow, see suppl. fig. 3). Using only the cross-links, the pilotin could be located in many positions and orientations around the OM ring (Suppl. fig. 4). However, none of the orientations could satisfy all the cross-links, generally indicating high flexibility of the complex. From all 5 obtained location clusters, only one cluster (cluster 3) placed the pilotin near the assumed area near the OM (Suppl. fig. 4). Yet, in

this cluster, the lipidated N-terminus of the pilotin is not oriented toward the OM but faces the other direction instead.

In the next step, the weak unassigned density around the exterior of the OM ring (Suppl. fig. 2), in which the pilotin was assumed, was included in the modeling approach. This generated three different clusters of orientations of 15 MxiM subunits around the secretin ring. Each of these clusters satisfies ≈ 65% of all secretin-pilotin cross-links (Suppl. fig. 5, 6). Thereby orientations of MxiM vary, with the main β-sheet adopting a skewed (cluster 0), a parallel (cluster 1), or a perpendicular (cluster 2) conformation in relation to the cylindrical axis of the secretin pore. Considering the orientation of the secretin, a rearrangement of the loops E545-P548 and S549-D556 could ensure connectivity between the two secretin structures in all of the clusters. However, cluster 2, is in the best agreement with the EM map (Fig. 16). Further, this cluster enables the lipidated N-terminus of the pilotin to be oriented towards the OM (Fig. 16 A, black arrow), suggesting this orientation to be the most stable during physiological conditions.

When comparing the cryo-EM density of the isolated T3SS needle complex to the cryo-ET maps of the T3SS within *Shigella* minicells [35], the density in which MxiM was modeled into overlaps with a likewise unassigned density (EMD-2667) (Fig. 16 D). The authors of this study speculated that this density adjoining the OM might originate from the pilotin. In contrast, no such density was found in in the cryo-ET map of *Salmonella* mini cells [36], indicating another organization of that region.

Taken together, the integrative modeling approach reveals that the pilotin, MxiM, remains bound to the *Shigella* T3SS needle complex upon isolation and displays high flexibility.
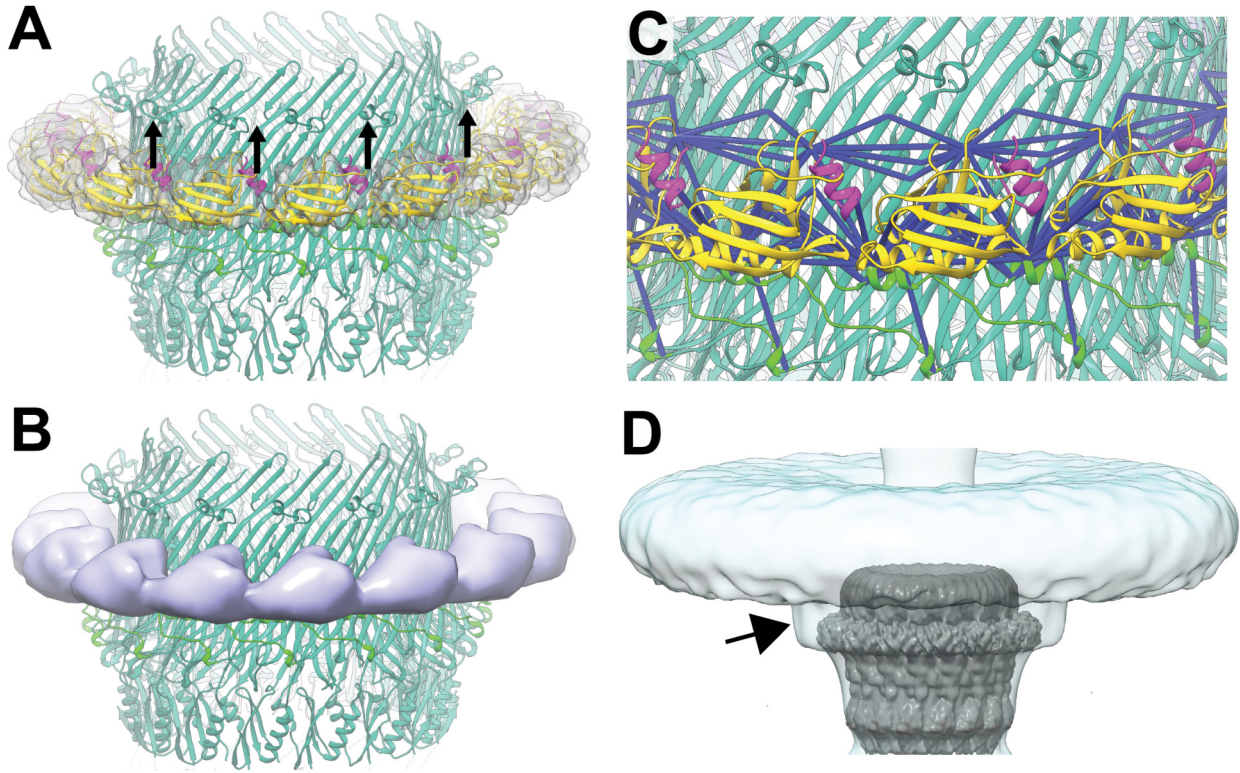
Figure 16: Structure of the isolated secretin-pilotin complex obtained by cross-links and cryo-EM. **(A)** A model of the secretin-pilotin complex (cluster 2 in suppl. fig. 5, 6) including MxiD$_{180-548}$ from cryo-EM (PDB ID: 8AXL [151]) as cartoon representation (green) with additional surrounding density (transparent gray) from our study and MxiM (yellow) in complex with MxiD$_{549-566}$ (purple, PDB ID: 2JW1 [121]) fitted into the EM density ring. Arrows indicate the orientation of the lipidated MxiM N-terminus. **(B)** The localization probability density (violet solid surface) representing the ensemble of the MxiM conformations in the cluster 2. **(C)** Cross-links satisfying the distance threshold of ≤30 Å (blue) between MxiD$_{180-548}$ and MxiM in complex with MxiD$_{549-566}$ (cluster 2). Violated cross-links to more distant regions of MxiD are not shown. **(D)** Superposition of the MxiD cryo-EM density from this study (gray solid surface) with the cryo-ET density (cyan, transparent, volume contorting threshold of 0.726) from *Shigella* T3SS in mini-cells [35] indicate similar putative MxiM location (arrow).

Integrative modeling was performed by Dr. Karol Kaszuba. Parts of this chapter will be published in [151].

2.2. Analysis of the *in vivo* cross-linked *Shigella* type III secretion system

The previous chapter focused on the analysis of the isolated needle complex. However, removing the complex from its native environment could have an impact on it structurally. Membrane-associated subunits might undergo disorder when deprived of their original

membrane surrounding. Further, components that are lost during the isolation process (such as e.g., the sorting platform) could contribute to structural changes. Therefore, analyzing the T3SS *in vivo* could provide additional information and give further insight into the physiological relevant structure. For this purpose, *Shigella* cells were cross-linked *in vivo*, followed by subsequent isolation of the needle complex. The resulting complexes were viewed with negative stain TEM and analyzed with a cross-linking MS approach.

2.2.1. *In vivo* cross-linking sample optimization

In order to compare the outcome to the isolated needle complex (chap. 2.1.4.2.), it was intended to cross-link the *Shigella* cells *in vivo* with a similar, albeit membrane-permeable cross-linker, followed by a subsequent needle complex isolation. Therefore, instead of BS3, its non-sulfonated form, disuccinimidyl suberate (DSS), was chosen. Both reagents have been successfully used in various applications and are standardly used in many laboratories [180].

2.2.1.1. Finding the optimal cell-to-cross-linker ratio

The relatively low abundance of T3SS for each cell (ranging from a few up to 100) [22,181] and its transmembrane localization led to the assumption that high amounts of cross-linker might be necessary to sufficiently cross-link the T3SS. On the contrary, exhaustive cross-linking of cells could hamper subsequent native purification of the T3SS needle complex and, therefore, should be avoided. To estimate an initial ratio, four concentrations of *Shigella* cells, each differing by the factor of ten (OD_{600nm} 0.03, 0.3, 3, 30), were tested with three final cross-linker concentrations (1 mM, 2 mM, 5 mM of DDS) that lie within the suggested range provided from the manufacturer.

As seen by western blot (Fig. 17 A), all samples treated with DSS display a shift of MxiG bands to higher molecular masses, indicating desired cross-linking. The optimal ratio for further experiments was chosen at a cell concertation OD_{600nm} 3 cross-linked with 2 mM DSS. This ratio showed a broad distribution of MxiG bands, from monomeric to higher molecular masses, at the highest concentration of cells. Further, there was no strong smear visible within the Coomassie-stained SDS-PAGE (Fig. 17B), indicating undesired extensive cross-linking.
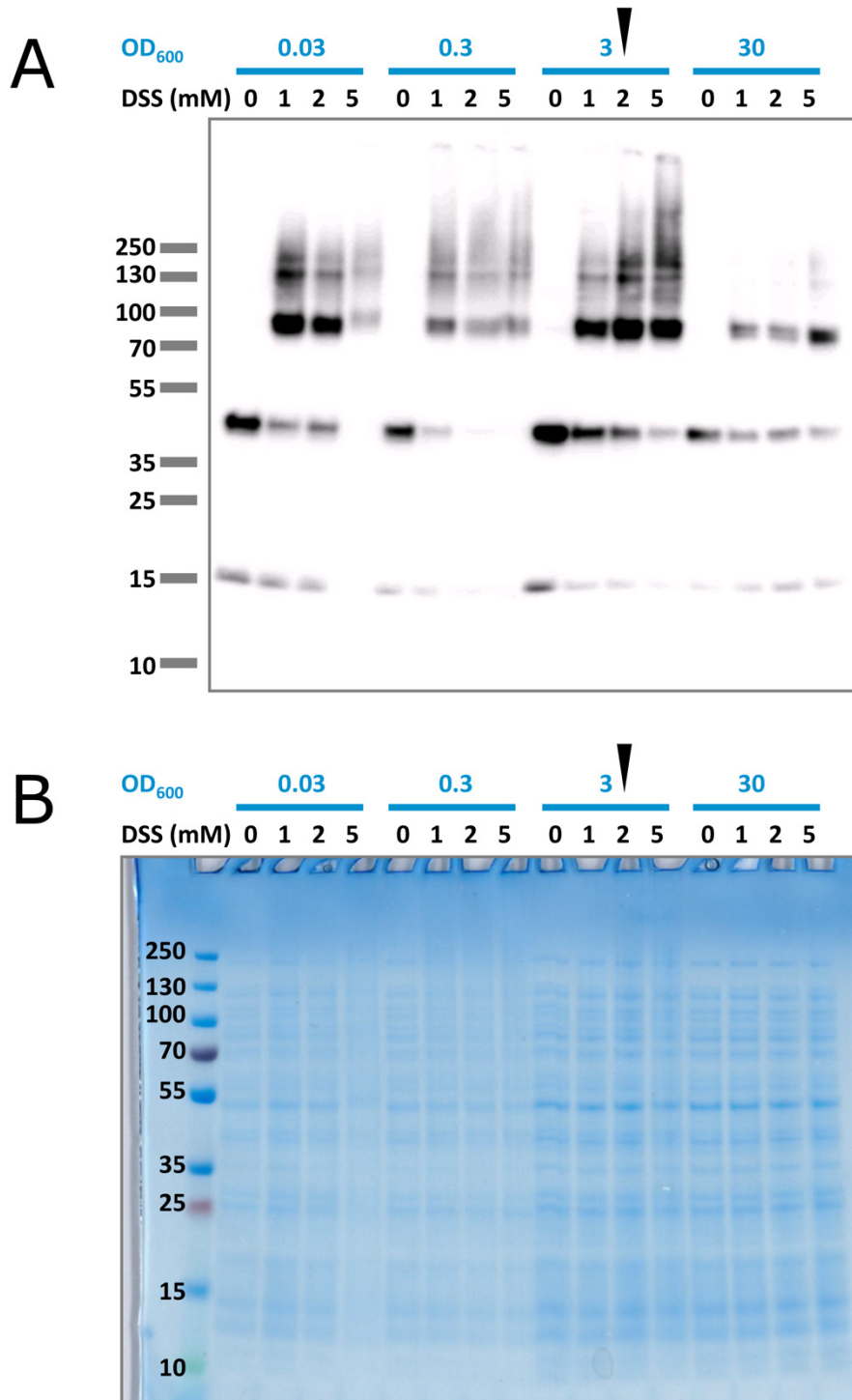
Figure 17: Finding the optimal ratio of *S. flexneri*-cells-to-DSS for *in vivo* cross-linking. Gradient (4-12%) SDS-PAGE of *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* cell lysates at differing cells concentrations (OD$_{600nm}$ 0.03, 0.3, 3, 30) after cross-linking procedure with increasing amounts of DSS (0, 1, 2 or 5 mM) followed by **(A)** a western blot with an anti-MxiG antibody or **(B)** a Coomassie staining. The optimal cell-to-cross-linker ratio was chosen at OD$_{600nm}$ 3 with 2 mM DSS and is highlighted with an arrow.

2.2.1.2. *In vivo* cross-linked, natively isolated needle complexes display additional irregular densities at the cytoplasmic side

To test whether the T3SS needle complexes could be purified after the in-cell cross-linking procedure, a small-scale purification was conducted and compared to a non-cross-linked control. Purification of *S. flexneri* M90T Δ*ipaD* Δ*mxiH strep-mxiH* needle complex was carried out as in the described protocol (chap. 2.1.1.4.), however, after harvest, cells were either cross-linked with the cells-to-cross-liker-ratio determined from the previous experiment (chap. 2.2.1.1.) or only treated with dimethyl sulfoxide (DMSO) as control. Since spheroplast formation (Suppl. fig. 1) is a critical step during needle complex isolation, cell shape was monitored with light microscopy. As seen in figure 18, desired spheroplast formation is visible in the control (Fig. 18 A), as well as in the cross-linked sample (Fig. 18 B). Thus, the cross-linking procedure at the chosen ratio does not interfere with spheroplast formation.

Additionally, the purification procedure was monitored by western blot with an antibody against the needle complex IM ring protein, MxiG (Fig. 18 C). As expected, the cross-linked sample exhibits MxiG bands with higher molecular weights (Fig. 18 C, marked as +), while the untreated sample does not show bands exceeding the molecular height of ≈ 43 kDa (Fig. 18 C, marked as -). However, both samples show similar behavior in terms of abundance in the individual purification steps, as far as this can be judged by this western blot. Nevertheless, in both purifications, elutions 2 and 3 contain MxiG (Fig 18 C).

To verify the integrity of the needle complexes, the elutions were monitored by TEM. Intact needle complexes were found in the control (Fig. 18 D) and also in the cross-linked sample (Fig. 18 E). Interestingly, some of the cross-linked needle complexes displayed an additional density at the cytoplasmic side (Fig. 18 E, indicated by arrows). By closer examination of these particles, the extra density appears to be very irregular. A selection of these particles with extra density can be viewed in figure 19.
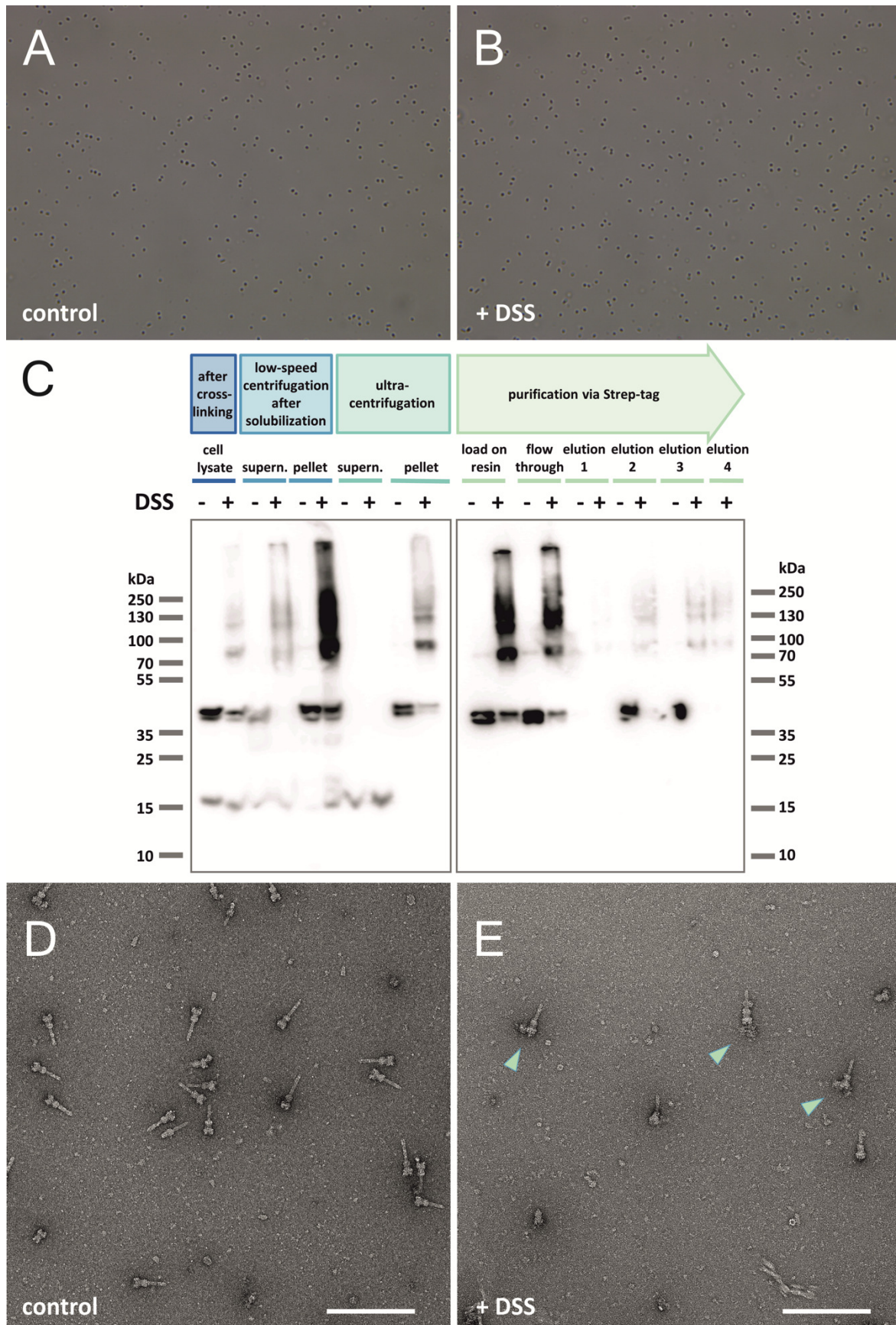
Figure 18: *S. flexneri* **M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* needle complex isolation procedure after *in vivo* cross-linking with DSS.** Light microscopy image (magnification: 400x) of *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* cells after spheroplast formation during

needle complex isolation. **(A)** The control and **(B)** previously DSS cross-linked *S. flexneri* exhibit spheroplast formation. For comparison to intact, rod-shaped bacteria, see suppl. fig. 1. **(C)** Gradient (4-12%) SDS-PAGE followed by a western blot, detecting the needle complex protein, MxiG (43 kDa). Samples monitoring isolation procedure are sequentially loaded from left to right. **(-)** indicates the control and **(+)** indicates the initially DSS cross-linked sample. A shift of MxiG to higher molecular masses is visible in DSS treated sample. In both samples, MxiG can be detected in the elution fractions. Electron micrographs of negatively stained isolated T3SS needle complexes of **(D)** the control or **(E)** the initially cross-linked *Shigella*. Some needle complexes of the cross-linked sample exhibit an additional irregular density on the cytoplasmic side (indicated by arrows). Scale bar: 200 nm.
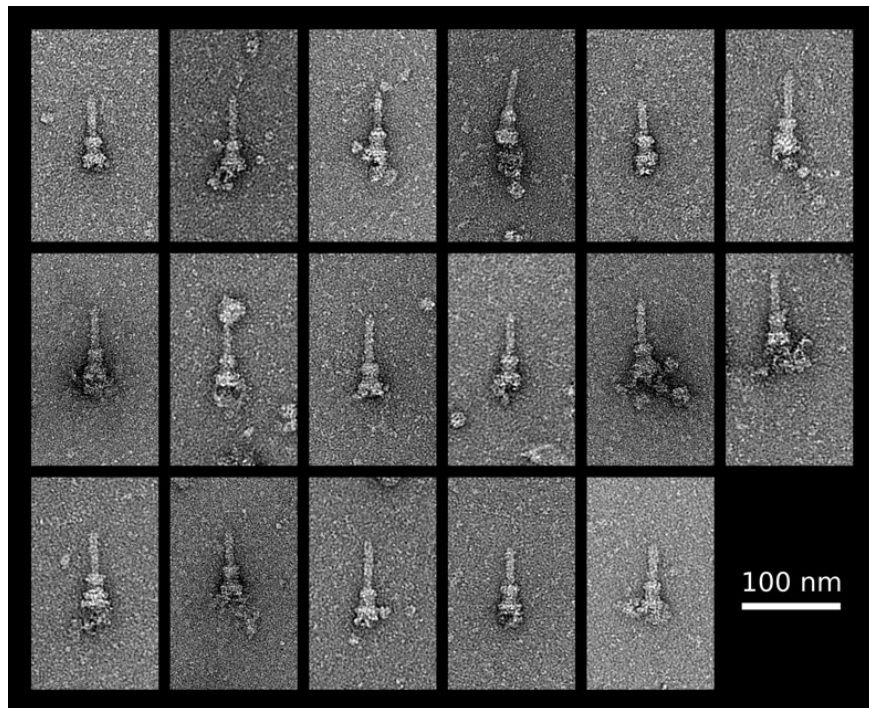
Figure 19: Selection of *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* needle complexes with extra density after *in vivo* cross-linking with DSS. TEM micrographs snippets of negatively stained isolated needle complexes that display an extra density on the cytoplasmic side. Scale bar 100 nm.

2.2.1.3. *In vivo* cross-linked, natively isolated substrate-trapped needle complexes exhibit high background with potential additional density

The previous experiment showed that it is possible to isolate needle complexes after in-cell cross-linking with DSS. Further, some of these complexes display an additional density on the cytoplasmic side. However, this density was very irregular. In order to generate a more

homogenous sample, a different needle complex construct was used. The mutant *S. flexneri* M90T Δ*ipaD ipaB*::*knot-strep* complemented with *strep-mxiH* contains a chimeric protein where an effector protein is fused to a tightly packed 'knotted' protein [26]. Since the effector proteins must be at least partially unfolded to transit through the narrow channel of the needle, this chimera effector gets stuck within the needle complex due to the not unfoldable knot and remains within the needle [26,27]. It was speculated that if the T3SS is jammed during the same time in secretion, this might lead to a more homogenous sample regarding the additional cytosolic density.

The procedure of cross-linking and subsequent needle complex isolation of the substrate-trapped mutant is based on the protocol as previously described (chap. 2.2.1.2.). Thereby, the isolation procedure was monitored by western blot detecting MxiG (43 kDa). As can be seen in figure 20 A, MxiG migrates with higher molecular masses, indicating successful cross-linking of the complex. Further, MxiG can be detected in all expected fractions during isolation. The positive signal in the elutions, especially in elution 1 and 2, suggests successful isolation of the needle complex.

Negative stain micrographs of the isolated samples exhibit some needle complexes with a reproducible high background (Fig. 20 B, C). Compared to previous isolations, the amount of needle complexes remained relatively low. The high background within the samples complicated the valuation of the needle complexes in terms of quality and integrity (Fig. 20 B, C). A selection of needle complex particles is displayed in figure 21. While some needle complexes might exhibit a putative additional density on the cytoplasmic side (Fig. 21 middle and bottom row), others may not (Fig. 21 top row). However, the high background prohibits clear statements about additional densities.
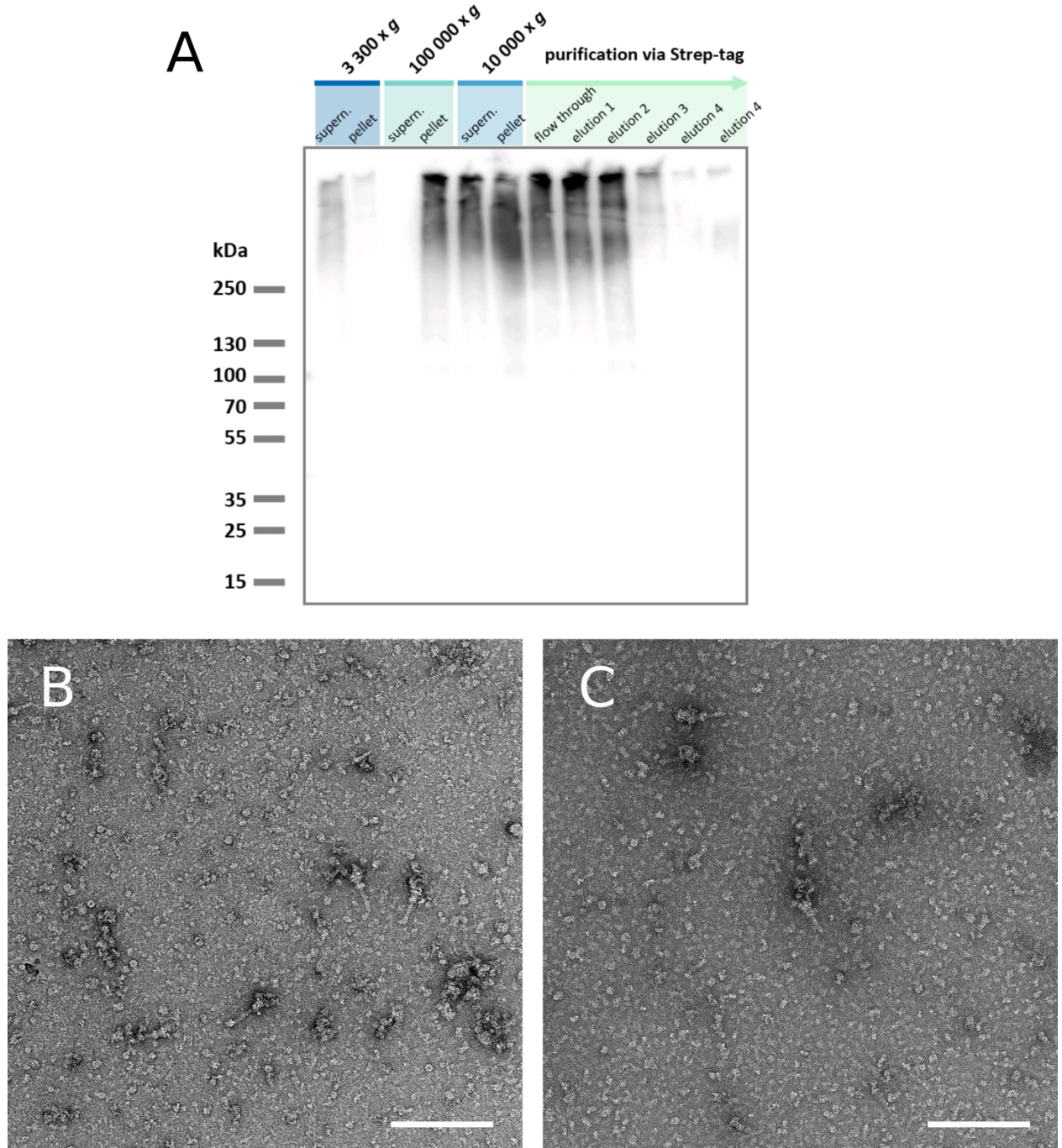
Figure 20: *S. flexneri* **M90T Δ*ipaD ipaB*::*knot-strep* complemented with *strep-mxiH* needle complex isolation procedure after *in vivo* cross-linking with DSS. (A)** Gradient (4-15%) SDS-PAGE followed by a western blot detecting MxiG (43 kDa) for monitoring the centrifugation and isolation steps. In all expected fractions, as well as the elutions, higher molecular masses of MxiG are detected. **(B, C)** Negatively stained TEM micrographs from two individual isolations after *in vivo* cross-linking. Scale bar: 200 nm. Both samples contain needle complexes and exhibit high background.
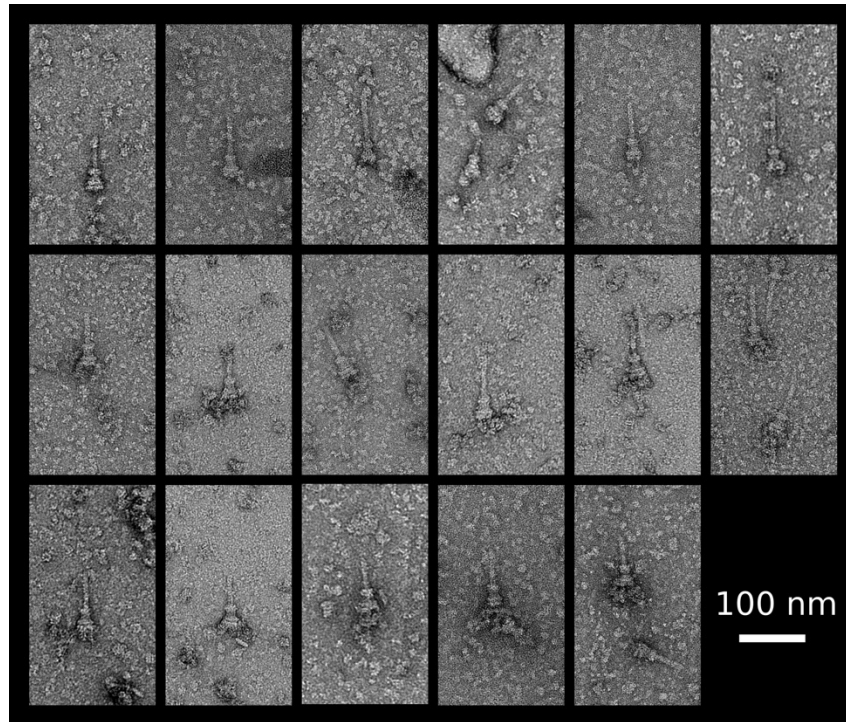
Figure 21: Selection of *S. flexneri* M90T Δ*ipaD ipaB*::*knot-strep* complemented with *strep-mxiH* needle complexes after *in vivo* cross-linking with DSS. Selected TEM micrograph sections of negatively stained isolated needle complexes. All samples exhibit high background, preventing clear statements of potential extra densities. However, some needle complexes seemingly display no or little extra density (top row) in comparison to others (middle & bottom row) which might exhibit a putative additional density on the cytoplasmic side. Scale bar 100 nm.

2.2.2. *In vivo* cross-linking mass spectrometry of the substrate-trapped type III secretion system

Cross-linking MS of the isolated needle complex confirmed the model obtained by cryo-EM, suggested an additional, more compact version of the secretin, and, most strikingly, revealed the presence of the flexibly attached pilotin (Fig. 14-16). To compare these findings to the situation within the cell, *Shigella* cells were cross-linked *in vivo*, followed by the isolation of the needle complex and cross-linking MS analysis. The substrate-trapped mutant strain was chosen because of the potential interesting additional interactions arising from the background and the putative extra density on the cytoplasmic side (Fig. 20-21).

2.2.2.2. *In vivo* cross-links agree with structures obtained by cryo-electron microscopy

The *in vivo* DSS cross-linked and subsequently isolated needle complex was further analyzed by MS. As in the previous experiment (chap. 2.1.4.2.), the generated peak list was searched against the 34 target proteins encoded by the 'entry region' of the *Shigella* virulence plasmid [69,75]. Thereby 345 cross-links were identified at a 3% FDR at linkage level [174,175] (Suppl. table 4). The network includes proteins from the needle complex (MxiD, MxiG, MxiJ, MxiI, MxiH-strep, SpaP, MxiM), sorting platform proteins (Spa33, Spa47, MxiN, Spa13, MxiK), effectors, chaperons and other cytosolic components (IpaB-Knot, IpaD, IpaA, IcsB, IpgD, IpgC, Ipg E, IpgF, VirB, Spa32, MxiE) (Fig. 22 A). Yet, many of these connections are solely based on one or two cross-links. Proteins with multiple heteromeric-cross links, however, were either associated with the strep-tagged substrate protein IpaB-knot linked with IpgC (35 cross-links), Spa33 (3 cross-links), MxiD (3 cross-links), MxiH (3 cross-links) or were found within proteins of the needle complex such as between MxiM-MxiD (14 cross-links), MxiG-MxiJ (3 cross-links), MxiD-MxiH (9 cross-links). From these cross-links, 40 could be plotted on the T3SS needle complex structure obtained by cryo-EM, with 38 cross-links satisfying the distance restraint below ≤30 Å (Fig. 22 B, C). The only two cross-links exceeding this threshold were found in MxiD and involved K222 (Fig. 22 B). It must be emphasized that IpaD was found among the cross-linked proteins even though its gene should have been deleted in the utilized mutant strain M90T Δ*ipaD ipaB*::*knot-strep* complemented with *strep-mxiH.*

When compared with the cross-linking approach from the isolated needle complex (chap. 2.1.4.2.), both experiments generated a similar number of overall cross-links (isolated: 364 (2% FDR), *in vivo*: 345 (3% FDR)). However, the *in vivo* approach generated more proteins within its cross-linking network, in return the number of cross-links between these proteins was lower (Fig. 14 A, Fig. 22 A). When plotting cross-links on the needle complex structure obtained by cryo-EM, the isolated sample could map 106 cross-links (includes additional filtering) (Fig. 14 B, C), while the *in vivo* sample could only map 40 cross-links (Fig. 22 B, C). From these, the percentage of satisfied distance restraints is higher in the *in vivo* cross-linked sample (satisfied (≤30 Å): 38 (95%), violated (>30Å): 2 (5%)) than in the isolated sample ((satisfied (≤30 Å): 90 (84.9%), violated (>30Å): 16 (15.1%)). The only distance-violating cross-links within the *in vivo* cross-linked sample involve the residue K222 within MxiD, which is also involved in violations within the isolated sample. Both samples exhibit extensive cross-links between the

secretin (MxiD) and the pilotin (MxiM), which will be further compared in the following chapter (2.2.2.3.).
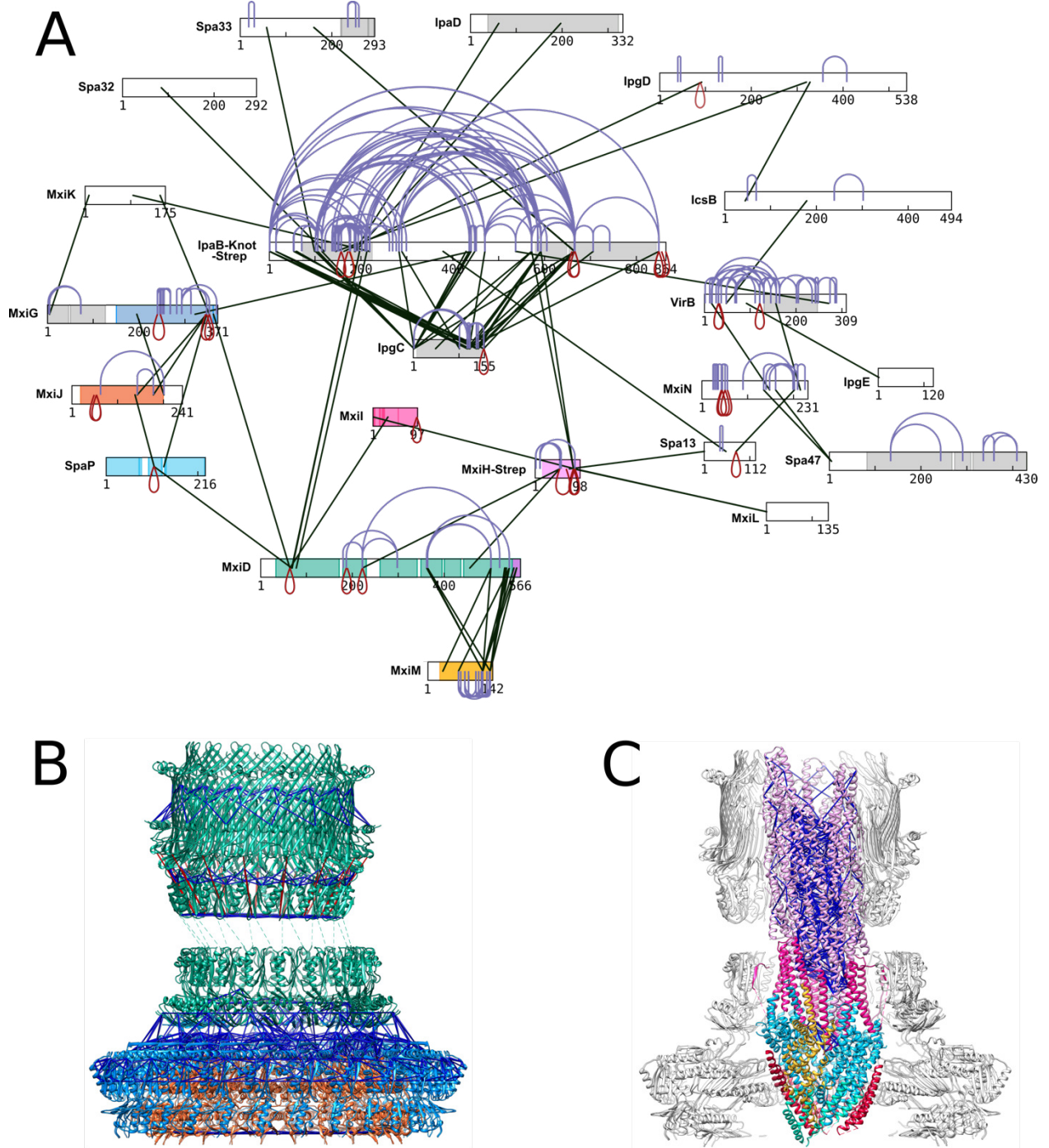
Figure 22: *In vivo* **cross-links of the T3SS from** *S. flexneri*. **(A)** Network of T3SS proteins cross-linked *in vivo*. Protein sequences are displayed as bars, and heteromeric cross-links are shown as black, self-

cross-links as violet, and homomultimeric cross-links as red lines. Regions of proteins that are structurally solved are indicated by color fill. Subunits of the needle complex that were solved by cryo-EM (PDB ID: 8AXK [151]) are vividly colored (MxiJ orange; MxiG medium blue; MxiD green, SpaP light blue; MxiI magenta; MxiH pink). Structures available from previous studies are indicated in either gray (cytoplasmic domain of MxiG (PDB ID: 4A4Y [177]), IpaB (PDB ID: 3U0C [182]), rrmA ('knot') (PDB ID: 1IPA [183]), IpgC (PDB ID: 3GYZ [184]), IpaD (PDB ID: 2J0O [185]), Spa33 (PDB ID: 4TT9 [186]), Spa47 (PDB ID: 5SW1 [147]), VirB (PDB ID: 3W3C [187]) or yellow (for MxiM) and purple (S domain of MxiD) (PDB ID: 2JW1 [121]) filling. Proteins not involved in the network and cross-links originating from the MxiH strep-tag, were removed for clarity. Cross-links mapped onto the atomic models obtained by cryo-EM of the **(B)** basal body rings (MxiJ orange; MxiG medium blue; MxiD green) and **(C)** export apparatus including the needle (SpaP light blue; SpaQ red; SpaR dark yellow; SpaS turquoise; MxiI magenta; MxiH pink). Structures are shown in the cartoon representation. Cross-links below the distance threshold of ≤30 Å are shown as dark blue sticks, and cross-links exceeding this threshold (>30Å) are shown as red sticks.

The cross-linking MS measurement was performed by Dr. Zhuo Angel Chen. The cryo-EM model of the needle complex was provided by Dr. Michele Lunelli.

2.2.2.3. *In vivo* cross-links suggest a more confined localization of the pilotin compared to the isolated needle complex

Analysis of cross-links from the isolated T3SS needle complex revealed that the pilotin, MxiM, remains bound even after isolation (chap. 2.1.4.2, Fig. 14 A). In addition, the broad distribution of cross-links from MxiM on the secretin ring suggest a wide localization range of the pilotin (Fig. 15). In docking experiments, no position of the pilotin could satisfy all cross-links, suggesting multiple poses or flexibility (Suppl. fig. 4). Because the pilotin possess an N-terminal lipid anchor [133], it was concluded that this flexibility might be caused by the isolation of the complex from the membrane [151]. If this is the case, cross-links indicating flexibility should not be present when the cross-linking procedure is conducted within the native environment of the complex. Therefore, the cross-links of the isolated complex were compared to cross-links that were generated *in vivo*.

As previously described (chap. 2.2.2.2.), more cross-links involving needle complex proteins were found in the isolated sample. This finding also applies to the cross-links involving pilotin,

MxiM (Fig. 23 A). From the 39 cross-links found *in vivo* within MxiM and MxiD (including hetero and self-cross-links) (Fig. 23A right), 28 were also present in the isolated sample (Fig. 23 A left). However, from these 11 absent cross-links, 9 of them contain K137 of MxiM.

For further comparison, cross-links were plotted on the NMR structure (PDB ID: 2JW1 [121]), which includes MxiM and the last few C-terminal residues of MxiD, previously termed MxiD-α2 (Fig. 23 B). Both samples exhibit similar cross-links, yet the isolated sample exhibits slightly more (Fig. 23 B left). From 34 cross-links, 31 were within the range below 30 Å, and 3 exceeded this limit. In the sample cross-linked *in vivo*, 18 cross-links were found, from which 17 are within the accepted range, and only one cross-link reaches beyond that (Fig. 23 B right). Noticeably, all length-exceeding cross-links in both samples originate from the same residue on MxiM (K121).

When visualizing residues on the secretin structure involved in the cross-links with the pilotin, the isolated sample displays a broad distribution on the OM ring (Fig. 23 C left). Further, residues linking to the pilotin also largely cross-link with MxiD-α2 (the last few residues of MxiD that are missing in the cryo-EM structure and are most likely in complex with the pilotin). In contrast, in the sample cross-linked *in vivo*, residues cross-linking to the pilotin are only found in the area between the OM and the S domain (Fig. 23 C right). Even though cross-links between MxiD-α2 and the pilotin were found, none were found between MxiD-α2 and the rest of MxiD.

Finally, the potential locations of the pilotin between the two samples were compared. In the isolated sample, previous docking experiments suggested multiple, differing locations of the pilotin on the secretin ring (Suppl. fig. 4). In the *in vivo* cross-linked sample, however, the pilotin could be placed manually around the OM ring as a rigid body in such manner, that the distance restraints of all found cross-links could be satisfied (Fig. 24). Thereby MxiM locates above the S domain on the distal half of the OM ring, from the cryo-EM model. Its α-helix is oriented parallel to the membrane and is facing outward, away from the secretin ring. In this orientation, the lipidated N-terminus of the pilotin is directed toward the OM. In addition, the secretin helix, α2, bound to the pilotin, is not far from the last few modeled residues of the S domain in the cryo-EM model. Overall, 13 cross-links were found between the pilotin and the secretin, of which 3 are within α2 of MxiD. 15 self-links were found within the pilotin of which one violated the distance restraint when plotted on the NMR structure (Fig. 23 B). Noteworthy,

in the pilotin docking around the secretin, this violated restraint (K121 to K68) could be satisfied by interacting with an adjacent pilotin.

Taken together, the cross-linking data demonstrates that the pilotin locates near the secretin in both samples. While in the isolated sample, the cross-links are widely distributed and indicate flexibility of the pilotin and MxiD-α2, the location from the *in vivo* cross-linked sample is narrower and limited to the distal upper half of the OM ring.
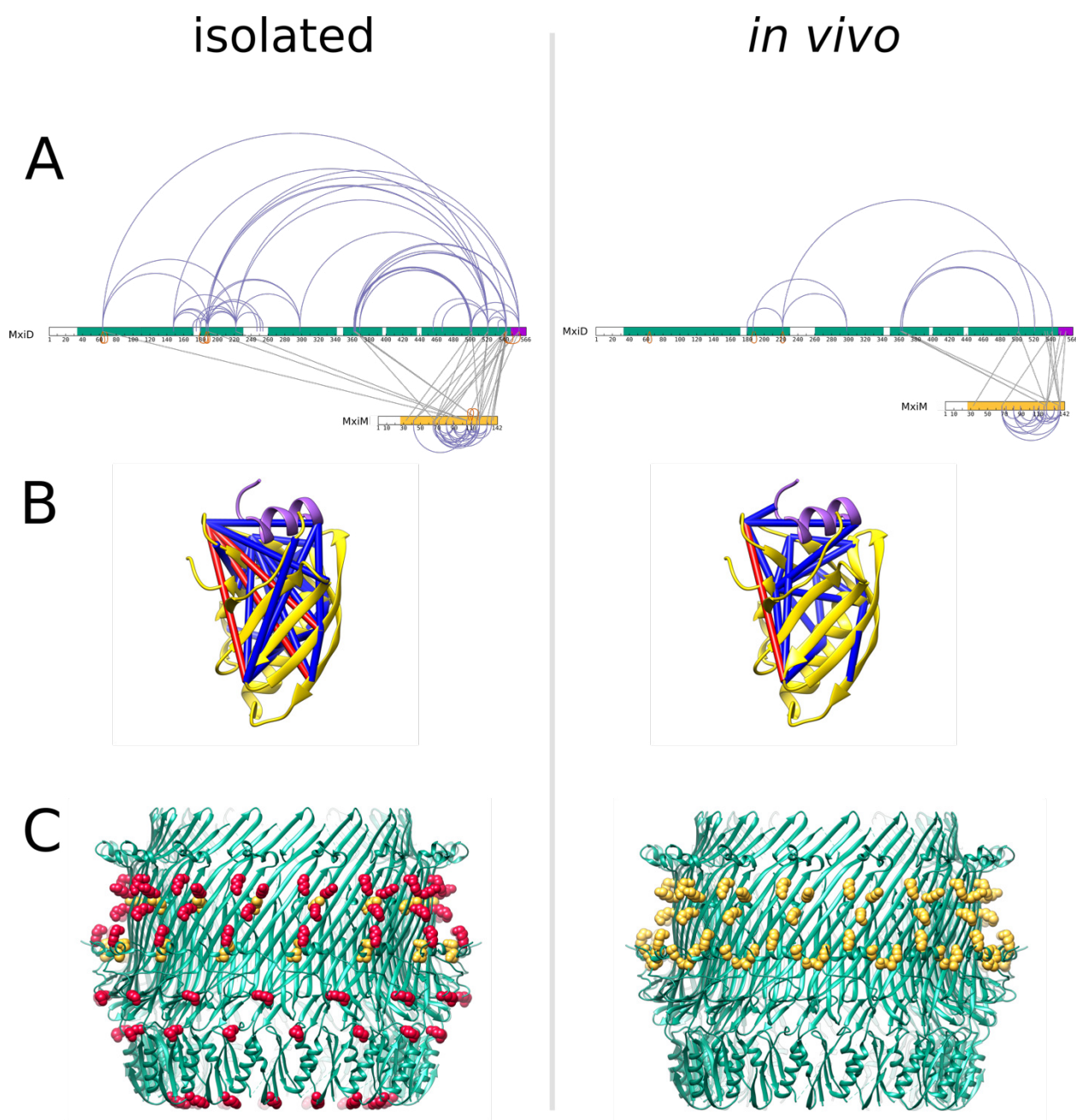
Figure 23: Comparison of pilotin-related cross-links obtained after isolation of the needle complex or generated directly *in vivo*. MxiM cross-links from the isolated needle complex (from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH*) cross-linked with BS3 (left) or from *Shigella* cells (*S. flexneri* M90T Δ*ipaD ipaB*::*knot-strep* complemented with *strep-mxiH*) cross-linked with DSS *in vivo* with subsequent needle complex isolation (right). **(A)** Cross-link network between the pilotin (MxiM) and secretin (MxiD). The protein sequences are displayed as bars, and color fill indicates structural coverage (PDB ID: 8AXL [151], PDB ID: 2JW1 [121]). Heteromeric cross-links are displayed in gray, self-cross-links in purple, and homomultimeric cross-links in orange. **(B)** Cross-links plotted on the atomic model of MxiM (yellow) in complex with C-terminal peptide of $MxiD_{549-566}$ (PDB ID: 2JW1 [121] (purple) as cartoon representation. Cross-links below the distance threshold of ≤30 Å are presented as blue sticks, and cross-links exceeding this threshold (>30Å) are shown as red sticks. **(C)** Distribution of residues involved in cross-links with the pilotin (MxiM) and C-terminal secretin α2 ($MxiD_{549-566}$) across the OM ring ($MxiD_{180-548}$). Atomic model of the OM ring as cartoon representation obtained by cryo-EM PDB ID: 8AXL [151] is shown in green. Residues involved in cross-links to MixM are highlighted as yellow spheres. Residues cross-linking to both, MxiM and MxiD-α2, are displayed as red spheres.

The cross-linking MS measurement was performed by Dr. Zhuo Angel Chen. The cryo-EM atomic model of the secretin was provided by Dr. Michele Lunelli.
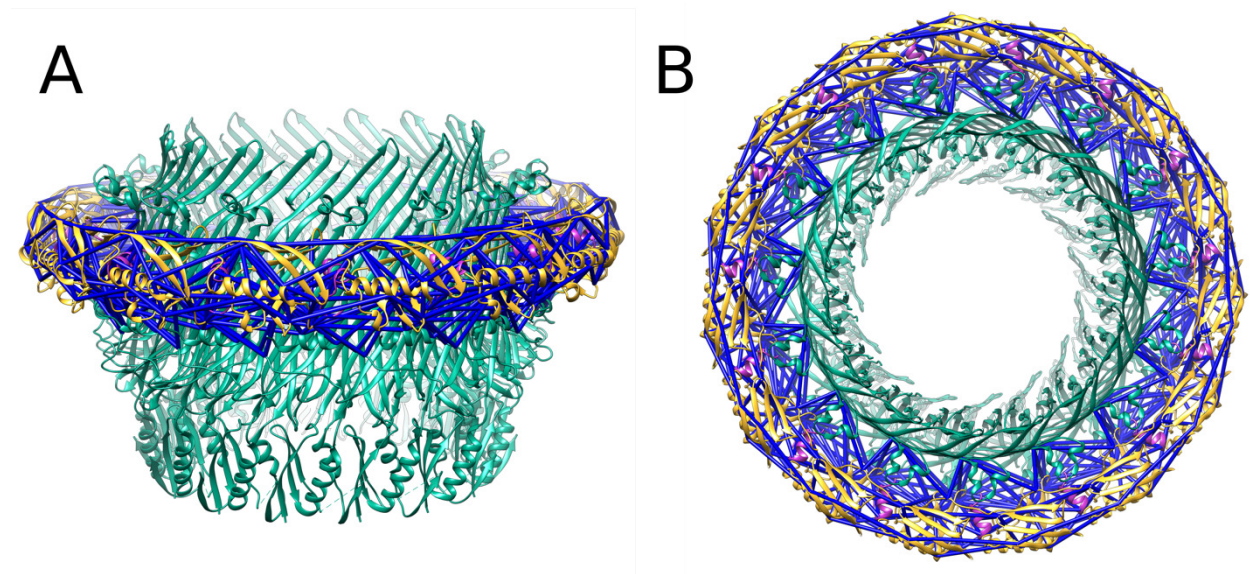
Figure 24: Docking of the pilotin based on cross-links obtained *in vivo*. Side **(A)** and top **(B)** view of the atomic model (cartoon representation) from the OM ring ($MxiD_{180-548}$, green) obtained by cryo-EM (PDB ID: 8AXL [151]), encircled by an equal number of pilotin subunits (MxiM, yellow) in complex with α2 ($MxiD_{549-566}$, purple) from (PDB ID: 2JW1 [121]). The manual rigid body docking of MxiM is based on cross-links obtained by DSS *in vivo*. Cross-links between MxiD and MxiM, and MxiM self-links are displayed. In this orientation, all found cross-links satisfy the distance threshold of ≤30 Å and are displayed as blue sticks.

The cross-linking MS measurement was performed by Dr. Zhuo Angel Chen. The atomic model of the secretin was provided by Dr. Michele Lunelli.

2.3. Analysis of the *Salmonella* (SPI-1) type III secretion system sorting platform proteins

The previous two chapters focused mainly on the transmembrane needle complex of the T3SS. However, the system also includes a large protein complex found within the cytosol, the so-called sorting platform [5]. It is responsible for effector recruitment hierarchy and is likely very dynamic [28,29].

In general, the sorting platform is lost upon needle complex isolation, and so far, it cannot be isolated as a whole in any other way. Therefore, most structural analyses comprise only single subunits [139] or are based on cryo-ET [36]. Recently, members of my research group carried out a reverse approach to characterize the sorting platform: they generated larger complexes by reconstitution and subsequently characterized these by native MS [137]. In line with this approach, the oligomerization behavior of the sorting platform ATPase InvCΔ79 was analyzed by native MS. Additionally, distance restraints of the reconstituted sorting platform protein complexes (SpaO, SpaO$_C$, OrgB and InvC) were generated by cross-linking MS, which might aid the modeling of these complexes in the future. Since the proteins of interest originate from the *Salmonella* (SPI-1) T3SS, protein names will be referred to according the SPI-1 nomenclature in this chapter (Fig. 1).

2.3.1. Native mass spectrometry of the ATPase InvCΔ79 reveals monomeric and dimeric states

InvC is the ATPase of the *Salmonella* (SPI-1) T3SS. It has an essential role during secretion and is conserved among other T3SS-depended pathogens [43,140]. Because no structural information on the *Salmonella* ATPase was available at the time, further characterization of this protein was necessary. Upon purification, the full-length of InvC was susceptible to proteolytic degradation, wherefore the first 79 residues were removed in the following construct [140]. It was shown that the N-terminus is involved in membrane association and interacts with the regulating stator protein (OrgB in *Salmonella*) [137,142,143]. Other T3SS ATPases which were successfully solved structurally, excluded the N-terminal residues (PDB ID: 2OBL [149], 4NPH [146], 5SWJ [147]) as well. The resulting InvCΔ79 construct, which included a C-terminal strep-tag, was recombinantly expressed and purified by affinity followed by size-exclusion chromatography (SEC). The SEC exhibited two elution peaks, of which both contained the

ATPase, as determined by Coomassie-stained SDS-PAGE (Fig. 25, inlet), [140] indicating different oligomeric states.

Native MS confirmed that the SEC fractions primarily contained either the monomeric (39,901.9 ± 0.6 Da) or dimeric (79,808 ± 2 Da) state of InvCΔ79 (Fig. 25, Suppl. table 5). The dimeric stoichiometry was confirmed by MS/MS (data not shown). These findings are in line with size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis, which obtained two elution peaks with weight-averaged molar masses of 40 kDa and 80 kDa [140]. These experiments demonstrate that InvCΔ79 can form monomers and dimers.
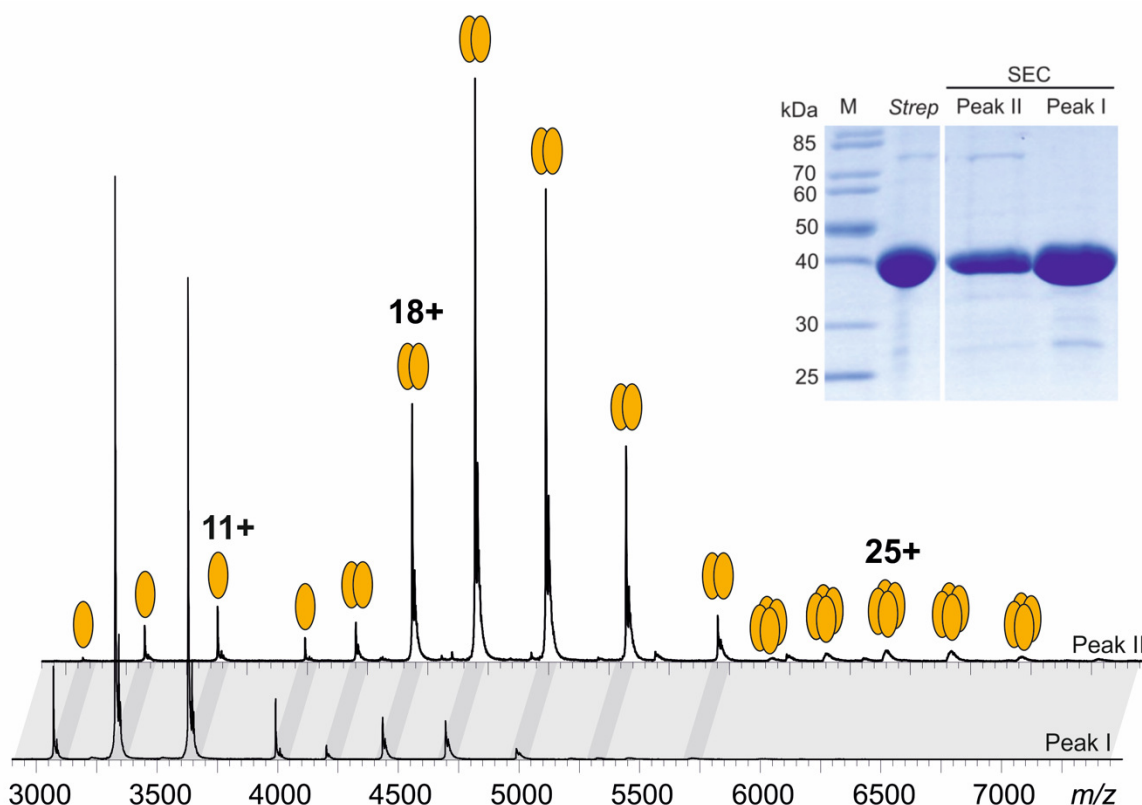
Figure 25: InvCΔ79 forms monomers and dimers. Inlet: Coomassie-stained SDS-PAGE of InvCΔ79 with C-terminal strep-tag during purification procedure from [140]. Lanes from left to right: molecular mass marker, strep-affinity purification, and SEC elution peak II and peak I. Native mass spectra of the two InvCΔ79 SEC elution peaks. While the SEC peak I displays the monomeric state with a mass of 39,901.9 ± 0.6 Da the SEC peak II predominantly exhibits the dimeric state with 79,808 ± 2 Da. Theoretical and experimental masses are summarized suppl. table 5.

Parts of this chapter are published [140]. Dr. Ivonne Bernal provided the protein sample for native MS and performed the SDS-PAGE shown in the inlet.

2.3.2. Cross-linking mass spectrometry of the SpaO/ SpaO$_C$/ OrgB/ InvC complex

Recently it was shown that the sorting platform proteins SpaO/SpaO$_C$/OrgB/InvC can form larger sub-complexes [137]. Native MS analysis revealed the stoichiometries of the stepwise increasing building blocks of these reconstituted proteins. The most prominent complexes, including all four proteins, displayed the 2(SpaO-2SpaO$_C$)-2OrgB-InvC (at 215 kDa) and SpaO-2SpaO$_C$-2OrgB-InvC (at 159 kDa) stoichiometry. In addition, SAXS suggested that SpaO/SpaO$_C$/OrgB/InvC adopts an extended L-shape. Restraints from cross-linking MS could provide additional structural insight. By combing all available information, such as the I) previously mentioned native MS and SAX data [137], II) structures obtained by X-ray crystallography [139,140] or by prediction [188] and III) the newly generated cross-linking restraints, might generate a more detailed model of the sorting platform building block.

2.3.2.1. Finding the optimal protein-to-cross-linker ratio

Prior to MS analysis, the optimal protein-to-cross-linker ratio must be determined. Therefore, a consistent amount of the isolated SpaO/SpaO$_C$/OrgB/InvC complex was cross-linked with increasing amounts of BS3 (table 1) for 30 min at room temperature (RT) in a small-scale experiment. After quenching the reaction, the samples were monitored with SDS-PAGE and subsequent silver staining (Fig. 26). With increasing amounts of cross-linker, blurry bands of higher molecular masses became visible. At the higher molar ratios, starting at 1:147 and above, protein bands were found near the loading pockets, which hardly migrated into the gel. This suggests that larger complexes are heavily cross-linked, which should be avoided for the MS sample preparation. The optimal ratio was therefore determined between the molar ratios of 1:59-1:147 (SpaO-2SpaO$_C$-2OrgB-InvC-complex : BS3). Here a wide range of bands are visible, indicating the onset of the cross-linking, however, excessive over-cross-linking has not been reached yet.
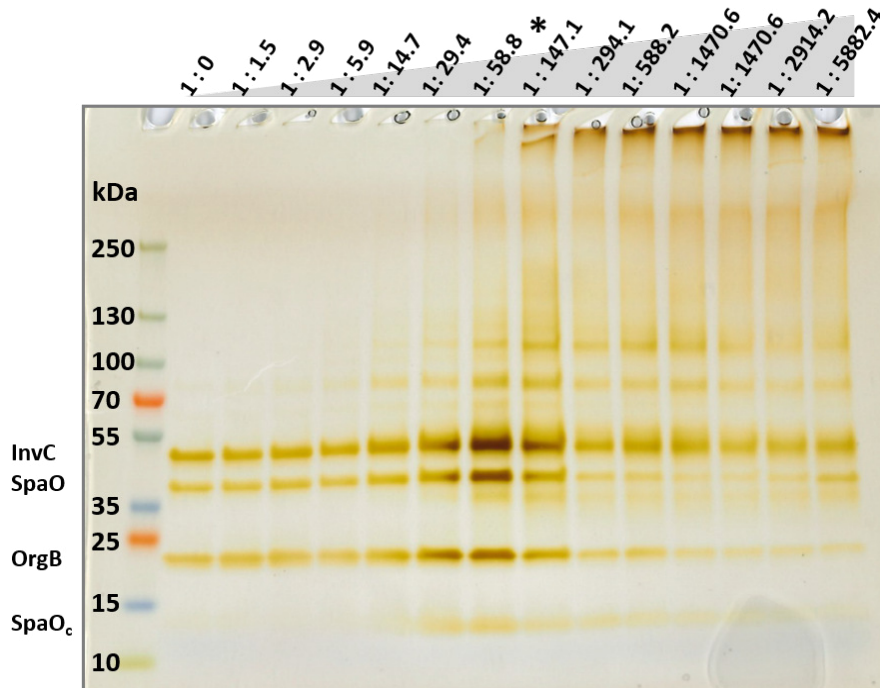
Figure 26: Finding the optimal protein-to-cross-linker-ratio for the sorting platform protein complex SpaO/SpaO$_C$/OrgB/InvC. Gradient (4-15%) SDS-PAGE followed by silver staining of isolated SpaO/SpaO$_C$/OrgB/InvC complex cross-linked with increasing amounts of BS3 (from left to right) (For pipetting scheme, see table 1). The molar ratio of protein complex to cross-linker labeled on top. With an increasing amount of cross-linker, fuzzy bands shifting to higher masses become visible. The optimal ratio was defined between 1:59 to 1:147 (marked with an asterisk).

2.3.2.2. Mass spectrometry reveals a dense network of cross-links between SpaO/ SpaO$_C$/ OrgB/ InvC

After MS analysis, the peak list was searched against the four sorting platform proteins SpaO, SpaO$_C$, OrgB and InvC. Thereby 605 cross-links at linkage level 2 % FDR were found (Suppl. table 6). Figure 27 displays the dense network of cross-links between the proteins. The most heteromeric cross-links were found between InvC & OrgB (162 cross-links) and OrgB & SpaO (108 cross-links), followed by SpaO & InvC (41 cross-links). Only a few cross-links were found between OrgB & SpaO$_C$ (8 cross-links), SpaO & SpaO$_C$ (8 cross-links), and InvC & SpaO$_C$ (2 cross-links).

Since SpaO$_C$ is the product of the internal translation initiation site of the *spaO* gene, SpaO$_C$ shares nearly the identical sequence with SpaO [136–138]. Therefore, some cross-links, including these two proteins, remain ambiguous and cannot be assigned to one or the other. However,

because of the initial methionine in SpaO$_C$, cross-links with peptides that include the aforementioned sequence region can clearly distinguish between the two proteins. The latter applies to a large part of the existing cross-links, and therefore these provide distinctive information about the two proteins originating from the same *spaO* gene.

In terms of self-links, InvC displays the highest amount (123 cross-links), followed by OrgB (110 cross-links) and SpaO (43 cross-links). None were found in SpaO$_C$. It remains unclear whether these self-links are from within the same subunit or are derived from multimers of the same protein. Only when the involved peptides overlap in sequence it is certain that these cross-links must arise from different subunits. Multiple of these homomultimeric cross-links were found in OrgB (14 cross-links) and a few in SpaO (2 cross-links) and InvC (2 cross-links).

The dense network of generated cross-links could be utilized in a future integrative modeling approach.
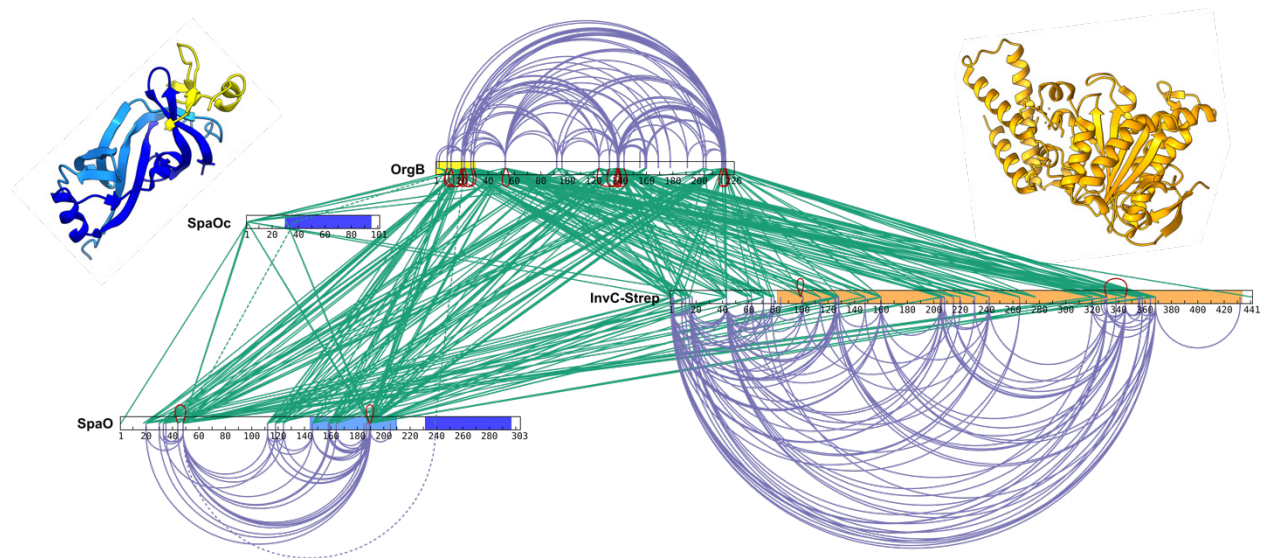
Figure 27: Cross-linking network of the sorting platform protein complex SpaO, SpaO$_C$, OrgB and InvC. Isolated SpaO/SpaO$_C$/OrgB/InvC complex cross-linked with BS3 and analyzed MS. Proteins sequences are indicated as bars. Color fill marks available structural information from X-ray crystallography. SpaO is colored in shades of blue, OrgB in yellow, and InvC in Orange. Corresponding atomic models are pictured in cartoon representation. Left: PDB ID: 4YX7 [139] (crystallization chaperone T4 lysozyme and second constituent complex from the crystallographic asymmetric unit were removed for clarity). Right: PDB ID: 6RAE [140]. Heteromeric cross-links are displayed in green, self-cross-links in purple and homomultimeric cross-links in dark red. Ambiguous cross-links are presented as dashed lines. For a list of cross-links, see suppl. table 6. All four proteins are involved in a dense cross-

linking network. These cross-links provide distance restraints that could facilitate an integrative modeling approach to provide novel structural information about the sorting platform proteins.

The protein sample was provided by the protein production core facility at the CSSB in Hamburg, Germany and the cross-linking MS analysis was performed by Dr. Zhuo Angel Chen.

3. Discussion

3.1. Analysis of the isolated *Shigella* type III secretion system needle complex

3.1.1. Purification adjustments for downstream analysis of the needle complex

The *Shigella* T3SS needle complex has been isolated and structurally analyzed several times before [23,26,33,134,152], yet isolation in sufficient amounts remains challenging. This work shows that it is possible to adapt the existing isolation protocol [33] for the *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* needle complex according to the needs of downstream analysis. This includes the exchange of detergent and buffer to Triton X-100 and HEPES (Fig. 9). Further, the isolation with the detergent LMNG is possible, however needs further improvement (Fig. 10). Because of its slow off-rate and the possibility to remove excess detergent micelles, LMNG might be beneficial for analysis without the obstructive micelles [163–165].

While looking for ways to improve the existing protocol, interesting findings were made. For example, that the strep-tagged needle protein MxiH is not detectable in western blot or SDS-containing dot blot. Closer examination revealed that denaturing conditions such as heat exposure and the addition of SDS loading dye led to this undetectability (Fig. 5). Since this is not the case for the control, it is likely specific to the needle protein itself. Even though it is unclear what might cause this behavior, whether its aggregation or something else, this work suggests a protocol to overcome this problem. This comprises a dot blot of the samples in the native condition. With this, two needle complex proteins (MxiG and MxiH-strep) could be detected among different isolation steps, therefore making it an applicable method for monitoring the isolation procedure while making adaptions (Fig. 6). The nativeness of the samples comes with additional downsides, such as they (and everything that comes in contact with them, such as washing buffers) must be treated and discarded according to their biological safety level, in this case S2. Further, because the samples are still biologically active, they must be used immediately or frozen until use. Besides that, the suggested dot blot is a reliable and quick method to detect the strep-tagged MxiH. Additionally, it would be interesting to check whether the detection-omitting properties of MxiH persist with other anti-strep-tag antibodies from different manufacturers.

3. Discussion

While monitoring if the choice growth medium (either LB or TSB) could affect needle complex isolation, it became apparent that the bacteria grow faster in TSB than LB medium. This was confirmed by additional experiments, describing the OD_{600nm} growth curves of the *Shigella* mutant in LB/ TSB and in inducing/ non-inducing conditions (Fig. 8, Suppl. table 1). The mutant grew faster in the TSB than in the LB medium, and unsurprisingly, plasmid induction led to a reduced growth in all cultures. When comparing the needle complex isolations from isolations grown in these two media, both exhibit similar sample quality (Fig. 7 C-D). However, it became clear that harvest at lower OD_{600nm} is crucial. It reduces background and generally improves sample quality as far as can be judged by native stain TEM (Fig. 7 A-D). Because of these findings, LB was continued as the medium of choice. Even though the sample quality was similar to TSB, due to the slower growth rate in LB, the widow in which cells display optimal OD_{600nm} is extended. Since harvest at lower OD_{600nm} is critical for sample quality, this extended time window facilitates harvesting at the right time, especially with large-scale isolations.

Mass photometry of the isolated needle complex generated a broad distribution of binding events within the MDa range with peaks at 3, 4.6, 6, and 8 MDa (Fig. 11). It should be considered that the measurement conditions may unavoidably not have been ideal and could affect the outcome. This includes the strong background of detergent, the lack of calibrant for the higher mass regions, and the need for further replicates. That being said, the most abundant peak around 4.6 MDa most likely represents the needle complex that has a calculated mass of 4.2 MDa (Suppl. table 2). The peak around 3 MDa could display the basal body without the needle, either because it has broken off or it was still in the process of assembling itself [169,170]. While lower mass species could be explained as damaged or incomplete needle complex components, the identity of the higher mass species remain unclear and only can be speculated about. Later discussed cross-linking experiments (Fig. 14 A) reveal the presence of additional subunits (e.g., MxiC, MxiA), which may add to higher masses. Since isolation is based merely on a single affinity step, the presence of contaminates might also be likely. For both cases, however, additional proteins must be in a rather low abundance or relatively small because they are not very noticeable by negative stain TEM. Even though the exact compositions of the peaks remain to be identified, the mass photometry analysis indicates heterogeneity within the sample. For improved native mass measurements of this complex, a reduction in heterogeneity could be achieved by adding an additional isolation step (e.g., gradient centrifugation) or by removal of the needle. Additionally, the role of the detergent during the mass photometry

measurements that generates a notable background and potentially impedes landing events due to its coating properties remains uncertain. Therefore, mass photometry measurements may be improved by utilizing the detergent LMNG and drop-diluting the sample as described in [165]. It is possible to solubilize the needle complex in LMNG (Fig. 10), and the isolation procedure described here could serve as a starting point to develop a suitable sample for successful mass analysis.

3.1.2. The secretin exhibits a novel S domain which might allow the pilotin to remain bound

Single-particle cryo-EM of the needle complex from the mutant *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* allowed for *de novo* modeling of many areas of the needle complex, including the OM ring. The OM is composed of the secretin (MxiD), which assembles into a large pentadecameric, doubly-walled, antiparallel β-barrel (Fig. 12 A-E). The obtained atomic model (PDB ID: 8AXL [151]) confirms the previously proposed structure, which was based on homology [33]. Novel and remarkable, however, is the structure of the C-terminal S domain (residues K521 to P548). It adopts the conformation of a long loop and an amphipathic helix (α1), which extends laterally across two following subunits (n+2) on the exterior of the secretin barrel (Fig. 12 D-E). Yet, the last eighteen C-terminal residues (S549 to Y566) of the S domain remain unobserved, indicating flexibility or disorder of that region.

When compared to the only other T3SS secretin structure available, InvG from *Salmonella* (PDB ID: 6DV6 [34]), the structural conservation is remarkable (RMSD 1.2 Å for 290 aligned residues). However, as indicated in previously published sequence alignments [32,33], the conservation is restricted to the two major domains of the OM ring, the N3, and the secretin domain. Strikingly, the relatively small, C-terminal S domains share no sequence conservation whatsoever [33]. Regardless of their small size and diversity, the S domains can be crucial for secretin assembly [78,101,113]. So much so that deleting only the last ten C-terminal residues of the *Salmonella* secretin leads to loss of formation and function of the entire needle complex, highlighting the importance of this domain [80].

In the OM ring of *Salmonella*, the S domain adopts a helix-turn-helix conformation with the helices oriented perpendicular to the membrane [34]. Thereby, the first helix (α1) interacts with

the following subunit (n+1) and the second helix (α2) with the subunit thereafter (n+2) (Fig. 12 F). In contrast, the *Shigella* S domain extends parallel to the membrane, and its amphipathic helix (α1), interacts directly with the hydrophobic patch of the second neighboring subunit (n+2) (Fig. 12 D-E). Seemingly, by skipping the directly adjacent subunit, the S domain conformation compensates for the absence of the second helix (α2). Intriguingly, these last eighteen residues of the S domain remain unobserved in the cryo-EM model and has been shown to interact with the pilotin (MxiM) [121,130]. In the NMR structure of this interaction, the secretin peptide (which is disordered in solution) rearranges into a short loop and an amphipathic helix (α2) across the hydrophobic cavity of the pilotin (PDB ID: 2JW1 [121]).

Excitingly, cross-linking experiments of the isolated needle complex revealed the presence of the pilotin. For these experiments, first, the protocol for isolating needle complexes had to be adjusted to a cross-linking compatible buffer, in this case, HEPES (Fig. 9). Then, the optimal protein-to-cross-linker ratio was determined (Fig. 13 A, B) and it was checked for cross-linking induced aggregation of particles at that chosen ratio (Fig. 13 C). The cross-link MS analysis revealed a dense network of needle complex proteins (Fig. 14 A, Suppl. table 3). When plotted on the structure obtained by single-particle cryo-EM [151], 90 (85%) cross-links satisfied the expected distance restraint of 30 Å and therefore agree with the EM model (Fig. 14 B). 16 (15%) cross-links, mainly present between the connector and OM ring of MxiD, exceeded this restraint (Fig. 14 C). It remains unclear if this more compact conformation is of biological relevance or just an artifact caused by the isolation of the complex. Three additional subunits, which are not present in the EM model, have also been detected. This includes the export gate of the export apparatus (MxiA), the gatekeeper (MxiC), and the pilotin (MxiM). The export gate MxiA locates beneath the export apparatus within the needle complex [35,94] (Fig. 1), which agrees with the obtained cross-links, yet these are not sufficient for an integrative modeling approach. The latter also applies to the substrate-regulating MxiC that can promiscuously interact with multiple subunits [189].

Fascinatingly, numerous cross-links were found between the pilotin and the secretin, demonstrating that the pilotin is isolated along with the needle complex (Fig. 14 A). This is in good agreement with earlier immuno-gold labeling experiments from [134] and the reported co-isolation of MxiM with the needle complex [81]. These experiments, however, only demonstrate the presence and not the location of the pilotin within the needle complex. In order to further define the location of the pilotin, the obtained cross-links were subjected to an integrative modeling approach (Suppl. fig. 3). Thereby, the pilotin could not be docked around the secretin

in such fashion that a majority of the cross-links could be fulfilled. Instead, docking based solely on cross-links generated multiple positions and orientations of the pilotin with respect to the secretin (Suppl. fig. 4). This led to the assumption that the pilotin must be in complex with the last few C-terminal residues of the secretin (α2) as described in the NMR structure [121], however they are both highly flexible together. This is supported by intra-MxiD cross-links. Almost all residues on the secretin barrel that cross-linked to the pilotin also cross-linked to the last C-terminal residues of the secretin (α2) (Fig. 15), indicating a broad but similar area of localization. It was concluded that this flexibility is caused by the loss of the OM in which the pilotin could anchor with it's lipidated N-terminus [123,133].

Nevertheless, single-particle cryo-EM of the needle complex generated an undefined density around the distal half of the secretin ring (Suppl. fig. 2) in which the pilotin was assumed. Because this location correlates with an undefined density visible in cryo-EM tomography of *Shigella* minicells [35], it was thought that this is likely the physiological native location of the pilotin. Further, Sani and collogues [81] suggested a similar location for the pilotin. Therefore, this cryo-EM density was included in an additional integrative modeling approach. It generated three location clusters (Suppl. fig. 5-6). However, in only one of the clusters, the lipidated N-terminus of the pilotin was directed toward the OM (Fig. 16). This displays the limitations of the available data from the isolated sample, which outcome most likely represents the optimal location, however not the physiological orientation of the pilotin. How this problem was overcome by *in vivo* cross-linking will be discussed in the following chapter.

3.2. *In vivo* cross-links suggest a more confined localization of the pilotin compared to the isolated needle complex

In order to obtain more structural information on the *Shigella* needle complex in physiological conditions, an *in vivo* cross-linking approach was chosen. Initially, the optimal cross-linker-to-cell-ratio had to be determined (Fig. 17). Then, the needle complex from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* was isolated after *in vivo* cross-linking in parallel to a non-cross-linked control (Fig. 18). Many of these cross-linked particles displayed an additional irregular density on the cytosolic side (Fig. 19). Because of its location, it seems likely that this density is composed of sorting platform proteins. In comparable experiments, isolated needle complexes from bacteria previously treated with a similar cross-linker, dithiobis(succinimidyl propionate) (DSP), likewise exhibited additional features on the

cytoplasmic side and demonstrably contained the sorting platform protein Spa33 [190]. Unfortunately, the obtained this density is very irregular. This might be either because the sorting platform complex is too fragile and not fully cross-linked to endure the isolation procedure or because it might reflect different stages of the dynamic sorting platform assembly [28–30]. To reduce the latter option, the procedure was repeated with a *Shigella* mutant (*S. flexneri* M90T Δ*ipaD ipaB*::*knot-strep* complemented with *strep-mxiH*) which expresses a knotted substrate protein [26] that remains trapped within the needle complex when secreted [25–27]. It was thought that jamming the complex in a similar state of secretion might generate a more heterogeneous sample. Unexpectedly, the substrate-trapped mutant generated a strong and reproducible background (Fig. 20 B-C). Some needle complexes might display some additional densities, especially on the cytoplasmic side; the strong background, however, prohibits clear statements according this (Fig. 21). It is uncertain what causes this background, yet two explanations, or a combination thereof, might be plausible. First, the background is caused by two strep-tag-possessing subunits, the needle protein (MxiH-strep) and the knotted substrate (IpaB-knot-strep), which, if cross-linked to larger complexes, cannot be separated from the needle complex by the ultracentrifugation step during the isolation procedure. Further, the comparison of the western blots might indicate a slightly higher amount of cross-linking in the substrate-trapped mutant (Fig. 20 A) than in the apo needle complex (Fig. 18 C), which is likely the result of using a fresh vail of cross-linking reagent. It is known that N-Hydroxysuccinimide (NHS) ester reagents are very moister sensitive [158,191,192], and therefore a fresh chemical may lead to increased cross-linking than the initially tested amount. Nevertheless, the sample was subjected to cross-linking MS analysis. This method has no difficulties dealing with complex heterogenous samples [173], and it could nicely complement the previous data of the isolated needle complex with physiological insights.

The obtained cross-link network comprises needle complex proteins, the previously assumed sorting platform proteins (including Spa33), and cytosolic components along with effectors (Fig. 22 A, Suppl. table 4). When compared to the isolated sample, unsurprisingly, more proteins were found within the *in vivo* cross-linking network, albeit the number of cross-links between these proteins was generally lower. From these cross-links generated *in vivo*, 40 could be mapped on the cryo-EM structure (Fig. 22 B-C), whereby 38 (95%) cross-links satisfied (≤30 Å) and 2 (5%) exceeded the given distance restraint (>30 Å) of the cross-linker. In comparison to the isolated sample, these are overall fewer cross-links, in terms of percentage, however, more of them satisfy the cryo-EM model indicating major structural agreement.

3. Discussion

Fourteen additional subunits that are not present in the cryo-EM model could be detected in the cross-linking network *in vivo* (Fig. 22 A). For most of these subunits, only one or two cross-links are present, which is inadequate for a structural modeling approach. However, two protein pairs are extensively cross-linked. First, this includes the substrate IpaB-knot and its cytosolic chaperone IpgC. Since substrates must be at least partially unfolded to fit through the narrow channel of the needle complex, extensive self-cross-links, as well as the interaction with its cytosolic chaperone demonstrate that the displayed IpaB-knot, is not only engaged within the needle complex, but also present in its cytosolic form. This is likely because the knotted substrate has a strep-tag, by which the cytoplasmic state could be isolated along with the needle complexes. Nonetheless, the (cytosolic) interaction between IpaB and IpgC has been thoroughly characterized previously [184,193–196] and will not be addressed further in this work. In addition, it must be pointed out that among the proteins cross-linking to the substrate, IpaD was identified even though it should have been deleted in the mutant strain (*S. flexneri* M90T Δ*ipaD ipaB*::*knot-strep* complemented with *strep-mxiH*). Before continuing the use of this strain, additional characterization and clarification is strongly advised. Nonetheless, the presence of IpaD should have no impact on the following analyzed interactions between the secretin and the pilotin.

The second pair of considerably cross-linked subunits, intriguingly, is the secretin and the pilotin. Previously, in the isolated sample, pilotin and α2 cross-links were broadly distributed across the OM, indicating high flexibility of both subunits (Fig. 15, Suppl. fig. 4, Fig. 23 A, C left column). In contrast, cross-linked residues on the OM ring generated *in vivo* only originated from the pilotin and conglomerate in a more defined region on the OM ring (Fig. 23 A, C, right column). Since no cross-links between the OM and α2 were found (Fig. 23 A, C left column), the S domain likely adopts a steadier conformation *in vivo* than in the isolated sample. Considering that cross-links between the pilotin and α2 of the S domain were still obtained, the stabilization is most likely achieved by the OM-anchored pilotin, with the interaction as described in the NMR model (PDB ID: 2JW1 [121]).

Excitingly, utilizing the cross-links generated *in vivo*, the pilotin can be manually placed around the secretin as a rigid body in such manner that all obtained cross-links satisfy the distance restraint below 30 Å (Fig. 24). Even the violated distance restraint within the pilotin NMR model itself (Fig. 23 B), could be satisfied through the docking by interacting with adjacent pilotin subunits. In the obtained assembly, the pilotin is located in the distal half of the OM ring, with its lipidated N-terminus oriented toward the OM. This location coincides with

undefined densities in cryo-ET of the *Shigella* T3SS from minicells [35,40] and confirms the previously suggested pilotin localization on the OM ring [81]. Further, it has been shown that MxiM is associated with the inner leaflet of the OM [133]. Therefore, this finding is in good agreement with all previous studies. Moreover, it confirms the previous hypothesis (chap. 3.1.2.) that the flexibility of the pilotin (and α2 of the S domain) is due to the isolation procedure of the needle complex [151].

Taken together, these previous findings and the experiments presented here clearly demonstrate that the *S. flexneri* pilotin remains part of the needle complex even after secretin assembly and that it localizes in the distal half of the OM ring. However, while pilotins are known to play essential roles during localization and assembly, it is not generally assumed that the pilotins persist along their secretins after assembly. So far, this behavior has only been shown for a T2SS [107]. In contrast, in the *Salmonella* (SPI-1) T3SS, which is currently the only T3SS secretin structure available other than the one present here, binding of the pilotin (PDB ID: 6XFK [116]) to the secretin (PDB ID: 6DV6 [34]) would require a conformational shift indicating a different mechanism of action. This difference is also visible in cryo-ET of the T3SSs in bacterial minicells. While the *Shigella* T3SS needle complex exhibits an undefined density at the same location in which the cross-links suggest the pilotin [35,40], there is no such extra density present in *Salmonella* [36]. Further, single-particle cryo-EM maps of isolated *Salmonella* needle complexes [27,34] could resolve most of the S domain and did not exhibit an extra, unresolved density around the OM ring that suggests the presence of an additional subunit like it was the case in *Shigella* (Suppl. fig. 2).

All these indications point out that there are differences in the persistence of the pilotin after needle complex assembly between these two pathogens. A suggested model is displayed in figure 28. How far this difference also applies to functions beyond the known initial assembly and localization of the secretins [101,113] remains to be investigated. Together, the secretin and pilotin demonstrate how conserved and unique features within the T3SS make it worthwhile to investigate these systems in a species-specific fashion to fully understand their underlying mechanisms.
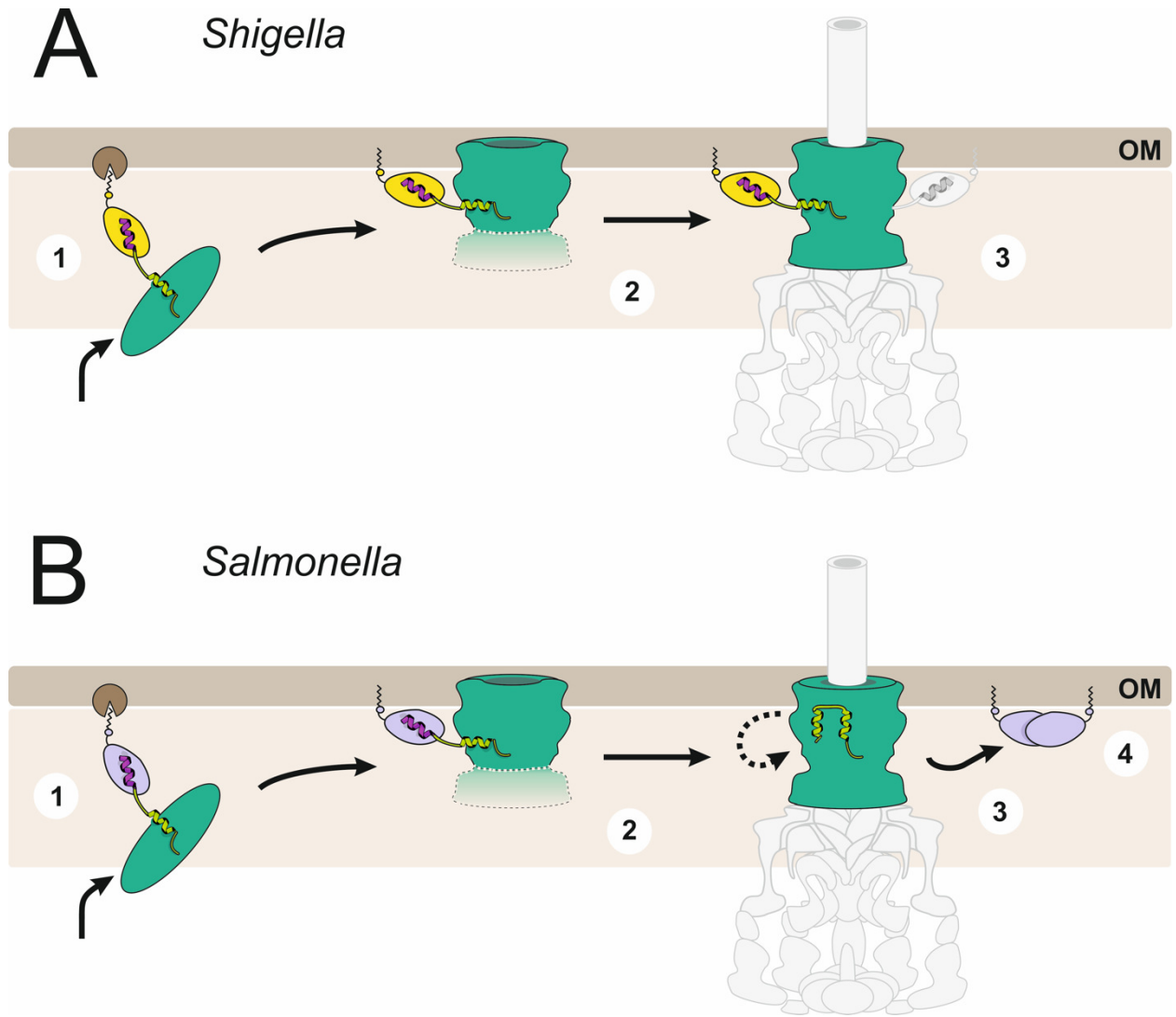
Figure 28: Model of potential pilotin-assisted secretin assembly mechanisms. (1) Monomeric secretin (dark green) is translocated across the IM into the periplasm where the pilotin (*Shigella*: yellow, *Salmonella*: lilac) forms a complex with the last C-terminal residues (α2, purple) of the secretins S domain (lime green). The pilotin is transferred to the OM by its lipidated N-terminus via the Lol pathway (LolB indicated in brown). **(2)** The pilotin anchors in the inner leaflet of the OM, and the secretin embeds within the membrane and oligomerizes to a pore. For complete needle complex assembly, a symmetry mismatch between the secretin pore and secretin connector (green faded) must be generated in order to connect with the IM ring. **(3)** In *Shigella*, the pilotin remains bound to the fully assembled T3SS to anchor it in the OM. Because the α2 remains in complex with the pilotin, the S domain adopts an extended conformation to reach the neighboring secretin subunit (n+2) for stabilization. Unlike *Salmonella,* where the pilotins dissociate from the secretin **(4)**. Thereby the dimerization of the pilotin aids in covering the hydrophobic interface where α2 was previously bound. The S domain, in turn, adopts a helix-turn-helix conformation and the α2 interacts with the hydrophobic patch of neighboring secretin subunit (n+2).

3.3. Analysis of the *Salmonella* (SPI-1) type III secretion system sorting platform proteins

The sorting platform remains the least understood unit of the T3SS. Since it cannot be isolated as a whole, subunits or sub-complexes must be analyzed individually to gain further structural insight. It is composed of multiple subunits (SpaO, SpaO$_C$, OrgA, OrgB, InvI, and InvC in the *Salmonella* (SPI-1) T3SS) arranged into a six L-shaped spoked wheel with the ATPase InvC as the central hub, when assembled with the needle complex [35,36,137].

T3SS ATPases are likely hexameric or dodecameric when associated with the transmembrane needle complex [29,35,36], and isolated full-length ATPases are able to adopt homo-oligomeric states [197,198]. In this work, however, the full-length InvC suffered proteolysis, wherefore the first 79 N-terminal residues were removed for structural characterization [140]. Even though InvCΔ79 lacked the N-terminus, which is involved in oligomerization and OrgB interaction [137,143], monomeric and dimeric states were determined by native MS (Fig. 25, Suppl. table 5). This was confirmed with SEC-MALS of InvCΔ79 [140], indicating that oligomerization of InvC is not solely dependent on the N-terminus or interaction with other T3SS components. Recently, hexameric formation of the *E. coli* ATPase (EscN) in complex with the stalk (EscO) in the presence of the transition state analog Mg^{2+}APD-AlF$_3$ was achieved, whereby EscN lacked the initial 28 N-terminal residues [150]. It would be tempting to test whether the hexameric assembly of InvC could be generated with the protocol successfully applied for EscN-EscO. However, further research will be necessary to fully comprehend oligomerization, function and interaction of the ATPase InvC with the substrates and sorting platform components.

In a second experiment regarding the sorting platform, the recombinantly expressed sorting platform protein complex SpaO, SpaO$_C$, OrgB, InvC generated a dense network of cross-links including all of the subunits (Fig. 26-27, Suppl. table 6).

Because SpaO$_C$ is the shorter variant of SpaO, they share nearly the identical sequence except for the initial methionine in SpaO$_C$ [136–138], making it potentially challenging to distinguish the SpaO$_C$ peptides. Unexpectedly, multiple peptides included this methionine and could therefore be clearly attributed to SpaO$_C$, providing some distinctive information about these similar proteins. Likewise interesting are the homomultimeric cross-links found within the subunits. These types of cross-links have sequence overlapping peptides and, therefore, demonstrably originate from different (multimeric) subunits of the same protein [173].

Many of these were found within OrgB, demonstrating multimeric occurrence within the complex. This is consistent with earlier native MS analysis that OrgB-containing complexes usually possess two OrgB molecules [137]. It would be exciting to investigate whether the homomultimeric and intra-OrgB cross-links suggest a similar asymmetric arrangement as it is found in the flagellar homolog FliH [199]. A few homomultimeric cross-links were also found in SpaO and InvC, which is constant with previous findings. For example, minor dimerization of the SpaO-2SpaO$_C$ trimer with and without OrgB has been previously reported [137]. Also, the ATPase, InvC, can form dimers (Fig. 25) [140] and most likely is hexameric when completely assembled [36,140,150].

The generated cross-linking data lays the groundwork for an integrative modeling approach; however, this complex analysis goes beyond the scope of this thesis. In future studies, the cross-linking information could be combined with cryo-ET maps [36], the stoichiometry of the complex generated by native MS [137], structural information from X-ray crystallography of the SpaO-OrgB interaction [139], the recently solved structure of InvC [140], as well as predicted structures from newly available computational tools such as AlphaFold [188,200].

In the future, an integrative modeling approach (including the missing subunits such as OrgA, InvI) might make up for the infeasibility to isolate or reconstitute the entire sorting platform and give insight into the process of effector recruitment and secretion.

3.4. Conclusion and Outlook

This study aimed to better understand the infection-essential T3SS by analyzing its architecture. For many structural methods, prior isolation is required. Therefore, the adaptations to the *S. flexneri* T3SS needle complex isolation procedure reported here will facilitate future needle complex isolation experiments. Additionally, the combination of multiple structural methods applied here led to a more precise localization of the pilotin subunit within the *Shigella* T3SS needle complex, demonstrating the benefits of integrative approaches for challenging subunits. Further, when these obtained findings were compared to the homologs of a closely related genus, small but striking structural differences were disclosed. This highlights the need to investigate the T3SS in a species-specific manner in order to fully comprehend the underlying molecular mechanisms of these systems. Finally, analysis of recombinantly expressed proteins of the *Salmonella* (SPI-1) T3SS sorting platform led to novel

structural insights of the T3SS ATPase. Furthermore, obtained distance restraints from the soluble sorting platform proteins might facilitate an integrative modeling approach in the future.

Taken together, this work increased the insight into the T3SS. Nonetheless, more research is necessary to fully comprehend this complex system, especially regarding the multitude of bacterial species utilizing it. Hopefully, further T3SS-aimed research will facilitate drug development to treat infections caused by T3SS-dependent, Gram-negative bacteria in the near future.

4. Materials & methods

4.1. Protein production

4.1.1. Needle complex isolation from *Shigella flexneri*

4.1.1.1. Adaption to HEPES and Triton X-100

The purification procedure is based on the protocol from my laboratory published in [33] with additional adaptions. A brief protocol is published in [151].

S. flexneri M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* were stored as glycerol stock at -80 °C. The overnight culture was prepared with LB (Luria/Miller) medium and antibiotics added in the following final concentrations: 100 µg/mL carbenicillin or ampicillin, 50 µg/mL kanamycin, 6 µg/mL chloramphenicol.

It was incubated shaking at 180 rpm (Multitron and Multitron Pro, Infors AG) at 37 °C overnight. For expression, *S. flexneri* M90T mutants were cultured in LB medium (Luria/Miller) with reduced amounts of antibiotics (50 µg/mL ampicillin and 25 µg/mL kanamycin) shaking with 180 rpm at 37 °C. At OD_{600nm} 0.02-0.03, the pASK-IBA5plus plasmid (IBA Lifesciences) was induced with anhydrotetracycline (AHT) from a stock solution (2 mg/mL AHT in dimethylformamide (DMF)) to a final concentration of 200 µg AHT/L bacterial culture. Cells were harvested when reaching OD_{600nm} 0.25-0.3 by centrifugation at 10 000 x *g* at 4 °C for 15 min. After that, cell pellets were washed with cold phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4; pH 7.4). The pellet from one liter of culture was resuspended in 20 mL containing 18% (w/v) sucrose, 100 mM Tris (pH 8.0), 100 mM NaCl, 1.25 mM EDTA, 1 mg/mL lysozyme, ¼ tablet protease inhibitor (SigmaFAST™ EDTA free, Sigma-Aldrich) and was incubated, rolling at 4 °C for 1.5 h. After this, cells were incubated for an additional 30 min, gently shaking at 37 °C, and the resulting spheroplast formation was monitored by light microscopy. Subsequently, the spheroplasts were lysed by addition of 5 mL lysis buffer (100 mM Tris (pH 8.0), 100 mM NaCl, 10 % (w/v) Triton X-100 as well as a spatula tip of DNase I, and the for its activation required $MgSO_4$ (10 mM final concentration). After a 30 min incubation of rolling at RT, unopened cells were removed by centrifugation at 3 300 x *g* at 4 °C for 10 min. The supernatant was subjected to ultracentrifugation at 100 000 x *g* overnight at 4 °C. The resulting pellet was washed and resuspended in 15-20 mL buffer (50 mM HEPES (pH 8.0), 100 mM NaCl, 5 mM

EDTA, 0.04 % (w/v) Triton X-100) + ¼ tablet protease inhibitor rolling at 4 °C for approximately 6 h, until fully dissolved. For removal of insolubilized fragments, the sample was subjected to a final centrifugation at 15 000 x *g* for 20 min at 4 °C. The supernatant was transferred to buffer equilibrated Strep-Tacin®Sepharose® (IBA Lifesciences) (100-200 µL slushy suspension) in a column (CHROMABOND® LV 15 mL PE-frits, Macherey-Nagel GmbH & Co. KG). The sample was extensively washed with buffer before being eluted in (50 mM HEPES (pH 8.0), 100 mM NaCl, 5 mM EDTA, 25 mM Biotin, 0.04 % (w/v) Triton X-100). Eluted fractions were monitored by negative stain TEM and flash-frozen in liquid nitrogen.

4.1.1.2. Adaption to lauryl maltose neopentyl glycol (LMNG)

Needle complexes from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* were isolated as previously described (chap. 4.1.1.1.) with the adaption that the detergent LMNG was used as detergent instead of Triton X-100 after the overnight ultracentrifugation step. Therefore, the buffer composition was 50 mM HEPES (pH 8.0), 100 mM NaCl, 5 mM EDTA, 0.02 mM LMNG.

4.1.1.3. Culturing in different media

Needle complexes from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* were isolated as described in chapter 4.1.1.1. with the following adaptions: In the first experiment, two isolations were run in parallel with the bacteria either grown in TSB or LB medium (including the overnight culture). The bacteria were harvested after the same time point but exhibited different OD_{600nm}, indicating different growth behavior between the two media. In the second experiment, bacteria were grown again in either TSB or LB, yet were harvested when reaching the same OD_{600nm} rather than after the same time.

4.1.2. Sorting platform proteins of *Salmonella* Typhimurium (SPI-1)

4.1.2.1. InvCΔ79

The ATPase InvCΔ79 was designed, recombinantly expressed, and purified by Dr. Ivonne Bernal as described in [140,201].

4.1.2.2. SpaO/SpaO$_C$/OrgB/InvC

The sorting platform SpaO/SpaO$_C$/OrgB/InvC protein complex was provided by the protein production core facility (PPCF) at the CSSB in Hamburg, Germany, based on the protocol provided in [137,202].

4.2. Bacterial growth curves

To test how the medium would affect the mutant *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH*, growth curves were obtained by measuring the hourly OD$_{600nm}$ of bacteria cultured either in TSB or LB. Further, the effect of plasmid induction on growth was monitored. Therefore, an overnight culture was prepared with LB medium supplemented with antibiotics (100 µg/mL carbenicillin, 50 µg/mL kanamycin, 6 µg/mL chloramphenicol) and inoculated with *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* from a frozen glycerol stock. It was cultured overnight at 37 °C shaking at 180 rpm (Multitron or Multitron Pro). For the experiment, 25 mL of medium (either LB or TSB) was substituted with a reduced amount of antibiotics (50 µg/mL ampicillin and 25 µg/mL kanamycin) and inoculated with 75 µL overnight culture. After 1 h of incubation at 37 °C at 180 rpm, the first OD$_{600nm}$ measurements were taken with a Spectrophotometer (BioPhotometer 6131, Eppendorf SE) equipped with single-use cuvettes (Rotilabo®, Carl Roth GmBH + Co. KG). Simultaneously, all samples that should be induced were supplemented with 200 µg/L AHT. After that, the OD$_{600nm}$ were obtained hourly for the following 4 h. For each condition, four biological and technical replicates were taken, and average values, as well as standard deviations, were calculated (see Suppl. table 1).

4.3. Dot blot

4.3.1. Detection of step-tag-MxiH after different pretreatments

To test whether different pretreatments could affect anti-strep-tag-MxiH dot blot outcome, native and denaturing conditions were tested on *Shigella* cells. Therefore, *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* was grown in TSB 37 °C at 180 rpm (Multitron/Multitron Pro), induced with AHT at OD$_{600nm}$ ≈ 0.12 and harvested at OD$_{600nm}$ ≈ 1.2. Cells were washed with PBS, normalized by OD$_{600nm}$ and treated with either 4x SDS

sample loading buffer (40% (v/v) glycerol, 240 mM Tris-HCl (pH 6.8), 8% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.04% (w/v) bromophenol blue), clear native or blue native sample buffer (SERVA Electrophoresis GmbH) or the same amount of water. In addition, some samples were heated for 10 min at 95 °C or frozen at -80 °C before being applied. Next, 15 µL of each sample was dotted onto a PBS-dampened nitrocellulose membrane (Amersham Protran ™ Premium ™ NC 0.45, GE Healthcare) in 2 µL steps with enough time between the steps to ensure the sample was absorbed by the membrane. After that, the membranes were blocked in 5% (w/v) skim milk powder in tris-buffered saline with Tween 20 (TBS-T) (20 mM Tris (pH 7.5), 150 mM NaCl, 0.1 % (v/v) Tween 20) for 1 h at RT and incubated with anti-strep-tag antibody (Qiagen) at a dilution of 1:1 000 in the previously mentioned solution at 4 °C overnight. After three washing steps with TBS-T, the membranes were incubated with horseradish peroxidase (HRP)-conjugated anti-mouse antibody (Jackson ImmunoResearch Laboratories Inc.) in a 1:10 000 dilution in TBS-T + 5% (w/v) skim milk powder for 1 h. After multiple thorough washing steps with TBS-T once again, the membranes were submitted to HPR substrate (SuperSignal™ West Pico PLUS, Thermo Fisher Scientific), and the chemiluminescence signal was detected with a LAS 4000 mini ImageQuant (GE Healthcare).

4.3.2. Strep-tag-MxiH and MxiG detection during needle complex isolation

To monitor if the previously described dot plot procedure (chap. 4.3.1.) is transferable to T3SS needle complex isolation samples and to anti-MxiG detection, a dot plot procedure was conducted with those samples. Therefore, T3SS needle complexes were isolated as described in (chap. 4.1.1.2.), and a dot blot with subsequent immunodetection of those samples was prepared as described above. Samples were not heated or treated with sample buffer, however, frozen until use. The anti-MxiG antibody (produced in-house, Protein Purification facility, Max-Planck Institute for Infection Biology) was used in a 1:10 000 dilution in TBS-T (20 mM Tris, 150 mM NaCl, pH 7.5 + 0.1 % (v/v) Tween 20) and downstream treatment was the same as described above for anti-strep-tag-MxiH.

4.4. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGEs were performed with precast gradient gels either using the Mini-PROTEAN TGX gels (4-15%) (Bio-Rad Laboratories Inc.) or the RunBlue Bis-Tris Precast Gels (4-12%) (Expedeon Ltd.). The gels were run on the corresponding electrophoresis systems, either the Mini-PROTEAN® Tetra Cell (Bio-Rad Laboratories Inc.) or the XCell SureLock™ Mini-Cell (Thermo Fischer Scientific Inc.), according to the companies' instructions. A prestained ladder (PageRuler™ Plus (10-250 kDa), Thermo Fischer Scientific Inc.) was used to monitor protein migration and estimate molecular weights.

4.4.1. Coomassie staining

After SDS-PAGE (chap. 4.4), gels were washed with water, incubated Coomassie staining solution (0.08 g/L Coomassie Brilliant Blue G-250, 35 mM HCL) [203] overnight, and destained in water until bands reached desired visibility.

4.4.2. Silver staining

After SDS-PAGE (chap. 4.4), a rapid silver staining was performed as developed by [204].

4.4.3 Western blot

After SDS-PAGE (chap. 4.4), the gels were briefly washed in transfer buffer (25 mM Tris (pH 8.3), 190 mM glycine, 20% (v/v) methanol) and assembled between a methanol-activated polyvinylidene difluoride (PVDF) membrane (Amersham ™ Hybond ™ P 0.45, GE Healthcare) and buffer-soaked filter papers. The samples were transferred from the gels to the membranes at $\approx$ 1.5 mA/cm^2 for 40-60 min with a semi-dry blotting system (Scie-Plas Ltd.). The membranes were blocked with 5% (w/v) skim milk powder in TBS-T (20 mM Tris, 150 mM NaCl, pH 7.5 + 0.1 % (v/v) Tween 20) for 1 h at RT and incubated with anti-MxiG antibody (Max-Planck Institute for Infection Biology, Protein Purification facility, Berlin) in a 1:10 000 dilution in the previously mentioned blocking solution for either 1 h at RT or overnight at 4 °C. Afterwards, the membranes were washed with TBS-T and incubated with an HRP-conjugated anti-mouse antibody (Jackson ImmunoResearch Laboratories Inc.) in a 1:10

000 dilution for 1 h at RT. The membrane was thoroughly washed three times for 15 min with TBS-T before applying the substrate (SuperSignal™ West Pico PLUS, Thermo Fisher Scientific Inc.). The chemiluminescence signal was imaged with a LAS 4000 mini ImageQuant (GE Healthcare).

4.5. Light microscopy

Bacterial cells and spheroplast formation were monitored with Leica DM IL LED Microscope (objective: 40x or 100x and eyepiece: 10x) (Leica Microsystems) equipped with a Leica MC 170 HD camera.

4.6. Electron microscopy

4.6.1. Negative stain

Carbon-coated copper grids CF400-CU (Electron Microscopy Sciences) were glow discharged for 2 min at the highest setting with the plasma cleaner PDC-002-CE (Harrick Plasma). The freshly glow-discarded grids were then placed on top of a 5-10 µL sample droplet on Parafilm® (Amcor plc.) and incubated for 10 min at RT. After incubation, grids were washed 3 times with water, and access liquid was removed with a damp filter paper after the last washing step. Subsequently, samples were stained with a 1% uranyl acetate solution (Merck) for 3 sec, and access liquid was again removed with first damp, then dry filter paper. Grids were allowed to air dry with the sample side oriented downwards (by being held with negative tweezers) for at least 10 min before further handling. Grids were either viewed directly or stored in a grid holder at RT until then. Samples were viewed with a Talos™ L120C transmission electron microscope (Thermo Fischer Scientific Inc.) equipped with a Ceta CMOS camera.

4.6.2. Single-particle cryo-electron microscopy

Single-particle cryo-EM, including sample preparation, data acquisition, model building, and refinement was performed by Dr. Michele Lunelli as described in [33,151].

4.6.2.1. Analysis and visualization

The obtained atomic model of the secretin PDB ID: 8AXL[151] was structurally analyzed, compared to PDB ID: 6DV6 [34], and visualized with UCSF Chimera [205] and ChimeraX [206,207]. The protein interfaces were calculated utilizing the 'Protein interfaces, surfaces and assemblies' (PISA) service from the European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html) [208].

4.7. Native mass spectrometry

The native MS procedure of InvCΔ79 has been published in [140,201]. In summary, the sample was exchanged into 150 mM ammonium acetate, 1 mM Dithiothreitol (DTT) pH 8.0 with Micro Bio-Spin™ 6 columns (Bio-Rad Laboratories) and measured at ≈ 10 μM with a QTof 2 mass spectrometer (Waters & MS Vision) modified for high mass [209] equipped with gold-coated glass capillaries produced in-house as described in [210]. The mass spectrometer was operated in positive ion mode, and the source pressure was set to 10 mbar. The collision cell was operated between 1.7-1.8 x 10^{-2} mbar with argon used as gas. For InvCΔ79 measurements, the following voltages were applied: 1 300 V to the capillary, 130 V to the cone, and 50 V as accelerating voltage. The raw data was calibrated with obtained cesium iodide clusters (25 mg/mL), and the spectra were analyzed with the MassLynx software (V.4.1). Average masses, standard deviations, full width half maxima (FWHM) were calculated from multiple measurements (Suppl. table 2). The stoichiometry of the obtained complexes was confirmed by MS/MS (data not shown). The protein sequence of InvCΔ79 is listed in suppl. table 7.

4.8. Mass photometry

Mass photometry was performed with a Refeyn OneMP (Refeyn Ltd.,) and analyzed with the corresponding program DiscoverMP 2.2.0. Contrast calibration was performed with an unstained NativeMark™ protein ladder (ThermoFischer Scientific Inc.) diluted 1:250 in PBS in overall 20 μL. Samples were pipetted on pretreated high-precision glass coverslips (No. 1.5H, 24 x 50 mm) provided by Refeyn Ltd., equipped with cropped CultureWell™ CW-50R-1.0 reusable gaskets (Grace Bio-Labs Inc.). The lack of surface tension in the detergent-containing samples only allowed a volume of up to 12 μL per gasket. Therefore, first 10 μL of

buffer (50 mM HEPES (pH 8.0), 100 mM NaCl, 5 mM EDTA, 0.04 % (w/v) Triton X-100) was measured. Then, 2 μL of T3SS needle complex isolated from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* (unknown protein concertation) was added to the buffer and measured again. It should be mentioned that only single measurements were performed for each sample, even though triplicates are recommended.

4.9. Cross-linking

4.9.1. Sample preparation

4.9.1.1. Isolated (apo) needle complex

A brief protocol of this cross-linking procedure of the isolated needle complex can be found in [151].

S. flexneri M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* was isolated as described in chapter 4.1.1.1.. To ensure consistent quality of the sample, a large amount of sample was generated by pooling many suitable elutions of multiple liters of bacterial culture. Then, the optimal protein-to-cross-linker ratio was tested on aliquots before being applied to the sample for subsequent cross-liking MS analysis. For this purpose, increasing amounts of freshly prepared BS3 was titrated to the needle complex sample to reach the final concentrations of 0 mM (control), 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 1 mM, 2 mM of BS3 and was incubated 30 min at RT. The cross-linking reaction was quenched with a final concentration of 100 mM ammonium bicarbonate which was incubated for 15 min at RT. Cross-linked samples were analyzed by SDS-PAGE followed by silver staining (chap. 4.4.2.) and western blot detecting the needle complex subunit MxiG (chap. 4.4.3). From these experiments (Fig. 13 A, B), the cross-linking at a BS3 concentration at 0.2 mM was considered as optimal. To monitor whether cross-linking induced aggregation occurs at this chosen concentration, the needle complex was cross-linked, quenched and diluted 1:10 before being viewed with negative stain TEM (chap. 4.6.1.). Since aggregation was not noticeable on the micrographs (Fig. 13 C) it was proceeded with this concentration.

For cross-linking MS analysis, the needle complex was cross-linked with 0.2-0.3 mM BS3 for 30 min RT and quenched with 100 mM ammonium bicarbonate for 15 min at RT as previously described. Finally, the sample was precipitated with acetone to remove undesired buffer

components such as detergent, as described in [211]. In brief, four times the sample volume of ice-cold acetone (-20 °C) was added to the sample and thoroughly mixed before being incubated for 1 h at -20 °C. After that, the sample was centrifuged 15 000 x *g* at 4 °C, and the supernatant was discarded. This procedure was repeated an additional two times. In the final round, the resulting protein pellet was left to dry for 30 min and stored at -80 °C until further use.

4.9.1.2. *Shigella* bacteria *in vivo*

4.9.1.2.1. The (apo) needle complex

Initially, the optimal cell-to-cross-linker needed to determined. For this purpose, the mutant *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* was cultured as described in chapter 4.1.1.1. In this case, the plasmid was induced with AHT at OD_{600nm} 0.014, and the bacteria were harvested at OD_{600nm} 0.308. After washing the cells with PBS, these were diluted to reach four concentrations differing by the factor of ten (OD_{600nm} 0.03, 0.3, 3, 30) with PBS. Then, these cells were cross-linked with three final concentrations of DSS (1 mM, 2 mM, 5 mM) from a freshly prepared DSS stock solution dissolved in DMSO. For the control, only DMSO was added to the cells. The cross-linking was allowed to react for 30 min at RT before being quenched with 100 mM ammonium bicarbonate for 15 min at RT. The degree of cross-linking was monitored by SDS-PAGE followed by Coomassie staining (chap. 4.4.1.) or western blot detecting MxiG (chap. 4.4.3). To ensure comparability, cells were normalized by OD_{600nm} for SDS-PAGE. The optimal ratio was chosen at OD_{600nm} 3 with 2 mM DSS (Fig. 17).

To test whether *in vivo* cross-linked needle complexes could be isolated, *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* mutants were cross-linked in the previously determined ratio (OD_{600nm} 3, with 2 mM DSS) and needle complexes were subsequently isolated according to the procedure described in chapter 4.1.1.1. For comparison, a non-cross-linking control was run in parallel. The proceeding isolation procedure was monitored by western blot (chap. 4.4.3), and spheroplast formation was monitored with light microscopy (chap. 4.5). The obtained needle complexes were visualized with negative stain TEM (chap. 4.6.1).

4.9.1.2.2. The substrate-trapped needle complex

S. flexneri M90T Δ*ipaD ipaB*::*rrmA("knot")-strep* [26,153] complemented with *strep-mxiH* was cross-linked and isolated as previously described (chap. 4.9.1.2.1) with minor adaptions. The final concentration of the cross-linker was lowered from 2 mM to 1.6 mM DSS because of the usage of a fresh cross-linker vial. Initial cross-linking at 2 mM from the fresh cross-linker lot led to excessive cross-linking as seen by western blot (data not shown) and hampered needle complex isolation. Therefore, the concentration was adapted to 1.6 mM DSS. Further, in the first approach, the plasmid expressing strep-MxiH was immediately induced with AHT during bacterial growth because the mutant strain still processes the wild-type MxiH (Fig. 20 C). However, due to the high background visible by TEM, the plasmid was induced at OD_{600nm} 0.029, as described in the standard protocol in the second approach (Fig. 20 A, B). Nevertheless, the background remained. For MS analysis, a large amount (6 L bacterial culture) from the later approach was produced, precipitated with acetone, and stored at -80 °C until further use.

4.9.1.3. Sorting platform protein complex SpaO/SpaO$_C$/OrgB/InvC

After SEC, the SpaO/SpaO$_C$/OrgB/InvC-strep complex was exchanged into cross-linker compatible buffer (10 mM HEPES (pH 8.0), 50 mM NaCl) with Micro Bio-Spin™ P-30 (Bio-Rad Laboratories Inc.) or Zeba™ Spin desalting columns (Thermo Fisher Scientific Inc.). The molar concentration of the protein complex was determined by measuring absorbance at 280 nm (A_{280nm}) and calculated with the molar extinction coefficient (ε 186 670 M^{-1} cm^{-1}) based on the SpaO-2SpaO$_C$-2OrgB-InvC stoichiometry [137]. First, the optimal protein-to-cross-linker ratio was determined by a small-scale cross-linker titration. Therefore, increasing amounts of freshly dissolved cross-linker BS3 (Thermo Fisher Scientific Inc.) were added to a consistent amount of protein. The mixture was allowed to react for 30 min at RT before being quenched with a final concertation of $\approx$ 100 mM ammonium bicarbonate for 15 min at RT. The titration scheme with applied volumes, concentrations, and ratios can be viewed in table 1. After the complete cross-linking and quenching reaction, samples were separated by SDS-PAGE as previously described (chap. 4.4.) and were subsequently visualized by silver staining (chap. 4.4.2.). Based on the SDS-PAGE, the optimal protein-to-cross-linker range was defined between the ratios 1:59 (lane 8) and 1:147 (lane 9) (Fig. 26). For subsequent MS analysis, a total of $\approx$ 1.6 mg of sorting platform proteins were cross-linked with a protein-to-cross-linker

ratio 1:78 with the same protocol as described above. After the cross-linking and quenching reaction, the sample was precipitated with acetone as described in (chap. 4.9.1.1.), dried, and stored at -80 °C until further use.

Table 1: Pipetting scheme of the sorting platform protein complex (SpaO/SpaO$_C$/OrgB/InvC) and the cross-linker BS3. Corresponding SDS-PAGE is shown in Fig. 26.

Volume protein sample [µL]	Volume buffer [µL]	Volume of BS3 stock solution [µL]	Concentration of BS3 stock solution [mM]	Final volume of reaction [µL]	Final concentration of protein complex in 40 µL [µM]	Final concentration of BS3 in 40 µL [mM]	Protein complex to BS3 molar ratio	Lane loaded on SDS-PAGE
32	8	-	-	40	3.4	0	1 : 0	2
32	6	2	0.1	40	3.4	0.005	1 : 1.5	3
32	4	4	0.1	40	3.4	0.01	1 : 2.9	4
32	0	8	0.1	40	3.4	0.02	1 : 5.9	5
32	6	2	1	40	3.4	0.05	1 : 14.7	6
32	4	4	1	40	3.4	0.1	1 : 29.4	7
32	0	8	1	40	3.4	0.2	1 : 58.8	8
32	6	2	10	40	3.4	0.5	1 : 147.1	9
32	4	4	10	40	3.4	1	1 : 294.1	10
32	0	8	10	40	3.4	2	1 : 588.2	11
32	6	2	100	40	3.4	5	1 : 1470.6	12,13
32	4	4	100	40	3.4	10	1 : 2941.1	14
32	0	8	100	40	3.4	20	1 : 5882.4	15

4.9.2. Mass spectrometry analysis

4.9.2.1. The isolated (apo) needle complex

Cross-linking MS analysis was performed by Dr. Zhuo Angel Chen and Dr. Francis J. O'Reilly as described in [151]. The protein sequence of MxiH-strep is listed in suppl. table 7.

4.9.2.2. The **i*n vivo* cross-linked, substrate-trapped needle complex**

The cross-linking MS analysis of the *in vivo* cross-linked *S. flexneri* was performed by Dr. Zhuo Angel Chen as previously described in [151] with minor adaptations as follows: Liquid

chromatography-tandem mass spectrometry (LC-MS/MS) measurements were conducted as duplicates, whereby the amount of 5 mAU*mL of peptides were used for the acquisitions. The following parameters were chosen for the cross-linked peptide identification search: MS accuracy = 3 ppm; MS2 accuracy = 5 ppm; allowed number of missed cleavages = 2. Finally, a FDR of 3% on the residue-pair level was applied. The protein sequence of IpaB-knot-strep is listed in suppl. table 7.

4.9.2.3. The sorting platform protein complex SpaO/SpaO$_C$/OrgB/InvC

The cross-linking MS analysis of the sorting platform proteins was performed by Dr. Zhuo Angel Chen as described previously in [151] with the following adaptions: each fraction analyzed by LC-MS/MS was measured in duplicates with 4 mAU*mL unit equivalent of peptides injected for each acquisition. For the identification of the cross-linked peptides, the peak lists were searched against the sequences and the reversed sequences of the four *S.* Typhimurium (SPI-1) T3SS sorting platform proteins SpaO, SpaO$_C$, OrgB, InvC [137,202]. These parameters were used for the search: MS accuracy = 3 ppm; MS2 accuracy = 5 ppm; allowed number of missed cleavages = 2. A FDR of 1% on residue-pair level was used. The proteins sequences of SpaO, SpaO$_C$, OrgB, InvC are listed in suppl. table. 7.

4.9.3. Structural analysis, visualization, and integrative modeling

4.9.3.1. The isolated (apo) needle complex

The integrative modeling approach of the isolated needle complex and the pilotin was performed by Dr. Karol Kaszuba as described in [151].

4.9.3.2. The *in vivo* cross-linked, substrate-trapped needle complex

The obtained cross-linking network of the *in vivo* cross-linked and subsequently isolated substrate-trapped needle complex was visualized with xiNET [212], including imported PDB files from: 8AXK [151], 2JW1 [121], 4A4Y [177], 3U0C [182], 1IPA [183], 3GYZ [184], 2J0O [185], 4TT9 [186], 5SW1 [147], and 3W3C [187].

Cross-links were plotted on the cryo-EM model of the needle complex 8AXK [151] with Xlink Analyzer [213] and UCSF Chimera [205]. The length threshold for the cross-linker was set to 30 Å.

Docking of the pilotin around the OM ring was achieved in UCSF Chimera with Xlink Analyzer by manually placing a single pilotin (MxiM) subunit (in complex with the C-terminal residues of MxiD) 2JW1 [121] on two secretin (MxiD) subunits 8AXL [151] and rotating the pilotin until all displayed cross-link restraints were satisfied. The generated location of the pilotin between the two secretin subunits was transferred around the OM ring by superposition so that for each MxiD subunit, a pilotin subunit was present.

4.9.3.2. The sorting platform protein complex SpaO/SpaO$_C$/OrgB/InvC

The generated cross-linking network of the sorting platform proteins SpaO/SpaO$_C$/OrgB/InvC was visualized with xiNET [212] with imported PDB files from 4YX7 [139] and 6RAE [140].

5. References

1. Smith, A. C. & Hussey, M. A. Gram stain protocols. *American Society for Microbiology* (2005) https://asm.org/Protocols/Gram-Stain-Protocols (accessed 23.08.2022).

2. Gram, H. C. Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschr Med* **2**, 185–189 (1884).

3. Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**, 1–15 (2010).

4. Costa, T. R. D. *et al.* Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nature Reviews Microbiology* **13**, 343–359 (2015).

5. Deng, W. *et al.* Assembly, structure, function and regulation of type III secretion systems. *Nature Reviews Microbiology* **15**, 323–337 (2017).

6. Galán, J. E. & Collmer, A. Type III Secretion Machines: Bacterial Devices for Protein Delivery into Host Cells. *Science* **284**, 1322–1328 (1999).

7. Staehelin, C. & Krishnan, H. B. Review Article: Nodulation outer proteins: Double-edged swords of symbiotic Rhizobia. *Biochem J* **470**, 263–274 (2015).

8. Zboralski, A., Biessy, A. & Filion, M. Bridging the Gap: Type III Secretion Systems in Plant-Beneficial Bacteria. *Microorganisms* **10**, 187 (2022).

9. Coburn, B., Sekirov, I. & Finlay, B. B. Type III secretion systems and disease. *Clin Microbiol Rev* **20**, 535–549 (2007).

10. Abby, S. S. & Rocha, E. P. C. The Non-Flagellar Type III Secretion System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted Systems. *PLoS Genet* **8**, 1002983 (2012).

11. Kotloff, K. L., Riddle, M. S., Platts-Mills, J. A., Pavlinac, P. & Zaidi, A. K. M. Shigellosis. *The Lancet* **391**, 801–812 (2018).

12. Kamanova, J. *Bordetella* Type III Secretion Injectosome and Effector Proteins. *Front Cell Infect Microbiol* **10**, 466 (2020).

13. Horna, G. & Ruiz, J. Type 3 secretion system of *Pseudomonas aeruginosa*. *Microbiol Res* **246**, (2021).

14. Lyons, B. J. E. & Strynadka, N. C. J. On the road to structure-based development of anti-virulence therapeutics targeting the type III secretion system injectisome. *Medchemcomm* **10**, 1273–1289 (2019).

5. References

15. World Health Organization. Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance (accessed: 01.10.2021).

16. Ventola, C. L. The antibiotic resistance crisis: causes and threats. *P & T journal* **40**, 277–83 (2015).

17. Dadgostar, P. Antimicrobial Resistance: Implications and Costs. *Infect Drug Resist* **12**, 3903 (2019).

18. World Health Organization. Prioritization Of Pathogens To Guide Discovery, Research And Development Of New Antibiotics For Drug-Resistant Bacterial Infections, Including Tuberculosis. *WHO/EMP/IAU/2017.12* 1–87 (2017).

19. Wagner, S. *et al.* Bacterial type III secretion systems: A complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS Microbiol Lett* **365**, 1–13 (2018).

20. Jenkins, J., Worrall, L. J. & Strynadka, N. C. J. Recent structural advances towards understanding of the bacterial type III secretion injectisome. *Trends Biochem Sci* (2022).

21. Cornelis, G. R. The *Yersinia* deadly kiss. *J Bacteriol* **180**, 5495–5504 (1998).

22. Kubori, T. *et al.* Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. *Science* **280**, 602–605 (1998).

23. Blocker, A. *et al.* Structure and composition of the *Shigella flexneri* 'needle complex', a part of its type III secreton. *Mol Microbiol* **39**, 652–663 (2001).

24. Blocker, A. *et al.* Target cell contact triggers expression and polarized transfer of *Yersinia* YopE cytotoxin into mammalian cells. *EMBO Journal* **13**, 964–972 (1994).

25. Radics, J., Königsmaier, L. & Marlovits, T. C. Structure of a pathogenic type 3 secretion system in action. *Nat Struct Mol Biol* **21**, 82–87 (2014).

26. Dohlich, K., Zumsteg, A. B., Goosmann, C. & Kolbe, M. A Substrate-Fusion Protein Is Trapped inside the Type III Secretion System Channel in *Shigella flexneri*. *PLoS Pathog* **10**, e1003881 (2014).

27. Miletic, S. *et al.* Substrate-engaged type III secretion system structures reveal gating mechanism for unfolded protein translocation. *Nat Commun* **12**, 1546 (2021).

28. Lara-Tejero, M., Kato, J., Wagner, S., Liu, X. & Galán, J. E. A sorting platform determines the order of protein secretion in bacterial type III systems. *Science* **331**, 1188–1191 (2011).

29. Diepold, A. *et al.* A dynamic and adaptive network of cytosolic interactions governs protein export by the T3SS injectisome. *Nat Commun* **8**, 15940 (2017).

30. Diepold, A., Kudryashev, M., Delalez, N. J., Berry, R. M. & Armitage, J. P. Composition, Formation, and Regulation of the Cytosolic C-ring, a Dynamic Component of the Type III Secretion Injectisome. *PLoS Biol* **13**, e1002039 (2015).

5. References

31.   Bergeron, J. R. C. & Marlovits, T. C. Cryo-EM of the injectisome and type III secretion systems. *Curr Opin Struct Biol* **75**, 102403 (2022).

32.   Worrall, L. J. *et al.* Near-atomic-resolution cryo-EM analysis of the *Salmonella* T3S injectisome basal body. *Nature* **540**, 597–601 (2016).

33.   Lunelli, M. *et al.* Cryo-EM structure of the *Shigella* type III needle complex. *PLoS Pathog* **16**, e1008263 (2020).

34.   Hu, J. *et al.* Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin. *Nat Commun* **9**, 3840 (2018).

35.   Hu, B. *et al.* Visualization of the type III secretion sorting platform of *Shigella flexneri*. *Proc Natl Acad Sci U S A* **112**, 1047–1052 (2015).

36.   Hu, B., Lara-Tejero, M., Kong, Q., Galán, J. E. & Liu, J. *In Situ* Molecular Architecture of the *Salmonella* Type III Secretion Machine. *Cell* **168**, 1065-1074.e10 (2017).

37.   Nans, A., Kudryashev, M., Saibil, H. R. & Hayward, R. D. Structure of a bacterial type III secretion system in contact with a host membrane *in situ*. *Nat Commun* **6**, 10114 (2015).

38.   Berger, C. *et al.* Structure of the *Yersinia* injectisome in intracellular host cell phagosomes revealed by cryo FIB electron tomography. *J Struct Biol* **213**, 107701 (2021).

39.   Tachiyama, S. *et al.* Composition and Biophysical Properties of the Sorting Platform Pods in the *Shigella* Type III Secretion System. *Front Cell Infect Microbiol* **11**, 682635 (2021).

40.   Tachiyama, S. *et al.* The cytoplasmic domain of MxiG interacts with MxiK and directs assembly of the sorting platform in the *Shigella* type III secretion system. *J Biol Chem* **294**, 19184–19196 (2019).

41.   Makino, F. *et al.* The Architecture of the Cytoplasmic Region of Type III Secretion Systems. *Sci Rep* **6**, 1–12 (2016).

42.   Lee, P. C. & Rietsch, A. Fueling type III secretion. *Trends Microbiol* **23**, 296–300 (2015).

43.   Akeda, Y. & Galán, J. E. Chaperone release and unfolding of substrates in type III secretion. *Nature* **437**, 911–915 (2005).

44.   Yoshida, Y. *et al.* Functional characterization of the type III secretion ATPase SsaN encoded by *Salmonella* pathogenicity island 2. *PLoS One* **9**, e94347 (2014).

45.   Lee, P. C., Zmina, S. E., Stopford, C. M., Toska, J. & Rietsch, A. Control of type III secretion activity and substrate specificity by the cytoplasmic regulator PcrG. *Proc Natl Acad Sci U S A* **111**, E2027–E2036 (2014).

46.   Shen, D. K. & Blocker, A. J. MxiA, MxiC and IpaD regulate substrate selection and secretion mode in the T3SS of *Shigella flexneri*. *PLoS One* **11**, e0155141 (2016).

47. Hu, J., Worrall, L. J. & Strynadka, N. C. Towards capture of dynamic assembly and action of the T3SS at near atomic resolution. *Curr Opin Struct Biol* **61**, 71–78 (2020).

48. Parsot, C., Hamiaux, C. & Page, A. L. The various and varying roles of specific chaperones in type III secretion systems. *Curr Opin Microbiol* **6**, 7–14 (2003).

49. Hueck, C. J. Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants. *Microbiology and Molecular Biology Reviews* **62**, 379–433 (1998).

50. Portaliou, A. G., Tsolis, K. C., Loos, M. S., Zorzini, V. & Economou, A. Type III Secretion: Building and Operating a Remarkable Nanomachine. *Trends in Biochemical Sciences* **41**, 175–189 (2016).

51. Wagner, S. & Diepold, A. A unified nomenclature for injectisome-type type III secretion systems. *Curr Top Microbiol Immunol* **427**, 1–10 (2020).

52. Schnupf, P. & Sansonetti, P. J. *Shigella* Pathogenesis: New Insights through Advanced Methodologies. *Microbiol Spectr* **7**, 1–24 (2019).

53. Khalil, I. A. *et al.* Morbidity and mortality due to *Shigella* and enterotoxigenic *Escherichia coli* diarrhoea: the Global Burden of Disease Study 1990–2016. *Lancet Infect Dis* **18**, 1229–1240 (2018).

54. Puzari, M., Sharma, M. & Chetia, P. Emergence of antibiotic resistant *Shigella* species: A matter of concern. *J Infect Public Health* **11**, 451–454 (2018).

55. World Health Organization. Immunization, Vaccines and Biologicals: *Shigella*. https://www.who.int/teams/immunization-vaccines-and-biologicals/diseases/shigella (accessed: 22.08.2022).

56. Muthuirulandi Sethuvel, D. P., Devanga Ragupathi, N. K., Anandan, S. & Veeraraghavan, B. Update on: *Shigella* new serogroups/serotypes and their antimicrobial resistance. *Lett Appl Microbiol* **64**, 8–18 (2016).

57. Kothary, M. H. & Babu, U. S. Infective dose of foodborne pathogens in volunteers: A review. *J Food Saf* **21**, 49–68 (2001).

58. Wassef, J. S., Keren, D. F. & Mailloux, J. L. Role of M cells in initial antigen uptake and in ulcer formation in the rabbit intestinal loop model of shigellosis. *Infect Immun* **57**, 858–863 (1989).

59. Neutra, M. R., Frey, A. & Kraehenbuhl, J. P. Epithelial M cells: Gateways for mucosal infection and immunization. *Cell* **86**, 345–348 (1996).

60. Zychlinsky, A., Prevost, M. C. & Sansonetti, P. J. *Shigella flexneri* induces apoptosis in infected macrophages. *Nature* **358**, 167–169 (1992).

61. High, N., Mounier, J., Prevost, M. C. & Sansonetti, P. J. IpaB of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole. *EMBO Journal* **11**, 1991–1999 (1992).

62. Mounier, J., Vasselon, T., Hellio, R., Lesourd, M. & Sansonetti, P. J. *Shigella flexneri* enters human colonic Caco-2 epithelial cells through the basolateral pole. *Infect Immun* **60**, 237–248 (1992).

63. Ray, K. *et al.* Tracking the dynamic interplay between bacterial and host factors during pathogen-induced vacuole rupture in real time. *Cell Microbiol* **12**, 545–556 (2010).

64. Cossart, P. & Sansonetti, P. J. Bacterial Invasion: The Paradigms of Enteroinvasive Pathogens. *Science* **304**, 242–248 (2004).

65. Bernardini, M. L., Mounier, J., D'Hauteville, H., Coquis-Rondon, M. & Sansonetti, P. J. Identification of icsA, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proc Natl Acad Sci U S A* **86**, 3867–3871 (1989).

66. Mattock, E. & Blocker, A. J. How do the virulence factors of *Shigella* work together to cause disease? *Front Cell Infect Microbiol* **7**, 1–24 (2017).

67. Perdomo, J. J., Gounon, P. & Sansonetti, P. J. Polymorphonuclear leukocyte transmigration promotes invasion of colonic epithelial monolayer by *Shigella flexneri*. *Journal of Clinical Investigation* **93**, 633–643 (1994).

68. Sansonetti, P. J., Arondel, J., Huerre, M., Harada, A. & Matsushima, K. Interleukin-8 controls bacterial transepithelial translocation at the cost of epithelial destruction in experimental shigellosis. *Infect Immun* **67**, 1471–1480 (1999).

69. Buchrieser, C. *et al.* The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. *Mol Microbiol* **38**, 760–771 (2000).

70. Venkatesan, M. M. *et al.* Complete DNA sequence and analysis of the large virulence plasmid of *Shigella flexneri*. *Infect Immun* **69**, 3271–3285 (2001).

71. Sansonnetti, P. J., Kopecko, D. J. & Formal, S. B. Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect Immun* **35**, 852–860 (1982).

72. Maurelli, A. T., Baudry, B., Hélène D'hauteville, H., Hale, T. L. & Sansonetti, P. J. Cloning of plasmid DNA sequences involved in invasion of HeLa cells by *Shigella flexneri*. *Infect Immun* **49**, 164–171 (1985).

73. Sasakawa, C. *et al.* Virulence-associated genetic regions comprising 31 kilobases of the 230-kilobase plasmid in *Shigella flexneri* 2a. *J Bacteriol* **170**, 2480–2484 (1988).

74. Onodera, N. T. *et al.* Genome Sequence of *Shigella flexneri* Serotype 5a Strain M90T Sm. *J Bacteriol* **194**, 3022 (2012).

75.   Cervantes-Rivera, R., Tronnet, S. & Puhar, A. Complete genome sequence and annotation of the laboratory reference strain *Shigella flexneri* serotype 5a M90T and genome-wide transcriptional start site determination. *BMC Genomics* **21**, 1–15 (2020).

76.   Tamano, K. *et al.* Supramolecular structure of the *Shigella* type III secretion machinery: The needle part is changeable in length and essential for delivery of effectors. *EMBO Journal* **19**, 3876–3887 (2000).

77.   Spreter, T. *et al.* A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nat Struct Mol Biol* **16**, 468–476 (2009).

78.   Majewski, D. D., Worrall, L. J. & Strynadka, N. C. Secretins revealed: structural insights into the giant gated outer membrane portals of bacteria. *Curr Opin Struct Biol* **51**, 61–72 (2018).

79.   Hu, J. *et al.* T3S injectisome needle complex structures in four distinct states reveal the basis of membrane coupling and assembly. *Nat Microbiol* **4**, 2010–2019 (2019).

80.   Goessweiner-Mohr, N. *et al.* Structural control for the coordinated assembly into functional pathogenic type-3 secretion systems. *bioRxiv* 714097 (doi:10.1101/714097) (2019).

81.   Sani, M. *et al.* Structural organization of the needle complex of the type III secretion apparatus of *Shigella flexneri*. *Micron* **38**, 291–301 (2007).

82.   Wagner, S. *et al.* Organization and coordinated assembly of the type III secretion export apparatus. *Proceedings of the National Academy of Sciences* **107**, 17745–17750 (2010).

83.   Johnson, S., Kuhlen, L., Deme, J. C., Abrusci, P. & Lea, S. M. 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**, 1–7 (2019).

84.   Kuhlen, L. *et al.* Structure of the core of the type III secretion system export apparatus. *Nat Struct Mol Biol* **25**, 583–590 (2018).

85.   Kuhlen, L. *et al.* The substrate specificity switch FlhB assembles onto the export gate to regulate type three secretion. *Nat Commun* **11**, 1296 (2020).

86.   Sorg, I. *et al.* YscU recognizes translocators as export substrates of the *Yersinia* injectisome. *EMBO J* **26**, 3015–3024 (2007).

87.   Deane, J. E. *et al.* Crystal structure of Spa40, the specificity switch for the *Shigella flexneri* type III secretion system. *Mol Microbiol* **69**, 267–276 (2008).

88.   Abrusci, P. *et al.* Architecture of the major component of the type III secretion system export apparatus. *Nat Struct Mol Biol* **20**, 99–104 (2013).

89.   Yuan, B. *et al.* Structural Dynamics of the Functional Nonameric Type III Translocase Export Gate. *J Mol Biol* **433**, 167188 (2021).

90. Kuhlen, L., Johnson, S., Cao, J., Deme, J. C. & Lea, S. M. Nonameric structures of the cytoplasmic domain of FlhA and SctV in the context of the full-length protein. *PLoS One* **16**, e0252800 (2021).

91. Majewski, D. D., Lyons, B. J. E., Atkinson, C. E. & Strynadka, N. C. J. Cryo-EM analysis of the SctV cytosolic domain from the enteropathogenic *E. coli* T3SS injectisome. *J Struct Biol* **212**, 107660 (2020).

92. Matthews-Palmer, T. R. S. *et al.* Structure of the cytoplasmic domain of SctV (SsaV) from the *Salmonella* SPI-2 injectisome and implications for a pH sensing mechanism. *J Struct Biol* **213**, 107729 (2021).

93. Jensen, J. L., Yamini, S., Rietsch, A. & Spiller, B. W. "The structure of the Type III secretion system export gate with CdsO, an ATPase lever arm". *PLoS Pathog* **16**, e1008923 (2020).

94. Butan, C., Lara-Tejero, M., Li, W., Liu, J. & Galán, J. E. High-resolution view of the type III secretion export apparatus *in situ* reveals membrane remodeling and a secretion pathway. *Proc Natl Acad Sci U S A* **116**, 24786–24795 (2019).

95. Marlovits, T. C. *et al.* Structural Insights into the Assembly of the Type III Secretion Needle Complex. *Science* **306**, 1040–1042 (2004).

96. Torres-Vargas, C. E. *et al.* The inner rod of virulence-associated type III secretion systems constitutes a needle adapter of one helical turn that is deeply integrated into the system's export apparatus. *Mol Microbiol* **112**, 918–931 (2019).

97. Zilkenat, S. *et al.* Determination of the stoichiometry of the complete bacterial type III secretion needle complex using a combined quantitative proteomic approach. *Molecular and Cellular Proteomics* **15**, 1598–1609 (2016).

98. Cordes, F. S. *et al.* Helical Structure of the Needle of the Type III Secretion System of *Shigella flexneri*. *J Biol Chem* **278** 17103–17107 (2003).

99. Demers, J. P. *et al.* The Common Structural Architecture of *Shigella flexneri* and *Salmonella typhimurium* Type Three Secretion Needles. *PLoS Pathog* **9**, e1003245 (2013).

100. Muthuramalingam, M., Whittier, S. K., Picking, W. L., Picking, W. D. & Ramos-Morales, F. The *Shigella* Type III Secretion System: An Overview from Top to Bottom. *Microorganisms* **9**, 1– 26 (2021).

101. Silva, Y. R. de O., Contreras-Martel, C., Macheboeuf, P. & Dessen, A. Bacterial secretins: Mechanisms of assembly and membrane targeting. *Protein Science* **29**, 893–904 (2020).

102. Korotkov, K. V., Sandkvist, M. & Hol, W. G. J. The type II secretion system: Biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* **10**, 336–351 (2012).

103. Howard, S. P. *et al.* Structure and assembly of pilotin-dependent and-independent secretins of the type II secretion system. *PLoS Pathog* **15**, e1007731 (2019).

104. Hay, I. D., Belousoff, M. J., Dunstan, R. A., Bamert, R. S. & Lithgow, T. Structure and membrane topography of the vibrio-type secretin complex from the type 2 secretion system of enteropathogenic *Escherichia coli*. *J Bacteriol* **200**, 1–15 (2018).

105. Hay, I. D., Belousoff, M. J. & Lithgow, T. Structural basis of type 2 secretion system engagement between the inner and outer bacterial membranes. *mBio* **8**, 1–6 (2017).

106. Yan, Z., Yin, M., Xu, D., Zhu, Y. & Li, X. Structural insights into the secretin translocation channel in the type II secretion system. *Nat Struct Mol Biol* **24**, 177–183 (2017).

107. Yin, M., Yan, Z. & Li, X. Structural insight into the assembly of the type II secretion system pilotin-secretin complex from enterotoxigenic *Escherichia coli*. *Nat Microbiol* **3**, 581–587 (2018).

108. Chernyatina, A. A. & Low, H. H. Core architecture of a bacterial type II secretion system. *Nat Commun* **10**, 5437 (2019).

109. Schraidt, O. & Marlovits, T. C. Three-dimensional model of *Salmonella's* needle complex at subnanometer resolution. *Science* **331**, 1192–1195 (2011).

110. Tosi, T. *et al.* Structural Similarity of Secretins from Type II and Type III Secretion Systems. *Structure* **22**, 1348–1355 (2014).

111. Kowal, J. *et al.* Structure of the dodecameric *Yersinia enterocolitica* secretin YscC and its trypsin-resistant core. *Structure* **21**, 2152–2161 (2013).

112. Filloux, A. & Voulhoux, R. Multiple structures disclose the secretins' secrets. *J Bacteriol* **200**, 1–5 (2018).

113. Koo, J., Burrows, L. L. & Lynne Howell, P. Decoding the roles of pilotins and accessory proteins in secretin escort services. *FEMS Microbiol Lett* **328**, 1–12 (2012).

114. Genin, S. & Boucher, C. A. A superfamily of proteins involved in different secretion pathways in gram-negative bacteria: modular structure and specificity of the N-terminal domain. *Molec & Gen Genet* **243**, 112–118 (1994).

115. Crago, A. M. & Koronakis, V. *Salmonella* InvG forms a ring-like multimer that requires the InvH lipoprotein for outer membrane localization. *Mol Microbiol* **30**, 47–56 (1998).

116. Majewski, D. D. *et al.* Characterization of the Pilotin-Secretin Complex from the *Salmonella enterica* Type III Secretion System Using Hybrid Structural Methods. *Structure* **29**, 125-138.e5 (2021).

117. Weaver, S. J. *et al.* CryoEM structure of the type IVa pilus secretin required for natural competence in *Vibrio cholerae*. *Nat Commun* **11**, 5080 (2020).

118. Conners, R. *et al.* CryoEM structure of the outer membrane secretin channel pIV from the f1 filamentous bacteriophage. *Nat Commun* **12**, 6316 (2021).

119. Okuda, S. & Tokuda, H. Lipoprotein Sorting in Bacteria. *Annu Rev Microbiol* **7**, 239–259 (2011).

120. Collin, S., Guilvout, I., Nickerson, N. N. & Pugsley, A. P. Sorting of an integral outer membrane protein via the lipoprotein-specific Lol pathway and a dedicated lipoprotein pilotin. *Mol Microbiol* **80**, 655–665 (2011).

121. Okon, M. *et al.* Structural Characterization of the Type-III Pilot-Secretin Complex from *Shigella flexneri*. *Structure* **16**, 1544–1554 (2008).

122. Carbonnelle, E., Hélaine, S., Prouvensier, L., Nassif, X. & Pelicic, V. Type IV pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occurring after pilus assembly, essential for fibre stability and function. *Mol Microbiol* **55**, 54–64 (2005).

123. Schuch, R. & Maurelli, A. T. MxiM and MxiJ, base elements of the Mxi-Spa type III secretion system of *Shigella*, interact with and stabilize the MxiD secretin in the cell envelope. *J Bacteriol* **183**, 6991–6998 (2001).

124. Koo, J. *et al.* PilF is an outer membrane lipoprotein required for multimerization and localization of the *Pseudomonas aeruginosa* type IV pilus secretin. *J Bacteriol* **190**, 6961–6969 (2008).

125. Viarre, V. *et al.* HxcQ Liposecretin Is Self-piloted to the Outer Membrane by Its N-terminal Lipid Anchor. *J Biol Chem* **284**, 33815–33823 (2009).

126. Trindade, M. B., Job, V., Contreras-Martel, C., Pelicic, V. & Dessen, A. Structure of a Widely Conserved Type IV Pilus Biogenesis Factor that Affects the Stability of Secretin Multimers. *J Mol Biol* **378**, 1031–1039 (2008).

127. Tosi, T. *et al.* Pilotin-secretin recognition in the type II secretion system of *Klebsiella oxytoca*. *Mol Microbiol* **82**, 1422–1432 (2011).

128. Korotkov, K. V. & Hol, W. G. J. Crystal structure of the pilotin from the enterohemorrhagic *Escherichia coli* type II secretion system. *J Struct Biol* **182**, 186–191 (2013).

129. Izoré, T. *et al.* Structural characterization and membrane localization of ExsB from the type III secretion system (T3SS) of *Pseudomonas aeruginosa*. *J Mol Biol* **413**, 236–246 (2011).

130. Lario, P. I. *et al.* Structure and biochemical analysis of a secretin pilot protein. *EMBO Journal* **24**, 1111–1121 (2005).

131. Dunstan, R. A. *et al.* Assembly of the Type II Secretion System such as Found in *Vibrio cholerae* Depends on the Novel Pilotin AspS. *PLoS Pathog* **9**, e1003117 (2013).

132. Rehman, S., Gu, S., Shevchik, V. E. & Pickersgill, R. W. Anatomy of secretin binding to the *Dickeya dadantii* type II secretion system pilotin. *Acta Crystallogr D Biol Crystallogr* **69**, 1381–1386 (2013).

133. Schuch, R. & Maurelli, A. T. The Mxi-Spa type III secretory pathway of *Shigella flexneri* requires an outer membrane lipoprotein, MxiM, for invasin translocation. *Infect Immun* **67**, 1982–1991 (1999).

134. Zenk, S. F. *et al.* Identification of minor inner-membrane components of the *Shigella* type III secretion system 'needle complex'. *Microbiology* **153**, 2405–2415 (2007).

135. Nouwen, N. *et al.* Secretin PulD: Association with pilot PulS, structure, and ion- conducting channel formation. *Proc Natl Acad Sci U S A* **96**, 8173–8177 (1999).

136. Song, M. *et al.* Control of type III protein secretion using a minimal genetic system. *Nat Commun* **8**, 14737 (2017).

137. Bernal, I. *et al.* Molecular Organization of Soluble Type III Secretion System Sorting Platform Complexes. *J Mol Biol* **431**, 3787–3803 (2019).

138. Lara-Tejero, M. *et al.* Role of SpaO in the assembly of the sorting platform of a *Salmonella* type III secretion system. *PLoS Pathog* **15**, e1007565 (2019).

139. Notti, R. Q., Bhattacharya, S., Lilic, M. & Stebbins, C. E. A common assembly module in injectisome and flagellar type III secretion sorting platforms. *Nat Commun* **6**, 7125 (2015).

140. Bernal, I. *et al.* Structural analysis of ligand-bound states of the *Salmonella* type III secretion system ATPase InvC. *Protein Sci* **28**, 1888–1901 (2019).

141. Eichelberg, K., Ginocchio, C. C. & Galan, J. E. Molecular and functional characterization of the *Salmonella typhimurium* invasion genes invB and invC: Homology of InvC to the F0F1 ATPase family of proteins. *J Bacteriol* **176**, 4501–4510 (1994).

142. Akeda, Y. & Galán, J. E. Genetic Analysis of the *Salmonella enterica* Type III Secretion-Associated ATPase InvC Defines Discrete Functional Domains. *J Bacteriol* **186**, 2402–2412 (2004).

143. Case, H. B. & Dickenson, N. E. MxiN Differentially Regulates Monomeric and Oligomeric Species of the *Shigella* Type Three Secretion System ATPase Spa47. *Biochemistry* **57**, 2266–2277 (2018).

144. Woestyn, S., Allaoui, A., Wattiau, P. & Cornelis, G. R. YscN, the putative energizer of the *Yersinia* Yop secretion machinery. *J Bacteriol* **176**, 1561–1569 (1994).

145. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1**, 945–951 (1982).

146. Allison, S. E. *et al.* Identification of the docking site between a type III secretion system ATPase and a chaperone for effector cargo. *J Biol Chem* **289**, 23734–23744 (2014).

147. Burgess, J. L. *et al.* Structural and biochemical characterization of Spa47 provides mechanistic insight into type III secretion system ATPase activation and *Shigella* virulence regulation. *J Biol Chem* **291**, 25837–25852 (2016).

148. Imada, K., Minamino, T., Tahara, A. & Namba, K. Structural similarity between the flagellar type III ATPase FliI and F 1-ATPase subunits. *Proc Natl Acad Sci U S A* **104**, 485–490 (2007).

149. Zarivach, R., Vuckovic, M., Deng, W., Finlay, B. B. & Strynadka, N. C. J. Structural analysis of a prototypical ATPase from the type III secretion system. *Nat Struct Mol Biol* **14**, 131–137 (2007).

150. Majewski, D. D. *et al.* Cryo-EM structure of the homohexameric T3SS ATPase-central stalk complex reveals rotary ATPase-like asymmetry. *Nat Commun* **10**, 626 (2019).

151. Flacht, L. *et al.* Integrative structural analysis of the type III secretion system needle complex from *Shigella flexneri*. *Submitted* (2022).

152. Kamprad, A. Expression, purification and structural analysis of key components of the *Shigella* T3SS basal body in two states. Doctoral dissertation, *Universität Hamburg*, Hamburg, Germany (2019).

153. Dohlich, K. S. Analysis of Type Three System Transport Mechanism in Gram-negative Bacteria. Doctoral dissertation, *Humboldt Universität zu Berlin*, Berlin, Germany (2013).

154. Wexler, A. S. Determination of Phenolic Substances by Ultraviolet Difference Spectroscopy. *Anal Chem* **35**, 1936–1943 (1963).

155. Sigma-Aldrich. *Triton X-100 product information sheet from distributer.* https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/product/documents/160/855/t85 32pis.pdf (accessed: 12.10.2022).

156. Sigma-Aldrich. *Triton X-100 product specification sheet from distributer.* https://www.sigmaaldrich.com/specification-sheets/313/327/X100-BULK.pdf (accessed: 12.10.2022).

157. Konijnenberg, A. *et al.* Global structural changes of an ion channel during its gating are followed by ion mobility mass spectrometry. *Proc Natl Acad Sci U S A* **111**, 17170–17175 (2014).

158. Thermo Fischer Scientific Inc. DSS and BS3 Crosslinkers. *Product instructions from manufacturer* **0418.6** 1–3 (2018).

159. Irvin, R. T., MacAlister, T. J. & Costerton, J. W. Tris(hydroxymethyl)aminomethane Buffer Modification of *Escherichia coli* Outer Membrane Permeability. *J Bacteriol* **145**, 1397–1403 (1981).

5. References

160. Birdsell, D. C. & Cota-Robles, E. H. Production and ultrastructure of lysozyme and ethylenediaminetetraacetate-lysozyme spheroplasts of *Escherichia coli*. *J Bacteriol* **93**, 427–437 (1967).

161. Chae, P. S. *et al.* Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* **7**, 1003–1008 (2010).

162. Breyton, C. *et al.* Assemblies of lauryl maltose neopentyl glycol (LMNG) and LMNG-solubilized membrane proteins. *Biochim Biophys Acta Biomembr* **1861**, 939–957 (2019).

163. Chung, K. Y. *et al.* Role of detergents in conformational exchange of a G protein-coupled receptor. *Journal of Biological Chemistry* **287**, 36305–36311 (2012).

164. Hauer, F. *et al.* GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. *Structure* **23**, 1769–1775 (2015).

165. Olerinyova, A. *et al.* Mass Photometry of Membrane Proteins. *Chem* **7**, 224–236 (2021).

166. Boeri Erba, E., Signor, L. & Petosa, C. Exploring the structure and dynamics of macromolecular complexes by native mass spectrometry. *J Proteomics* **222**, 103799 (2020).

167. Sonn-Segev, A. *et al.* Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nat Commun* **11**, 1772 (2020).

168. Young, G. *et al.* Quantitative mass imaging of single biological macromolecules. *Science* **360**, 423–427 (2018).

169. Poyraz, Ö. *et al.* Protein refolding is required for assembly of the type three secretion needle. *Nature Structural & Molecular Biology* **17**, 788–792 (2010).

170. Kubori, T., Sukhan, A., Aizawa, S. I. & Galán, J. E. Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. *PNAS* **97**, 10225–10230 (2000).

171. Cheng, Y. Single-particle cryo-EM-How did it get here and where will it go. *Science* **361**, 876–880 (2018).

172. Lyumkis, D. Challenges and opportunities in cryo-EM single-particle analysis. *J Biol Chem* **294**, 5181–5197 (2019).

173. O'Reilly, F. J. & Rappsilber, J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nature Structural and Molecular Biology* **25**, 1000–1008 (2018).

174. Fischer, L. & Rappsilber, J. Quirks of Error Estimation in Cross-Linking/Mass Spectrometry. *Anal Chem* **89**, 3829–3833 (2017).

175. Mendes, M. L. *et al.* An integrated workflow for crosslinking mass spectrometry. *Mol Syst Biol* **15**, 1–13 (2019).

176. Deane, J. E., Roversi, P., King, C., Johnson, S. & Lea, S. M. Structures of the *Shigella flexneri* Type 3 Secretion System Protein MxiC Reveal Conformational Variability Amongst Homologues. *J Mol Biol* **377**, 985–992 (2008).

177. Barison, N., Lambers, J., Hurwitz, R. & Kolbe, M. Interaction of MxiG with the cytosolic complex of the type III secretion system controls Shigella virulence. *The FASEB Journal* **26**, 1717–1726 (2012).

178. Webb, B. *et al.* Integrative structure modeling with the Integrative Modeling Platform. *Protein Science* **27**, 245–258 (2018).

179. Rantos, V., Karius, K. & Kosinski, J. Integrative structural modelling of macromolecular complexes using Assembline Nature Protocols **17,** 152–176 (2022).

180. Belsom, A. & Rappsilber, J. Anatomy of a crosslinker. *Curr Opin Chem Biol* **60**, 39–46 (2021).

181. Perdu, C. *et al.* ExsB is required for correct assembly of the *Pseudomonas aeruginosa* Type III secretion apparatus in the bacterial membrane and full virulence *in vivo*. *Infect Immun* **83**, 1789–1798 (2015).

182. Barta, M. L. *et al.* The Structures of Coiled-Coil Domains from Type III Secretion System Translocators Reveal Homology to Pore-Forming Toxins. *J Mol Biol* **417**, 395–405 (2012).

183. Nureki, O. *et al.* An enzyme with a deep trefoil knot for the active-site architecture. *Acta Crystallogr D Biol Crystallogr* **58**, 1129–1137 (2002).

184. Lunelli, M., Lokareddy, R. K., Zychlinsky, A. & Kolbe, M. IpaB–IpgC interaction defines binding motif for type III secretion translocator. *Proceedings of the National Academy of Sciences* **106**, 9661–9666 (2009).

185. Johnson, S. *et al.* Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD. *J Biol Chem* **282**, 4035–4044 (2007).

186. Mcdowell, M. A. *et al.* Characterisation of *Shigella* Spa33 and *Thermotoga* FliM/N reveals a new model for C-ring assembly in T3SS. *Mol Microbiol* **99**, 749–766 (2016).

187. Gao, X. *et al.* Structural insights into VirB-DNA complexes reveal mechanism of transcriptional activation of virulence genes. *Nucleic Acids Res* **41**, 10529–10541 (2013).

188. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

189. Bajunaid, W. *et al.* The T3SS of *Shigella*: Expression, Structure, Function, and Role in Vacuole Escape. *Microorganisms* **8**, 1–27 (2020).

190. Morita-Ishihara, T. *et al.* Shigella Spa33 is an essential C-ring component of type III secretion machinery. *J Biol Chem* **281**, 599–607 (2006).

191. Matzinger, M. & Mechtler, K. Cleavable Cross-Linkers and Mass Spectrometry for the Ultimate Task of Profiling Protein-Protein Interaction Networks *in Vivo*. *Journal of Proteome Research* **20**, 78–93 (2021).

192. Klykov, O. & Weller, M. G. Quantification of N-hydroxysuccinimide and N-hydroxysulfosuccinimide by hydrophilic interaction chromatography (HILIC). *Analytical Methods* **7**, 6443–6448 (2015).

193. Adam, P. R. *et al.* Binding affects the tertiary and quaternary structures of the *Shigella* translocator protein IpaB and its chaperone IpgC. *Biochemistry* **51**, 4062–4071 (2012).

194. Ménard, R., Sansonetti, P., Parsot, C. & Vasselon, T. Extracellular association and cytoplasmic partitioning of the IpaB and IpaC invasins of *S. flexneri*. *Cell* **79**, 515–525 (1994).

195. Picking, W. L. & Picking, W. D. The many faces of IpaB. *Front Cell Infect Microbiol* **6**, 1–7 (2016).

196. Ferrari, M. L., Charova, S. N., Sansonetti, P. J., Mylonas, E. & Gazi, A. D. Structural Insights of *Shigella* Translocator IpaB and Its Chaperone IpgC in Solution. *Front Cell Infect Microbiol* **11**, 353 (2021).

197. Burgess, J. L. *et al.* Spa47 is an oligomerization-activated type three secretion system (T3SS) ATPase from *Shigella flexneri*. *Protein Science* **25**, 1037–1048 (2016).

198. Pozidis, C. *et al.* Type III protein translocase. HrcN is a peripheral membrane ATPase that is activated by oligomerization. *J of Biol Chem* **278**, 25816–25824 (2003).

199. Imada, K. *et al.* Insight into the flagella type III export revealed by the complex structure of the type III ATPase and its regulator. *Proc Natl Acad Sci U S A* **113**, 3633–3638 (2016).

200. Varadi, M. *et al.* AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439–D444 (2022).

201. Bernal, I. Structural-functional characterization of soluble components of the *Salmonella* type III secretion system sorting platform. Unpublished Doctoral Thesis, *Humboldt-Universität zu Berlin*, Berlin, Germany (2019).

202. Börnicke, C. J. *In vitro* Reconstitution and Characterization of Soluble Complexes of the *Salmonella* Type III Secretion System Sorting Platform. Doctoral dissertation, *Universität Hamburg*, Hamburg, Germany (2020).

203. Lawrence, A. M. & Besir, H. Staining of proteins in gels with Coomassie G-250 without organic solvent and acetic acid. *Journal of Visualized Experiments* **30,** 1–3 (2009).

204. Nesterenko, M. V., Tilley, M. & Upton, S. J. A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in polyacrylamide gels. *J Biochem and Biophys Methods* **28,** 239–242 (1994).

205. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).

206. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science* **30**, 70–82 (2021).

207. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science* **27**, 14–25 (2018).

208. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* **372**, 774–797 (2007).

209. Van Den Heuvel, R. H. H. *et al.* Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem* **78**, 7473–7483 (2006).

210. Pogan, R., Schneider, C., Reimer, R., Hansman, G. & Uetrecht, C. Norovirus-like VP1 particles exhibit isolate dependent stability profiles. *Journal of Physics Condensed Matter* **30**, 1–8 (2018).

211. Thermo Fischer Scientific Inc. *Tech Tip #49: Acetone precipitation of proteins (manual TR0049.1)* (2009) http://www.integral-process.com/iso_album/h-046-002371-00(2.0)_imec_service_manual_en.pdf (accessed: 29.05.2022).

212. Combe, C. W., Fischer, L. & Rappsilber, J. xiNET: Cross-link Network Maps With Residue Resolution. *Mol Cell Proteomics* **14**, 1137–1147 (2015).

213. Kosinski, J. *et al.* Xlink analyzer: Software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J Struct Biol* **189**, 177–183 (2015).

214. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* **49**, D480–D489 (2021).

215. Tamano, K., Katayama, E., Toyotome, T. & Sasakawa, C. *Shigella* Spa32 is an essential secretory protein for functional type III secretion machinery and uniformity of its needle length. *J Bacteriol* **184**, 1244–1252 (2002).

216. Kotov, V., Lunelli, M., Wald, J., Kolbe, M. & Marlovits, T. C. Helical reconstruction of *Salmonella* and *Shigella* needle filaments attached to type 3 basal bodies. *Biochem Biophys Rep* **27**, 1– 8 (2021).

217. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**, 3784–3788 (2003).

218. Kim, S. *et al.* PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res* **49**, D1388-D1395 (2021).

219. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242 (2000).

220. Burley, S. K. *et al.* RCSB Protein Data Bank: Powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res* **49**, D437–D451 (2021).

6. Appendix

6.1. Supplementary figures
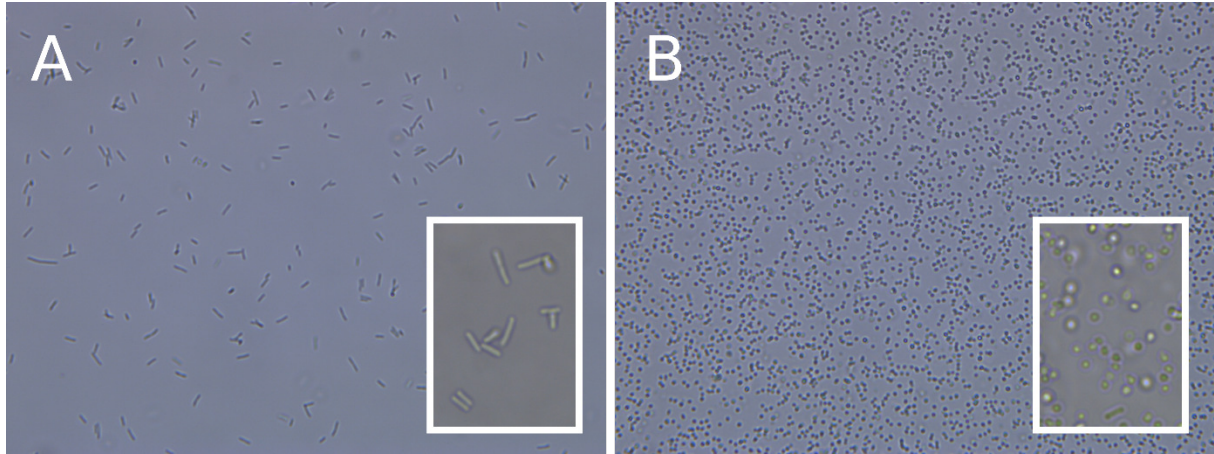
Supplementary figure 1: Spheroplast formation during needle complex isolation. Light microscopy image of *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* before **(A)** and after **(B)** spheroplast formation during T3SS needle complex isolation procedure at 400x magnification. Inlet represents the same samples at 1000x magnification.
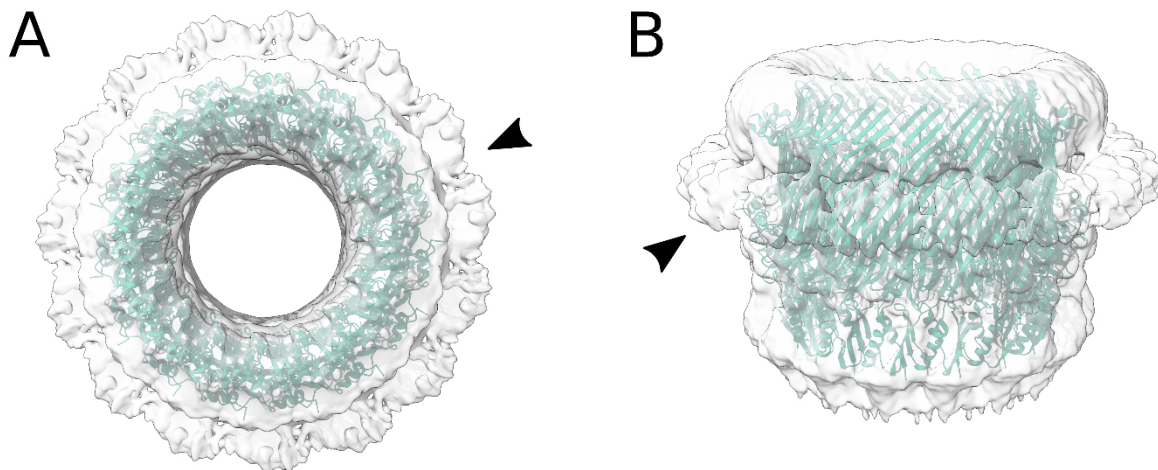
Supplementary figure 2: The unresolved density around the secretin. Bottom **(A)** and side **(B)** view of the atomic model from the MxiD$_{180\text{-}548}$ OM ring as cartoon representation (green) within the cryo-EM density (transparent gray surface) PDB ID: 8AXL, EMD-15701 [151]. An arrow indicates the additional, unresolvable density visible in the single-particle cryo-EM map of the OM ring within the isolated T3SS needle complex from *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH*.

The data collection and atomic modeling of the OM ring were performed by Dr. Michele Lunelli.
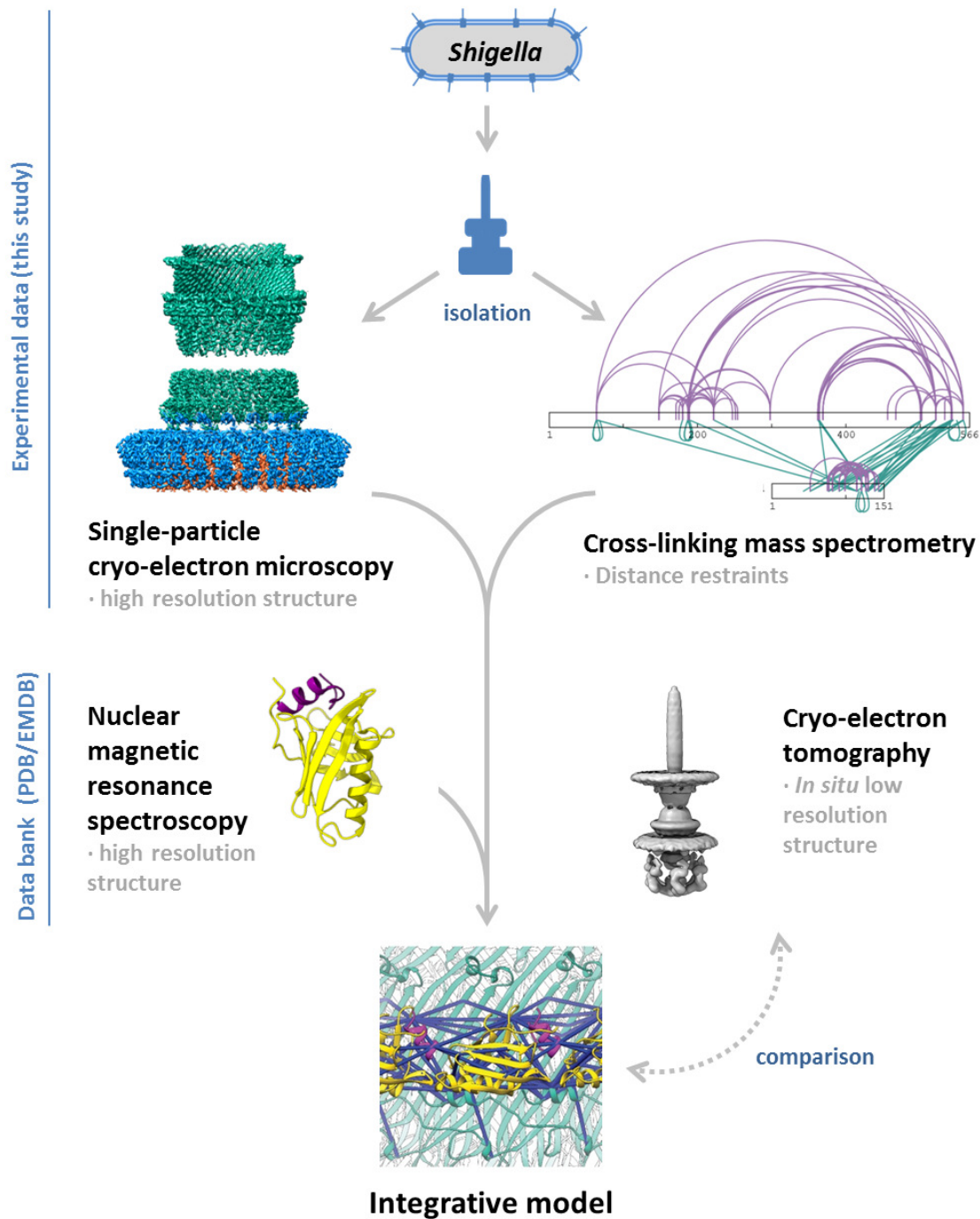
Supplementary figure 3: Integrative modeling workflow of the isolated T3SS needle complex. Schematic illustration of the integrative modeling workflow, which combines single-particle cryo-EM and cross-linking MS of isolated *Shigella* T3SS needle complexes with an NMR structure of the pilotin MxiM PDB ID: 2JW1 [121]. The obtained integrative model is subsequently compared to cryo-ET [35] structures of the T3SS needle complex.

This figure will be published in [151].
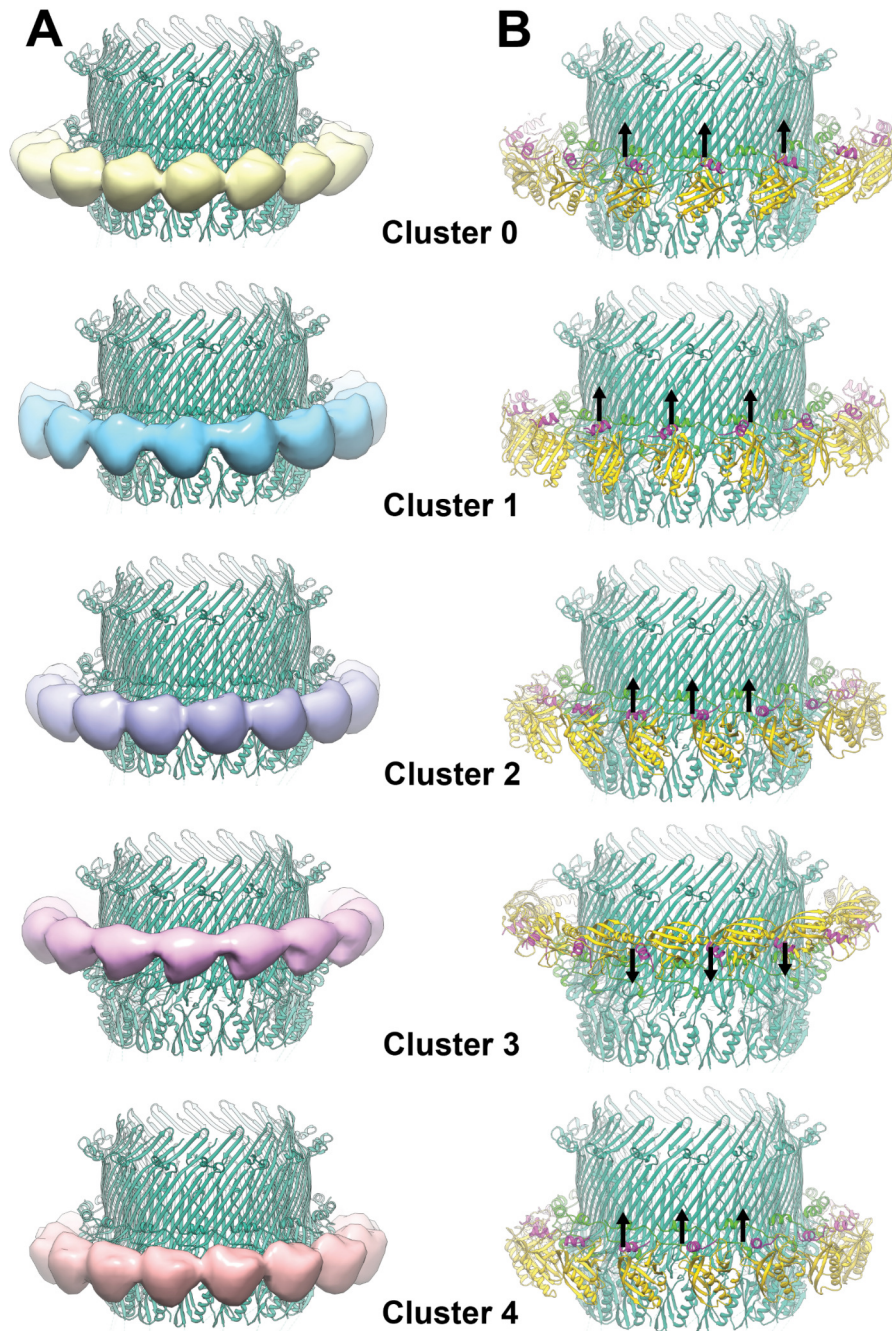
Supplementary figure 4: Docking of the pilotin based on cross-links of the isolated needle complex. The cryo-EM structure of the $MxiD_{180-548}$ OM ring (PDB ID: 8AXL[151]) in cartoon representation (green) with **(A)** localization probability densities (solid surfaces) representing the ensemble of the MxiM conformations of the five clusters. **(B)** Cartoon model of the respective clusters with a representative MxiM (yellow) in complex with $MxiD_{549-566}$ (purple) from [121], PDB ID: 2JW1. Arrows indicate the orientation of the lipidated MxiM N-terminus.

The docking of the pilotin was performed by Dr. Karol Kaszuba and will be published in [151].
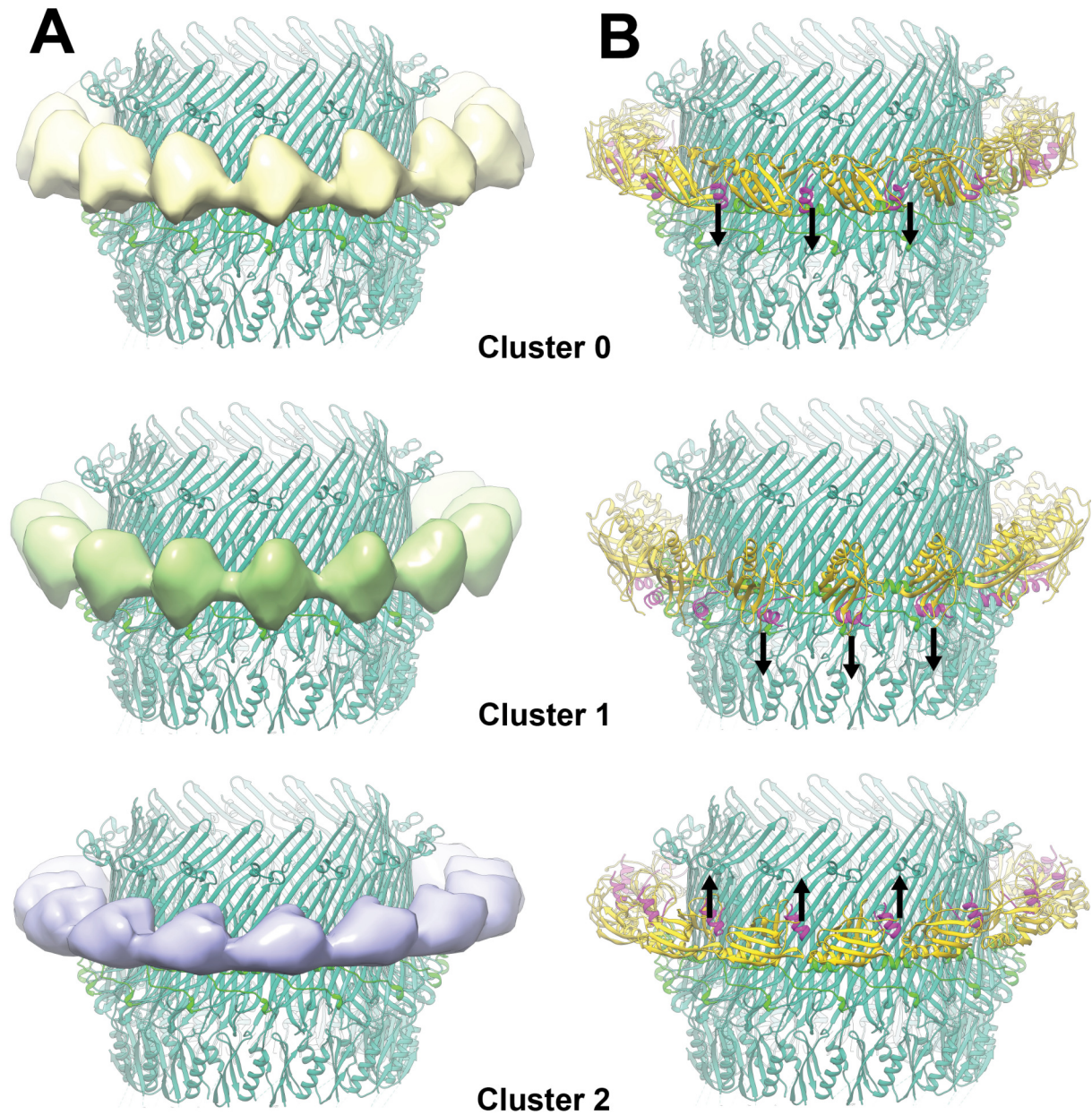
Supplementary figure 5: Docking of the pilotin based on cross-links and the cryo-EM map of the isolated needle complex. The cryo-EM structure of the MxiD$_{180-548}$ OM ring (PDB ID: 8AXL[151]) in cartoon representation (green) and **(A)** localization probability densities (solid surfaces) representing the ensemble of the MxiM conformations of the three clusters obtained by docking based on cross-links and the cryo-EM map. **(B)** Cartoon representation of the respective clusters with a representative MxiM conformation (yellow) in complex with MxiD$_{549-566}$ (purple) from [121], PDB ID 2JW1. Arrows indicate the orientation of the lipidated MxiM N-terminus.

The docking of the pilotin was performed by Dr. Karol Kaszuba and will be published in [151].
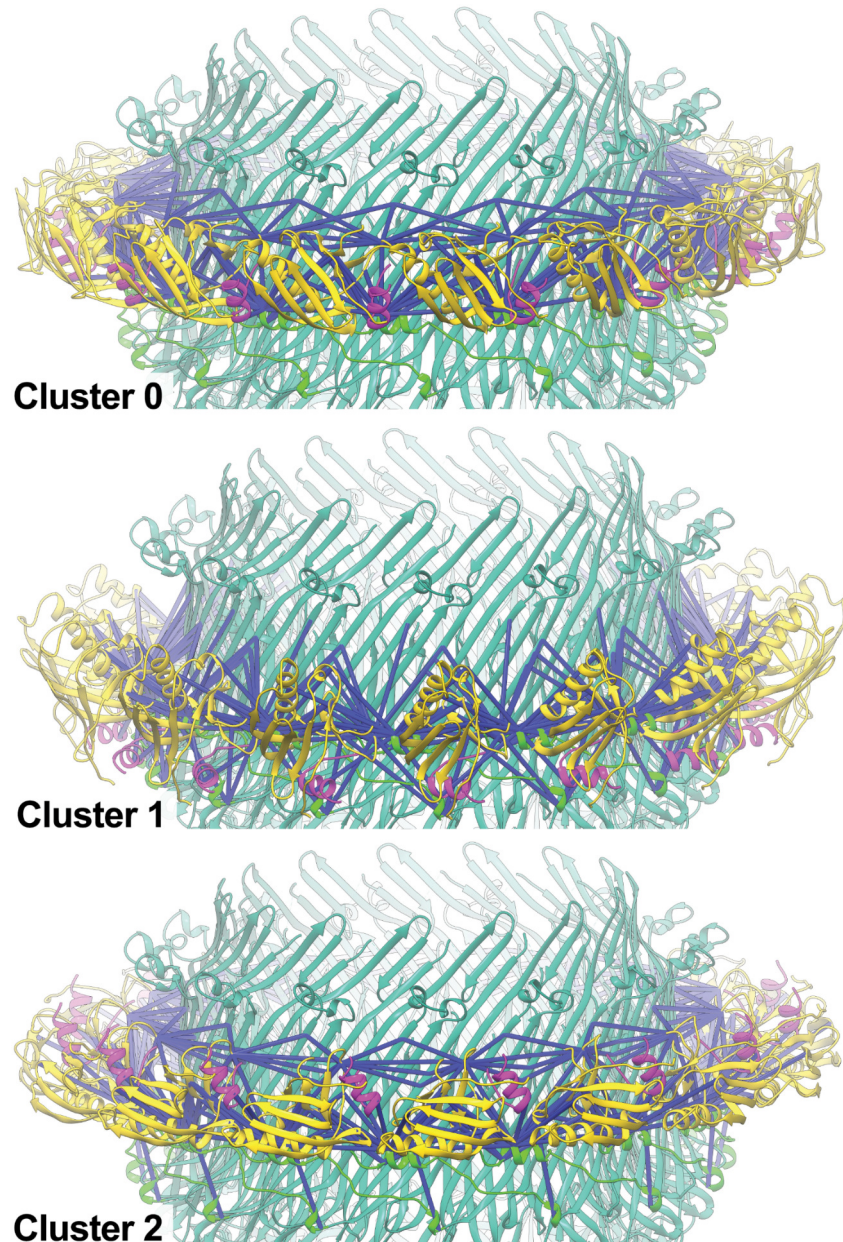
Supplementary figure 6: Docking of the pilotin based on the cross-links and cryo-EM map with highlighted cross-links (≤30 Å). Representative MxiM orientations of the three clusters obtained by docking based on cross-links and the cryo-EM map (Sup. Fig. 5). $MxiD_{180\text{-}548}$ from cryo-EM (PDB ID: 8AXL[151]) as cartoon (green) with docked MxiM (yellow) in complex with $MxiD_{549\text{-}566}$ (purple) from [121] (PDB ID: 2JW1). Cross-links satisfying the distance threshold of ≤30 Å are displayed in dark blue. The violated (>30Å) cross-links to more distant regions of MxiD are not shown for clarity.

The docking of the pilotin was performed by Dr. Karol Kaszuba and will be published in [151].

6.2. Supplementary tables

Supplementary table 1: Growth of the *S. flexneri* mutant under different conditions. Hourly measured optical densities at 600 nm of *S. flexneri* M90T Δ*ipaD* Δ*mxiH* complemented with *strep-mxiH* cultured in either TSB or LB media under inducing or non-inducing conditions. Average values and standard deviations were calculated from biological quadruplets.

Hours post-inoculation	OD_{600nm} in TSB					
	1	2	3	4	Average	Standard deviation
1	0.023	0.022	0.022	0.026	0.023	0.002
2	0.054	0.053	0.054	0.069	0.058	0.008
3	0.216	0.204	0.202	0.228	0.213	0.012
4	0.772	0.748	0.721	0.798	0.760	0.033
5	1.772	1.703	1.684	1.79	1.737	0.052

Hours post-inoculation	OD_{600nm} in LB					
	1	2	3	4	Average	Standard deviation
1	0.024	0.023	0.023	0.024	0.024	0.001
2	0.049	0.047	0.05	0.05	0.049	0.001
3	0.166	0.157	0.165	0.17	0.165	0.005
4	0.542	0.52	0.543	0.548	0.538	0.012
5	0.879	0.872	0.89	0.893	0.884	0.01

Hours post-inoculation	OD_{600nm} in TSB + induction					
	1	2	3	4	Average	Standard deviation
1	0.022	0.022	0.017	0.022	0.021	0.003
2	0.046	0.047	0.044	0.045	0.046	0.001
3	0.151	0.148	0.144	0.145	0.147	0.003
4	0.516	0.502	0.494	0.504	0.504	0.009
5	1.188	1.165	1.137	1.152	1.161	0.022

Hours post-inoculation	OD_{600nm} in LB + induction					
	1	2	3	4	Average	Standard deviation
1	0.023	0.023	0.026	0.026	0.025	0.002
2	0.041	0.041	0.041	0.043	0.042	0.001
3	0.122	0.121	0.124	0.131	0.125	0.005
4	0.38	0.372	0.375	0.395	0.381	0.01
5	0.705	0.702	0.706	0.725	0.710	0.01

Supplementary table 2: Theoretical masses of the *S. flexneri* M90T T3SS needle complex subunits. Masses were obtained from UniProtKnowledgebase [214]. If not noted otherwise, the numbers of copies are derived from single-particle cryo-EM analysis [33,151].

Component	UniProtKB accession number	Theoretical mass of single copy [Da]	Number of copies in complex	Total theoretical mass in complex [Da]
SpaP	P0A1L3	24 200	5	121 000
SpaQ	P0A1M4	9 429	4	37 716
SpaR	P0A1M6	28 498	1	28 498
SpaS	P0A1M8	39 852	1	39 852
MxiI	P0A225	10 633	6	63 798
MxiG	P0A221	43 002	24	1 032 048
MxiJ	Q06081	25 485*	24	611 640
MxiD (N0, N1)	Q04641	17 152*	16	274 432
MxiD (N3, Secretin, S)	Q04641	43 461	15	651 915
MxiM	P0A1X2	13 302*	15**	199 530
Basel body				**3 060 429**
MxiH	P0A223	11 050***	105****	1 160 250
Needle complex				**4 220 679**

* N-terminal signal sequence not included

** Based on their interaction, an equal number of MxiM to MxiD is assumed [121,151]

*** Includes strep-tag and cleavage site

**** The needle can vary in length [76,215]. The reported needle lenght of 45.4 nm [76], and the helical rise of 4.31 Å from the needle structure [216] were used to estimate an average amount of MxiH subunits per isolated needle complex.

Supplementary table 3: List of cross-links found in the isolated T3SS needle complex. The table will be published in [151]. Link type: Homomultimeric cross-link (homo), heteromeric cross-link (hetero) and self-cross-link (self). Used for modeling: true (TR) or false (FL). *The cross-linking MS analysis was performed by Dr. Zhuo Angel Chen and Dr. Francis O'Reilly.*

Protein 1	Protein name 1	SeqPos 1	LinkedRes 1	Protein 2	Protein name 2	SeqPos 2	LinkedRes 2	Highest score	Match count	Decoy type	Link type	Cα–Cα distance in the 3D model	From chain	To chain	Used for modeling
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	11	K	21.2	4	TT	Homo				FL
TagQCC34415.1	TagMxiH	84	K	TagQCC34415.1	TagMxiH	87	K	19.92	9	TT	Homo	15.91	u	v	TR
QCC34408.1	mxiD	364	K	QCC34408.1	mxiD	542	K	19.27	6	TT	Self	17.98	Z15	Z1	TR
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	353	K	19.1	5	TT	Self	12.26	X22	X22	TR
QCC34408.1	mxiD	364	K	QCC34408.1	mxiD	502	K	18.64	8	TT	Self	11.62	Z1	Z0	TR
QCC34410.1	mxiM	82	K	QCC34410.1	mxiM	99	K	18.24	7	TT	Self				TR
QCC34408.1	mxiD	187	K	TagQCC34415.1	TagMxiH	2	A	18	5	TT	Hetero				FL
QCC34413.1	mxiJ	143	K	QCC34413.1	mxiJ	178	K	17.96	6	TT	Self	11.26	Y5	Y5	TR
QCC34416.1	mxiG	252	K	QCC34416.1	mxiG	268	K	17.49	54	TT	Self	13.17	X11	X11	TR
QCC34408.1	mxiD	222	K	TagQCC34415.1	TagMxiH	2	A	17.45	4	TT	Hetero				FL
QCC34414.1	mxiI	1	M	QCC34414.1	mxiI	33	K	17.29	3	TT	Self				TR
QCC34408.1	mxiD	187	K	QCC34408.1	mxiD	222	K	17.29	3	TT	Self	26.33	Z14	Z13	TR
QCC34416.1	mxiG	252	K	TagQCC34415.1	TagMxiH	2	A	17.26	2	TT	Hetero				FL
QCC34413.1	mxiJ	49	K	QCC34413.1	mxiJ	54	K	16.71	4	TT	Homo	21.57	Y6	Y5	TR
QCC34413.1	mxiJ	54	K	QCC34413.1	mxiJ	54	K	16.59	8	TT	Homo	14.33	Y1	Y22	TR
QCC34416.1	mxiG	254	K	QCC34416.1	mxiG	268	K	16.55	9	TT	Self	8.75	X23	X23	TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	40	T	16.47	4	TT	Self				FL
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	54	K	16.44	7	TT	Self				FL
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	68	K	16.29	1	TT	Self				FL
QCC34408.1	mxiD	222	K	QCC34408.1	mxiD	521	K	16.06	2	TT	Self	46.24	Z10	Z10	TR
QCC34414.1	mxiI	1	M	TagQCC34415.1	TagMxiH	87	K	15.94	4	TT	Hetero				TR
QCC34408.1	mxiD	364	K	QCC34408.1	mxiD	558	K	15.93	3	TT	Self				TR
QCC34408.1	mxiD	362	K	QCC34408.1	mxiD	502	K	15.89	4	TT	Self	9.75	Z3	Z1	TR
TagQCC34415.1	TagMxiH	54	K	TagQCC34415.1	TagMxiH	68	K	15.85	1	TT	Homo	13.8	p	u	TR
QCC34408.1	mxiD	189	K	TagQCC34415.1	TagMxiH	2	A	15.85	6	TT	Hetero				FL
QCC34408.1	mxiD	187	K	QCC34408.1	mxiD	298	K	15.82	3	TT	Self	13.68	Z8	Z7	TR
QCC34410.1	mxiM	82	K	TagQCC34415.1	TagMxiH	2	A	15.8	4	TT	Hetero				FL
QCC34410.1	mxiM	99	K	TagQCC34415.1	TagMxiH	2	A	15.66	2	TT	Hetero				FL
QCC34408.1	mxiD	171	K	TagQCC34415.1	TagMxiH	2	A	15.54	2	TT	Hetero				FL
QCC34416.1	mxiG	5	K	QCC34416.1	mxiG	17	K	15.33	1	TT	Self				TR
QCC34408.1	mxiD	521	K	QCC34410.1	mxiM	52	K	15.29	2	TT	Hetero				TR
QCC34414.1	mxiI	33	K	QCC34414.1	mxiI	86	K	15.21	1	TT	Self	18.14	Q	Q	TR
QCC34416.1	mxiG	5	K	QCC34416.1	mxiG	73	K	15.1	5	TT	Self				TR
TagQCC34415.1	TagMxiH	54	K	TagQCC34415.1	TagMxiH	87	K	15.01	3	TT	Self	16.64	T	e	TR
QCC34414.1	mxiI	33	K	TagQCC34415.1	TagMxiH	2	A	14.99	3	TT	Hetero				FL
QCC34410.1	mxiM	121	K	TagQCC34415.1	TagMxiH	2	A	14.92	2	TT	Hetero				FL
QCC34408.1	mxiD	180	K	TagQCC34415.1	TagMxiH	2	A	14.86	4	TT	Hetero				FL
QCC34408.1	mxiD	189	K	TagQCC34415.1	TagMxiH	11	K	14.85	3	TT	Hetero				FL
QCC34416.1	mxiG	240	K	QCC34416.1	mxiG	268	K	14.63	2	TT	Self	14.7	X10	X10	TR
QCC34408.1	mxiD	502	K	QCC34410.1	mxiM	130	K	14.61	2	TT	Hetero				TR
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	222	K	14.57	4	TT	Self	31.85	Z14	Z13	TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	28	S	14.54	1	TT	Self				FL
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	19	T	14.53	2	TT	Self				FL
QCC34410.1	mxiM	115	K	QCC34410.1	mxiM	139	K	14.49	3	TT	Self				TR
QCC34408.1	mxiD	187	K	QCC34408.1	mxiD	521	K	14.44	2	TT	Self	22.31	Z1	Z0	TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	87	K	14.44	2	TT	Self				FL
QCC34408.1	mxiD	521	K	QCC34408.1	mxiD	542	K	14.42	1	TT	Self	22.96	Z8	Z9	TR
QCC34400.1	Spa24	105	K	QCC34414.1	mxiI	1	M	14.42	2	TT	Hetero				TR
QCC34414.1	mxiI	1	M	TagQCC34415.1	TagMxiH	84	K	14.42	1	TT	Hetero				TR
QCC34408.1	mxiD	171	K	QCC34408.1	mxiD	189	K	14.38	2	TT	Self	43.1	Z0	Z0	TR
QCC34414.1	mxiI	1	M	QCC34414.1	mxiI	86	K	14.35	2	TT	Self				TR
QCC34414.1	mxiI	25	S	TagQCC34415.1	TagMxiH	2	A	14.29	1	TT	Hetero				FL
QCC34408.1	mxiD	148	K	QCC34408.1	mxiD	189	K	14.22	1	TT	Self	47.55	Z13	Z14	TR
QCC34408.1	mxiD	542	K	QCC34410.1	mxiM	130	K	14.19	4	TT	Hetero				TR
QCC34408.1	mxiD	364	K	QCC34410.1	mxiM	130	K	14.19	4	TT	Hetero				TR
QCC34408.1	mxiD	148	K	QCC34408.1	mxiD	180	K	14.11	2	TT	Self	21.37	Z13	Z14	TR
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	362	K	14.02	3	TT	Self	34.25	X13	X13	TR
QCC34408.1	mxiD	456	K	TagQCC34415.1	TagMxiH	2	A	14.02	2	TT	Hetero				FL
QCC34408.1	mxiD	222	K	TagQCC34415.1	TagMxiH	54	K	14.01	1	TT	Hetero	6.71	Z15	T	TR
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	369	K	13.99	3	TT	Self				TR
QCC34408.1	mxiD	467	K	QCC34408.1	mxiD	521	K	13.96	1	TT	Self	12.56	Z1	Z3	TR
QCC34416.1	mxiG	2	S	QCC34416.1	mxiG	73	K	13.92	8	TT	Self				TR
QCC34410.1	mxiM	127	K	TagQCC34415.1	TagMxiH	2	A	13.89	1	TT	Hetero				FL
QCC34400.1	Spa24	105	K	TagQCC34415.1	TagMxiH	2	A	13.81	2	TT	Hetero				FL

QCC34408.1	mxiD	521	K	TagQCC34415.1	TagMxiH	2	A	13.81	2	TT	Hetero				FL
QCC34408.1	mxiD	222	K	QCC34410.1	mxiM	115	K	13.74	2	TT	Hetero				TR
QCC34408.1	mxiD	222	K	QCC34408.1	mxiD	298	K	13.72	1	TT	Self	34.75	Z0	Z0	TR
QCC34408.1	mxiD	180	K	QCC34408.1	mxiD	222	K	13.71	2	TT	Self	10.56	Z14	Z13	TR
QCC34416.1	mxiG	293	K	TagQCC34415.1	TagMxiH	2	A	13.71	1	TT	Hetero				FL
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	38	T	13.67	2	TT	Self				FL
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	65	Y	13.66	1	TT	Self				FL
QCC34408.1	mxiD	521	K	QCC34408.1	mxiD	558	K	13.65	1	TT	Self				TR
QCC34410.1	mxiM	99	K	QCC34410.1	mxiM	121	K	13.65	2	TT	Self				TR
QCC34408.1	mxiD	542	K	QCC34410.1	mxiM	77	K	13.64	1	TT	Hetero				TR
QCC34408.1	mxiD	298	K	TagQCC34415.1	TagMxiH	2	A	13.6	5	TT	Hetero				FL
QCC34414.1	mxiI	1	M	TagQCC34415.1	TagMxiH	54	K	13.58	2	TT	Hetero				TR
QCC34408.1	mxiD	64	K	QCC34416.1	mxiG	362	K	13.58	3	TT	Hetero	14.8	Z7	X0	TR
QCC34408.1	mxiD	222	K	QCC34414.1	mxiI	1	M	13.58	1	TT	Hetero				TR
TagQCC34415.1	TagMxiH	68	K	TagQCC34415.1	TagMxiH	87	K	13.57	1	TT	Self	10.81	b	q	TR
QCC34400.1	Spa24	126	K	QCC34416.1	mxiG	347	K	13.56	1	TT	Hetero				TR
QCC34408.1	mxiD	148	K	QCC34408.1	mxiD	521	K	13.46	1	TT	Self	59.43	Z15	Z15	TR
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	313	K	13.44	2	TT	Self	16.54	X5	X4	TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	84	K	13.42	4	TT	Self				FL
QCC34410.1	mxiM	77	K	TagQCC34415.1	TagMxiH	2	A	13.41	2	TT	Hetero				FL
QCC34408.1	mxiD	64	K	QCC34408.1	mxiD	521	K	13.4	1	TT	Self	92.38	Z1	Z0	TR
QCC34414.1	mxiI	1	M	QCC34414.1	mxiI	20	S	13.34	1	TT	Self				TR
QCC34416.1	mxiG	246	K	QCC34416.1	mxiG	252	K	13.33	6	TT	Homo	26.53	X16	X17	TR
QCC34413.1	mxiJ	178	K	QCC34416.1	mxiG	347	K	13.32	8	TT	Hetero	13.28	Y4	X8	TR
QCC34410.1	mxiM	121	K	QCC34410.1	mxiM	139	K	13.31	4	TT	Self				TR
QCC34408.1	mxiD	363	K	QCC34410.1	mxiM	130	K	13.17	1	TT	Hetero				TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	57	S	13.15	9	TT	Self				FL
QCC34414.1	mxiI	1	M	QCC34416.1	mxiG	293	K	13.13	2	TT	Hetero				TR
QCC34408.1	mxiD	64	K	TagQCC34415.1	TagMxiH	2	A	13.12	3	TT	Hetero				FL
QCC34408.1	mxiD	542	K	QCC34410.1	mxiM	115	K	13.11	2	TT	Hetero				TR
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	298	K	13.05	4	TT	Self	11	Z9	Z9	TR
QCC34410.1	mxiM	82	K	QCC34410.1	mxiM	130	K	13.02	1	TT	Self				TR
TagQCC34415.1	TagMxiH	84	K	TagQCC34415.1	TagMxiH	84	K	12.99	3	TT	Homo	15.63	s	t	TR
QCC34408.1	mxiD	521	K	QCC34410.1	mxiM	130	K	12.96	1	TT	Hetero				TR
QCC34416.1	mxiG	246	K	TagQCC34415.1	TagMxiH	2	A	12.94	2	TT	Hetero				FL
QCC34414.1	mxiI	16	S	TagQCC34415.1	TagMxiH	2	A	12.94	1	TT	Hetero				FL
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	521	K	12.91	2	TT	Self	19.73	Z9	Z9	TR
QCC34408.1	mxiD	558	K	QCC34410.1	mxiM	82	K	12.89	2	TT	Hetero				TR
QCC34410.1	mxiM	99	K	QCC34410.1	mxiM	130	K	12.87	1	TT	Self				TR
QCC34410.1	mxiM	91	K	QCC34410.1	mxiM	121	K	12.86	4	TT	Self				TR
QCC34408.1	mxiD	189	K	QCC34414.1	mxiI	1	M	12.79	2	TT	Hetero				TR
QCC34429.1	VirB	48	K	QCC34429.1	VirB	121	K	12.72	3	TT	Self				TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	51	K	12.7	1	TT	Self				FL
QCC34408.1	mxiD	298	K	QCC34408.1	mxiD	521	K	12.7	3	TT	Self	12.32	Z1	Z1	TR
QCC34416.1	mxiG	246	K	QCC34416.1	mxiG	268	K	12.69	5	TT	Self	15.29	X22	X22	TR
QCC34408.1	mxiD	502	K	QCC34408.1	mxiD	542	K	12.62	1	TT	Self	26.11	Z14	Z1	TR
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	355	S	12.61	2	TT	Self	12.01	X11	X11	TR
QCC34408.1	mxiD	542	K	QCC34410.1	mxiM	146	K	12.6	3	TT	Hetero				TR
QCC34416.1	mxiG	2	S	QCC34416.1	mxiG	72	S	12.55	3	TT	Self				TR
QCC34410.1	mxiM	115	K	TagQCC34415.1	TagMxiH	2	A	12.51	3	TT	Hetero				FL
QCC34408.1	mxiD	502	K	QCC34408.1	mxiD	558	K	12.49	2	TT	Self				TR
QCC34416.1	mxiG	345	K	QCC34416.1	mxiG	347	K	12.47	2	TT	Homo	5.76	X7	X8	TR
QCC34410.1	mxiM	77	K	QCC34410.1	mxiM	130	K	12.43	1	TT	Self				TR
QCC34416.1	mxiG	268	K	TagQCC34415.1	TagMxiH	2	A	12.41	1	TT	Hetero				FL
QCC34416.1	mxiG	265	K	TagQCC34415.1	TagMxiH	2	A	12.35	2	TT	Hetero				FL
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	542	K	12.35	1	TT	Self	33.81	Z1	Z3	TR
QCC34408.1	mxiD	542	K	QCC34410.1	mxiM	52	K	12.26	1	TT	Hetero				TR
QCC34429.1	VirB	48	K	QCC34429.1	VirB	70	K	12.25	1	TT	Self				TR
QCC34429.1	VirB	48	K	QCC34429.1	VirB	157	K	12.22	2	TT	Self				TR
QCC34408.1	mxiD	148	K	QCC34408.1	mxiD	222	K	12.15	1	TT	Self	23.59	Z9	Z9	TR
QCC34408.1	mxiD	542	K	TagQCC34415.1	TagMxiH	2	A	12.15	2	TT	Hetero				FL
QCC34414.1	mxiI	33	K	QCC34416.1	mxiG	347	K	12.11	1	TT	Hetero	33.05	N	X12	TR
QCC34414.1	mxiI	1	M	QCC34414.1	mxiI	16	S	12.07	1	TT	Self				TR
QCC34414.1	mxiI	24	S	TagQCC34415.1	TagMxiH	2	A	12.07	1	TT	Hetero				FL
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	67	S	12.07	1	TT	Self				FL
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	247	S	12.05	2	TT	Self				TR
QCC34413.1	mxiJ	143	K	QCC34416.1	mxiG	347	K	12.04	4	TT	Hetero	18.89	Y16	X19	TR
QCC34410.1	mxiM	115	K	QCC34410.1	mxiM	144	K	12	8	TT	Self				TR
QCC34408.1	mxiD	456	K	TagQCC34415.1	TagMxiH	54	K	11.96	1	TT	Hetero	12.93	Z0	d	TR
QCC34416.1	mxiG	1	M	QCC34416.1	mxiG	7	S	11.95	2	TT	Homo				TR
TagQCC34415.1	TagMxiH	68	K	TagQCC34415.1	TagMxiH	84	K	11.92	1	TT	Self	11.85	S	k	TR
QCC34408.1	mxiD	363	K	QCC34408.1	mxiD	502	K	11.92	3	TT	Self	9.9	Z3	Z1	TR
QCC34410.1	mxiM	82	K	QCC34410.1	mxiM	115	K	11.9	1	TT	Self				TR
QCC34408.1	mxiD	175	S	TagQCC34415.1	TagMxiH	2	A	11.9	1	TT	Hetero				FL
QCC34416.1	mxiG	240	K	QCC34416.1	mxiG	252	K	11.87	7	TT	Homo	21.1	X23	X20	TR
QCC34429.1	VirB	17	K	QCC34429.1	VirB	130	K	11.76	6	TT	Self				TR
QCC34416.1	mxiG	151	K	QCC34416.1	mxiG	195	K	11.71	1	TT	Self	9.97	X19	X19	TR
QCC34410.1	mxiM	77	K	QCC34410.1	mxiM	99	K	11.68	4	TT	Self				TR
QCC34414.1	mxiI	46	S	TagQCC34415.1	TagMxiH	2	A	11.67	1	TT	Hetero				FL
QCC34408.1	mxiD	187	K	QCC34408.1	mxiD	189	K	11.67	3	TT	Homo	20.15	Z8	Z7	TR

QCC34410.1	mxiM	115	K	QCC34410.1	mxiM	130	K	11.62	2	TT	Self				TR
QCC34416.1	mxiG	347	K	TagQCC34415.1	TagMxiH	11	K	11.61	2	TT	Hetero				FL
QCC34416.1	mxiG	5	K	QCC34416.1	mxiG	72	S	11.6	2	TT	Self				TR
QCC34410.1	mxiM	99	K	QCC34410.1	mxiM	115	K	11.55	3	TT	Self				TR
TagQCC34415.1	TagMxiH	54	K	TagQCC34415.1	TagMxiH	84	K	11.48	2	TT	Self	19.77	i	d	TR
QCC34416.1	mxiG	282	K	QCC34416.1	mxiG	293	K	11.46	2	TT	Self	10.05	X2	X2	TR
TagQCC34415.1	TagMxiH	11	K	TagQCC34415.1	TagMxiH	54	K	11.45	4	TT	Self				FL
TagQCC34415.1	TagMxiH	40	T	TagQCC34415.1	TagMxiH	54	K	11.43	1	TT	Self	9.57	o	c	TR
QCC34414.1	mxiI	1	M	QCC34414.1	mxiI	31	S	11.38	1	TT	Self				TR
QCC34416.1	mxiG	347	K	QCC34416.1	mxiG	362	K	11.38	2	TT	Self	11.12	X10	X10	TR
QCC34408.1	mxiD	542	K	QCC34410.1	mxiM	127	K	11.26	5	TT	Hetero				TR
QCC34429.1	VirB	31	K	QCC34429.1	VirB	130	K	11.2	5	TT	Self				TR
QCC34414.1	mxiI	33	K	TagQCC34415.1	TagMxiH	11	K	11.17	1	TT	Hetero				FL
QCC34413.1	mxiJ	201	K	QCC34416.1	mxiG	195	K	11.14	3	TT	Hetero				TR
QCC34408.1	mxiD	558	K	QCC34410.1	mxiM	130	K	11.13	2	TT	Hetero				TR
QCC34408.1	mxiD	64	K	TagQCC34415.1	TagMxiH	11	K	11.13	1	TT	Hetero				FL
QCC34408.1	mxiD	64	K	QCC34408.1	mxiD	189	K	11.09	1	TT	Self	79.39	Z0	Z0	TR
QCC34429.1	VirB	17	K	QCC34429.1	VirB	31	K	11.04	2	TT	Self				TR
QCC34408.1	mxiD	369	S	QCC34408.1	mxiD	501	Y	11.02	1	TT	Self	28.26	Z1	Z0	TR
QCC34408.1	mxiD	369	S	QCC34408.1	mxiD	502	K	11.02	3	TT	Self	25	Z1	Z0	TR
QCC34414.1	mxiI	31	S	TagQCC34415.1	TagMxiH	2	A	11.01	1	TT	Hetero				FL
QCC34408.1	mxiD	502	K	QCC34410.1	mxiM	43	K	11	1	TT	Hetero				TR
QCC34416.1	mxiG	268	K	QCC34416.1	mxiG	282	K	10.93	2	TT	Self	29.18	X2	X2	TR
QCC34406.1	mxiA	151	K	QCC34406.1	mxiA	170	K	10.9	2	TT	Self				TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	72	Y	10.87	1	TT	Self				FL
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	558	K	10.82	2	TT	Self				TR
QCC34414.1	mxiI	86	K	TagQCC34415.1	TagMxiH	84	K	10.82	1	TT	Hetero	14.16	R	S	TR
QCC34416.1	mxiG	240	K	QCC34416.1	mxiG	246	K	10.81	2	TT	Homo	23.12	X16	X17	TR
QCC34408.1	mxiD	364	K	TagQCC34415.1	TagMxiH	2	A	10.78	1	TT	Hetero				FL
QCC34410.1	mxiM	52	K	QCC34410.1	mxiM	115	K	10.77	1	TT	Self				TR
QCC34416.1	mxiG	254	K	TagQCC34415.1	TagMxiH	2	A	10.77	1	TT	Hetero				FL
QCC34408.1	mxiD	187	K	TagQCC34415.1	TagMxiH	11	K	10.73	2	TT	Hetero				FL
QCC34408.1	mxiD	362	K	QCC34410.1	mxiM	130	K	10.64	3	TT	Hetero				TR
QCC34414.1	mxiI	33	K	TagQCC34415.1	TagMxiH	54	K	10.62	1	TT	Hetero	42.37	R	S	TR
QCC34410.1	mxiM	91	K	QCC34410.1	mxiM	127	K	10.62	3	TT	Self				TR
QCC34408.1	mxiD	180	K	TagQCC34415.1	TagMxiH	54	K	10.6	1	TT	Hetero	15.58	Z0	T	TR
QCC34408.1	mxiD	558	K	QCC34410.1	mxiM	77	K	10.59	2	TT	Hetero				TR
QCC34408.1	mxiD	64	K	QCC34414.1	mxiI	1	M	10.56	2	TT	Hetero				TR
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	347	K	10.56	2	TT	Self	20.15	X5	X5	TR
QCC34408.1	mxiD	189	K	TagQCC34415.1	TagMxiH	54	K	10.53	1	TT	Hetero	24.23	Z1	j	TR
QCC34408.1	mxiD	187	K	QCC34408.1	mxiD	187	K	10.49	2	TT	Homo	23.2	Z8	Z7	TR
QCC34410.1	mxiM	115	K	QCC34410.1	mxiM	127	K	10.33	2	TT	Self				TR
QCC34416.1	mxiG	195	K	QCC34416.1	mxiG	265	K	10.32	2	TT	Self	21.13	X8	X8	TR
QCC34413.1	mxiJ	63	K	QCC34413.1	mxiJ	201	K	10.21	1	TT	Self				TR
QCC34414.1	mxiI	1	M	QCC34416.1	mxiG	347	K	10.19	2	TT	Hetero				TR
QCC34408.1	mxiD	64	K	QCC34408.1	mxiD	222	K	10.13	1	TT	Self	54.35	Z12	Z12	TR
QCC34414.1	mxiI	86	K	QCC34414.1	mxiI	96	K	10.13	1	TT	Homo	14.02	M	R	TR
QCC34416.1	mxiG	271	Y	QCC34416.1	mxiG	282	K	10.12	1	TT	Self	20.72	X10	X10	TR
QCC34416.1	mxiG	265	K	QCC34416.1	mxiG	282	K	10.1	1	TT	Self	35.65	X20	X23	TR
QCC34416.1	mxiG	240	K	QCC34416.1	mxiG	273	S	10.07	2	TT	Self	5.43	X16	X16	TR
QCC34416.1	mxiG	5	K	QCC34416.1	mxiG	62	K	10.04	2	TT	Self				TR
QCC34406.1	mxiA	170	K	QCC34416.1	mxiG	5	K	10.04	2	TT	Hetero				TR
QCC34414.1	mxiI	14	K	TagQCC34415.1	TagMxiH	2	A	10	1	TT	Hetero				FL
TagQCC34415.1	TagMxiH	11	K	TagQCC34415.1	TagMxiH	87	K	10	1	TT	Self				FL
QCC34408.1	mxiD	521	K	QCC34410.1	mxiM	77	K	9.99	1	TT	Hetero				TR
QCC34408.1	mxiD	542	K	QCC34410.1	mxiM	91	K	9.77	1	TT	Hetero				TR
QCC34410.1	mxiM	77	K	QCC34410.1	mxiM	115	K	9.73	1	TT	Self				TR
QCC34408.1	mxiD	148	K	TagQCC34415.1	TagMxiH	2	A	9.71	1	TT	Hetero				FL
QCC34408.1	mxiD	558	K	TagQCC34415.1	TagMxiH	2	A	9.7	2	TT	Hetero				FL
QCC34429.1	VirB	70	K	QCC34429.1	VirB	157	K	9.68	1	TT	Self				TR
QCC34408.1	mxiD	189	K	QCC34410.1	mxiM	115	K	9.64	1	TT	Hetero				TR
QCC34429.1	VirB	70	K	QCC34429.1	VirB	130	K	9.64	1	TT	Self				TR
QCC34406.1	mxiA	170	K	QCC34413.1	mxiJ	54	K	9.55	2	TT	Hetero				TR
QCC34416.1	mxiG	268	K	QCC34416.1	mxiG	287	K	9.55	2	TT	Self	28.14	X20	X23	TR
QCC34408.1	mxiD	502	K	QCC34410.1	mxiM	77	K	9.49	1	TT	Hetero				TR
QCC34408.1	mxiD	521	K	QCC34414.1	mxiI	1	M	9.45	1	TT	Hetero				TR
QCC34410.1	mxiM	130	K	TagQCC34415.1	TagMxiH	2	A	9.42	1	TT	Hetero				FL
QCC34408.1	mxiD	171	K	TagQCC34415.1	TagMxiH	11	K	9.39	1	TT	Hetero				FL
QCC34416.1	mxiG	160	K	QCC34416.1	mxiG	265	K	9.36	1	TT	Self	16.67	X22	X22	TR
QCC34413.1	mxiJ	143	K	TagQCC34415.1	TagMxiH	11	K	9.34	1	TT	Hetero				FL
QCC34408.1	mxiD	222	K	QCC34408.1	mxiD	254	S	9.26	1	TT	Self				TR
QCC34408.1	mxiD	566	Y	TagQCC34415.1	TagMxiH	2	A	9.24	1	TT	Hetero				FL
QCC34413.1	mxiJ	178	K	TagQCC34415.1	TagMxiH	11	K	9.23	1	TT	Hetero				FL
QCC34416.1	mxiG	300	K	QCC34416.1	mxiG	313	K	9.21	1	TT	Self	16.2	X4	X3	TR
QCC34406.1	mxiA	170	K	QCC34416.1	mxiG	17	K	9.19	1	TT	Hetero				TR
QCC34408.1	mxiD	191	T	TagQCC34415.1	TagMxiH	2	A	9.17	2	TT	Hetero				FL
QCC34424.1	ipaB	157	K	QCC34424.1	ipaB	188	K	9.12	2	TT	Self				TR
QCC34406.1	mxiA	170	K	QCC34416.1	mxiG	268	K	9.03	1	TT	Hetero				TR
QCC34408.1	mxiD	456	K	QCC34408.1	mxiD	521	K	9.02	1	TT	Self	25	Z13	Z14	TR
QCC34410.1	mxiM	82	K	QCC34410.1	mxiM	127	K	8.97	1	TT	Self				TR

ID 1	Name 1	Pos 1	Res 1	ID 2	Name 2	Pos 2	Res 2	Score	N	TT	Type	Val	C1	C2	TR/FL
QCC34408.1	mxiD	521	K	QCC34410.1	mxiM	127	K	8.96	2	TT	Hetero				TR
QCC34408.1	mxiD	64	K	QCC34414.1	mxiI	33	K	8.95	1	TT	Hetero	16.18	Z6	P	TR
QCC34429.1	VirB	121	K	QCC34429.1	VirB	157	K	8.91	1	TT	Self				TR
QCC34410.1	mxiM	91	K	QCC34410.1	mxiM	139	K	8.88	1	TT	Self				TR
QCC34413.1	mxiJ	54	K	QCC34413.1	mxiJ	58	S	8.78	3	TT	Homo	16.15	Y5	Y6	TR
QCC34408.1	mxiD	362	K	QCC34408.1	mxiD	558	K	8.78	2	TT	Self				TR
QCC34400.1	Spa24	105	K	QCC34416.1	mxiG	362	K	8.76	2	TT	Hetero	16.69	B	X0	TR
QCC34408.1	mxiD	180	K	QCC34408.1	mxiD	189	K	8.74	1	TT	Self	28.32	Z13	Z13	TR
QCC34400.1	Spa24	84	T	TagQCC34415.1	TagMxiH	87	K	8.7	1	TT	Hetero				TR
TagQCC34415.1	TagMxiH	54	K	TagQCC34415.1	TagMxiH	70	S	8.67	1	TT	Self	14.96	p	u	TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	17	S	8.62	1	TT	Self				FL
QCC34408.1	mxiD	222	K	TagQCC34415.1	TagMxiH	11	K	8.62	2	TT	Hetero				FL
QCC34408.1	mxiD	174	S	TagQCC34415.1	TagMxiH	2	A	8.61	1	TT	Hetero				FL
QCC34416.1	mxiG	5	K	QCC34416.1	mxiG	7	S	8.6	1	TT	Homo				TR
QCC34406.1	mxiA	170	K	QCC34416.1	mxiG	195	K	8.58	1	TT	Hetero				TR
QCC34413.1	mxiJ	133	Y	QCC34413.1	mxiJ	143	K	8.56	2	TT	Self	10.25	Y22	Y22	TR
QCC34408.1	mxiD	298	K	TagQCC34415.1	TagMxiH	11	K	8.55	3	TT	Hetero				FL
QCC34408.1	mxiD	247	S	TagQCC34415.1	TagMxiH	2	A	8.52	1	TT	Hetero				FL
QCC34408.1	mxiD	180	K	TagQCC34415.1	TagMxiH	11	K	8.45	1	TT	Hetero				FL
TagQCC34415.1	TagMxiH	54	K	TagQCC34415.1	TagMxiH	72	Y	8.38	1	TT	Self	11.89	p	u	TR
QCC34416.1	mxiG	273	S	QCC34416.1	mxiG	282	K	8.25	2	TT	Self	14.67	X2	X2	TR
QCC34408.1	mxiD	38	Y	QCC34416.1	mxiG	293	K	8.21	7	TT	Hetero	22.55	Z10	X3	TR
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	251	S	8.2	1	TT	Self				TR
QCC34400.1	Spa24	100	K	TagQCC34415.1	TagMxiH	2	A	8.2	1	TT	Hetero				FL
QCC34410.1	mxiM	75	Y	QCC34410.1	mxiM	115	K	8.13	1	TT	Self				TR
QCC34429.1	VirB	130	K	QCC34429.1	VirB	157	K	8.13	1	TT	Self				TR
QCC34400.1	Spa24	86	S	TagQCC34415.1	TagMxiH	2	A	8.08	1	TT	Hetero				FL
QCC34410.1	mxiM	121	K	QCC34410.1	mxiM	130	K	8.07	1	TT	Self				TR
QCC34408.1	mxiD	64	K	QCC34410.1	mxiM	115	K	7.95	1	TT	Hetero				TR
TagQCC34415.1	TagMxiH	72	Y	TagQCC34415.1	TagMxiH	84	K	7.94	1	TT	Self	13.74	f	m	TR
QCC34400.1	Spa24	100	K	QCC34408.1	mxiD	64	K	7.91	1	TT	Hetero	13.32	A	Z10	TR
QCC34408.1	mxiD	222	K	QCC34408.1	mxiD	558	K	7.9	2	TT	Self				TR
QCC34410.1	mxiM	77	K	QCC34410.1	mxiM	127	K	7.81	1	TT	Self				TR
TagQCC34415.1	TagMxiH	38	T	TagQCC34415.1	TagMxiH	54	K	7.8	1	TT	Self	5.2	u	p	TR
QCC34408.1	mxiD	64	K	QCC34408.1	mxiD	64	K	7.77	2	TT	Homo	18.22	Z12	Z11	TR
QCC34408.1	mxiD	523	S	TagQCC34415.1	TagMxiH	2	A	7.72	3	TT	Hetero				FL
QCC34408.1	mxiD	180	K	QCC34408.1	mxiD	187	K	7.71	1	TT	Homo	29.76	Z8	Z7	TR
QCC34410.1	mxiM	118	S	QCC34410.1	mxiM	121	K	7.56	1	TT	Homo				TR
QCC34408.1	mxiD	533	S	QCC34408.1	mxiD	542	K	7.56	1	TT	Self	14.94	Z1	Z3	TR
QCC34400.1	Spa24	100	K	QCC34414.1	mxiI	33	K	7.55	1	TT	Hetero	19.95	E	N	TR
QCC34408.1	mxiD	175	S	QCC34408.1	mxiD	222	K	7.47	1	TT	Self				TR
QCC34408.1	mxiD	38	Y	QCC34416.1	mxiG	300	K	7.46	4	TT	Hetero	33.46	Z15	X11	TR
QCC34406.1	mxiA	170	K	QCC34413.1	mxiJ	49	K	7.45	1	TT	Hetero				TR
QCC34410.1	mxiM	82	K	QCC34410.1	mxiM	121	K	7.37	1	TT	Self				TR
QCC34408.1	mxiD	544	T	QCC34410.1	mxiM	144	K	7.28	3	TT	Hetero				TR
QCC34408.1	mxiD	542	K	QCC34408.1	mxiD	549	S	7.26	1	TT	Homo				TR
QCC34408.1	mxiD	362	K	QCC34408.1	mxiD	500	K	7.25	1	TT	Self	11.17	Z1	Z0	TR
QCC34416.1	mxiG	160	K	QCC34416.1	mxiG	195	K	7.24	1	TT	Self	11.49	X23	X20	TR
QCC34408.1	mxiD	542	K	QCC34408.1	mxiD	558	K	7.11	1	TT	Homo				TR
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	352	S	7	2	TT	Self	15.47	X22	X22	TR
QCC34414.1	mxiI	38	K	TagQCC34415.1	TagMxiH	2	A	6.96	1	TT	Hetero				FL
QCC34413.1	mxiJ	49	K	QCC34413.1	mxiJ	58	S	6.94	1	TT	Homo	14.3	Y22	Y1	TR
QCC34408.1	mxiD	259	S	TagQCC34415.1	TagMxiH	2	A	6.87	1	TT	Hetero				FL
QCC34413.1	mxiJ	49	K	QCC34413.1	mxiJ	49	K	6.87	1	TT	Homo	18.21	Y1	Y22	TR
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	26	T	6.78	1	TT	Self				FL
QCC34400.1	Spa24	84	T	TagQCC34415.1	TagMxiH	2	A	6.74	1	TT	Hetero				FL
QCC34408.1	mxiD	369	S	QCC34408.1	mxiD	542	K	6.72	2	TT	Self	14.39	Z1	Z4	TR
QCC34408.1	mxiD	130	K	TagQCC34415.1	TagMxiH	2	A	6.68	1	TT	Hetero				FL
QCC34400.1	Spa24	105	K	QCC34400.1	Spa24	126	K	6.67	1	TT	Self				TR
QCC34414.1	mxiI	46	S	TagQCC34415.1	TagMxiH	11	K	6.67	1	TT	Hetero				FL
QCC34408.1	mxiD	364	K	QCC34410.1	mxiM	77	K	6.64	1	TT	Hetero				TR
TagQCC34415.1	TagMxiH	11	K	TagQCC34415.1	TagMxiH	57	S	6.57	3	TT	Self				FL
QCC34414.1	mxiI	1	M	QCC34414.1	mxiI	25	S	6.57	1	TT	Self				TR
QCC34400.1	Spa24	126	K	TagQCC34415.1	TagMxiH	2	A	6.42	1	TT	Hetero				FL
QCC34416.1	mxiG	252	K	QCC34416.1	mxiG	265	K	6.42	1	TT	Self	15.64	X4	X4	TR
QCC34408.1	mxiD	369	S	QCC34408.1	mxiD	543	T	6.4	1	TT	Self	10.61	Z1	Z4	TR
QCC34408.1	mxiD	64	K	QCC34408.1	mxiD	68	K	6.38	2	TT	Homo	17.03	Z4	Z3	TR
QCC34408.1	mxiD	544	T	QCC34410.1	mxiM	146	K	6.32	5	TT	Hetero				TR
QCC34408.1	mxiD	543	T	QCC34410.1	mxiM	144	K	6.32	2	TT	Hetero				TR
QCC34408.1	mxiD	543	T	QCC34410.1	mxiM	146	K	6.32	3	TT	Hetero				TR
QCC34410.1	mxiM	121	K	QCC34410.1	mxiM	127	K	6.25	1	TT	Homo				TR
QCC34407.1	MxiC	66	K	QCC34416.1	mxiG	268	K	6.25	2	TT	Hetero				TR
QCC34408.1	mxiD	544	T	QCC34410.1	mxiM	139	K	6.22	3	TT	Hetero				TR
QCC34406.1	mxiA	170	K	QCC34416.1	mxiG	254	K	6.17	1	TT	Hetero				TR
QCC34414.1	mxiI	1	M	QCC34416.1	mxiG	282	K	6.13	1	TT	Hetero				TR
QCC34416.1	mxiG	353	K	QCC34416.1	mxiG	369	K	6.12	1	TT	Self				TR
QCC34408.1	mxiD	523	S	TagQCC34415.1	TagMxiH	11	K	6.09	2	TT	Hetero				FL
QCC34408.1	mxiD	187	K	TagQCC34415.1	TagMxiH	54	K	6.04	1	TT	Hetero	22.12	Z1	j	TR
QCC34408.1	mxiD	148	K	QCC34408.1	mxiD	187	K	6.04	1	TT	Self	42.55	Z13	Z14	TR
QCC34416.1	mxiG	204	K	QCC34416.1	mxiG	246	K	5.97	1	TT	Self	29.15	X9	X9	FL

QCC34420.1	icsB	179	Y	TagQCC34415.1	TagMxiH	1	M	5.91	2	TT	Hetero				FL
TagQCC34415.1	TagMxiH	2	A	TagQCC34415.1	TagMxiH	82	T	5.89	1	TT	Self				FL
QCC34410.1	mxiM	127	K	QCC34410.1	mxiM	146	K	5.8	2	TT	Self				FL
QCC34406.1	mxiA	543	S	QCC34419.1	ipgD	40	K	5.8	2	TT	Hetero				FL
QCC34408.1	mxiD	363	K	QCC34408.1	mxiD	542	K	5.78	2	TT	Self	21.69	Z15	Z1	FL
QCC34416.1	mxiG	195	K	QCC34416.1	mxiG	293	K	5.71	1	TT	Self	54.65	X20	X23	FL
TagQCC34415.1	TagMxiH	11	K	TagQCC34415.1	TagMxiH	82	T	5.71	2	TT	Self				FL
QCC34429.1	VirB	116	T	TagQCC34415.1	TagMxiH	1	M	5.7	3	TT	Hetero				FL
QCC34410.1	mxiM	134	T	QCC34410.1	mxiM	144	K	5.62	1	TT	Homo				FL
QCC34408.1	mxiD	543	T	QCC34410.1	mxiM	139	K	5.55	2	TT	Hetero				FL
QCC34408.1	mxiD	222	K	QCC34408.1	mxiD	467	K	5.5	1	TT	Self	51.06	Z1	Z0	FL
QCC34416.1	mxiG	282	K	QCC34416.1	mxiG	300	K	5.34	1	TT	Self	9.43	X23	X20	FL
TagQCC34415.1	TagMxiH	11	K	TagQCC34415.1	TagMxiH	33	T	5.22	1	TT	Self				FL
QCC34408.1	mxiD	187	K	QCC34408.1	mxiD	198	Y	5.2	1	TT	Homo	22.54	Z8	Z7	FL
QCC34416.1	mxiG	282	K	QCC34416.1	mxiG	292	Y	5.19	1	TT	Self	8.72	X1	X1	FL
QCC34416.1	mxiG	246	K	QCC34416.1	mxiG	293	K	5.15	1	TT	Self	35.61	X20	X20	FL
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	356	Y	5.08	1	TT	Self	10.87	X11	X11	FL
QCC34408.1	mxiD	148	K	QCC34408.1	mxiD	174	S	5.06	1	TT	Self				FL
QCC34416.1	mxiG	304	S	QCC34416.1	mxiG	353	K	5.05	1	TT	Self	11.78	X2	X2	FL
QCC34410.1	mxiM	99	K	QCC34410.1	mxiM	127	K	4.96	1	TT	Self				FL
QCC34410.1	mxiM	127	K	QCC34410.1	mxiM	144	K	4.94	2	TT	Self				FL
QCC34416.1	mxiG	293	K	QCC34416.1	mxiG	304	S	4.89	2	TT	Self	5.21	X15	X15	FL
QCC34413.1	mxiJ	178	K	QCC34416.1	mxiG	353	K	4.87	2	TT	Hetero	18.92	Y13	X17	FL
QCC34410.1	mxiM	70	K	QCC34410.1	mxiM	130	K	4.82	1	TT	Self				FL
QCC34416.1	mxiG	245	S	QCC34416.1	mxiG	248	T	4.67	1	TT	Self	8.15	X17	X17	FL
QCC34429.1	VirB	48	K	QCC34429.1	VirB	145	K	4.61	1	TT	Self				FL
QCC34416.1	mxiG	282	K	QCC34416.1	mxiG	353	K	4.49	1	TT	Self	18.6	X4	X5	FL
QCC34408.1	mxiD	189	K	QCC34408.1	mxiD	467	K	4.32	1	TT	Self	19.46	Z1	Z0	FL
QCC34417.1	ipgF	130	Y	TagQCC34415.1	TagMxiH	1	M	4.29	5	TT	Hetero				FL
QCC34416.1	mxiG	310	T	QCC34416.1	mxiG	347	K	4.22	1	TT	Self	14.95	X4	X5	FL
QCC34410.1	mxiM	76	S	QCC34410.1	mxiM	115	K	4.2	1	TT	Self				FL
QCC34406.1	mxiA	641	K	QCC34424.1	ipaB	289	K	4.13	9	TT	Hetero				FL
QCC34406.1	mxiA	641	K	QCC34424.1	ipaB	290	S	4.13	7	TT	Hetero				FL
QCC34424.1	ipaB	172	S	TagQCC34415.1	TagMxiH	1	M	4.08	3	TT	Hetero				FL
QCC34425.1	ipaC	195	S	QCC34425.1	ipaC	238	S	3.8	1	TT	Self				FL
TagQCC34415.1	TagMxiH	72	Y	TagQCC34415.1	TagMxiH	87	K	3.71	1	TT	Self	10.73	U	i	FL
QCC34416.1	mxiG	292	Y	QCC34416.1	mxiG	304	S	3.68	1	TT	Self	5.55	X6	X6	FL
QCC34420.1	icsB	457	K	QCC34420.1	icsB	457	K	3.67	1	TT	Homo				FL
QCC34397.1	Spa40	229	K	QCC34397.1	Spa40	245	K	3.67	1	TT	Self				FL
QCC34427.1	ipaA	36	S	QCC34427.1	ipaA	231	S	3.6	1	TT	Self				FL
QCC34420.1	icsB	61	K	QCC34420.1	icsB	229	K	3.52	1	TT	Self				FL
QCC34420.1	icsB	61	K	QCC34420.1	icsB	228	K	3.52	1	TT	Self				FL
TagQCC34415.1	TagMxiH	11	K	TagQCC34415.1	TagMxiH	84	K	3.49	2	TT	Self				FL
QCC34424.1	ipaB	157	K	QCC34424.1	ipaB	429	K	3.48	1	TT	Self				FL
QCC34407.1	MxiC	70	T	TagQCC34415.1	TagMxiH	1	M	3.4	6	TT	Hetero				FL
QCC34416.1	mxiG	204	K	QCC34416.1	mxiG	347	K	3.37	1	TT	Self	42.6	X2	X2	FL
QCC34406.1	mxiA	133	S	QCC34406.1	mxiA	321	K	3.36	1	TT	Self				FL
QCC34413.1	mxiJ	54	K	QCC34413.1	mxiJ	63	K	3.2	2	TT	Homo	24.42	Y5	Y6	FL
QCC34416.1	mxiG	73	K	QCC34416.1	mxiG	112	K	3.09	1	TT	Self				FL
QCC34429.1	VirB	145	K	QCC34429.1	VirB	243	K	3.07	1	TT	Self				FL
QCC34425.1	ipaC	70	S	QCC34425.1	ipaC	355	S	3.05	1	TT	Self				FL
QCC34419.1	ipgD	129	K	QCC34419.1	ipgD	137	K	2.94	1	TT	Self				FL
TagQCC34415.1	TagMxiH	75	Y	TagQCC34415.1	TagMxiH	80	S	2.89	1	TT	Self	8.54	V	V	FL
QCC34462.1	mxiK	126	S	QCC34462.1	mxiK	162	Y	2.83	1	TT	Self				FL
QCC34419.1	ipgD	137	K	QCC34419.1	ipgD	478	K	2.79	1	TT	Self				FL
QCC34419.1	ipgD	119	S	QCC34419.1	ipgD	308	K	2.75	1	TT	Self				FL
QCC34427.1	ipaA	146	S	QCC34427.1	ipaA	453	S	2.69	1	TT	Self				FL
QCC34402.1	Spa32	72	K	QCC34402.1	Spa32	220	T	2.68	2	TT	Self				FL
QCC34402.1	Spa32	30	S	QCC34402.1	Spa32	121	Y	2.58	1	TT	Self				FL
QCC34429.1	VirB	48	K	QCC34429.1	VirB	282	T	2.56	1	TT	Self				FL

Supplementary table 4: List of T3SS-related cross-links from *S. flexneri* cross-linked *in vivo*. *The cross-linking MS analysis was performed by Dr. Zhuo Angel Chen.*

Protein 1	Protein name 1	SeqPos 1	LinkedRes 1	Protein 2	protein name 2	SeqPos 2	LinkedRes 2	Highest score	Match count	Decoy type	Self-link	Cα-Cα distance in the 3D model	From chain	To chain
IpaBrrmA	IpaBrrmA	157	K	IpaBrrmA	IpaBrrmA	188	K	23.9	14	TT	true			
Tag-QCC34415.1	MxiH-Strep	84	K	Tag-QCC34415.1	MxiH-Strep	87	K	23.7	5	TT	true	15.91	u	v
IpaBrrmA	IpaBrrmA	849	K	IpaBrrmA	IpaBrrmA	864	K	22.6	2	TT	true			
QCC34413.1	MxiJ	54	K	QCC34413.1	MxiJ	54	K	22.2	9	TT	true	14.33	Y1	Y22
IpaBrrmA	IpaBrrmA	429	K	QCC34423.1	IpgC	2	S	21.6	2	TT	false			
Tag-QCC34415.1	MxiH-Strep	2	A	Tag-QCC34415.1	MxiH-Strep	11	K	20.7	2	TT	true			
IpaBrrmA	IpaBrrmA	139	K	IpaBrrmA	IpaBrrmA	202	K	20.5	18	TT	true			
QCC34408.1	MxiD	364	K	QCC34408.1	MxiD	502	K	19.7	3	TT	true	11.62	Z1	Z0
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	106	K	19.6	3	TT	true			
QCC34416.1	MxiG	293	K	QCC34416.1	MxiG	313	K	19.4	4	TT	true	16.54	X5	X4
IpaBrrmA	IpaBrrmA	139	K	IpaBrrmA	IpaBrrmA	209	K	19.2	13	TT	true			
IpaBrrmA	IpaBrrmA	1	M	QCC34423.1	IpgC	138	K	18.9	2	TT	false			
QCC34408.1	MxiD	456	K	Tag-QCC34415.1	MxiH-Strep	2	A	18.6	2	TT	false			
IpaBrrmA	IpaBrrmA	157	K	IpaBrrmA	IpaBrrmA	177	K	18.5	6	TT	true			
IpaBrrmA	IpaBrrmA	202	K	IpaBrrmA	IpaBrrmA	213	K	18.4	7	TT	true			
Tag-QCC34415.1	MxiH-Strep	2	A	Tag-QCC34415.1	MxiH-Strep	54	K	18.3	3	TT	true			
IpaBrrmA	IpaBrrmA	188	K	IpaBrrmA	IpaBrrmA	194	K	18	5	TT	true			
IpaBrrmA	IpaBrrmA	666	K	QCC34423.1	IpgC	2	S	17.9	1	TT	false			
IpaBrrmA	IpaBrrmA	436	K	QCC34423.1	IpgC	2	S	17.9	2	TT	false			
Tag-QCC34415.1	MxiH-Strep	2	A	Tag-QCC34415.1	MxiH-Strep	2	A	17.9	3	TT	true			
QCC34408.1	MxiD	362	K	QCC34408.1	MxiD	502	K	17.8	4	TT	true	9.75	Z3	Z1
QCC34404.1	Spa47	336	K	QCC34404.1	Spa47	409	K	17.6	2	TT	true			
QCC34408.1	MxiD	364	K	QCC34408.1	MxiD	542	K	17.6	2	TT	true	17.98	Z15	Z1
IpaBrrmA	IpaBrrmA	150	K	IpaBrrmA	IpaBrrmA	188	K	17.6	17	TT	true			
IpaBrrmA	IpaBrrmA	106	K	IpaBrrmA	IpaBrrmA	470	K	17.3	2	TT	true			
IpaBrrmA	IpaBrrmA	115	K	IpaBrrmA	IpaBrrmA	219	K	17.3	4	TT	true			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	572	K	17.3	2	TT	true			
QCC34429.1	VirB	48	K	QCC34429.1	VirB	121	K	17.2	5	TT	true			
IpaBrrmA	IpaBrrmA	143	K	IpaBrrmA	IpaBrrmA	202	K	17.2	23	TT	true			
QCC34413.1	MxiJ	49	K	QCC34413.1	MxiJ	54	K	17.1	1	TT	true	21.57	Y6	Y5
QCC34416.1	MxiG	252	K	QCC34416.1	MxiG	268	K	17	8	TT	true	13.17	X11	X11
IpaBrrmA	IpaBrrmA	592	K	IpaBrrmA	IpaBrrmA	666	K	17	3	TT	true			
QCC34420.1	IcsB	50	K	QCC34420.1	IcsB	69	K	17	1	TT	true			
IpaBrrmA	IpaBrrmA	289	K	QCC34423.1	IpgC	101	K	17	1	TT	false			
IpaBrrmA	IpaBrrmA	122	K	IpaBrrmA	IpaBrrmA	219	K	17	1	TT	true			
IpaBrrmA	IpaBrrmA	106	K	QCC34423.1	IpgC	2	S	16.8	2	TT	false			
QCC34401.1	Spa33	19	K	QCC34401.1	Spa33	31	K	16.8	2	TT	true			
IpaBrrmA	IpaBrrmA	440	K	QCC34423.1	IpgC	2	S	16.7	2	TT	false			
IpaBrrmA	IpaBrrmA	194	K	IpaBrrmA	IpaBrrmA	202	K	16.7	7	TT	true			
Tag-QCC34415.1	MxiH-Strep	11	K	Tag-QCC34415.1	MxiH-Strep	54	K	16.7	2	TT	true			
QCC34408.1	MxiD	502	K	QCC34410.1	MxiM	130	K	16.6	2	TT	false			
QCC34410.1	MxiM	82	K	QCC34410.1	MxiM	99	K	16.5	4	TT	true			
QCC34429.1	VirB	31	K	QCC34429.1	VirB	130	K	16.5	3	TT	true			
QCC34416.1	MxiG	254	K	QCC34416.1	MxiG	268	K	16.4	3	TT	true	8.75	X23	X23
IpaBrrmA	IpaBrrmA	666	K	IpaBrrmA	IpaBrrmA	706	K	16.2	2	TT	true			
QCC34423.1	IpgC	124	K	QCC34423.1	IpgC	154	K	16.2	2	TT	true			
IpaBrrmA	IpaBrrmA	470	K	IpaBrrmA	IpaBrrmA	537	K	16.2	1	TT	true			
Tag-QCC34415.1	MxiH-Strep	5	S	Tag-QCC34415.1	MxiH-Strep	11	K	16.2	1	TT	true			
QCC34429.1	VirB	17	K	QCC34429.1	VirB	130	K	16.1	4	TT	true			
IpaBrrmA	IpaBrrmA	15	K	QCC34423.1	IpgC	124	K	16.1	2	TT	false			
IpaBrrmA	IpaBrrmA	1	M	QCC34423.1	IpgC	122	K	15.9	4	TT	false			
QCC34416.1	MxiG	2	S	QCC34416.1	MxiG	73	K	15.9	1	TT	true			
IpaBrrmA	IpaBrrmA	1	M	QCC34423.1	IpgC	154	K	15.8	1	TT	false			
IpaBrrmA	IpaBrrmA	106	K	IpaBrrmA	IpaBrrmA	440	K	15.7	1	TT	true			
IpaBrrmA	IpaBrrmA	572	K	IpaBrrmA	IpaBrrmA	665	K	15.7	2	TT	true			
IpaBrrmA	IpaBrrmA	150	K	IpaBrrmA	IpaBrrmA	202	K	15.6	4	TT	true			
IpaBrrmA	IpaBrrmA	470	K	QCC34423.1	IpgC	122	K	15.4	4	TT	false			
QCC34416.1	MxiG	5	K	QCC34416.1	MxiG	73	K	15.3	1	TT	true			
IpaBrrmA	IpaBrrmA	1	M	QCC34423.1	IpgC	142	K	15.2	1	TT	false			
QCC34408.1	MxiD	222	K	Tag-QCC34415.1	MxiH-Strep	54	K	15.1	1	TT	false	6.71	Z15	T
IpaBrrmA	IpaBrrmA	219	K	IpaBrrmA	IpaBrrmA	277	S	14.9	1	TT	true			
QCC34410.1	MxiM	115	K	QCC34410.1	MxiM	139	K	14.9	3	TT	true			
IpaBrrmA	IpaBrrmA	152	K	IpaBrrmA	IpaBrrmA	189	K	14.9	4	TT	true			
QCC34429.1	VirB	48	K	QCC34429.1	VirB	157	K	14.8	4	TT	true			
IpaBrrmA	IpaBrrmA	145	K	IpaBrrmA	IpaBrrmA	152	K	14.8	9	TT	true			
QCC34408.1	MxiD	187	K	QCC34408.1	MxiD	298	K	14.7	1	TT	true	13.68	Z8	Z7
QCC34416.1	MxiG	240	K	QCC34416.1	MxiG	252	K	14.7	3	TT	true	18.58	X20	X20
QCC34423.1	IpgC	2	S	QCC34423.1	IpgC	122	K	14.6	1	TT	true			

QCC34413.1	MxiJ	178	K	QCC34416.1	MxiG	347	K	14.5	5	TT	false	13.28	Y4	X8
QCC34408.1	MxiD	364	K	QCC34410.1	MxiM	146	K	14.5	1	TT	false			
QCC34408.1	MxiD	364	K	QCC34410.1	MxiM	130	K	14.4	1	TT	false			
IpaBrrmA	IpaBrrmA	139	K	IpaBrrmA	IpaBrrmA	150	K	14.4	2	TT	true			
IpaBrrmA	IpaBrrmA	706	K	IpaBrrmA	IpaBrrmA	742	S	14.4	1	TT	true			
IpaBrrmA	IpaBrrmA	71	K	IpaBrrmA	IpaBrrmA	106	K	14.3	1	TT	true			
QCC34408.1	MxiD	456	K	Tag-QCC34415.1	MxiH-Strep	54	K	14.2	1	TT	false	12.93	Z0	d
IpaBrrmA	IpaBrrmA	849	K	IpaBrrmA	IpaBrrmA	855	S	14.2	1	TT	true			
IpaBrrmA	IpaBrrmA	15	K	QCC34423.1	IpgC	154	K	14	3	TT	false			
QCC34416.1	MxiG	240	K	QCC34416.1	MxiG	268	K	14	2	TT	true	14.7	X10	X10
IpaBrrmA	IpaBrrmA	101	K	QCC34423.1	IpgC	2	S	14	2	TT	false			
IpaBrrmA	IpaBrrmA	440	K	IpaBrrmA	IpaBrrmA	451	K	13.9	1	TT	true			
IpaBrrmA	IpaBrrmA	106	K	IpaBrrmA	IpaBrrmA	572	K	13.9	2	TT	true			
Tag-QCC34415.1	MxiH-Strep	84	K	Tag-QCC34415.1	MxiH-Strep	84	K	13.9	3	TT	true	15.63	s	t
QCC34404.1	Spa47	134	K	QCC34404.1	Spa47	243	K	13.9	2	TT	true			
QCC34408.1	MxiD	222	K	Tag-QCC34415.1	MxiH-Strep	2	A	13.8	1	TT	false			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	665	K	13.8	1	TT	true			
IpaBrrmA	IpaBrrmA	139	K	IpaBrrmA	IpaBrrmA	145	K	13.8	2	TT	true			
QCC34429.1	VirB	67	K	QCC34429.1	VirB	121	K	13.5	2	TT	true			
IpaBrrmA	IpaBrrmA	666	K	QCC34423.1	IpgC	138	K	13.5	1	TT	false			
QCC34429.1	VirB	26	K	QCC34429.1	VirB	70	K	13.5	2	TT	true			
IpaBrrmA	IpaBrrmA	115	K	QCC34423.1	IpgC	2	S	13.4	1	TT	false			
QCC34413.1	MxiJ	143	K	QCC34413.1	MxiJ	178	K	13.4	2	TT	true	11.26	Y5	Y5
IpaBrrmA	IpaBrrmA	269	K	IpaBrrmA	IpaBrrmA	666	K	13.4	2	TT	true			
QCC34408.1	MxiD	180	K	Tag-QCC34415.1	MxiH-Strep	2	A	13.3	1	TT	false			
QCC34423.1	IpgC	101	K	QCC34423.1	IpgC	122	K	13.3	1	TT	true			
IpaBrrmA	IpaBrrmA	59	T	QCC34423.1	IpgC	118	K	13.2	2	TT	false			
QCC34408.1	MxiD	532	K	QCC34410.1	MxiM	146	K	13.1	3	TT	false			
IpaBrrmA	IpaBrrmA	219	K	IpaBrrmA	IpaBrrmA	666	K	13	2	TT	true			
QCC34429.1	VirB	121	K	QCC34429.1	VirB	157	K	12.9	3	TT	true			
IpaBrrmA	IpaBrrmA	572	K	IpaBrrmA	IpaBrrmA	666	K	12.9	2	TT	true			
QCC34408.1	MxiD	542	K	QCC34410.1	MxiM	130	K	12.9	2	TT	false			
QCC34408.1	MxiD	521	K	Tag-QCC34415.1	MxiH-Strep	2	A	12.8	1	TT	false			
Tag-QCC34415.1	MxiH-Strep	2	A	Tag-QCC34415.1	MxiH-Strep	57	S	12.8	4	TT	true			
IpaBrrmA	IpaBrrmA	666	K	QCC34423.1	IpgC	154	K	12.6	1	TT	false			
QCC34408.1	MxiD	222	K	QCC34408.1	MxiD	222	K	12.6	1	TT	true	16.96	Z8	Z7
IpaBrrmA	IpaBrrmA	150	K	IpaBrrmA	IpaBrrmA	194	K	12.6	3	TT	true			
IpaBrrmA	IpaBrrmA	143	K	IpaBrrmA	IpaBrrmA	150	K	12.6	7	TT	true			
IpaBrrmA	IpaBrrmA	122	K	IpaBrrmA	IpaBrrmA	213	K	12.6	5	TT	true			
IpaBrrmA	IpaBrrmA	194	K	IpaBrrmA	IpaBrrmA	262	K	12.5	1	TT	true			
QCC34416.1	MxiG	293	K	QCC34416.1	MxiG	353	K	12.5	2	TT	true	12.26	X22	X22
QCC34412.1	MxiN	140	K	QCC34412.1	MxiN	199	K	12.4	2	TT	true			
IpaBrrmA	IpaBrrmA	54	T	QCC34423.1	IpgC	142	K	12.3	1	TT	false			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	115	K	12.3	1	TT	true			
IpaBrrmA	IpaBrrmA	592	K	QCC34423.1	IpgC	118	K	12.3	1	TT	false			
QCC34416.1	MxiG	5	K	QCC34462.1	MxiK	9	K	12.3	2	TT	false			
QCC34423.1	IpgC	138	K	QCC34423.1	IpgC	142	K	12.2	1	TT	true			
IpaBrrmA	IpaBrrmA	106	K	IpaBrrmA	IpaBrrmA	115	K	12.2	2	TT	true			
IpaBrrmA	IpaBrrmA	1	M	QCC34423.1	IpgC	118	K	12.2	2	TT	false			
IpaBrrmA	IpaBrrmA	143	K	IpaBrrmA	IpaBrrmA	194	K	12.2	5	TT	true			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	283	K	12.1	6	TT	true			
IpaBrrmA	IpaBrrmA	666	K	QCC34423.1	IpgC	142	K	12.1	1	TT	false			
QCC34410.1	MxiM	115	K	QCC34410.1	MxiM	144	K	12	1	TT	true			
QCC34401.1	Spa33	235	K	QCC34401.1	Spa33	252	K	12	2	TT	true			
Tag-QCC34415.1	MxiH-Strep	11	K	Tag-QCC34415.1	MxiH-Strep	87	K	12	2	TT	true			
QCC34410.1	MxiM	121	K	QCC34410.1	MxiM	139	K	11.9	1	TT	true			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	470	K	11.9	1	TT	true			
QCC34408.1	MxiD	502	K	QCC34410.1	MxiM	43	K	11.9	2	TT	false			
QCC34408.1	MxiD	180	K	QCC34408.1	MxiD	222	K	11.8	1	TT	true	10.56	Z14	Z13
QCC34412.1	MxiN	51	K	QCC34412.1	MxiN	56	K	11.7	3	TT	true			
IpaBrrmA	IpaBrrmA	592	K	IpaBrrmA	IpaBrrmA	665	K	11.7	1	TT	true			
IpaBrrmA	IpaBrrmA	470	K	IpaBrrmA	IpaBrrmA	572	K	11.6	2	TT	true			
IpaBrrmA	IpaBrrmA	283	K	QCC34423.1	IpgC	2	S	11.6	1	TT	false			
QCC34408.1	MxiD	542	K	QCC34410.1	MxiM	77	K	11.6	1	TT	false			
QCC34423.1	IpgC	118	K	QCC34423.1	IpgC	154	K	11.5	3	TT	true			
QCC34425.1	IpaC	183	K	QCC34425.1	IpaC	188	K	11.5	1	TT	true			
QCC34408.1	MxiD	222	K	QCC34408.1	MxiD	521	K	11.4	1	TT	true	46.24	Z10	Z10
IpaBrrmA	IpaBrrmA	152	K	IpaBrrmA	IpaBrrmA	262	K	11.4	1	TT	true			
QCC34423.1	IpgC	2	S	QCC34423.1	IpgC	118	K	11.3	1	TT	true			
QCC34420.1	IcsB	239	K	QCC34420.1	IcsB	302	K	11.3	1	TT	true			
IpaBrrmA	IpaBrrmA	283	K	IpaBrrmA	IpaBrrmA	572	K	11.3	1	TT	true			
QCC34416.1	MxiG	240	K	QCC34416.1	MxiG	246	K	11.3	3	TT	true	23.12	X16	X17
QCC34410.1	MxiM	77	K	QCC34410.1	MxiM	130	K	11.2	1	TT	true			
QCC34410.1	MxiM	127	K	QCC34410.1	MxiM	146	K	11.2	2	TT	true			
QCC34412.1	MxiN	29	K	QCC34412.1	MxiN	36	K	11.2	3	TT	true			
QCC34400.1	SpaP	126	K	QCC34416.1	MxiG	347	K	11.2	1	TT	false			
QCC34410.1	MxiM	77	K	QCC34410.1	MxiM	115	K	11.2	1	TT	true			
IpaBrrmA	IpaBrrmA	283	K	IpaBrrmA	IpaBrrmA	289	K	11.2	4	TT	true			
IpaBrrmA	IpaBrrmA	470	K	IpaBrrmA	IpaBrrmA	666	K	11.1	2	TT	true			
QCC34408.1	MxiD	542	K	QCC34410.1	MxiM	146	K	11.1	3	TT	false			

Tag-QCC34415.1	MxiH-Strep	87	K	Tag-QCC34415.1	MxiH-Strep	87	K	11	1	TT	true	17.17	s	t
QCC34404.1	Spa47	134	K	QCC34404.1	Spa47	291	K	11	3	TT	true			
IpaBrrmA	IpaBrrmA	15	K	IpaBrrmA	IpaBrrmA	106	K	10.9	3	TT	true			
QCC34408.1	MxiD	64	K	QCC34416.1	MxiG	362	K	10.9	1	TT	false	14.8	Z7	X0
QCC34410.1	MxiM	130	K	QCC34410.1	MxiM	146	K	10.9	2	TT	true			
QCC34410.1	MxiM	77	K	QCC34410.1	MxiM	146	K	10.9	2	TT	true			
QCC34416.1	MxiG	293	K	QCC34416.1	MxiG	369	K	10.9	1	TT	true			
QCC34406.1	MxiA	562	K	QCC34406.1	MxiA	562	K	10.8	1	TT	true			
IpaBrrmA	IpaBrrmA	101	K	IpaBrrmA	IpaBrrmA	299	K	10.7	1	TT	true			
IpaBrrmA	IpaBrrmA	60	S	QCC34423.1	IpgC	118	K	10.7	2	TT	false			
QCC34413.1	MxiJ	63	K	QCC34413.1	MxiJ	201	K	10.6	1	TT	true			
QCC34423.1	IpgC	122	K	QCC34423.1	IpgC	154	K	10.6	1	TT	true			
IpaBrrmA	IpaBrrmA	537	K	IpaBrrmA	IpaBrrmA	666	K	10.6	1	TT	true			
QCC34408.1	MxiD	187	K	Tag-QCC34415.1	MxiH-Strep	2	A	10.6	1	TT	false			
QCC34412.1	MxiN	45	K	QCC34412.1	MxiN	51	K	10.5	1	TT	true			
QCC34429.1	VirB	37	K	QCC34429.1	VirB	70	K	10.5	1	TT	true			
IpaBrrmA	IpaBrrmA	150	K	IpaBrrmA	IpaBrrmA	157	K	10.5	2	TT	true			
QCC34429.1	VirB	70	K	QCC34429.1	VirB	130	K	10.5	1	TT	true			
QCC34412.1	MxiN	202	K	QCC34412.1	MxiN	209	K	10.5	2	TT	true			
QCC34429.1	VirB	130	K	QCC34429.1	VirB	157	K	10.5	1	TT	true			
QCC34410.1	MxiM	91	K	QCC34410.1	MxiM	127	K	10.5	1	TT	true			
IpaBrrmA	IpaBrrmA	106	K	IpaBrrmA	IpaBrrmA	448	K	10.4	2	TT	true			
IpaBrrmA	IpaBrrmA	451	K	IpaBrrmA	IpaBrrmA	470	K	10.4	1	TT	true			
IpaBrrmA	IpaBrrmA	157	K	IpaBrrmA	IpaBrrmA	176	K	10.4	4	TT	true			
QCC34429.1	VirB	17	K	QCC34429.1	VirB	31	K	10.4	2	TT	true			
IpaBrrmA	IpaBrrmA	177	K	IpaBrrmA	IpaBrrmA	188	K	10.3	3	TT	true			
IpaBrrmA	IpaBrrmA	194	K	IpaBrrmA	IpaBrrmA	572	K	10.2	1	TT	true			
QCC34403.1	Spa13	35	K	QCC34403.1	Spa13	41	K	10.2	1	TT	true			
IpaBrrmA	IpaBrrmA	572	K	QCC34423.1	IpgC	2	S	10.2	1	TT	false			
QCC34408.1	MxiD	552	T	QCC34410.1	MxiM	130	K	9.93	1	TT	false			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	666	K	9.76	1	TT	true			
QCC34422.1	IpgB	39	K	QCC34422.1	IpgB	45	K	9.72	1	TT	true			
QCC34408.1	MxiD	64	K	Tag-QCC34415.1	MxiH-Strep	2	A	9.71	1	TT	false			
IpaBrrmA	IpaBrrmA	106	K	IpaBrrmA	IpaBrrmA	592	K	9.7	2	TT	true			
IpaBrrmA	IpaBrrmA	54	T	IpaBrrmA	IpaBrrmA	106	K	9.67	1	TT	true			
QCC34423.1	IpgC	118	K	QCC34423.1	IpgC	124	K	9.67	2	TT	true			
QCC34416.1	MxiG	345	K	QCC34416.1	MxiG	347	K	9.63	1	TT	true	5.76	X7	X8
QCC34410.1	MxiM	121	K	QCC34410.1	MxiM	146	K	9.61	1	TT	true			
IpaBrrmA	IpaBrrmA	665	K	QCC34423.1	IpgC	118	K	9.6	1	TT	false			
IpaBrrmA	IpaBrrmA	592	K	QCC34423.1	IpgC	138	K	9.59	1	TT	false			
QCC34416.1	MxiG	246	K	QCC34416.1	MxiG	254	K	9.59	1	TT	true	12.74	X2	X2
IpaBrrmA	IpaBrrmA	143	K	IpaBrrmA	IpaBrrmA	188	K	9.57	2	TT	true			
QCC34413.1	MxiJ	143	K	QCC34416.1	MxiG	347	K	9.47	1	TT	false	18.89	Y16	X19
QCC34400.1	SpaP	105	K	QCC34413.1	MxiJ	138	K	9.44	1	TT	false	10.57	E	Y7
IpaBrrmA	IpaBrrmA	283	K	IpaBrrmA	IpaBrrmA	592	K	9.43	5	TT	true			
QCC34408.1	MxiD	64	K	Tag-QCC34415.1	MxiH-Strep	11	K	9.42	1	TT	false			
IpaBrrmA	IpaBrrmA	188	K	IpaBrrmA	IpaBrrmA	592	K	9.42	1	TT	true			
QCC34404.1	Spa47	359	K	QCC34404.1	Spa47	409	K	9.35	1	TT	true			
QCC34429.1	VirB	164	K	QCC34429.1	VirB	235	K	9.35	4	TT	true			
QCC34408.1	MxiD	187	K	QCC34408.1	MxiD	189	K	9.32	1	TT	true	20.15	Z8	Z7
IpaBrrmA	IpaBrrmA	219	K	IpaBrrmA	IpaBrrmA	592	K	9.23	1	TT	true			
IpaBrrmA	IpaBrrmA	666	K	IpaBrrmA	IpaBrrmA	849	K	9.08	1	TT	true			
QCC34408.1	MxiD	222	K	QCC34408.1	MxiD	298	K	9.05	1	TT	true	34.75	Z0	Z0
QCC34429.1	VirB	205	K	QCC34429.1	VirB	248	K	9	2	TT	true			
IpaBrrmA	IpaBrrmA	172	S	IpaBrrmA	IpaBrrmA	176	K	8.98	1	TT	true			
QCC34429.1	VirB	17	K	QCC34429.1	VirB	164	K	8.95	1	TT	true			
QCC34429.1	VirB	274	K	QCC34429.1	VirB	288	K	8.89	1	TT	true			
QCC34429.1	VirB	130	K	QCC34429.1	VirB	164	K	8.89	1	TT	true			
Tag-QCC34415.1	MxiH-Strep	54	K	Tag-QCC34415.1	MxiH-Strep	68	K	8.77	1	TT	true	13.8	p	u
IpaBrrmA	IpaBrrmA	209	K	IpaBrrmA	IpaBrrmA	592	K	8.75	1	TT	true			
IpaBrrmA	IpaBrrmA	143	K	IpaBrrmA	IpaBrrmA	209	K	8.75	1	TT	true			
IpaBrrmA	IpaBrrmA	115	K	IpaBrrmA	IpaBrrmA	572	K	8.71	1	TT	true			
IpaBrrmA	IpaBrrmA	115	K	IpaBrrmA	IpaBrrmA	666	K	8.69	1	TT	true			
QCC34412.1	MxiN	135	K	QCC34412.1	MxiN	161	K	8.69	1	TT	true			
IpaBrrmA	IpaBrrmA	15	K	IpaBrrmA	IpaBrrmA	592	K	8.68	1	TT	true			
IpaBrrmA	IpaBrrmA	849	K	QCC34423.1	IpgC	118	K	8.64	1	TT	false			
QCC34412.1	MxiN	25	K	QCC34412.1	MxiN	33	K	8.6	2	TT	true			
QCC34409.1	MxiE	137	K	QCC34409.1	MxiE	144	K	8.58	1	TT	true			
IpaBrrmA	IpaBrrmA	706	K	QCC34423.1	IpgC	118	K	8.53	1	TT	false			
IpaBrrmA	IpaBrrmA	106	K	IpaBrrmA	IpaBrrmA	436	K	8.5	1	TT	true			
QCC34429.1	VirB	2	V	QCC34429.1	VirB	37	K	8.44	1	TT	true			
IpaBrrmA	IpaBrrmA	592	K	Tag-QCC34415.1	MxiH-Strep	84	K	8.43	2	TT	false			
IpaBrrmA	IpaBrrmA	440	K	QCC34416.1	MxiG	322	S	8.41	1	TT	false			
QCC34408.1	MxiD	64	K	QCC34414.1	MxiI	33	K	8.4	1	TT	false	16.18	Z6	P
IpaBrrmA	IpaBrrmA	436	K	IpaBrrmA	IpaBrrmA	448	K	8.36	1	TT	true			
IpaBrrmA	IpaBrrmA	283	K	IpaBrrmA	IpaBrrmA	849	K	8.26	1	TT	true			
QCC34429.1	VirB	121	K	QCC34429.1	VirB	121	K	8.24	1	TT	true			
QCC34410.1	MxiM	77	K	QCC34410.1	MxiM	127	K	8.22	2	TT	true			
QCC34425.1	IpaC	188	K	QCC34425.1	IpaC	204	K	8.15	2	TT	true			
IpaBrrmA	IpaBrrmA	53	T	IpaBrrmA	IpaBrrmA	106	K	8.12	1	TT	true			

IpaBrrmA	IpaBrrmA	172	S	IpaBrrmA	IpaBrrmA	175	K	8.11	1	TT	true			
IpaBrrmA	IpaBrrmA	572	K	QCC34423.1	IpgC	49	K	8.09	1	TT	false			
QCC34416.1	MxiG	282	K	QCC34416.1	MxiG	293	K	8.05	2	TT	true	10.05	X2	X2
IpaBrrmA	IpaBrrmA	448	K	QCC34423.1	IpgC	2	S	7.99	1	TT	false			
IpaBrrmA	IpaBrrmA	213	K	IpaBrrmA	IpaBrrmA	592	K	7.92	1	TT	true			
IpaBrrmA	IpaBrrmA	660	S	IpaBrrmA	IpaBrrmA	666	K	7.89	2	TT	true			
QCC34417.1	IpgF	135	Y	Tag-QCC34415.1	MxiH-Strep	1	M	7.85	1	TT	true			
IpaBrrmA	IpaBrrmA	592	K	QCC34423.1	IpgC	124	K	7.78	2	TT	false			
QCC34412.1	MxiN	43	K	QCC34412.1	MxiN	44	K	7.78	1	TT	true			
QCC34427.1	IpaA	226	K	QCC34427.1	IpaA	238	K	7.77	1	TT	true			
QCC34411.1	MxiL	1	M	Tag-QCC34415.1	MxiH-Strep	87	K	7.75	1	TT	false			
IpaBrrmA	IpaBrrmA	186	T	IpaBrrmA	IpaBrrmA	194	K	7.75	2	TT	true			
QCC34414.1	MxiI	1	M	Tag-QCC34415.1	MxiH-Strep	87	K	7.72	1	TT	false			
QCC34429.1	VirB	242	K	QCC34429.1	VirB	248	K	7.66	1	TT	true			
QCC34408.1	MxiD	558	K	QCC34410.1	MxiM	146	K	7.65	1	TT	false			
QCC34429.1	VirB	145	K	QCC34429.1	VirB	201	K	7.65	1	TT	true			
IpaBrrmA	IpaBrrmA	537	K	QCC34429.1	VirB	271	K	7.6	2	TT	false			
QCC34421.1	IpgA	4	K	QCC34421.1	IpgA	16	K	7.56	2	TT	true			
QCC34410.1	MxiM	91	K	QCC34410.1	MxiM	146	K	7.48	1	TT	true			
QCC34429.1	VirB	31	K	QCC34429.1	VirB	31	K	7.45	2	TT	true			
QCC34429.1	VirB	16	K	QCC34429.1	VirB	111	K	7.44	1	TT	true			
QCC34414.1	MxiI	96	K	QCC34414.1	MxiI	96	K	7.36	1	TT	true	9.84	M	N
IpaBrrmA	IpaBrrmA	283	K	IpaBrrmA	IpaBrrmA	665	K	7.29	1	TT	true			
IpaBrrmA	IpaBrrmA	186	T	QCC34426.1	IpaD	62	K	7.27	1	TT	false			
IpaBrrmA	IpaBrrmA	115	K	IpaBrrmA	IpaBrrmA	592	K	7.26	2	TT	true			
QCC34404.1	Spa47	5	K	QCC34412.1	MxiN	135	K	7.24	1	TT	false			
IpaBrrmA	IpaBrrmA	157	K	IpaBrrmA	IpaBrrmA	157	K	7.19	1	TT	true			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	219	K	7.15	1	TT	true			
QCC34408.1	MxiD	536	S	QCC34410.1	MxiM	146	K	7.14	1	TT	false			
QCC34408.1	MxiD	533	S	QCC34410.1	MxiM	146	K	7.14	1	TT	false			
QCC34408.1	MxiD	64	K	QCC34408.1	MxiD	64	K	7.14	1	TT	true	18.22	Z12	Z11
IpaBrrmA	IpaBrrmA	572	K	IpaBrrmA	IpaBrrmA	592	K	7.1	1	TT	true			
Tag-QCC34415.1	MxiH-Strep	11	K	Tag-QCC34415.1	MxiH-Strep	57	S	7.05	1	TT	true			
QCC34429.1	VirB	31	K	QCC34429.1	VirB	33	K	7.04	1	TT	true			
QCC34429.1	VirB	48	K	QCC34429.1	VirB	145	K	7.04	1	TT	true			
IpaBrrmA	IpaBrrmA	176	K	IpaBrrmA	IpaBrrmA	188	K	7.01	3	TT	true			
QCC34419.1	IpgD	40	K	QCC34419.1	IpgD	46	K	7	1	TT	true			
QCC34429.1	VirB	70	K	QCC34429.1	VirB	157	K	6.99	1	TT	true			
IpaBrrmA	IpaBrrmA	145	K	IpaBrrmA	IpaBrrmA	592	K	6.91	1	TT	true			
QCC34413.1	MxiJ	201	K	QCC34416.1	MxiG	195	K	6.87	1	TT	false			
QCC34429.1	VirB	201	K	QCC34429.1	VirB	243	K	6.86	1	TT	true			
QCC34412.1	MxiN	25	K	QCC34412.1	MxiN	29	K	6.84	1	TT	true			
QCC34409.1	MxiE	209	K	QCC34427.1	IpaA	123	S	6.79	1	TT	false			
QCC34401.1	Spa33	235	K	QCC34401.1	Spa33	259	K	6.79	1	TT	true			
IpaBrrmA	IpaBrrmA	849	K	IpaBrrmA	IpaBrrmA	849	K	6.67	2	TT	true			
IpaBrrmA	IpaBrrmA	585	S	IpaBrrmA	IpaBrrmA	592	K	6.66	1	TT	true			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	194	K	6.65	1	TT	true			
IpaBrrmA	IpaBrrmA	219	K	QCC34462.1	MxiK	105	K	6.63	1	TT	false			
IpaBrrmA	IpaBrrmA	176	K	IpaBrrmA	IpaBrrmA	186	T	6.54	1	TT	true			
IpaBrrmA	IpaBrrmA	592	K	QCC34423.1	IpgC	142	K	6.47	1	TT	false			
QCC34412.1	MxiN	215	K	QCC34429.1	VirB	156	K	6.47	2	TT	false			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	175	K	6.46	1	TT	false			
IpaBrrmA	IpaBrrmA	572	K	Tag-QCC34415.1	MxiH-Strep	84	K	6.41	1	TT	false			
IpaBrrmA	IpaBrrmA	470	K	IpaBrrmA	IpaBrrmA	605	S	6.25	1	TT	true			
QCC34412.1	MxiN	140	K	QCC34412.1	MxiN	202	K	6.25	1	TT	true			
QCC34408.1	MxiD	362	K	QCC34410.1	MxiM	130	K	6.23	2	TT	false			
QCC34416.1	MxiG	293	K	QCC34416.1	MxiG	347	K	6.23	1	TT	true	20.15	X5	X5
IpaBrrmA	IpaBrrmA	665	K	QCC34401.1	Spa33	161	T	6.22	2	TT	false			
IpaBrrmA	IpaBrrmA	666	K	QCC34401.1	Spa33	161	T	6.22	2	TT	false			
QCC34412.1	MxiN	90	K	QCC34412.1	MxiN	199	K	6.21	2	TT	true			
QCC34404.1	Spa47	5	K	QCC34412.1	MxiN	161	K	6.17	2	TT	false			
QCC34410.1	MxiM	99	K	QCC34410.1	MxiM	121	K	6.16	1	TT	true			
QCC34420.1	IcsB	179	Y	QCC34429.1	VirB	48	K	6.15	1	TT	true			
IpaBrrmA	IpaBrrmA	1	M	IpaBrrmA	IpaBrrmA	269	K	6.14	1	TT	true			
QCC34400.1	SpaP	104	K	QCC34400.1	SpaP	105	K	6.09	1	TT	true	35.78	D	C
IpaBrrmA	IpaBrrmA	157	K	IpaBrrmA	IpaBrrmA	194	K	6.05	2	TT	true			
IpaBrrmA	IpaBrrmA	188	K	QCC34426.1	IpaD	197	K	6	2	TT	false			
QCC34408.1	MxiD	558	K	QCC34410.1	MxiM	130	K	5.97	1	TT	false			
QCC34411.1	MxiL	1	M	Tag-QCC34415.1	MxiH-Strep	84	K	5.83	1	TT	false			
IpaBrrmA	IpaBrrmA	145	K	QCC34419.1	IpgD	320	S	5.68	1	TT	false			
QCC34412.1	MxiN	148	K	QCC34412.1	MxiN	209	K	5.62	1	TT	true			
QCC34412.1	MxiN	33	K	QCC34412.1	MxiN	45	K	5.58	1	TT	true			
QCC34418.1	IpgE	2	E	QCC34429.1	VirB	92	T	5.47	1	TT	false			
QCC34404.1	Spa47	336	K	QCC34404.1	Spa47	359	K	5.44	1	TT	false			
IpaBrrmA	IpaBrrmA	586	T	Tag-QCC34415.1	MxiH-Strep	2	A	5.43	1	TT	false			
QCC34412.1	MxiN	43	K	QCC34412.1	MxiN	51	K	5.4	1	TT	true			
QCC34417.1	IpgF	41	K	QCC34417.1	IpgF	145	S	5.34	1	TT	true			
QCC34400.1	SpaP	105	K	QCC34408.1	MxiD	68	K	5.31	1	TT	false	14.31	A	Z10
IpaBrrmA	IpaBrrmA	15	K	QCC34423.1	IpgC	118	K	5.31	1	TT	false			
QCC34419.1	IpgD	328	S	QCC34420.1	IcsB	45	S	5.3	1	TT	false			

IpaBrrmA	IpaBrrmA	15	K	QCC34423.1	IpgC	122	K	5.25	1	TT	false			
IpaBrrmA	IpaBrrmA	188	K	IpaBrrmA	IpaBrrmA	666	K	5.22	1	TT	true			
QCC34416.1	MxiG	353	K	QCC34462.1	MxiK	164	K	5.21	1	TT	false			
IpaBrrmA	IpaBrrmA	213	K	QCC34408.1	MxiD	78	S	5.13	1	TT	false			
QCC34403.1	Spa13	70	S	QCC34412.1	MxiN	202	K	5.13	1	TT	false			
QCC34412.1	MxiN	148	K	QCC34429.1	VirB	16	K	5.09	1	TT	false			
QCC34425.1	IpaC	257	K	QCC34425.1	IpaC	275	K	5.07	1	TT	true			
QCC34429.1	VirB	92	T	Tag-QCC34415.1	MxiH-Strep	2	A	5.05	1	TT	false			
QCC34412.1	MxiN	209	K	QCC34412.1	MxiN	225	K	4.97	1	TT	true			
IpaBrrmA	IpaBrrmA	120	S	QCC34402.1	Spa32	85	S	4.96	1	TT	false			
IpaBrrmA	IpaBrrmA	99	T	QCC34401.1	Spa33	58	S	4.92	1	TT	false			
IpaBrrmA	IpaBrrmA	197	T	QCC34408.1	MxiD	68	K	4.88	1	TT	false			
IpaBrrmA	IpaBrrmA	197	T	QCC34408.1	MxiD	69	K	4.88	1	TT	false			
QCC34403.1	Spa13	1	M	Tag-QCC34415.1	MxiH-Strep	75	Y	4.88	1	TT	false			
QCC34423.1	IpgC	154	K	QCC34423.1	IpgC	154	K	4.87	1	TT	true			
QCC34409.1	MxiE	111	K	Tag-QCC34415.1	MxiH-Strep	1	M	4.84	2	TT	false			
QCC34406.1	MxiA	151	K	QCC34406.1	MxiA	170	K	4.73	1	TT	true			
QCC34416.1	MxiG	245	S	QCC34416.1	MxiG	248	T	4.72	1	TT	true	8.15	X17	X17
IpaBrrmA	IpaBrrmA	143	K	IpaBrrmA	IpaBrrmA	157	K	4.59	1	TT	true			
QCC34429.1	VirB	248	K	QCC34429.1	VirB	285	K	4.58	1	TT	true			
IpaBrrmA	IpaBrrmA	163	T	QCC34419.1	IpgD	92	K	4.5	2	TT	false			
QCC34419.1	IpgD	88	T	QCC34419.1	IpgD	89	S	4.45	1	TT	true			
IpaBrrmA	IpaBrrmA	101	K	IpaBrrmA	IpaBrrmA	470	K	4.41	1	TT	true			
QCC34419.1	IpgD	129	K	QCC34419.1	IpgD	137	K	4.31	1	TT	true			
QCC34427.1	IpaA	524	K	QCC34427.1	IpaA	530	K	4.09	1	TT	true			
QCC34429.1	VirB	12	K	QCC34429.1	VirB	31	K	4.09	1	TT	true			
QCC34419.1	IpgD	356	K	QCC34419.1	IpgD	408	K	4.07	1	TT	true			
IpaBrrmA	IpaBrrmA	665	K	IpaBrrmA	IpaBrrmA	666	K	3.96	1	TT	true			
QCC34410.1	MxiM	91	K	QCC34410.1	MxiM	121	K	3.94	1	TT	true			
QCC34412.1	MxiN	43	K	QCC34412.1	MxiN	56	K	3.73	1	TT	true			
IpaBrrmA	IpaBrrmA	379	S	QCC34403.1	Spa13	48	S	3.72	2	TT	false			
IpaBrrmA	IpaBrrmA	188	K	IpaBrrmA	IpaBrrmA	202	K	3.64	1	TT	true			
QCC34409.1	MxiE	113	K	QCC34409.1	MxiE	118	K	3.61	1	TT	true			
QCC34416.1	MxiG	353	K	QCC34416.1	MxiG	353	K	3.42	1	TT	true	18.19	X13	X14
IpaBrrmA	IpaBrrmA	137	Y	IpaBrrmA	IpaBrrmA	586	T	3.23	1	TT	true			
QCC34422.1	IpgB	27	K	QCC34422.1	IpgB	31	K	3.08	1	TT	true			
QCC34403.1	Spa13	70	S	QCC34403.1	Spa13	70	S	3.01	2	TT	true			

Supplementary table 5: Theoretical and measured masses of InvCΔ79 from [140]**.** The protein sequence of InvCΔ79 is listed in suppl. table. 7.

		Theoretical mass* [Da]	Experimental mass** [Da]			FWHM** [Da]
Fraction I	Monomer	39901.5	39901.9	±	0.6	30
	Dimer	79802.9	79900	±	200	400
Fraction II	Monomer	39901.5	39906	±	4	30
	Dimer	79802.9	79808	±	2	70
	Trimer	119704.4	119730	±	40	600
	Tetramer	159605.9	159880	±	70	800

* Reflects N-terminal methionine excision and C-terminal strep-tag.

** Average mass, standard deviation, and FWHM were calculated from triplicates.

Supplementary Table 6: List of cross-links from SpaO/SpaO_c/OrgB/InvC-strep complexes. *The cross-linking MS analysis was performed by Dr. Zhuo Angel Chen.*

Protein1	SeqPos1	LinkedRes1	Protein2	SeqPos2	LinkedRes2	Highest score	Match count	Decoy type	Self
OrgB	138	K	OrgB	159	K	25.06	9	TT	TRUE
InvC-Strep	43	K	OrgB	53	K	24.65	5	TT	FALSE
OrgB	138	K	SpaO	118	K	24.25	4	TT	FALSE
InvC-Strep	43	K	InvC-Strep	331	K	23.25	3	TT	TRUE
InvC-Strep	160	S	InvC-Strep	345	K	22.62	5	TT	TRUE
OrgB	53	K	OrgB	203	T	22.36	1	TT	TRUE
OrgB	53	K	SpaO	146	K	22.03	4	TT	FALSE
OrgB	219	K	SpaO	146	K	21.97	4	TT	FALSE
InvC-Strep	160	S	OrgB	138	K	21.67	2	TT	FALSE
OrgB	124	K	OrgB	138	K	21.11	2	TT	TRUE
OrgB	138	K	SpaO	146	K	21.04	2	TT	FALSE
OrgB	92	K	OrgB	138	K	20.77	6	TT	TRUE
OrgB	138	K	OrgB	205	S	20.46	2	TT	TRUE
OrgB	138	K	OrgB	219	K	20.42	5	TT	TRUE
InvC-Strep	331	K	OrgB	139	K	20.35	2	TT	FALSE
SpaO	189	K	SpaOc	1	M	20.23	1	TT	FALSE
OrgB	53	K	OrgB	138	K	20.14	4	TT	TRUE
InvC-Strep	43	K	InvC-Strep	364	S	20.01	3	TT	TRUE
InvC-Strep	43	K	InvC-Strep	101	K	19.89	2	TT	TRUE
InvC-Strep	320	S	InvC-Strep	331	K	19.86	2	TT	TRUE
OrgB	138	K	OrgB	203	T	19.81	1	TT	TRUE
InvC-Strep	331	K	OrgB	138	K	19.77	3	TT	FALSE
OrgB	138	K	SpaO	20	T	19.61	3	TT	FALSE
SpaO	146	K	SpaO	189	K	19.52	4	TT	TRUE
InvC-Strep	43	K	OrgB	219	K	19.30	4	TT	FALSE
OrgB	92	K	OrgB	219	K	19.14	4	TT	TRUE
InvC-Strep	160	S	OrgB	53	K	19.09	2	TT	FALSE
OrgB	53	K	OrgB	159	K	19.00	3	TT	TRUE
OrgB	159	K	OrgB	219	K	18.99	4	TT	TRUE
OrgB	53	K	SpaO	189	K	18.94	3	TT	FALSE
OrgB	53	K	SpaOc	1	M	18.92	1	TT	FALSE
OrgB	138	K	SpaO	189	K	18.64	2	TT	FALSE
OrgB	22	K	SpaO	146	K	18.64	3	TT	FALSE
OrgB	22	K	SpaOc	1	M	18.63	7	TT	FALSE
OrgB	53	K	OrgB	92	K	18.46	4	TT	TRUE
OrgB	22	K	OrgB	138	K	18.36	7	TT	TRUE
OrgB	138	K	OrgB	217	K	18.36	2	TT	TRUE
InvC-Strep	43	K	InvC-Strep	345	K	18.33	7	TT	TRUE
OrgB	138	K	SpaO	159	S	17.92	1	TT	FALSE
InvC-Strep	15	K	InvC-Strep	43	K	17.69	6	TT	TRUE
InvC-Strep	1	M	InvC-Strep	43	K	17.69	18	TT	TRUE
InvC-Strep	331	K	OrgB	22	K	17.64	4	TT	FALSE
InvC-Strep	331	K	OrgB	53	K	17.53	2	TT	FALSE
InvC-Strep	43	K	InvC-Strep	160	S	17.29	7	TT	TRUE
OrgB	138	K	SpaO	124	S	17.26	2	TT	FALSE
InvC-Strep	101	K	OrgB	219	K	17.25	1	TT	FALSE
InvC-Strep	331	K	OrgB	219	K	17.24	2	TT	FALSE
OrgB	53	K	OrgB	217	K	17.19	2	TT	TRUE
SpaO	118	K	SpaO	190	K	17.18	6	TT	TRUE
SpaO	124	S	SpaO	189	K	17.04	2	TT	TRUE
InvC-Strep	15	K	OrgB	138	K	16.99	2	TT	FALSE
InvC-Strep	345	K	OrgB	219	K	16.88	2	TT	FALSE
OrgB	20	K	OrgB	124	K	16.86	1	TT	TRUE
InvC-Strep	71	S	OrgB	138	K	16.82	2	TT	FALSE
OrgB	155	T	OrgB	160	Y	16.78	6	TT	TRUE
OrgB	20	K	OrgB	219	K	16.70	2	TT	TRUE
OrgB	203	T	OrgB	219	K	16.70	1	TT	TRUE
InvC-Strep	160	S	SpaO	48	K	16.65	3	TT	FALSE
OrgB	138	K	SpaO	190	K	16.63	3	TT	FALSE

InvC-Strep	43	K	OrgB	124	K	16.63	4	TT	FALSE
OrgB	20	K	OrgB	92	K	16.60	3	TT	TRUE
InvC-Strep	331	K	SpaO	48	K	16.51	2	TT	FALSE
InvC-Strep	15	K	InvC-Strep	62	S	16.48	1	TT	TRUE
InvC-Strep	43	K	OrgB	138	K	16.43	9	TT	FALSE
OrgB	20	K	OrgB	27	Y	16.38	1	TT	TRUE
SpaO	190	K	SpaOc	1	M	16.34	2	TT	FALSE
OrgB	219	K	SpaO	48	K	16.24	2	TT	FALSE
OrgB	138	K	OrgB	143	K	16.22	4	TT	TRUE
OrgB	124	K	OrgB	219	K	16.20	1	TT	TRUE
OrgB	53	K	SpaO	118	K	16.18	4	TT	FALSE
InvC-Strep	43	K	OrgB	217	K	16.14	2	TT	FALSE
OrgB	22	K	OrgB	217	K	16.14	5	TT	TRUE
InvC-Strep	62	S	OrgB	138	K	16.14	2	TT	FALSE
OrgB	143	K	OrgB	219	K	15.97	2	TT	TRUE
OrgB	139	K	OrgB	219	K	15.94	4	TT	TRUE
OrgB	124	K	OrgB	139	K	15.93	3	TT	TRUE
OrgB	139	K	OrgB	191	K	15.82	5	TT	TRUE
OrgB	22	K	OrgB	124	K	15.66	6	TT	TRUE
SpaO	159	S	SpaO	189	K	15.63	2	TT	TRUE
InvC-Strep	59	T	OrgB	138	K	15.61	1	TT	FALSE
InvC-Strep	77	Y	OrgB	138	K	15.56	2	TT	FALSE
OrgB	92	K	OrgB	124	K	15.46	3	TT	TRUE
InvC-Strep	15	K	OrgB	219	K	15.45	1	TT	FALSE
OrgB	132	T	OrgB	138	K	15.36	1	TT	TRUE
InvC-Strep	15	K	InvC-Strep	160	S	15.35	1	TT	TRUE
InvC-Strep	345	K	OrgB	22	K	15.30	4	TT	FALSE
InvC-Strep	101	K	OrgB	53	K	15.30	1	TT	FALSE
InvC-Strep	368	K	InvC-Strep	432	S	15.30	3	TT	TRUE
InvC-Strep	9	Y	OrgB	138	K	15.26	2	TT	FALSE
OrgB	219	K	OrgB	220	T	15.25	2	TT	TRUE
InvC-Strep	99	T	InvC-Strep	99	T	15.19	1	TT	TRUE
OrgB	20	K	SpaO	190	K	15.17	2	TT	FALSE
InvC-Strep	364	S	OrgB	138	K	15.17	1	TT	FALSE
InvC-Strep	345	K	OrgB	20	K	15.15	2	TT	FALSE
InvC-Strep	79	T	OrgB	138	K	15.10	1	TT	FALSE
SpaO	48	K	SpaO	189	K	15.10	4	TT	TRUE
InvC-Strep	43	K	InvC-Strep	59	T	15.08	4	TT	TRUE
SpaO	190	K	SpaO	205	T	15.07	1	TT	TRUE
OrgB	139	K	OrgB	159	K	14.98	5	TT	TRUE
OrgB	96	Y	OrgB	138	K	14.96	2	TT	TRUE
OrgB	92	K	OrgB	217	K	14.96	3	TT	TRUE
SpaO	112	S	SpaO	146	K	14.93	1	TT	TRUE
SpaO	189	K	SpaOc	31	K	14.92	1	TT	FALSE
OrgB	53	K	OrgB	219	K	14.91	2	TT	TRUE
InvC-Strep	345	K	InvC-Strep	358	T	14.90	3	TT	TRUE
OrgB	139	K	OrgB	143	K	14.80	2	TT	TRUE
InvC-Strep	101	K	OrgB	139	K	14.79	2	TT	FALSE
OrgB	124	K	SpaO	190	K	14.79	2	TT	FALSE
InvC-Strep	15	K	InvC-Strep	345	K	14.78	2	TT	TRUE
InvC-Strep	43	K	InvC-Strep	114	S	14.77	2	TT	TRUE
InvC-Strep	43	K	InvC-Strep	208	K	14.73	2	TT	TRUE
OrgB	219	K	SpaO	118	K	14.70	5	TT	FALSE
OrgB	53	K	SpaO	190	K	14.69	4	TT	FALSE
OrgB	22	K	OrgB	203	T	14.65	2	TT	TRUE
OrgB	138	K	SpaO	112	S	14.58	3	TT	FALSE
InvC-Strep	43	K	SpaO	118	K	14.51	2	TT	FALSE
InvC-Strep	43	K	InvC-Strep	71	S	14.49	4	TT	TRUE
InvC-Strep	15	K	OrgB	53	K	14.48	2	TT	FALSE
InvC-Strep	160	S	InvC-Strep	331	K	14.43	5	TT	TRUE
InvC-Strep	345	K	OrgB	92	K	14.43	2	TT	FALSE
InvC-Strep	345	K	OrgB	138	K	14.41	3	TT	FALSE
OrgB	27	Y	OrgB	217	K	14.37	1	TT	TRUE
InvC-Strep	43	K	InvC-Strep	125	S	14.37	2	TT	TRUE
InvC-Strep	43	K	InvC-Strep	62	S	14.19	4	TT	TRUE

OrgB	139	K	SpaO	20	T	14.19	2	TT	FALSE
InvC-Strep	320	S	InvC-Strep	345	K	14.15	2	TT	TRUE
InvC-Strep	15	K	InvC-Strep	128	S	14.14	1	TT	TRUE
InvC-Strep	2	K	InvC-Strep	331	K	14.13	2	TT	TRUE
OrgB	53	K	SpaO	48	K	14.13	2	TT	FALSE
OrgB	217	K	SpaO	118	K	14.13	2	TT	FALSE
InvC-Strep	43	K	InvC-Strep	128	S	14.11	3	TT	TRUE
InvC-Strep	15	K	InvC-Strep	331	K	14.10	2	TT	TRUE
OrgB	53	K	OrgB	220	T	14.04	1	TT	TRUE
InvC-Strep	2	K	InvC-Strep	15	K	14.01	6	TT	TRUE
OrgB	139	K	SpaO	146	K	13.94	4	TT	FALSE
OrgB	217	K	SpaO	190	K	13.89	2	TT	FALSE
InvC-Strep	59	T	OrgB	22	K	13.88	1	TT	FALSE
SpaO	20	T	SpaO	48	K	13.86	2	TT	TRUE
InvC-Strep	277	S	OrgB	138	K	13.85	1	TT	FALSE
OrgB	53	K	SpaO	124	S	13.85	2	TT	FALSE
InvC-Strep	43	K	InvC-Strep	79	T	13.80	3	TT	TRUE
SpaO	48	K	SpaO	190	K	13.73	4	TT	TRUE
OrgB	29	S	OrgB	138	K	13.72	2	TT	TRUE
OrgB	22	K	OrgB	159	K	13.70	2	TT	TRUE
OrgB	205	S	OrgB	219	K	13.62	2	TT	TRUE
SpaO	48	K	SpaO	146	K	13.59	3	TT	TRUE
InvC-Strep	43	K	OrgB	203	T	13.59	1	TT	FALSE
InvC-Strep	15	K	OrgB	22	K	13.50	4	TT	FALSE
InvC-Strep	43	K	SpaO	189	K	13.46	2	TT	FALSE
OrgB	12	S	OrgB	138	K	13.43	2	TT	TRUE
InvC-Strep	2	K	OrgB	20	K	13.40	1	TT	FALSE
OrgB	29	S	OrgB	53	K	13.39	2	TT	TRUE
InvC-Strep	160	S	OrgB	219	K	13.38	2	TT	FALSE
InvC-Strep	126	Y	OrgB	138	K	13.35	2	TT	FALSE
OrgB	220	T	SpaO	48	K	13.32	1	TT	FALSE
OrgB	12	S	OrgB	22	K	13.30	6	TT	TRUE
SpaO	20	T	SpaO	190	K	13.27	2	TT	TRUE
SpaO	48	K	SpaO	124	S	13.25	3	TT	TRUE
SpaO	146	K	SpaO	185	Y	13.24	1	TT	TRUE
InvC-Strep	160	S	OrgB	22	K	13.24	7	TT	FALSE
OrgB	20	K	OrgB	139	K	13.23	3	TT	TRUE
OrgB	139	K	OrgB	203	T	13.18	2	TT	TRUE
InvC-Strep	43	K	InvC-Strep	77	Y	13.15	2	TT	TRUE
OrgB	22	K	OrgB	53	K	13.14	6	TT	TRUE
OrgB	9	S	OrgB	22	K	13.12	5	TT	TRUE
InvC-Strep	2	K	OrgB	138	K	13.10	3	TT	FALSE
SpaO	159	S	SpaO	190	K	13.09	3	TT	TRUE
SpaO	48	K	SpaO	164	S	13.09	1	TT	TRUE
InvC-Strep	2	K	OrgB	124	K	13.05	2	TT	FALSE
OrgB	219	K	SpaO	20	T	13.05	1	TT	FALSE
OrgB	22	K	SpaO	20	T	13.04	3	TT	FALSE
InvC-Strep	331	K	OrgB	217	K	13.02	1	TT	FALSE
InvC-Strep	43	K	OrgB	20	K	13.00	4	TT	FALSE
OrgB	53	K	SpaO	159	S	13.00	1	TT	FALSE
InvC-Strep	125	S	OrgB	22	K	13.00	1	TT	FALSE
InvC-Strep	77	Y	OrgB	53	K	12.99	2	TT	FALSE
OrgB	217	K	OrgB	219	K	12.94	2	TT	TRUE
OrgB	53	K	OrgB	124	K	12.93	1	TT	TRUE
InvC-Strep	320	S	OrgB	138	K	12.90	1	TT	FALSE
InvC-Strep	320	S	OrgB	22	K	12.87	2	TT	FALSE
InvC-Strep	2	K	OrgB	217	K	12.86	1	TT	FALSE
OrgB	219	K	SpaO	189	K	12.85	3	TT	FALSE
OrgB	138	K	OrgB	220	T	12.81	1	TT	TRUE
InvC-Strep	320	S	OrgB	219	K	12.78	1	TT	FALSE
OrgB	92	K	SpaO	48	K	12.77	4	TT	FALSE
InvC-Strep	43	K	SpaO	164	S	12.77	1	TT	FALSE
OrgB	53	K	SpaO	187	Y	12.71	1	TT	FALSE
OrgB	20	K	OrgB	29	S	12.68	1	TT	TRUE
OrgB	138	K	SpaO	158	S	12.66	1	TT	FALSE

OrgB	27	Y	OrgB	53	K	12.66	2	TT	TRUE
InvC-Strep	1	M	InvC-Strep	15	K	12.65	7	TT	TRUE
OrgB	139	K	OrgB	217	K	12.59	3	TT	TRUE
InvC-Strep	331	K	SpaO	189	K	12.56	2	TT	FALSE
InvC-Strep	2	K	SpaOc	1	M	12.55	1	TT	FALSE
OrgB	220	T	SpaO	190	K	12.54	1	TT	FALSE
InvC-Strep	331	K	OrgB	20	K	12.53	2	TT	FALSE
OrgB	124	K	SpaO	118	K	12.53	1	TT	FALSE
InvC-Strep	15	K	SpaO	190	K	12.50	2	TT	FALSE
InvC-Strep	71	S	OrgB	219	K	12.50	2	TT	FALSE
OrgB	20	K	OrgB	138	K	12.48	3	TT	TRUE
InvC-Strep	345	K	OrgB	124	K	12.45	1	TT	FALSE
InvC-Strep	15	K	InvC-Strep	71	S	12.43	2	TT	TRUE
InvC-Strep	358	T	OrgB	219	K	12.42	1	TT	FALSE
InvC-Strep	331	K	OrgB	92	K	12.41	2	TT	FALSE
InvC-Strep	1	M	InvC-Strep	345	K	12.40	4	TT	TRUE
InvC-Strep	9	Y	InvC-Strep	43	K	12.39	2	TT	TRUE
OrgB	138	K	SpaO	35	T	12.34	1	TT	FALSE
OrgB	124	K	SpaO	48	K	12.32	1	TT	FALSE
SpaO	118	K	SpaO	189	K	12.31	5	TT	TRUE
InvC-Strep	345	K	OrgB	139	K	12.23	2	TT	FALSE
SpaO	44	S	SpaO	48	K	12.23	2	TT	TRUE
OrgB	138	K	SpaO	164	S	12.20	1	TT	FALSE
InvC-Strep	160	S	SpaO	190	K	12.20	2	TT	FALSE
InvC-Strep	345	K	OrgB	53	K	12.18	2	TT	FALSE
InvC-Strep	15	K	SpaO	189	K	12.14	1	TT	FALSE
InvC-Strep	2	K	OrgB	3	K	12.13	2	TT	FALSE
OrgB	53	K	OrgB	169	S	12.12	1	TT	TRUE
InvC-Strep	43	K	SpaO	20	T	12.08	2	TT	FALSE
SpaO	146	K	SpaO	187	Y	12.05	1	TT	TRUE
OrgB	20	K	OrgB	22	K	12.02	4	TT	TRUE
InvC-Strep	345	K	SpaO	189	K	12.01	2	TT	FALSE
InvC-Strep	368	K	OrgB	138	K	11.98	2	TT	FALSE
InvC-Strep	43	K	SpaO	48	K	11.97	2	TT	FALSE
OrgB	92	K	SpaO	118	K	11.92	2	TT	FALSE
InvC-Strep	331	K	InvC-Strep	345	K	11.91	2	TT	TRUE
InvC-Strep	15	K	InvC-Strep	77	Y	11.91	1	TT	TRUE
OrgB	53	K	SpaO	164	S	11.86	2	TT	FALSE
OrgB	9	S	OrgB	219	K	11.85	1	TT	TRUE
InvC-Strep	43	K	SpaO	159	S	11.85	1	TT	FALSE
InvC-Strep	265	S	OrgB	22	K	11.84	2	TT	FALSE
OrgB	219	K	SpaO	190	K	11.84	3	TT	FALSE
OrgB	138	K	OrgB	160	Y	11.81	1	TT	TRUE
OrgB	220	T	SpaO	189	K	11.79	1	TT	FALSE
OrgB	22	K	SpaO	159	S	11.79	1	TT	FALSE
SpaO	33	Y	SpaO	48	K	11.77	2	TT	TRUE
InvC-Strep	114	S	OrgB	22	K	11.76	2	TT	FALSE
SpaO	48	K	SpaO	158	S	11.76	2	TT	TRUE
InvC-Strep	215	T	InvC-Strep	345	K	11.75	1	TT	TRUE
SpaO	48	K	SpaOc	1	M	11.75	1	TT	FALSE
OrgB	12	S	OrgB	53	K	11.72	1	TT	TRUE
InvC-Strep	62	S	OrgB	219	K	11.72	1	TT	FALSE
InvC-Strep	128	S	OrgB	219	K	11.70	1	TT	FALSE
InvC-Strep	215	T	OrgB	53	K	11.66	1	TT	FALSE
OrgB	22	K	SpaOc	2	E	11.66	1	TT	FALSE
SpaO	158	S	SpaO	190	K	11.65	1	TT	TRUE
OrgB	20	K	SpaO	118	K	11.65	2	TT	FALSE
OrgB	139	K	OrgB	169	S	11.62	2	TT	TRUE
InvC-Strep	1	M	OrgB	138	K	11.61	2	TT	FALSE
SpaO	164	S	SpaO	189	K	11.60	1	TT	TRUE
InvC-Strep	15	K	OrgB	124	K	11.60	1	TT	FALSE
InvC-Strep	15	K	SpaO	118	K	11.59	1	TT	FALSE
InvC-Strep	125	S	SpaO	190	K	11.58	1	TT	FALSE
InvC-Strep	43	K	InvC-Strep	358	T	11.57	4	TT	TRUE
InvC-Strep	160	S	OrgB	139	K	11.56	2	TT	FALSE

InvC-Strep	331	K	InvC-Strep	368	K	11.54	3	TT	TRUE
InvC-Strep	2	K	InvC-Strep	101	K	11.52	3	TT	TRUE
OrgB	138	K	SpaO	33	Y	11.51	2	TT	FALSE
OrgB	138	K	SpaO	30	T	11.51	2	TT	FALSE
OrgB	20	K	OrgB	53	K	11.51	2	TT	TRUE
InvC-Strep	345	K	SpaO	190	K	11.50	3	TT	FALSE
InvC-Strep	208	K	InvC-Strep	242	K	11.49	4	TT	TRUE
OrgB	53	K	OrgB	54	T	11.47	1	TT	TRUE
OrgB	12	S	OrgB	219	K	11.40	1	TT	TRUE
OrgB	22	K	SpaO	118	K	11.35	1	TT	FALSE
OrgB	22	K	OrgB	205	S	11.34	2	TT	TRUE
InvC-Strep	17	T	InvC-Strep	43	K	11.33	3	TT	TRUE
OrgB	22	K	SpaO	124	S	11.31	1	TT	FALSE
InvC-Strep	1	M	SpaO	48	K	11.30	3	TT	FALSE
InvC-Strep	15	K	InvC-Strep	368	K	11.27	1	TT	TRUE
OrgB	132	T	OrgB	219	K	11.26	1	TT	TRUE
InvC-Strep	331	K	SpaO	190	K	11.24	2	TT	FALSE
InvC-Strep	208	K	SpaO	48	K	11.22	1	TT	FALSE
OrgB	22	K	SpaO	161	T	11.21	1	TT	FALSE
InvC-Strep	2	K	InvC-Strep	345	K	11.20	3	TT	TRUE
OrgB	20	K	SpaO	48	K	11.19	1	TT	FALSE
OrgB	139	K	SpaOc	1	M	11.19	1	TT	FALSE
InvC-Strep	320	S	SpaO	48	K	11.18	1	TT	FALSE
OrgB	138	K	OrgB	138	K	11.17	2	TT	TRUE
OrgB	53	K	SpaO	20	T	11.16	2	TT	FALSE
InvC-Strep	43	K	OrgB	220	T	11.16	1	TT	FALSE
InvC-Strep	15	K	OrgB	139	K	11.11	2	TT	FALSE
InvC-Strep	345	K	InvC-Strep	368	K	11.09	2	TT	TRUE
SpaO	118	K	SpaO	146	K	11.08	1	TT	TRUE
InvC-Strep	231	T	OrgB	53	K	11.06	1	TT	FALSE
InvC-Strep	71	S	OrgB	53	K	11.05	2	TT	FALSE
OrgB	92	K	SpaO	112	S	11.03	1	TT	FALSE
OrgB	138	K	SpaO	48	K	11.02	3	TT	FALSE
InvC-Strep	1	M	InvC-Strep	331	K	10.98	4	TT	TRUE
InvC-Strep	345	K	InvC-Strep	356	T	10.98	1	TT	TRUE
InvC-Strep	77	Y	OrgB	219	K	10.97	2	TT	FALSE
InvC-Strep	15	K	OrgB	20	K	10.97	1	TT	FALSE
OrgB	219	K	SpaO	187	Y	10.96	1	TT	FALSE
OrgB	219	K	SpaO	112	S	10.94	3	TT	FALSE
InvC-Strep	15	K	InvC-Strep	79	T	10.93	1	TT	TRUE
OrgB	92	K	OrgB	139	K	10.91	1	TT	TRUE
InvC-Strep	1	M	OrgB	219	K	10.88	4	TT	FALSE
InvC-Strep	358	T	OrgB	22	K	10.87	1	TT	FALSE
OrgB	22	K	OrgB	219	K	10.87	4	TT	TRUE
InvC-Strep	12	Y	InvC-Strep	43	K	10.86	2	TT	TRUE
InvC-Strep	71	S	OrgB	22	K	10.84	2	TT	FALSE
InvC-Strep	242	K	OrgB	20	K	10.83	1	TT	FALSE
OrgB	53	K	SpaO	44	S	10.83	2	TT	FALSE
OrgB	22	K	SpaO	189	K	10.81	6	TT	FALSE
OrgB	53	K	OrgB	143	K	10.79	2	TT	TRUE
OrgB	53	K	OrgB	139	K	10.77	3	TT	TRUE
SpaO	159	S	SpaO	164	S	10.74	1	TT	TRUE
SpaO	48	K	SpaO	187	Y	10.70	1	TT	TRUE
InvC-Strep	242	K	OrgB	138	K	10.64	3	TT	FALSE
InvC-Strep	2	K	OrgB	219	K	10.64	4	TT	FALSE
InvC-Strep	43	K	SpaO	112	S	10.62	2	TT	FALSE
InvC-Strep	205	K	InvC-Strep	345	K	10.59	1	TT	TRUE
InvC-Strep	327	Y	InvC-Strep	345	K	10.58	1	TT	TRUE
InvC-Strep	71	S	SpaO	190	K	10.54	1	TT	FALSE
InvC-Strep	43	K	InvC-Strep	320	S	10.54	1	TT	TRUE
InvC-Strep	2	K	InvC-Strep	160	S	10.54	4	TT	TRUE
SpaO	189	K	SpaO	190	K	10.54	3	TT	TRUE
InvC-Strep	43	K	SpaOc	1	M	10.53	1	TT	FALSE
OrgB	219	K	SpaO	164	S	10.52	1	TT	FALSE
InvC-Strep	15	K	InvC-Strep	126	Y	10.50	1	TT	TRUE

OrgB	53	K	OrgB	160	Y	10.50	2	TT	TRUE
InvC-Strep	15	K	InvC-Strep	358	T	10.48	2	TT	TRUE
OrgB	22	K	SpaO	30	T	10.47	2	TT	FALSE
SpaO	164	S	SpaO	185	Y	10.47	1	TT	TRUE
InvC-Strep	15	K	SpaO	48	K	10.45	1	TT	FALSE
InvC-Strep	2	K	OrgB	53	K	10.37	4	TT	FALSE
InvC-Strep	9	Y	OrgB	219	K	10.36	1	TT	FALSE
InvC-Strep	364	S	InvC-Strep	368	K	10.32	1	TT	TRUE
InvC-Strep	79	T	OrgB	22	K	10.32	2	TT	FALSE
InvC-Strep	71	S	InvC-Strep	331	K	10.31	1	TT	TRUE
InvC-Strep	71	S	OrgB	139	K	10.29	2	TT	FALSE
OrgB	53	K	SpaO	30	T	10.26	2	TT	FALSE
OrgB	219	K	SpaO	124	S	10.24	1	TT	FALSE
InvC-Strep	43	K	OrgB	139	K	10.23	4	TT	FALSE
OrgB	139	K	SpaO	118	K	10.23	5	TT	FALSE
OrgB	22	K	OrgB	220	T	10.21	2	TT	TRUE
OrgB	12	S	SpaO	190	K	10.13	2	TT	FALSE
InvC-Strep	368	K	OrgB	53	K	10.13	3	TT	FALSE
InvC-Strep	1	M	OrgB	53	K	10.10	4	TT	FALSE
InvC-Strep	345	K	OrgB	217	K	10.09	1	TT	FALSE
SpaO	30	T	SpaO	48	K	10.08	4	TT	TRUE
InvC-Strep	62	S	OrgB	53	K	10.05	2	TT	FALSE
OrgB	22	K	OrgB	92	K	10.02	1	TT	TRUE
InvC-Strep	1	M	InvC-Strep	160	S	10.01	3	TT	TRUE
InvC-Strep	128	S	SpaO	190	K	10.01	1	TT	FALSE
SpaO	48	K	SpaO	118	K	9.99	4	TT	TRUE
OrgB	12	S	OrgB	92	K	9.98	1	TT	TRUE
InvC-Strep	441	K	OrgB	22	K	9.97	1	TT	FALSE
OrgB	20	K	SpaO	164	S	9.96	1	TT	FALSE
InvC-Strep	62	S	InvC-Strep	77	Y	9.96	1	TT	TRUE
InvC-Strep	160	S	SpaO	189	K	9.93	2	TT	FALSE
SpaO	44	S	SpaO	189	K	9.93	2	TT	TRUE
InvC-Strep	368	K	OrgB	9	S	9.92	1	TT	FALSE
OrgB	9	S	OrgB	92	K	9.84	1	TT	TRUE
InvC-Strep	208	K	OrgB	22	K	9.84	1	TT	FALSE
InvC-Strep	277	S	OrgB	22	K	9.80	1	TT	FALSE
InvC-Strep	9	Y	OrgB	53	K	9.77	2	TT	FALSE
InvC-Strep	2	K	InvC-Strep	71	S	9.77	2	TT	TRUE
InvC-Strep	62	S	OrgB	139	K	9.76	1	TT	FALSE
OrgB	139	K	OrgB	205	S	9.75	1	TT	TRUE
InvC-Strep	148	T	InvC-Strep	345	K	9.74	2	TT	TRUE
InvC-Strep	1	M	InvC-Strep	114	S	9.73	1	TT	TRUE
InvC-Strep	79	T	OrgB	53	K	9.72	1	TT	FALSE
InvC-Strep	208	K	OrgB	138	K	9.72	1	TT	FALSE
InvC-Strep	231	T	InvC-Strep	345	K	9.71	1	TT	TRUE
OrgB	53	K	OrgB	178	S	9.70	1	TT	TRUE
OrgB	22	K	OrgB	169	S	9.70	3	TT	TRUE
OrgB	159	K	SpaO	48	K	9.68	2	TT	FALSE
SpaO	48	K	SpaO	185	Y	9.68	1	TT	TRUE
InvC-Strep	2	K	InvC-Strep	62	S	9.67	2	TT	TRUE
OrgB	219	K	SpaO	158	S	9.65	1	TT	FALSE
InvC-Strep	320	S	InvC-Strep	368	K	9.65	1	TT	TRUE
OrgB	53	K	SpaO	35	T	9.61	1	TT	FALSE
InvC-Strep	160	S	InvC-Strep	368	K	9.61	1	TT	TRUE
OrgB	27	Y	OrgB	219	K	9.60	1	TT	TRUE
OrgB	124	K	OrgB	217	K	9.59	1	TT	TRUE
InvC-Strep	205	K	OrgB	138	K	9.57	2	TT	FALSE
OrgB	219	K	SpaO	33	Y	9.57	2	TT	FALSE
InvC-Strep	128	S	InvC-Strep	160	S	9.56	1	TT	TRUE
InvC-Strep	59	T	OrgB	53	K	9.55	1	TT	FALSE
OrgB	22	K	SpaO	33	Y	9.55	1	TT	FALSE
SpaO	48	K	SpaOc	2	E	9.54	1	TT	FALSE
OrgB	20	K	SpaO	112	S	9.53	1	TT	FALSE
OrgB	139	K	OrgB	218	T	9.49	1	TT	TRUE
InvC-Strep	231	T	OrgB	138	K	9.48	1	TT	FALSE

InvC-Strep	43	K	OrgB	9	S	9.46	1	TT	FALSE
InvC-Strep	331	K	OrgB	124	K	9.46	1	TT	FALSE
InvC-Strep	2	K	SpaO	146	K	9.44	1	TT	FALSE
OrgB	217	K	SpaO	189	K	9.44	2	TT	FALSE
InvC-Strep	1	M	SpaO	190	K	9.43	2	TT	FALSE
InvC-Strep	358	T	OrgB	138	K	9.43	1	TT	FALSE
OrgB	29	S	SpaOc	1	M	9.42	1	TT	FALSE
OrgB	53	K	OrgB	191	K	9.40	2	TT	TRUE
InvC-Strep	1	M	SpaO	189	K	9.38	2	TT	FALSE
InvC-Strep	114	S	OrgB	139	K	9.36	1	TT	FALSE
SpaO	35	T	SpaO	190	K	9.35	2	TT	TRUE
InvC-Strep	329	S	InvC-Strep	345	K	9.34	2	TT	TRUE
OrgB	9	S	OrgB	53	K	9.33	2	TT	TRUE
OrgB	20	K	SpaO	124	S	9.32	2	TT	FALSE
InvC-Strep	114	S	InvC-Strep	242	K	9.32	1	TT	TRUE
OrgB	53	K	SpaO	33	Y	9.30	2	TT	FALSE
InvC-Strep	368	K	OrgB	124	K	9.29	1	TT	FALSE
OrgB	53	K	SpaO	158	S	9.28	1	TT	FALSE
InvC-Strep	368	K	OrgB	203	T	9.25	1	TT	FALSE
InvC-Strep	43	K	InvC-Strep	205	K	9.25	1	TT	TRUE
InvC-Strep	62	S	OrgB	22	K	9.24	2	TT	FALSE
OrgB	96	Y	SpaO	48	K	9.20	1	TT	FALSE
InvC-Strep	77	Y	OrgB	22	K	9.18	1	TT	FALSE
OrgB	29	S	OrgB	219	K	9.14	1	TT	TRUE
InvC-Strep	2	K	SpaO	20	T	9.13	1	TT	FALSE
OrgB	92	K	SpaO	189	K	9.13	2	TT	FALSE
OrgB	138	K	SpaOc	1	M	9.13	1	TT	FALSE
InvC-Strep	2	K	InvC-Strep	126	Y	9.08	1	TT	TRUE
InvC-Strep	368	K	SpaO	48	K	9.07	2	TT	FALSE
InvC-Strep	126	Y	InvC-Strep	338	Y	9.05	1	TT	TRUE
SpaO	185	Y	SpaO	190	K	9.05	1	TT	TRUE
InvC-Strep	265	S	OrgB	138	K	9.01	1	TT	FALSE
InvC-Strep	368	K	OrgB	219	K	8.99	4	TT	FALSE
InvC-Strep	126	Y	InvC-Strep	345	K	8.98	1	TT	TRUE
InvC-Strep	128	S	OrgB	138	K	8.98	2	TT	FALSE
InvC-Strep	206	K	InvC-Strep	242	K	8.96	1	TT	TRUE
OrgB	9	S	OrgB	20	K	8.96	1	TT	TRUE
InvC-Strep	355	T	InvC-Strep	368	K	8.92	2	TT	TRUE
InvC-Strep	128	S	OrgB	53	K	8.91	1	TT	FALSE
InvC-Strep	43	K	OrgB	12	S	8.91	1	TT	FALSE
InvC-Strep	160	S	InvC-Strep	205	K	8.91	4	TT	TRUE
InvC-Strep	364	S	OrgB	219	K	8.86	1	TT	FALSE
OrgB	160	Y	OrgB	219	K	8.82	1	TT	TRUE
InvC-Strep	1	M	OrgB	96	Y	8.81	3	TT	FALSE
OrgB	1	M	OrgB	22	K	8.80	3	TT	TRUE
SpaO	48	K	SpaO	159	S	8.78	1	TT	TRUE
OrgB	138	K	OrgB	169	S	8.78	1	TT	TRUE
OrgB	139	K	SpaO	189	K	8.77	1	TT	FALSE
OrgB	191	K	OrgB	219	K	8.77	2	TT	TRUE
OrgB	96	Y	OrgB	124	K	8.76	1	TT	TRUE
InvC-Strep	1	M	OrgB	92	K	8.73	1	TT	FALSE
OrgB	143	K	SpaO	48	K	8.69	2	TT	FALSE
InvC-Strep	1	M	InvC-Strep	368	K	8.68	6	TT	TRUE
OrgB	217	K	SpaO	48	K	8.67	1	TT	FALSE
OrgB	53	K	SpaO	112	S	8.64	3	TT	FALSE
SpaO	159	S	SpaO	187	Y	8.62	1	TT	TRUE
OrgB	53	K	OrgB	96	Y	8.59	1	TT	TRUE
InvC-Strep	160	S	OrgB	20	K	8.58	1	TT	FALSE
OrgB	22	K	OrgB	132	T	8.57	2	TT	TRUE
InvC-Strep	148	T	OrgB	22	K	8.57	1	TT	FALSE
InvC-Strep	15	K	OrgB	9	S	8.56	1	TT	FALSE
InvC-Strep	331	K	OrgB	29	S	8.55	1	TT	FALSE
InvC-Strep	43	K	InvC-Strep	148	T	8.53	1	TT	TRUE
InvC-Strep	126	Y	OrgB	22	K	8.52	2	TT	FALSE
InvC-Strep	331	K	InvC-Strep	364	S	8.49	1	TT	TRUE

InvC-Strep	62	S	InvC-Strep	101	K	8.49	1	TT	TRUE
InvC-Strep	329	S	OrgB	22	K	8.48	1	TT	FALSE
SpaO	48	K	SpaO	239	Y	8.47	1	TT	TRUE
SpaO	48	K	SpaOc	37	Y	8.47	1	TT	FALSE
InvC-Strep	160	S	InvC-Strep	242	K	8.47	3	TT	TRUE
InvC-Strep	43	K	OrgB	96	Y	8.46	1	TT	FALSE
OrgB	139	K	OrgB	220	T	8.45	1	TT	TRUE
OrgB	96	Y	OrgB	217	K	8.44	1	TT	TRUE
OrgB	219	K	SpaO	159	S	8.44	1	TT	FALSE
OrgB	217	K	SpaO	124	S	8.43	1	TT	FALSE
OrgB	96	Y	OrgB	219	K	8.41	2	TT	TRUE
InvC-Strep	62	S	InvC-Strep	242	K	8.41	2	TT	TRUE
OrgB	53	K	SpaO	161	T	8.39	1	TT	FALSE
OrgB	139	K	SpaO	48	K	8.39	1	TT	FALSE
OrgB	22	K	SpaO	48	K	8.38	3	TT	FALSE
InvC-Strep	2	K	InvC-Strep	9	Y	8.37	3	TT	TRUE
InvC-Strep	101	K	InvC-Strep	205	K	8.36	1	TT	TRUE
OrgB	20	K	SpaO	30	T	8.35	1	TT	FALSE
InvC-Strep	2	K	InvC-Strep	59	T	8.34	1	TT	TRUE
InvC-Strep	71	S	OrgB	92	K	8.33	1	TT	FALSE
InvC-Strep	17	T	InvC-Strep	242	K	8.33	1	TT	TRUE
InvC-Strep	71	S	InvC-Strep	77	Y	8.32	1	TT	TRUE
InvC-Strep	208	K	OrgB	53	K	8.30	1	TT	FALSE
InvC-Strep	265	S	OrgB	53	K	8.28	1	TT	FALSE
InvC-Strep	345	K	InvC-Strep	355	T	8.26	1	TT	TRUE
SpaO	48	K	SpaO	112	S	8.23	2	TT	TRUE
InvC-Strep	71	S	SpaO	48	K	8.23	2	TT	FALSE
InvC-Strep	205	K	OrgB	53	K	8.20	1	TT	FALSE
OrgB	139	K	SpaO	35	T	8.18	1	TT	FALSE
InvC-Strep	1	M	OrgB	22	K	8.17	6	TT	FALSE
InvC-Strep	208	K	OrgB	139	K	8.16	1	TT	FALSE
InvC-Strep	1	M	OrgB	139	K	8.16	2	TT	FALSE
InvC-Strep	1	M	InvC-Strep	9	Y	8.15	2	TT	TRUE
OrgB	92	K	SpaO	44	S	8.14	1	TT	FALSE
InvC-Strep	15	K	InvC-Strep	125	S	8.14	1	TT	TRUE
InvC-Strep	79	T	InvC-Strep	101	K	8.13	1	TT	TRUE
InvC-Strep	15	K	InvC-Strep	242	K	8.13	1	TT	TRUE
OrgB	138	K	OrgB	139	K	8.12	4	TT	TRUE
InvC-Strep	2	K	SpaO	189	K	8.10	2	TT	FALSE
SpaO	112	S	SpaO	189	K	8.09	2	TT	TRUE
InvC-Strep	2	K	InvC-Strep	12	Y	8.08	1	TT	TRUE
InvC-Strep	160	S	OrgB	217	K	8.07	1	TT	FALSE
InvC-Strep	12	Y	InvC-Strep	368	K	8.07	1	TT	TRUE
OrgB	20	K	SpaO	159	S	8.06	1	TT	FALSE
InvC-Strep	2	K	InvC-Strep	358	T	8.06	1	TT	TRUE
InvC-Strep	205	K	SpaO	146	K	8.06	1	TT	FALSE
OrgB	29	S	OrgB	217	K	8.00	1	TT	TRUE
InvC-Strep	231	T	OrgB	22	K	8.00	1	TT	FALSE
OrgB	22	K	SpaO	112	S	7.99	3	TT	FALSE
OrgB	203	T	SpaO	48	K	7.97	1	TT	FALSE
InvC-Strep	43	K	InvC-Strep	265	S	7.97	1	TT	TRUE
InvC-Strep	160	S	InvC-Strep	364	S	7.95	1	TT	TRUE
InvC-Strep	15	K	InvC-Strep	205	K	7.93	1	TT	TRUE
SpaO	112	S	SpaO	190	K	7.90	2	TT	TRUE
InvC-Strep	79	T	OrgB	219	K	7.90	1	TT	FALSE
InvC-Strep	220	S	OrgB	22	K	7.89	2	TT	FALSE
InvC-Strep	43	K	InvC-Strep	126	Y	7.85	1	TT	TRUE
OrgB	138	K	SpaO	161	T	7.84	1	TT	FALSE
InvC-Strep	364	S	OrgB	22	K	7.78	1	TT	FALSE
InvC-Strep	43	K	InvC-Strep	231	T	7.78	1	TT	TRUE
InvC-Strep	215	T	InvC-Strep	331	K	7.77	1	TT	TRUE
OrgB	22	K	OrgB	29	S	7.76	1	TT	TRUE
InvC-Strep	2	K	OrgB	205	S	7.75	1	TT	FALSE
OrgB	159	K	SpaO	112	S	7.75	1	TT	FALSE
InvC-Strep	128	S	OrgB	139	K	7.73	1	TT	FALSE

OrgB	9	S	SpaO	118	K	7.73	1	TT	FALSE
OrgB	53	K	SpaO	18	T	7.71	1	TT	FALSE
InvC-Strep	128	S	InvC-Strep	368	K	7.70	1	TT	TRUE
InvC-Strep	265	S	OrgB	139	K	7.70	1	TT	FALSE
OrgB	12	S	OrgB	139	K	7.68	1	TT	TRUE
OrgB	22	K	OrgB	27	Y	7.68	2	TT	TRUE
SpaO	30	T	SpaO	190	K	7.67	1	TT	TRUE
OrgB	22	K	SpaO	239	Y	7.67	1	TT	FALSE
OrgB	22	K	SpaOc	37	Y	7.67	1	TT	FALSE
OrgB	9	S	OrgB	12	S	7.67	2	TT	TRUE
InvC-Strep	1	M	InvC-Strep	364	S	7.65	1	TT	TRUE
InvC-Strep	15	K	InvC-Strep	265	S	7.65	1	TT	TRUE
OrgB	3	K	OrgB	22	K	7.64	3	TT	TRUE
InvC-Strep	206	K	InvC-Strep	345	K	7.64	1	TT	TRUE
InvC-Strep	43	K	InvC-Strep	220	S	7.64	1	TT	TRUE
InvC-Strep	9	Y	OrgB	20	K	7.62	1	TT	FALSE
InvC-Strep	9	Y	InvC-Strep	368	K	7.61	2	TT	TRUE
InvC-Strep	206	K	OrgB	138	K	7.59	1	TT	FALSE
InvC-Strep	2	K	InvC-Strep	208	K	7.57	1	TT	TRUE
InvC-Strep	77	Y	InvC-Strep	368	K	7.56	1	TT	TRUE
InvC-Strep	208	K	InvC-Strep	345	K	7.56	1	TT	TRUE
OrgB	9	S	SpaO	189	K	7.50	1	TT	FALSE
OrgB	220	T	SpaOc	1	M	7.50	1	TT	FALSE
InvC-Strep	331	K	SpaO	20	T	7.49	1	TT	FALSE
InvC-Strep	2	K	SpaO	159	S	7.49	1	TT	FALSE
InvC-Strep	9	Y	OrgB	22	K	7.48	1	TT	FALSE
InvC-Strep	231	T	OrgB	139	K	7.45	1	TT	FALSE
InvC-Strep	364	S	OrgB	53	K	7.41	1	TT	FALSE
InvC-Strep	364	S	OrgB	139	K	7.40	1	TT	FALSE
InvC-Strep	114	S	OrgB	53	K	7.40	1	TT	FALSE
InvC-Strep	15	K	OrgB	12	S	7.39	1	TT	FALSE
InvC-Strep	277	S	OrgB	53	K	7.37	1	TT	FALSE
InvC-Strep	2	K	InvC-Strep	114	S	7.37	1	TT	TRUE
OrgB	22	K	SpaO	35	T	7.37	1	TT	FALSE
OrgB	20	K	OrgB	217	K	7.36	1	TT	TRUE
OrgB	219	K	SpaO	44	S	7.35	1	TT	FALSE
SpaO	190	K	SpaOc	31	K	7.34	1	TT	FALSE
InvC-Strep	1	M	OrgB	23	T	7.32	1	TT	FALSE
InvC-Strep	128	S	InvC-Strep	345	K	7.31	1	TT	TRUE
InvC-Strep	125	S	InvC-Strep	368	K	7.31	1	TT	TRUE
SpaO	33	Y	SpaO	118	K	7.30	1	TT	TRUE
OrgB	217	K	SpaO	164	S	7.29	1	TT	FALSE
OrgB	203	T	SpaO	189	K	7.28	1	TT	FALSE
InvC-Strep	331	K	OrgB	9	S	7.28	1	TT	FALSE
InvC-Strep	1	M	InvC-Strep	208	K	7.26	1	TT	TRUE
InvC-Strep	148	T	SpaO	48	K	7.26	1	TT	FALSE
InvC-Strep	1	M	OrgB	124	K	7.25	1	TT	FALSE
InvC-Strep	205	K	InvC-Strep	331	K	7.22	1	TT	TRUE
InvC-Strep	128	S	OrgB	22	K	7.22	2	TT	FALSE
OrgB	12	S	OrgB	20	K	7.21	1	TT	TRUE
InvC-Strep	128	S	InvC-Strep	215	T	7.18	1	TT	TRUE
InvC-Strep	15	K	InvC-Strep	114	S	7.15	1	TT	TRUE
OrgB	12	S	OrgB	124	K	7.15	1	TT	TRUE
OrgB	20	K	SpaO	187	Y	7.13	1	TT	FALSE
InvC-Strep	1	M	InvC-Strep	242	K	7.12	1	TT	TRUE
InvC-Strep	215	T	OrgB	219	K	7.10	1	TT	FALSE
InvC-Strep	128	S	SpaO	118	K	7.10	1	TT	FALSE
OrgB	124	K	SpaO	189	K	7.08	1	TT	FALSE
InvC-Strep	17	T	OrgB	22	K	7.04	1	TT	FALSE
InvC-Strep	125	S	SpaO	48	K	7.03	1	TT	FALSE
OrgB	203	T	SpaO	118	K	7.02	1	TT	FALSE
InvC-Strep	242	K	OrgB	53	K	6.93	2	TT	FALSE
InvC-Strep	2	K	OrgB	9	S	6.92	1	TT	FALSE
InvC-Strep	160	S	OrgB	124	K	6.90	1	TT	FALSE
InvC-Strep	364	S	SpaO	48	K	6.85	1	TT	FALSE

OrgB	20	K	SpaO	189	K	6.83	1	TT	FALSE
OrgB	159	K	SpaO	189	K	6.82	1	TT	FALSE
InvC-Strep	277	S	OrgB	219	K	6.82	1	TT	FALSE
InvC-Strep	355	T	OrgB	22	K	6.81	1	TT	FALSE
InvC-Strep	43	K	InvC-Strep	84	S	6.79	2	TT	TRUE
SpaO	33	Y	SpaO	44	S	6.78	2	TT	TRUE
InvC-Strep	242	K	InvC-Strep	331	K	6.75	2	TT	TRUE
InvC-Strep	43	K	SpaO	33	Y	6.75	1	TT	FALSE
InvC-Strep	12	Y	OrgB	53	K	6.73	1	TT	FALSE
OrgB	22	K	OrgB	178	S	6.73	2	TT	TRUE
InvC-Strep	3	T	InvC-Strep	160	S	6.70	2	TT	TRUE
InvC-Strep	242	K	OrgB	124	K	6.69	1	TT	FALSE
InvC-Strep	205	K	OrgB	219	K	6.69	1	TT	FALSE
OrgB	22	K	OrgB	139	K	6.67	4	TT	TRUE
SpaO	1	M	SpaOc	2	E	6.66	1	TT	FALSE
OrgB	160	Y	SpaO	161	T	6.65	1	TT	FALSE
InvC-Strep	12	Y	OrgB	139	K	6.65	1	TT	FALSE
InvC-Strep	368	K	SpaO	33	Y	6.60	1	TT	FALSE
InvC-Strep	329	S	OrgB	53	K	6.57	1	TT	FALSE
InvC-Strep	148	T	OrgB	138	K	6.56	1	TT	FALSE
OrgB	9	S	SpaO	190	K	6.55	2	TT	FALSE
InvC-Strep	331	K	InvC-Strep	358	T	6.52	2	TT	TRUE
OrgB	139	K	SpaO	112	S	6.50	2	TT	FALSE
InvC-Strep	368	K	OrgB	1	M	6.43	2	TT	FALSE
InvC-Strep	205	K	InvC-Strep	208	K	5.99	2	TT	TRUE
SpaO	112	S	SpaO	124	S	5.68	3	TT	TRUE

Supplementary Table 7: Protein sequences used in this study. Constructs were designed in: MxiH-strep (includes a factor-Xa protease cleavage site) [33], IpaB-knot-strep [26], InvCΔ79 [140,201], SpaO/SpaO$_C$/OrgB/InvC complex [137,202]. Identical sequences are highlighted in the same color (Strep-tag: green, InvC: orange, SpaO: blue).

MxiH-strep

MASWSHPQFEKIEGRMSVTVPNDDWTLSSLSETFDDGTQTLQGELTLALDKLAKNPSNPQLLAEYQSKLSEYTLYRNAQSNTVKVIKDVDA
AIIQNFR

IpaB-knot-strep

MHNVSTTTTGFPLAKILTSTELGDNTIQAANDAANKLFSLTIADLTANQNINTTNAHSTSNILIPELKAPKSLNASSQLTLLIGNLIQILGEKSLTA
LTNKITAWKSQQQARQQKNLEFSDKINTLLSETEGLTRDYEKQINKLKNADSKIKDLENKINQIQTRLSELDPESPEKKKLSREEIQLTIKKDAAV
KDRTLIEQKTLSIHSKLTDKSMQLEKEIDSFSAFSNTASAEQLSTQQKSLTGLASVTQLMATFIQLVGKNNEESLKNDLALFQSLQESRKTEMER
KSDEYAAEVRKAEELNRVMGCVGKILGALLTIVSVVAAAFSGGASLALAAVGLALMVTDAIVQAATGNSFMEQALNPIMKAVIEPLIKLLSDA
FTKMLEGLGVDSKKAKMIGSILGAIAGALVLVAAVVLVATVGKQAAAKLAENIGKIIGKTLTDLIPKFLKNFSSQLDDLITNAVARLNKFLGAAG
DEVISKQIISTHLNQAVLLGESVNSATQAGGSVASAVFQNSASTNLADLTLSKYQVEQLSKYISEAIEKFGQLQEVIADLLASMSNSQANRTDV
AKAILQQTTAMRITSTANPRIKELARLLERKHRDSQRRFLIEGAREIERALQAGIELEQALVWEGGLNPEEQQVYAALGRVGRLALLEVSEAVL
KKLSVRDNPAGLIALARMPERTLEEYRPSPDALILVAVGLEKPGNLGAVLRSADAAGAEAVLVAGGVDLYSPQVIRNSTGVVFSLRTLAASESE
VLDWIKQHNLPLVATTPHAEALYWEANLRPPVAIAVGPEHEGLRAAWLEAAQTQVRIPMQGQADSLNVSVSAALLLYEALRQRLLRDRLTK
THSTLSAWSHPQFEK

InvCΔ79

MGRALSAWVGYSVLGAVLDPTGKIVERFTPEVAPISEERVIDVAPPSYASRVGVREPLITGVRAIDGLLTCGVGQRMGIFASAGCGKTMLMH
MLIEQTEADVFVIGLIGERGREVTEFVDMLRASHKKEKCVLVFATSDFPSVDRCNAAQLATTVAEYFRDQGKRVVLFIDSMTRYARALRDVAL
ASGERPARRGYPASVFDNLPRLLERPGATSEGSITAFYTVLLESEEEADPMADEIRSILDGHLYLSRKLAGQGHYPAIDVLKSVSRVFGQVTTPT
HAEQASAVRKLMTRLEELQLFIDLGEYRPGENIDNDRAMQMRDSLKAWLCQPVAQYSSFDDTLSGMNAFADQNSAWSHPQFEK

SpaO/SpaO$_C$/OrgB/InvC complex

SpaO

MSLRVRQIDRREWLLAQTATECQRHGREATLEYPTRQGMWVRLSDAEKRWSAWIKPGDWLEHVSPALAGAAVSAGAEHLVVPWLAATE
RPFELPVPHLSCRRLCVENPVPGSALPEGKLLHIMSDRGGLWFEHLPELPAVGGGRPKMLRWPLRFVIGSSDTQRSLLGRIGIGDVLLIRTSRA
EVYCYAKKLGHFNRVEGGIIVETLDIQHIEEENNTTETAETLPGLNQLPVKLEFVLYRKNVTLAELEAMGQQQLLSLPTNAELNVEIMANGVLL
GNGELVQMNDTLGVEIHEWLSESGNGE

SpaO$_C$

METLDIQHIEEENNTTETAETLPGLNQLPVKLEFVLYRKNVTLAELEAMGQQQLLSLPTNAELNVEIMANGVLLGNGELVQMNDTLGVEIHE
WLSESGNGE

OrgB

MLKNIPIPSPLSPVEGILIKRKTLERYFSIERLEQQAHQRAKRILREAEEEAKTLRMYAYQEGYEQGMIDALQQVAAYLTDNQTMAWKWMEK
IQIYARELFSAAVDHPETLLTVLDEWLRDFDKPEGQLFLTLPVNAKKDHQKLMVLLMENWPGTFNLKYHQEQRFIMSCGDQIAEFSPEQFVE
TAVGVIKHHLDELPQDCRTISDNAINALIDEWKTKTQAEVIR

InvC

MKTPRLLQYLAYPQKITGPIIEAELRDVAIGELCEIRRGWHQKQVVARAQVVGLQRERTVLSLIGNAQGLSRDVVLYPTGRALSAWVGYSVLG
AVLDPTGKIVERFTPEVAPISEERVIDVAPPSYASRVGVREPLITGVRAIDGLLTCGVGQRMGIFASAGCGKTMLMHMLIEQTEADVFVIGLIG
ERGREVTEFVDMLRASHKKEKCVLVFATSDFPSVDRCNAAQLATTVAEYFRDQGKRVVLFIDSMTRYARALRDVALASGERPARRGYPASVF
DNLPRLLERPGATSEGSITAFYTVLLESEEEADPMADEIRSILDGHLYLSRKLAGQGHYPAIDVLKSVSRVFGQVTTPTHAEQASAVRKLMTRLE
ELQLFIDLGEYRPGENIDNDRAMQMRDSLKAWLCQPVAQYSSFDDTLSGMNAFADQNSAWSHPQFEK

6.3. Lists of used materials

6.3.1. Chemicals and biomaterials

Compound	Abbreviation	CAS RN	Distributor
2-Mercaptoethanol		60-24-2	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Acetic acid		64-19-7	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Acetone		67-64-1	Merk KGaA (Darmstadt, Germany)
Ammonium acetate		631-61-8	Sigma-Aldrich (St. Louis, MO, USA)
Ammonium bicarbonate		1066-33-7	Sigma-Aldrich (St. Louis, MO, USA)
Ammonium hydroxide solution		1336-21-6	Honeywell International Inc. (Charlotte, NC, USA)
Ampicillin sodium salt		69-53-4	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Anhydrotetracycline	AHT	13803-65-1	Cayman Chemical Company (Ann Arbor, MI, USA)
Argon		7440-37-1	SOL Germany (Gersthofen, Germany)
Biotin		58-85-5	Sigma-Aldrich (St. Louis, MO, USA)
Bis(sulfosuccinimidyl)suberate	BS3	82436-77-9	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Bromophenol blue sodium salt		34725-61-6	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Carbenicillin disodium salt		4800-94-6	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Cesium iodide	CsI	7789-17-5	Sigma-Aldrich (St. Louis, MO, USA)
Chloramphenicol		56-75-7	Sigma-Aldrich (St. Louis, MO, USA)
Coomassie brilliant blue G-250		6104-58-1	Bio-Rad Laboratories, Inc. (Hercules, CA, USA)

Dimethyl sulfoxide	DMSO	67-68-5	Sigma-Aldrich (St. Louis, MO, USA)
Dimethylformamide	DMF	68-12-2	Acros Organics/ Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Disodium hydrogen phosphate dihydrate	$Na_2HPO_4 \cdot 2H_2O$	10028-24-7	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Disuccinimidyl suberate	DSS	68528-80-3	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Dithiothreitol	DTT	3483-12-3	Sigma-Aldrich (St. Louis, MO, USA)
Ethylenediaminetetraacetic acid	EDTA	6381-92-6	SERVA Electrophoresis GmbH (Heidelberg, Germany)
Formaldehyde solution (37%)	HCHO	50-00-0	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Glycerol		56-81-5	MP Biomedicals (Santa Ana, CA, USA)
Glycine		56-40-6	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
HEPES		7365-45-9	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Hydrochloric acid fuming (37%)	HCl	7647-01-0	Merk KGaA (Darmstadt, Germany)
Kanamycin sulfate		25389-94-0	Sigma-Aldrich (St. Louis, MO, USA)
Lauryl maltose neopentyl glycol	LMNG	1257852-96-2	Anatrace (Maumee, OH, USA)
Lysogeny broth (Luria/ Miller)	LB	-	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Lysozyme from chicken egg white		12650-88-3	Sigma-Aldrich (St. Louis, MO, USA)
Magnesium sulfate heptahydrate	$MgSO_4 \cdot 7H_2O$	10034-99-8	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Methanol	MeOH	67-56-1	VWR Chemicals (Radnor, PA, USA)
Nitrogen (liquid)	N_2	7727-37-9	Linde plc (Dublin, Ireland)
Potassium dihydrogen phosphate	KH_2PO_4	7778-77-0	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)

Silver nitrate	$AgNO_3$	7761-88-8	Sigma-Aldrich (St. Louis, MO, USA)
Skim milk powder		-	(Sucofin) TSI GmbH + Co. KG (Zeven, Germany)
Sodium carbonate	Na_2CO_3	497-19-8	Sigma-Aldrich (St. Louis, MO, USA)
Sodium chloride	NaCl	7647-14-5	Merck Millipore (Billerica, MA, USA)
Sodium dodecyl sulfate	SDS	151-21-3	Sigma-Aldrich (St. Louis, MO, USA)
Sodium hydroxide	NaOH	1310-73-2	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Sodium thiosulfate pentahydrate	$Na_2S_2O_3 \cdot 5H_2O$	10102-17-7	Merk KGaA (Darmstadt, Germany)
Sucrose		57-50-1	SERVA Electrophoresis GmbH (Heidelberg, Germany)
Trichloroacetic acid	TCA	76-03-9	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Tris base	Tris	77-86-1	Sigma-Aldrich (St. Louis, MO, USA)
Tris hydrochloride	Tris HCl	1185-53-1	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Triton^{TM} X-100		9036-19-5	Sigma-Aldrich (St. Louis, MO, USA)
Tryptic soy broth (CASO-Bouillon)	TSB	-	Carl Roth GmbH + Co. KG (Karlsruhe, Germany)
Tween® 20		9005-64-5	Sigma-Aldrich (St. Louis, MO, USA)
Uranyl acetate		541-09-3	Merk KGaA (Darmstadt, Germany)

6.3.2. Kits and consumables

Consumable	Manufacture	Catalog number
Borosilicate glass capillaries (length: 100 mm; outer diameter: 1.2 mm; inner diameter: 0.68 mm; with filament)	World Precision Instruments	1B120F-4
Carbon film, 400 mesh, copper grids CF400-CU	Electron Microscopy Sciences (Hatfield, PA, USA)	CF400-CU
Chemiluminescent substrate kit SuperSignal^{TM} West Pico PLUS	Thermo Fisher Scientific Inc. (Waltham, MA, USA)	34580
CHROMABOND® LV columns 15 ml PE-frits	Macherey-Nagel GmbH & Co. KG (Düren, Germany)	732501
CultureWell^{TM} CW-50R-1.0 reusable gaskets	Grace Bio-Labs Inc. (Bend, OR, USA)	103250
DNase I	Roche Holding AG (Basel, Switzerland)	10104159001
Falcons 15 ml, PP 50 ml, PP	Sarsted AG &Co. KG (Nümbrecht, Germany)	 62.554.502 62.547.254
Gold target for sputtering	Baltic Präparation (Wetter, Germany)	BP 2224-05
Membrane filters, pore size 0.1 µM	Whatman® / Cytiva (Malborough, MA, US)	6809-5012
Micro Bio-Spin^{TM} P-30 Gel Columns, Tris Buffer, 10-75 µl, 40 000 MW	Bio-Rad Laboratories, Inc. (Hercules, CA, USA)	7326250
Mini-PROTEAN© TGX^{TM} Gels (4-15%)	Bio-Rad Laboratories, Inc. (Hercules, CA, USA)	456-1086
MxiG antibody, mouse	Max-Planck Institute for Infection Biology, Protein Purification facility (Berlin, Germany)	-
NativeMark^{TM} unstained protein ladder	Thermo Fischer Scientific Inc. (Waltham, MA, USA)	LC0725
Nitrocellulose blotting membrane Amersham Protran^{TM} Premium^{TM} NC 0.45	GE Healthcare (Chicago, IL, USA)	10600003
PageRuler^{TM} Plus Prestained Protein Ladder (10-250 kDa)	Thermo Fisher Scientific Inc. (Waltham, MA, USA)	26619
Parafilm®	Amcor plc. (Zürich, Switzerland)	PM-996

Peroxidase AffiniPure Goat Anti-Mouse IgG (H+L)	Jackson ImmunoResearch Laboratories Inc. (West Grove, PA, USA)	115-035-003
Protein LoBind Eppendorf tubes ® 0.5 ml 2.0 ml	Eppendorf (Hamburg, Germany)	022431064 022431102
PVDF blotting membrane Amersham ™ Hybond ™ P 0.45	GE Healthcare (Chicago, IL, USA)	10600023
Reaction tubes 0.5 ml 1.5 ml 2 ml	Brand GmbH & Co. KG (Wertheim, Germany)	72.699 780502 72.691
Rotilabo® single-use cuvettes	Carl Roth GmBH + Co. KG (Karlsruhe, Germany)	XK20
RunBlue™ Bis-Tris Precast Gels (4-12%)	Expedeon Ltd. (Cambridge, UK)	NBT41227
Sample buffer for Blue Native (2x)	SERVA Electrophoresis GmbH (Heidelberg, Germany)	42533
Sample buffer for Clear Native (2x)	SERVA Electrophoresis GmbH (Heidelberg, Germany)	42534
SigmaFAST™ Protease Inhibitor Cocktail Tablet, EDTA free	Sigma-Aldrich (St. Louis, MO, USA)	S8830
Strep-Tacin®Sepharose®	IBA Lifesciences (Göttingen, Germany)	2-1201-010
Strep-tag antibody, mouse, monoclonal	Qiagen (Hilden, Germany)	34850
Ultracentrifugation tubes, 26.3 ml polycarbonate bottle with cap assembly, 25 x 89 mm for rotor 70 Ti	Beckman Coulter Inc. (Brea, CA, USA)	355618
Zeba™ Spin desalting columns, 7k MWCO, 5 ml	Thermo Fisher Scientific Inc. (Waltham, MA, USA)	89891

6.3.3. Instruments

Instrument	Model	Manufacture
Analytical Balance	QUINTIX224-1CEU	Sartorius Lab Instruments GmbH & Co. KG (Göttingen, Germany)
Balance	TE3102S	Sartorius Lab Instruments GmbH & Co. KG (Göttingen, Germany)
Coffee Machine	Dinamica ECAM350.15.B	De'Longhi Group (Treviso, Italy)
Freezer (-20 °C)	1500 Index 21R/ 001	Liebherr-Hausgeräte GmbH (Ochsenhausen, Germany)
Fridge (4 °C)	KT14R121/04	Siemens Electrogeräte GmbH (München, Germany)
Gold/ Carbon sputter Coater	Q150RES	Quorum technologies (Lewes, UK)
Imaging system (chemiluminescence)	ImageQuant LAS 4000 mini	GE Healthcare (Chicago, IL, USA)
Light microscope	Leica DM IL LED Microscope with Leica MC 170 HD camera	Leica Microsystems (Wetzlar, Germany)
Mass photometer	Refeyn OneMP	Refeyn Ltd. (Oxford, UK)
Mass spectrometer	Q-Tof 2	Waters Corporation (Milford, MA, USA) & MS Vision (Almere, the Netherlands)
Micropipette puller	P-1000	Sutter Instruments (Novato, CA, USA)
PH meter	S20 SevenEasy™ pH	Mettler Toledo (Greifensee, Switzerland)
Pipette controller	PipetBoy acu 2	Integra Biosciences AG (Zizers, Switzerland)
Plasma cleaner	Plasma cleaner PDC-002-CE	Harrick Plasma (Ithaca, NY, USA)
Reciprocal shaker	IKA® HS 260 control	IKA (Staufen, Germany)
Reverse action tweezers	5X.SA.1	Ideal-tek (Balerna, Switzerland)
SDS-PAGE Electrophoresis System	Invitrogen™ XCell SureLock™ Mini-Cell Electrophoresis System	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
SDS-PAGE Electrophoresis System	Mini-PROTEAN© Tetra Cell	Bio-Rad Laboratories, Inc. (Hercules, CA, USA)

SDS-PAGE Electrophoresis System	PowerPac^{TM} Basic Power Supply	Bio-Rad Laboratories, Inc. (Hercules, CA, USA)
Semi-dry blotter	V20-SDB	Scie-Plas Ltd. (Cambridge, UK)
Shaking incubator	Multitron/Multitron Pro	Infors AG (Bottmingen, Switzerland)
Single-channel pipettes	Ultra-High Performance 0.5-10 µl (613-1489) 2-20 µl (613-1490) 20-200 µl (613-1492) 100-1000 µl (613-1493)	VWR (Radnor, PA, USA)
Spectrophotometer	NanoDrop^{TM} 2000	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Spectrophotometer	BioPhotometer 6131	Eppendorf (Hamburg, Germany)
Tabletop centrifuge	5804 R (rotor S-4-72)	Eppendorf (Hamburg, Germany)
Tabletop centrifuge	5424 R (rotor FA-45-24-11)	Eppendorf (Hamburg, Germany)
TEM	Talos^{TM} L120C with a Ceta CMOS camera	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Tilt/roller mixer	RS-TR05	Phoenix Instrument GmbH (Garbsen, Germany)
Ultra-low temperature freezer (-80 °C)	UF V 700 E2.1	Binder GmbH (Tuttlingen, Germany)
Ultracentrifuge	Optima XPN-90 (fixed-angle rotor 70 Ti)	Beckman Coulter Inc. (Brea, CA, USA)
Ultrapure Water Dispenser	Milli-Q® Q-POD®	Merck KGaA (Darmstadt, Germany)

6.3.4. Software, web applications and databases

Name	Software version	Address of web application/ database	Developers
ChimeraX	0.9 - 1.3	-	[206,207]
CorelDRAW 2018	20.1.0.707		Corel Corporation (Ottawa, Canada)
ExPASy	-	https://www.expasy.org	[217]
Inkscape	0.9.2.3.	-	The Inkscape Project
MassLynx	4.1	-	Waters Corporation (Milford, MA, USA)
Microsoft Office Professional Plus 2016	16.0.5278.1000	-	Microsoft Corporation (Redmond, WA, USA)
PDBePISA	-	http://www.ebi.ac.uk/pdbe/prot_int/pistart.html	[208]
PubChem	-	https://pubchem.ncbi.nlm.nih.gov	[218]
RCSB PDB	-	https://www.rcsb.org	[219,220]
UCSF Chimera	1.13 - 1.16	-	[205]
UniProt	-	https://www.uniprot.org	[214]
xiNET	-	http://crosslinkviewer.org	[212]
Xlink Analyzer	1.1.4	-	[213]

6.4. Hazardous substances according to the Globally Harmonized System (GHS)

The hazardous substances used in this work declared according to the Globally Harmonized System of Classification and Labelling Chemicals (GHS). Information was retrieved from PubChem [218] (https://pubchem.ncbi.nlm.nih.gov/, accessed: 9-10.06.2022) or safety data sheets provided by manufacturers. The GHS pictograms were obtained from the United Nations Economic Commission for Europe (UNECE) webpage (https://unece.org/transportdangerous-goods/ghs-pictograms, accessed 09.06.2022) with permission.

Substance	Pictogram	Hazard (H) statements	Precautionary (P) statements
2-Mercaptoethanol		H301, H302, H310, H311, H314, H315, H317, H318, H331, H332, H361, H373, H400, H410, H411	P203, P260, P261, P262, P264, P264+P265, P270, P271, P272, P273, P280, P301+P316, P301+P317, P301+P330+P331, P302+P352, P302+P361+P354, P304+P340, P305+P354+P338, P316, P317, P318, P319, P321, P330, P332+P317, P333+P313, P361+P364, P362+P364, P363, P391, P403+P233, P405, P501
Acetic acid		H226, H314	P210, P233, P240, P241, P242, P243, P260, P264, P280, P301+P330+P331, P302+P361+P354, P303+P361+P353, P304+P340, P305+P354+P338, P316, P321, P363, P370+P378, P403+P235, P405, P501
Acetone		H225, H319, H336	P210, P233, P240, P241, P242, P243, P261, P264+P265, P271, P280, P303+P361+P353, P304+P340, P305+P351+P338, P319, P337+P317, P370+P378, P403+P233, P403+P235, P405, P501
Ammonium bicarbonate		H302	P264, P270, P301+P317, P330, P501

Ammonium hydroxide		H314, H400	P260, P264, P273, P280, P301+P330+P331, P302+P361+P354, P304+P340, P305+P354+P338, P316, P321, P363, P391, P405, P501
Ampicillin		H315, H317, H319, H334, H335	P261, P264, P264+P265, P271, P272, P280, P284, P302+P352, P304+P340, P305+P351+P338, P319, P321, P332+P317, P333+P313, P337+P317, P342+P316, P362+P364, P403+P233, P405, P501
Anhydrotetracycline (AHT)		H301, H311, H315, H319, H331, H361	P203, P261, P264, P264+P265, P270, P271, P280, P301+P316, P302+P352, P304+P340, P305+P351+P338, P316, P318, P321, P330, P332+P317, P337+P317, P361+P364, P362+P364, P403+P233, P405, P501
Argon		H280, H281	P403+P410
Bis(sulfosuccinimidyl) suberate (BS3)		H315, H319, H335	P261, P264, P264+P265, P271, P280, P302+P352, P304+P340, P305+P351+P338, P319, P321, P332+P317, P337+P317, P362+P364, P403+P233, P405, P501
Carbenicillin disodium		H317, H334	P261, P272, P280, P284, P302+P352, P304+P340, P321, P333+P313, P342+P316, P362+P364, P501
Cesium iodide (CsI)		H302, H315, H317, H319, H335, H361, H400	P203, P261, P264, P264+P265, P270, P271, P272, P273, P280, P301+P317, P302+P352, P304+P340, P305+P351+P338, P318, P319, P321, P330, P332+P317, P333+P313, P337+P317, P362+P364, P391, P403+P233, P405, P501
Chloramphenicol		H317, H318, H350, H351, H360, H361	P203, P261, P264+P265, P272, P280, P302+P352, P305+P354+P338, P317, P318, P321, P333+P313, P362+P364, P405, P501

	☠ (GHS08)		
Dimethyl sulfoxide (DMSO)	! (GHS07)	H315, H319, H335	P261, P264, P264+P265, P271, P280, P302+P352, P304+P340, P305+P351+P338, P319, P321, P332+P317, P337+P317, P362+P364, P403+P233, P405, P501
Dimethylformamide (DMF)	! (GHS07) ☠ (GHS08)	H312, H319, H332, H360D	P203, P261, P264+P265, P271, P280, P302+P352, P304+P340, P305+P351+P338, P317, P318, P321, P337+P317, P362+P364, P405, P501
Disodium phosphate dihydrate	! (GHS07)	H319	P264+P265, P280, P305+P351+P338, P337+P317
Disuccinimidyl suberate (DSS)	! (GHS07)	H302, H312, H332	P261, P264, P270, P271, P280, P301+P317, P302+P352, P304+P340, P317, P321, P330, P362+P364, P501
Dithiothreitol (DTT)	! (GHS07)	H302, H315, H319, H335	P261, P264, P264+P265, P270, P271, P280, P301+P317, P302+P352, P304+P340, P305+P351+P338, P319, P321, P330, P332+P317, P337+P317, P362+P364, P403+P233, P405, P501
DNase I	☠ (GHS08)	H317, H334	P261, P280, P284, P304+P340, P333+P313, P342+P311
Ethylenediaminetetraacetic acid (EDTA)	! (GHS07) ☠ (GHS08)	H302, H312, H315, H319, H332, H335, H373, H412	P260, P261, P264, P264+P265, P270, P271, P273, P280, P301+P317, P302+P352, P304+P340, P305+P351+P338, P317, P319, P321, P330, P332+P317, P337+P317, P362+P364, P403+P233, P405, P501
Formaldehyde	⟨corrosion⟩ (GHS05) ☠ (GHS06)	H301, H311, H314, H317, H318, H330, H331, H341, H350, H351	P203, P260, P261, P264, P264+P265, P270, P271, P272, P280, P281, P284, P301+P316, P301+P330+P331, P302+P352, P302+P361+P354, P304+P340, P305+P354+P338, P316, P317,

			P318, P320, P321, P330, P333+P313, P361+P364, P362+P364, P363, P403+P233, P405, P501
Hydrochloric acid		H314, H331	P260, P261, P264, P271, P280, P301+P330+P331, P302+P361+P354, P304+P340, P305+P354+P338, P316, P321, P363, P403+P233, P405, P501
Kanamycin sulfate		H360	P203, P280, P318, P405, P501
Lysozyme (egg white)		H334	P261, P284, P304+P340, P342+P316, P501
Methanol		H225, H301, H311, H331, H370	P210, P233, P240, P241, P242, P243, P260, P261, P264, P270, P271, P280, P301+P316, P302+P352, P303+P361+P353, P304+P340, P308+P316, P316, P321, P330, P361+P364, P370+P378, P403+P233, P403+P235, P405, P501
Nitrogen (liquid)		H280, H281	P282, P336+P317, P403, P410+P403
Protease inhibitor Cocktail Tablets, EDTA-free SigmaFast™		H315, H319	P264, P280, P302+P352, P305+P351+P338, P332+P313, P337+P313
Silver nitrate		H272, H314, H400, H410	P210, P220, P260, P264, P273, P280, P301+P330+P331, P302+P361+P354, P304+P340, P305+P354+P338, P316, P321, P363, P370+P378, P391, P405, P501
Sodium carbonate		H319	P264+P265, P280, P305+P351+P338, P337+P317
Sodium chloride		H319	P264+P265, P280, P305+P351+P338, P337+P317

Sodium dodecyl sulfate (SDS)		H228, H302, H315, H318, H319, H332, H335, H412	P210, P240, P241, P261, P264, P264+P265, P270, P271, P273, P280, P301+P317, P302+P352, P304+P340, P305+P351+P338, P305+P354+P338, P317, P319, P321, P330, P332+P317, P337+P317, P362+P364, P370+P378, P403+P233, P405, P501
Sodium hydroxide (NaOH)		H314	P260, P264, P280, P301+P330+P331, P302+P361+P354, P304+P340, P305+P354+P338, P316, P321, P363, P405, P501
Trichloroacetic acid (TCA)		H314, H400, H410	P260, P264, P273, P280, P301+P330+P331, P302+P361+P354, P304+P340, P305+P354+P338, P316, P321, P363, P391, P405, P501
Tris base		H315, H319, H335	P261, P264, P264+P265, P271, P280, P302+P352, P304+P340, P305+P351+P338, P319, P321, P332+P317, P337+P317, P362+P364, P403+P233, P405, P501
Tris hydrochloride (Tris HCl)		H315, H319, H335	P261, P264, P264+P265, P271, P280, P302+P352, P304+P340, P305+P351+P338, P319, P321, P332+P317, P337+P317, P362+P364, P403+P233, P405, P501
Triton X-100		H302, H315, H318, H319, H400, H410, H411, H412	P264, P264+P265, P270, P273, P280, P301+P317, P302+P352, P305+P351+P338, P305+P354+P338, P317, P321, P330, P332+P317, P337+P317, P362+P364, P391, P501
Uranyl acetate		H300+H330, H300, H330, H373, H411	P260, P264, P270, P271, P273, P284, P301+P316, P304+P340, P316, P319, P320, P321, P330, P391, P403+P233, P405, P501

6.5. List of figures

6.6. List of tables

Acknowledgments

First, I would like to thank my supervisors, Prof. Dr. Michael Kolbe and Dr. Charlotte Uetrecht, who have enabled this work. It was a great pleasure to perform experiments in the lab and acquire skills in cutting-edge techniques. I have learned a lot during this time and am deeply grateful for this opportunity.

In addition, I want to thank the great scientists working with me on our collaborative projects. I could not have undertaken this journey without their valuable input. This includes Dr. Michele Lunelli, Dr. Zhuo Angel Chen, Dr. Karol Kaszuba, Dr. Jan Kosinski, Dr. Ivonne Bernal, Jonas Römermann, Dr. Francis J. O'Reilly and Prof. Dr. Juri Rappsilber.

Further, I would like to acknowledge the technical support by the PPCF at the CSSB in Hamburg. In particular, I want to thank Dr. Susanne Witt and Maike Eiben for providing me with a protein sample. Moreover, I would like to extend my thanks to Dr. Katharina Häußermann for the great support with the mass photometry analysis.

Special thanks to Dr. Juana de Diego for the helpful scientific discussions and excellent management of our lab. Additionally, I would like to thank Dr. Rudolph Reimer, a great teacher who sparked my interest in electron microscopy.

Additionally, I would like to thank Cornelia Cazey and Dr. Carolin Seuring for my training at the Multi-User Cryo-EM facility at the CSSB.

Eventually, I want to thank all former and recent lab members from both groups for making the time worthwhile.

Furthermore, I want to express my gratitude to the Joachim Herz Foundation in Hamburg for my funding in the course of the 'Promotion of young CSSB scientists' program.

Last but not least, I would like to mention my reviewers and defense committee: Prof. Dr. Michael Kolbe, Dr. Christian Löw, Prof. Dr. Kay Grünewald and Prof. Dr. Jens Bosse. Thank you for considering this work!

Declaration of authorship/ Eidestattliche Versicherung

I hereby declare upon oath that I have written the present dissertation independently and have not used further resources and aids than those stated in the dissertation.

I, the undersigned, declare that this bound copy of the dissertation and the dissertation submitted in electronic form and the printed bound copy of the dissertation submitted to the faculty for archiving are identical.

This dissertation was not previously submitted to another doctoral examination board.

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Ich versichere, dass dieses gebundene Exemplar der Dissertation und das in elektronischer Form eingereichte Dissertationsexemplar und das bei der Fakultät zur Archivierung eingereichte gedruckte gebundene Exemplar der Dissertationsschrift identisch sind.

Diese Dissertation wurde nicht in einem früheren Promotionsverfahren eingereicht.

Hamburg, den 20.10.2022   Lara Flacht