Approaches to replace the HlyA T1SS by homologs

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Olivia Spitz
aus Duisburg

Düsseldorf, Oktober 2021

aus dem Institut für Biochemie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Lutz Schmitt

2. Prof. Dr. Johannes Hegemann

Tag der mündlichen Prüfung: 12.10.2021

Table of Contents

Abstract

Type 1 secretion systems (T1SSs) are wide spread among Gram-negative bacteria due to their simple tripartite architecture. Two proteins located in the inner membrane, an ABC transporter and a membrane fusion protein (MFP), form one continuous channel across both membranes together with an outer membrane protein. Channel formation is induced by the substrate, which is transported in a one-step mechanism and C-terminus first to the extracellular space. Differences concerning the ABC transporter domains led to the identification of four subgroups of T1SSs of which group 2 holds the most identified members so far. The most investigated system of this group 2, the HlyA T1SS from *E. coli*, is the main focus of this work.

Based on *in silico* analysis and similarity to the HlyA T1SS components 25 group 2 T1SSs from different Gram-negative bacteria were identified and analyzed to different extents. Five ABC transporters from these homologous systems were successfully expressed in different *E. coli* strains and purification protocols for three of these were established. Secretion experiments with these homologous transporters and chimeric ABC transporters highlighted the importance of the nucleotide binding and transmembrane domains for complex formation and subsequent secretion. These experiments were supported by domain- and group-specific alignments of both inner membrane components, which revealed additional similarities within the groups of T1SSs and allowed the identification of two unique motifs in the MFPs of group 2. Furthermore, two already proposed amphipathic helices were confirmed: one in the cytosolic domain of the MFP HlyD and one in the secretion signal of the substrate HlyA. For the latter, two possible binding pockets were identified in the nucleotide binding domain of the dedicated ABC transporter HlyB.

Zusammenfassung

Typ I Sekretionssysteme (T1SS) sind weit verbreitet unter Gram-negativen Bakterien aufgrund ihrer einfachen dreiteiligen Architektur. Zwei Proteine, die in der inneren Membran lokalisiert sind, ein ABC Transporter und ein Membranfusionsprotein (MFP), formen einen durchgängigen Kanal über beide Membranen zusammen mit einem Protein der äußeren Membran. Die Bildung des Kanals wird durch das Substrat induziert, welches in einem Ein-Schritt Mechanismus und C-Terminus zuerst in den extrazellulären Raum transportiert wird. Unterschiede, die die ABC Transporterdomänen betreffen, führten zu der Identifikation von vier Untergruppen von T1SS von welchen Gruppe 2 bisher die meisten identifizierten Mitglieder hat. Das am besten untersuchte System aus der Gruppe 2, das HlyA T1SS von *E. coli,* ist der Hauptfokus dieser Arbeit.

Basierend auf *in silico* Analysen und Ähnlichkeit zu den HlyA T1SS Komponenten wurden 25 Gruppe 2 T1SS von verschiedenen Gram-negativen Bakterien identifiziert und in unterschiedlichem Ausmaß analysiert. Fünf ABC Transporter von diesen homologen Systemen wurden erfolgreich in verschiedenen *E. coli* Stämmen exprimiert und für drei von diesen wurden Reinigungsprotokolle etabliert. Sekretionsexperimente mit diesen homologen Transportern und chimären ABC Transportern haben die Wichtigkeit der Nukleotidbinde- und Transmembrandomäne für die Komplexbildung und darauffolgende Sekretion hervorgehoben. Diese Experimente wurden unterstützt durch domänen- und gruppenspezifische Sequenzalignments von beiden Komponenten in der inneren Membran, welche zusätzliche Ähnlichkeiten innerhalb der Gruppen von T1SS aufdeckten und die Identifikation von zwei einzigartigen Motiven in den MFPs von Gruppe 2 erlaubten. Außerdem wurden zwei bereits vorgeschlagene amphipathische Helices bestätigt: Eine in der zytosolischen Domäne des MFPs HlyD und eine im Sekretionssignal des Substrates HlyA. Für die Letztere wurden zwei mögliche Bindetaschen in der Nukleotidbindedomäne des dazugehörigen ABC Transporters HlyB identifiziert.

1. Introduction

1.1 Secretion systems in Gram-negative bacteria

Secretion is an essential process of life, which is involved in many cellular functions. For example, organisms can protect themselves or attack others by secreting toxins, they can communicate amongst each other by secreting messenger signals and increase nutrient availability by lysing surrounding cells or increase nutrient uptake by secreting molecules that help nutrients to cross the membrane (Costa et al., 2015). Many secretion systems are involved in the pathogenesis of the respective organism and some also contribute to the rising resistance against antibiotics, which makes them a promising drug target (Nikaido, 1998, Baron, 2010, Boudaher and Shaffer, 2019). Apart from the pharmaceutical applications, there is also a rising interest in secretion systems as biotechnological platforms, since purification of a protein or peptide is simplified by first secreting it into the surrounding medium. Usually a portion of the original substrate, that is sufficient to promote secretion, is fused to the protein of interest to allow translocation of the fusion protein (Hess et al., 1990, Ryu et al., 2015).

Secretion systems are ubiquitously present in all domains of life and thus, several different strategies and systems have evolved. The focus of this thesis is on Gram-negative bacteria so systems can be grouped based on the number of membranes that they cross. The following sections will give a short overview over single membrane crossing systems (1.1.1) as well as double or triple membrane crossing systems (1.1.2).

1.1.1 Single membrane crossing systems

In the context of Gram-negative bacteria, single membrane crossing systems are further subdivided into systems that cross the inner membrane (IM) or outer membrane (OM) with the OM crossing systems being dependent on IM crossing systems to deliver their substrate to the periplasm.

One of the best researched IM crossing systems is the Sec translocon. It is present in the endoplasmic reticulum of eukaryotes, in the chloroplasts of plants and the cytoplasmic membrane of prokaryotes and archaea (Kinch et al., 2002). In *E. coli* the system is often referred to as the SecYEG translocon since these three proteins, SecY, SecE and SecG, form the core complex of the translocon in the IM: SecY forms the actual channel through the membrane with SecE wrapping around it and protecting SecY from degradation (Meyer et al., 1999, van den Berg et al., 2004). In the absence of SecE, SecY is rapidly degraded by the membrane protease FtsH (Kihara et al., 1995). SecG is actually non essential to many transport processes but increases their efficiency (Nishiyama et al., 1994).

The system transports a broad variety of substrates and is therefore highly modular and dynamic, which is reflected by the many accessory proteins, mainly chaperons, targeting factors and signal peptidases, that can interact with the translocon (Denks et al., 2014). In Gram-negative bacteria there are two main pathways by which the translocon functions: i) The post-translational pathway and ii) the co-translational pathway (Figure 1.1). The latter is mostly used by proteins of the IM and dependent on the signal recognition particle (SRP) (Lycklama a Nijeholt and Driessen, 2012). SRP rapidly scans ribosomes and binds to their tunnel exit only in the presence of a hydrophobic transmembrane segment of a nascent protein chain (Holtkamp et al., 2012). SRP then targets the whole ribosome nascent chain complex to the membrane receptor FtsY, which transfers the complex to the SecYEG translocon (Figure 1.1) (Luirink et al., 1994, Denks et al., 2014). Upon GTP hydrolysis, FtsY-SRP dissociates from the complex and translation at the ribosome provides energy for the translocation (Luirink and Sinning, 2004, Lycklama a Nijeholt and Driessen, 2012). The protein chain is threaded into SecY and the hydrophobic transmembrane segments leave the translocon via a lateral gate, where their folding may be assisted by the insertase YidC (Kater et al., 2019, Tsukazaki, 2019).

The post-translational pathway is commonly used by periplasmic or OM proteins and dependent on the motor protein SecA (Denks et al., 2014). Most of the substrates are synthesized as pre-proteins with a hydrophobic signal sequence at

their N-terminus, which slows folding of the pre-protein in the cytoplasm, therefore providing a time window for interaction with chaperons such as SecB and/or trigger factor (von Heijne, 1990, Randall et al., 1998). The motor protein SecA is essential for this pathway and has multiple interaction partners: It can interact with chaperons such as SecB (Randall et al., 2004), the SecYEG translocon (Mori and Ito, 2006), the signal peptide of a substrate (Akita et al., 1990) and the ribosome, thus also being involved in co-translational transport (Huber et al., 2011). Although the exact mechanism of SecA:SecYEG mediated translocation remains intensively debated, SecA is able to thread the pre-protein into the SecYEG translocon and energize the transport by hydrolyzing ATP (Figure 1.1) (Chatzi et al., 2014). The process happens stepwise with 20-40 amino acids being translocated in each step (Schiebel et al., 1991, Tomkiewicz et al., 2006). The membrane proteins SecDF act in a proton motive force (pmf)-dependent manner to support the translocation by pulling on the (pre-)protein chain from the periplasmic side (Tsukazaki et al., 2011). The hydrophobic signal sequence acquires transmembrane topology and is then cleaved off and recycled by different signal peptidases, so the mature translocated protein is released into the periplasm (von Heijne, 1990, Paetzel et al., 2002).
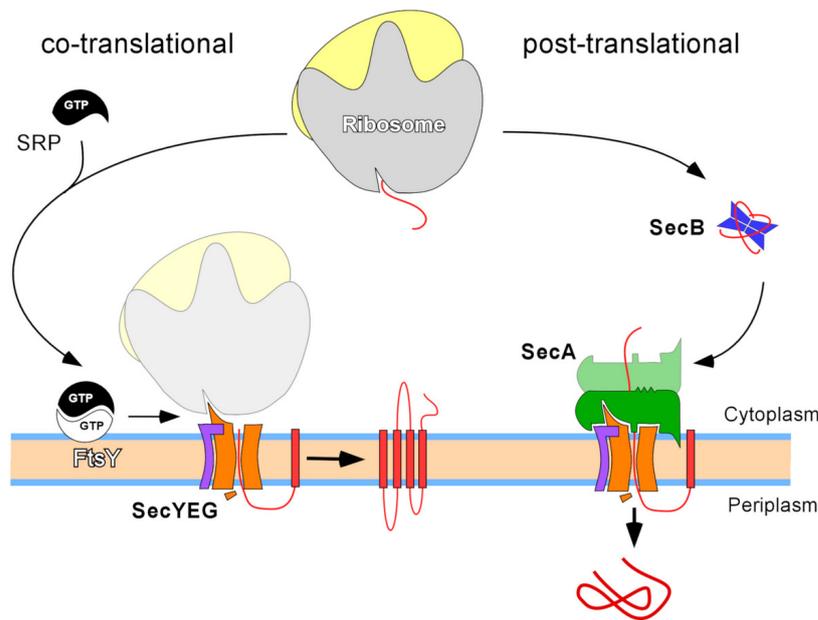
Figure 1.1 The co- and post-translational pathway of the Sec translocon (adapted from (Koch et al., 2021)). In the co-translational pathway SRP (black) binds to hydrophobic segments of a nascent protein chain (red line) and targets the ribosome complex (grey and yellow) to SecYEG (orange and purple) via FtsY (white). A chaperon such as SecB (blue) targets proteins that utilize the post-translational pathway to the translocon. The motor protein SecA (green) threads the pre-protein into the translocon. The signal peptide (red box) acquires membrane topology, is cleaved and later recycled by specific peptidases in the periplasm.

Independent of post- or co-translational translocation, all proteins that are transported by the SecYEG translocon are transported unfolded. Proteins, that need to fold prior to translocation in the cytoplasm, utilize the twin-arginine translocation (TAT) pathway. This pathway is less common and can be found in some bacteria and archaea as well as in thylakoid membranes of plants and algae, but is essential only in a few organisms (Palmer and Berks, 2012). Interestingly, in pathogens that use the TAT pathway it is almost always involved in the virulence of these pathogens (De Buck et al., 2008).

The name of the TAT pathway refers to a conserved motif (SRRxFLK) in the signal peptide of proteins transported by this pathway (Berks, 1996). The signal peptide is recognized by and binds to a complex of TatB and TatC in the IM, which leads to the pmf-dependent recruitment of TatA, which is proposed to form the actual translocation channel (Figure 1.2) (Cline and Mori, 2001, Mori and Cline, 2002, Gohlke et al., 2005, Dabney-Smith et al., 2006). So far no high-resolution structures of an assembled TAT system are available. However, according to the

'bespoke channel model', the size of the pore formed by polymerization of TatA protomers likely accommodates for the substrate size, which is between 20 Å and 70 Å in diameter for TAT-substrates in *E. coli* (Berks et al., 2000, Gohlke et al., 2005). Due to the range of size among the folded substrates, the size of the pore has to be adjustable to prevent ion leakage from the periplasm. After the transport of one substrate, which takes between one and a few minutes, the complex disassembles again (Mori and Cline, 2002, Whitaker et al., 2012). The signal peptide is usually cleaved by a peptidase in the periplasm (Lüke et al., 2009, Palmer and Berks, 2012) although there are some substrates, like the Rieske iron-sulfur proteins from *Paracoccus denitrificans* and *Legionella pneumophilia* that are anchored to the IM most likely by their signal peptide (Bachmann et al., 2006, De Buck et al., 2007).
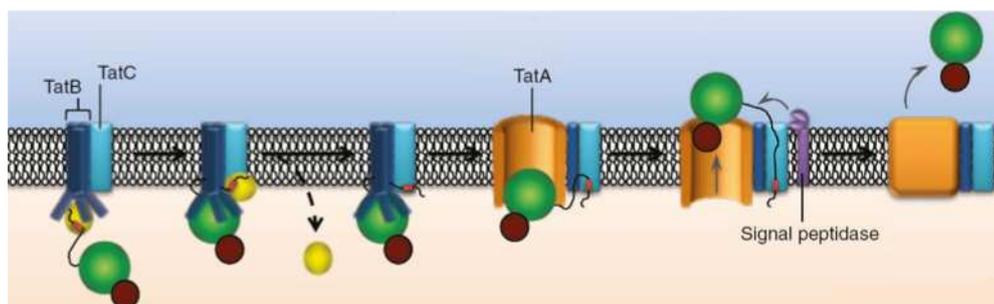
Figure 1.2 Schematic representation of TAT mediated translocation adapted from (Cherak and Turner, 2017). A chaperon (yellow circle) binds the signal peptide (red box) of a folded substrate or substrate complex (green and red circle). The chaperon-signal peptide complex is recognized by TatB (dark blue), which forms a complex with TatC (light blue). The chaperon gets removed and TatA (orange) is recruited into the complex. After translocation a signal peptidase (purple) cleaves of the signal peptide, thereby releasing the substrate (complex) into the periplasm (blue background).

While some substrates reach their final destination after one translocation step by either the Sec translocon or TAT pathway, there are other substrates that require an additional translocation step. For example all outer membrane proteins (OMPs) need to be assembled into the outer membrane after translocation via the Sec translocon. In *E. coli* and other Gram-negative bacteria this insertion is mostly performed by the β-barrel assembly machinery (BAM). In 2016 the structure of the BAM complex was solved in the resting and post-insertion state (Gu et al., 2016). Recently, in 2020 a structure during a late stage of substrate assembly followed (Figure 1.3 A) (Tomasek et al., 2020).

The BAM complex consists out of the OMP BamA and four lipoproteins, BamBCDE of which only BamA and BamD are essential (Kim et al., 2012). BamA itself is a β-barrel-containing protein with a unique feature: The β-barrel has a lateral gate that can open to the OM thereby forming the substrate exit channel (Tomasek et al., 2020). The periplasmic domains of BamA, termed POTRA (polypeptide transport-associated) domains, form a ring together with the periplasmic domains of the lipoproteins, which is thought to guide the substrate to the correct position (Gu et al., 2016). In the classical two-step model an OMP is transported by the Sec translocon across the IM, transferred to a periplasmic chaperon, like SurA or Skp, and then transferred to the BAM complex for OM insertion (Kim et al., 2012). However, there are also indications that the BAM complex can form a double membrane crossing supercomplex with the N-terminal domain of the motor protein SecA (SecA[N]), several periplasmic chaperons and SecYE allowing one-step translocation of OMPs from the cytoplasm to the OM (Figure 1.3 B) (Wang et al., 2016, Jin, 2020). Additionally, the BAM complex is involved in the chaperon-usher pathway and in the assembly of the type 5 secretion system (T5SS), which has been termed the auto-transporter system (Leo et al., 2012).
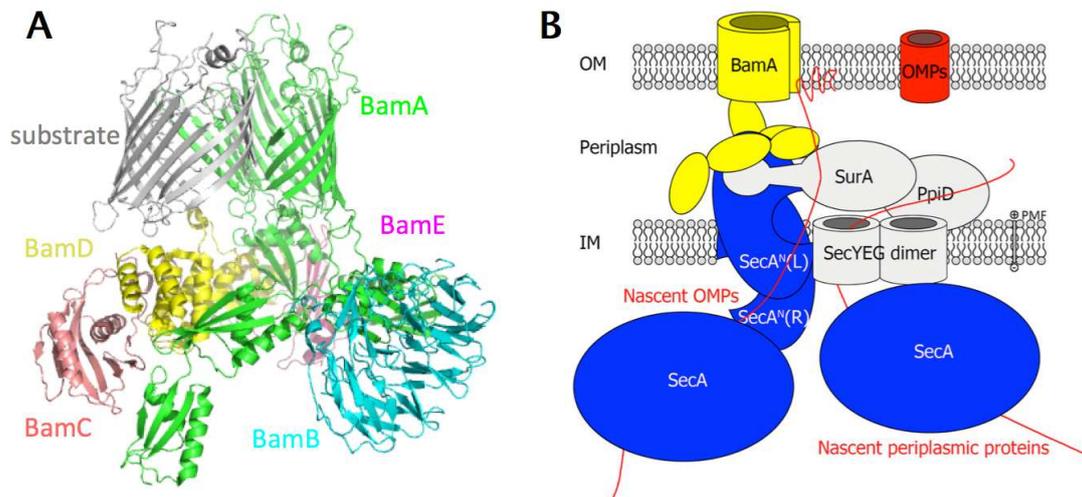
Figure 1.3 The BAM complex. A) Structure of the BAM complex during substrate assembly. BamA (green), BamB (cyan), BamC (salmon), BamD (yellow), BamE (light pink), substrate (grey) (PDB 6V05, drawn with PyMOL). **B)** Schematic representation of a BAM supercomplex adapted from (Jin, 2020). SecA targets periplasmic proteins to the Sec translocon, which translocates them to the periplasm. However, SecA and SecA^N (N-terminal part of SecA) can form oligomers in the IM and translocate nascent OMPs. SecA^N can contact the BAM complex of which only BamA is shown.

The T5SS has been termed the auto-transporter system, because in its simplest form it consists only out of one single self-secreting, self-processing protein (Jose et al., 1995). There are, however, some differences concerning the exact mechanism and directionality of translocation (C- or N-terminal), the oligomeric state of the auto-transporter and the dependence on the BAM complex in the secretion process, which lead to five subgroups of T5SS termed a-e (Figure 1.4) (Leo et al., 2012). All subgroups share common features such as the presence of a transporter and a passenger domain, often but not always present on a single polypeptide chain, and the dependence on the Sec translocon to reach the periplasm. Once they reached the periplasm, the auto-transporters are kept unfolded or partially folded with the help of periplasmic chaperons and in some cases due to their low intrinsic folding properties (Hartmann et al., 2009, Roman-Hernandez et al., 2014). The transporter domain is inserted into the OM with the help of the BAM complex and forms a β-barrel (Ieva et al., 2011). The passenger domain is transported most likely directly through this β-barrel and folds outside of the cell, which provides energy for the transport process (Peterson et al., 2010, van den Berg, 2010). Adhesin forming passenger domains are retained at the cell surface by their transporter domain, while most other passenger

domains are cleaved by their own peptidase domain and are released into the extracellular space (Tsai et al., 2010, Costa et al., 2015). The passenger domains are diverse in size and function including lipases, proteases and many other virulence factors (Leo et al., 2012).

Although there is experimental evidence that folding of the passenger domain at the outside of the cell does provide energy for the transport process (Junker et al., 2009, Peterson et al., 2010, Roman-Hernandez et al., 2014), there are also systems where folding is not necessary for secretion (Oliver et al., 2003, Skillman et al., 2005). At least in these systems the energy might be provided by an unidentified IM protein, which could use ATP hydrolysis or the pmf (Peterson et al., 2010). This interaction might not happen directly but could be facilitated by the BAM complex (Costa et al., 2015). It is well accepted that the BAM complex interacts with the transporter domain to facilitate its OM insertion (Leo et al., 2012). However, BamA has also been cross-linked to or co-purified with several passenger domains suggesting an active role in the secretion process (Ieva and Bernstein, 2009, Sauri et al., 2009, Peterson et al., 2010). Furthermore some transporter domains display homology to BamA by sharing its periplasmic POTRA domains (subgroup b and d), while others do not (subgroups a, c, e, Figure 1.4) (Leo et al., 2012). This also points to a role of BamA in the secretion process, at least in subgroups a, c and e. Taking into consideration, that the BAM complex might form a double membrane crossing complex with parts of the Sec translocon and especially SecA (Wang et al., 2016), the energy from ATP hydrolysis at the cytoplasmic site of the IM could be transferred through this supercomplex (Jin, 2020) and aid in the secretion of the passenger domain. This, together with the dependence on and interaction with periplasmic chaperons, challenges the simple view of one single polypeptide chain that is sufficient to cross the OM barrier (Costa et al., 2015).
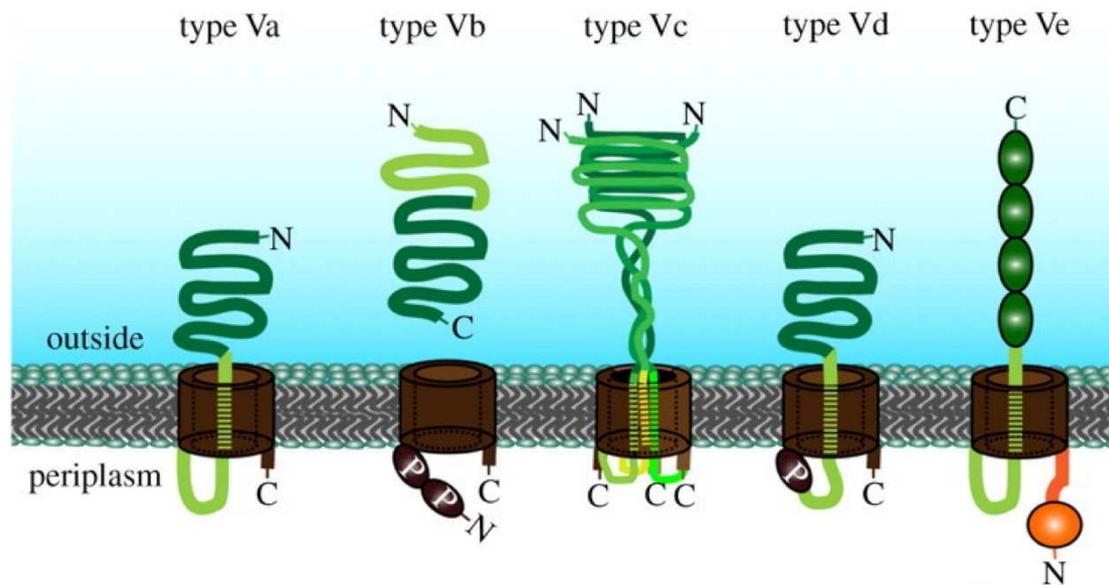
Figure 1.4 Schematic representation of T5SS subgroups adapted from (Leo et al., 2012). Transporter domains are shown in brown, passenger domains in dark green and linkers in light green. The POTRA domains of subgroup b and d are shown in black and labeled with P. The periplasmic domain of subgroup e is shown in orange. Termini of the proteins are indicated with N and C respectively.

The third and last OM crossing system that will be introduced here is the chaperon-usher pathway. It is used for the secretion and assembly of pili or fimbriae and therefore involved in host cell recognition and biofilm formation (Busch and Waksman, 2012). The usher protein is again a β-barrel protein of the OM that is transported to the periplasm by the Sec translocon and inserted into the OM by the BAM complex (Palomino et al., 2011). It displays N- as well as C-terminal periplasmic domains, which harbor specific chaperon binding sites, as well as a plug domain that can close the channel from the periplasmic site in the apo-state and is displaced to the periplasm during pilus assembly (Figure 1.5) (Phan et al., 2011). The secreted pilus is made up of a defined tip structure and a stalk with the tip being assembled first and the stalk growing underneath it by controlled polymerization of the monomeric stalk units, which pushes the emerging pilus through the translocation pore (Busch and Waksman, 2012). A specific chaperon, commonly encoded in the same gene cluster as the usher protein and the pilus subunits (Nuccio and Bäumler, 2007), targets these subunits from the Sec translocon to the usher protein (Ng et al., 2006). Different affinities between the single subunits, the chaperon and the usher

protein likely control the ordered assembly of the pilus subunits (Dodson et al., 1993).

All pilus subunits display an immunoglobin-like domain with six instead of the typical seven β-strands, leaving a hydrophobic groove on the subunits where the seventh β-strand would be. The respective chaperon of the system can complement this groove *in trans* with one of its own β-strands; a mechanism termed 'donor strand complementation' (DSC), leading to a non-canonical topology of the respective β-strands (Choudhury et al., 1999, Sauer et al., 1999). At the usher protein, the donor strand of the chaperon is exchanged with the N-terminal extension of another subunit in a 'zip-in-zip-out' manner (Rose et al., 2008). This 'donor strand exchange' (DSE) leads to a canonical topology of the respective β-strands, which is energetically favored and thought to drive the reaction in the direction of pilus assembly (Figure 1.5) (Zavialov et al., 2005). In the case of P pili, a terminator subunit has been identified, whose hydrophobic groove lacks the DSE-initiation site, therefore terminating the pilus assembly and controlling the length of the pilus (Verger et al., 2006). Such a subunit remains to be identified for type I pili (Busch and Waksman, 2012, Costa et al., 2015).
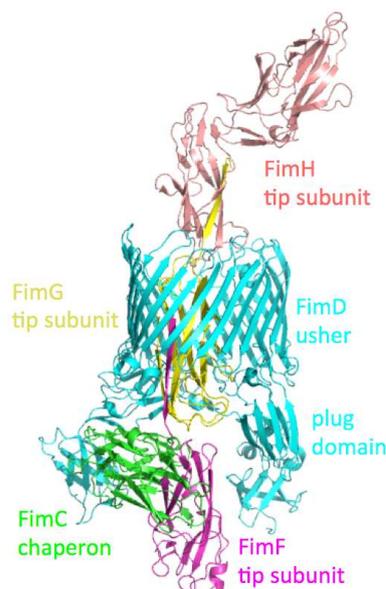
Figure 1.5 Structure of the usher protein FimD (cyan) from *E. coli* during pilus tip assembly (PDB 4J3O, drawn with PyMOL**).** FimH (salmon), FimG (yellow) and FimF (pink) form the pilus tip. The pilus stalk would be provided by FimA (not part of the structure). FimF is brought to the complex by the chaperon FimC (green), which can interact with the periplasmic domains of the usher FimD. FimH:FimG and FimG:FimF are engaged in DSC: One β-strand is incorporated into the interacting subunit thereby complementing the immunoglobin-like fold.

In summary, proteins that need to cross the IM either use the Sec translocon (unfolded substrates) or the TAT pathway (folded substrates) and display the respective N-terminal signal sequence. This signal sequence either anchors the substrate to the IM or is cleaved, thereby releasing the substrate to the periplasm (Lycklama a Nijeholt and Driessen, 2012, Palmer and Berks, 2012). If further targeting to or across the OM is required, substrates usually interact with periplasmic chaperons, which guide them to the respective OM crossing machinery (Costa et al., 2015). These machineries often display a β-barrel in the OM and are therefore dependent on the BAM complex for assembly (Kim et al., 2012). Transport across the IM can be energized by the pmf or NTP hydrolysis (Costa et al., 2015). Transport across the OM can be energized by the folding of the substrate or by coupling the OM crossing system to an IM crossing system (Costa et al., 2015, Jin, 2020).

1.1.2 Double membrane crossing systems

At least five double membrane crossing systems have been identified so far, termed type 1 secretion system (T1SS) to T4SS and T6SS. The T1SS will be introduced in detail in section 1.2.2.

The T2SS is present in pathogenic and non-pathogenic Gram-negative bacteria and can transport various folded proteins from the periplasm to the extracellular space (Costa et al., 2015). The substrates of T2SSs are diverse in size and function including toxins, adhesins or cytochromes but are mainly hydrolyzing enzymes that can degrade biopolymers thereby increasing nutrient availability (Nivaskumar and Francetic, 2014). The substrates are transported to the periplasm either by the Sec translocon or the TAT pathway (see section 1.1.1 for details on these systems).

The transport channel of T2SSs is a large multimeric complex consisting of 12-15 components, most of which form oligomers themselves (Nivaskumar and Francetic, 2014). The system can be divided into four parts: An OM complex, an IM complex, a periplasmic pseudopilus and a cytoplasmic ATPase (Figure 1.6 A). The pseudopilus shows homology to the type IV pilus but does not reach the cell surface, hence the prefix 'pseudo' (Pugsley, 1993, Vignon et al., 2003). The OM and IM

complex form the channel around the pseudopilus and the cytoplasmic ATPase is thought to provide the energy for its (dis-)assembly (Vignon et al., 2003).

Substrates enter the T2SS from the periplasm and are transported folded (Hirst and Holmgren, 1987, Pugsley, 1992). So far it is unclear how substrates enter the transport channel and how they are recognized. Neither a signal sequence nor a conserved structural motif has been identified among the T2 substrates so far (Pineau et al., 2014) and no lateral gate has been found in structures of OM complexes (Hay et al., 2017, Yan et al., 2017). However, in 2019 Chernyatina and Low were able to isolate a ~2.4 MDa complex from *Klebsiella pneumonia*, corresponding to seven of its fourteen T2SS proteins (Chernyatina and Low, 2019). Their cryogenic electron microscopy (cryo-EM) analysis allowed reconstitution of the full OM complex (Figure 1.6 B) and modeling of the cytoplasmic components. The authors argue that a weak interaction between OM and IM complex is a requirement for substrate uptake into the channel. There the substrate is pushed through the channel in a piston-like manner by extension of the pseudopilus (Vignon et al., 2003, Douzi et al., 2011).
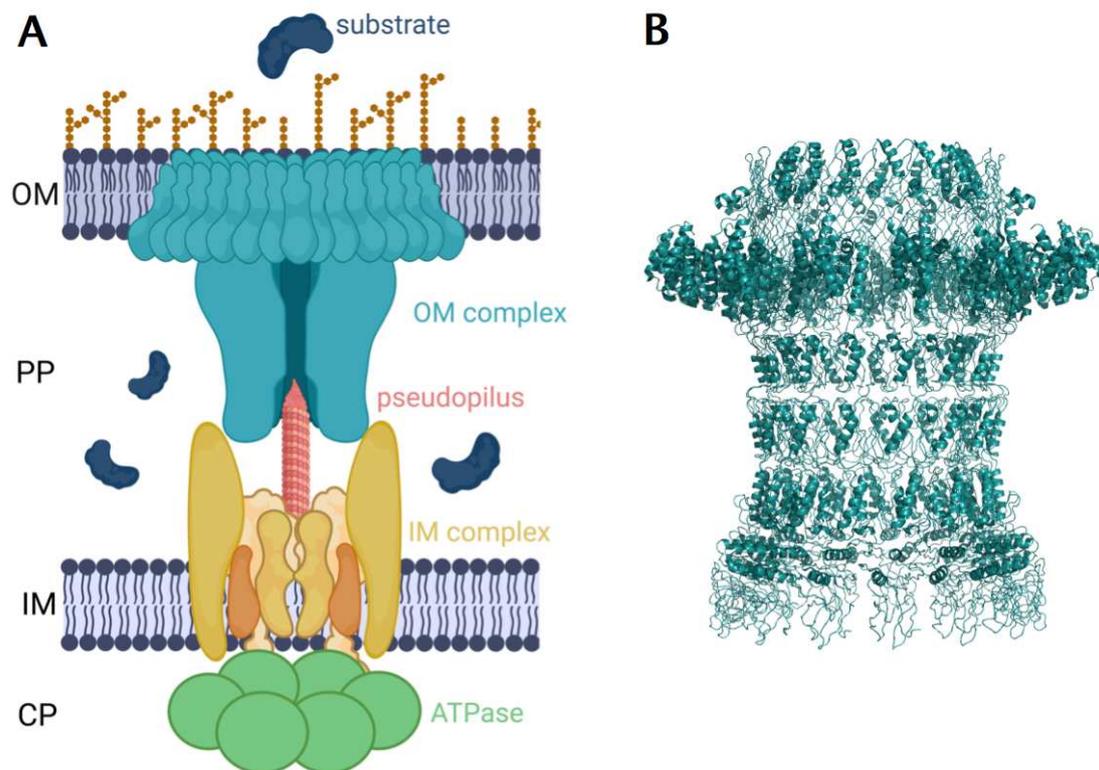
Figure 1.6 A) Schematic representation of a T2SS (created with BioRender.com). The T2SS transports folded substrates (dark blue) from the periplasm (PP) to the extracellular space. A complex in the OM (petrol) and the IM (shades of yellow and orange) surround a pseudopilus (red), which pushes substrates through the OM complex. The process is energized by a cytoplasmic ATPase (green). CP = cytoplasm. **B)** Structure of the OM complex from *Klebsiella pneumonia* showcasing size and complexity of this complex (PDB 6HCG, drawn with PyMOL).

T1 and T2SSs are double membrane crossing systems that deliver their substrate to the extracellular space or, in the case of adhesins, to the cell surface. In contrast to that, T3, T4 and T6SSs are able to cross a third membrane and deliver their substrates directly into the membrane or the cytosol of a host cell (Costa et al., 2015).

The T3SS is found mostly in pathogenic Gram-negative bacteria and has been termed 'the needle complex' due to its unique architecture (Figure 1.7). Around 20 different proteins with copy numbers ranging from 1 to over 100 form a large complex of ~6 MDa that can be separated into a double membrane spanning base and a 20-150 nm long needle filament (Wagner et al., 2018). The assembly of the base structure is itself a complex process involving local peptidoglycan reconstruction and different models are discussed in the literature (Deng et al., 2017). Substrates of T3SSs can be separated into early, intermediate and late substrates, which all have

different functions. The early substrates are involved in building the needle filament and controlling its length (Kubori et al., 2000). In contrast to the pseudopilus of T2SSs, the needle filament grows at its distal end and not at the cell-proximal end. Needle filament subunits have to pass the whole needle before being incorporated in the tip of it (Crepin et al., 2005). The needle tip consists of specialized tip proteins that are able to sense contact to a host cell (Veenendaal et al., 2007). The exact mechanisms of sensing and subsequent signal transduction to the cytoplasm of the bacterium are unknown so far and might differ for different species (Wagner et al., 2018). However, host cell contact triggers the secretion of intermediate substrates, so-called translocator proteins, which are able to form a pore in the host cell membrane allowing entry of the late substrates, the effector proteins (Figure 1.7 A) (Lara-Tejero and Galán, 2009, Montagner et al., 2011). These effectors trigger host cell responses, which support the life cycle of the respective pathogen and are therefore diverse in function (Raymond et al., 2013).

Substrates of T3SSs contain a N-terminal signal sequence of roughly 20 amino acids followed by a chaperon-binding domain (Stebbins and Galán, 2001, Samudrala et al., 2009). Chaperon and substrate are often distinct pairs and are adjacent in the T3 operon (Wattiau et al., 1996). Apart from guiding the substrates to the cytoplasmic sorting platform of the base structure, the chaperons also prevent toxic functions like insertion of translocator proteins in the bacterial IM (Krampen et al., 2018). The sorting platform itself is a modular complex that presents different chaperon-binding sites during the secretion process thereby contributing to secretion regulation (Portaliou et al., 2017). At the sorting platform the chaperon is removed, the substrate is unfolded and transported unfolded through the needle-complex in a pmf-dependent manner (Akeda and Galán, 2005, Lee et al., 2014, Radics et al., 2014). The sorting platform also contains a highly conserved ATPase whose involvement in the de-chaperoning and secretion process is heavily discussed with some publications claiming that secretion is dependent on ATP hydrolysis (Lorenz and Büttner, 2009), others claiming ATP hydrolysis is non essential (Paul et al., 2008) or that it is dispensable under certain conditions (Erhardt et al., 2014). Although

many questions remain unanswered, progress towards understanding the assembly and secretion mechanism of T3SSs is supported by crystal structures of single components or smaller subassemblies and high-resolution cryo-EM structures of the entire system (Figure 1.7 B) (Radics et al., 2014).
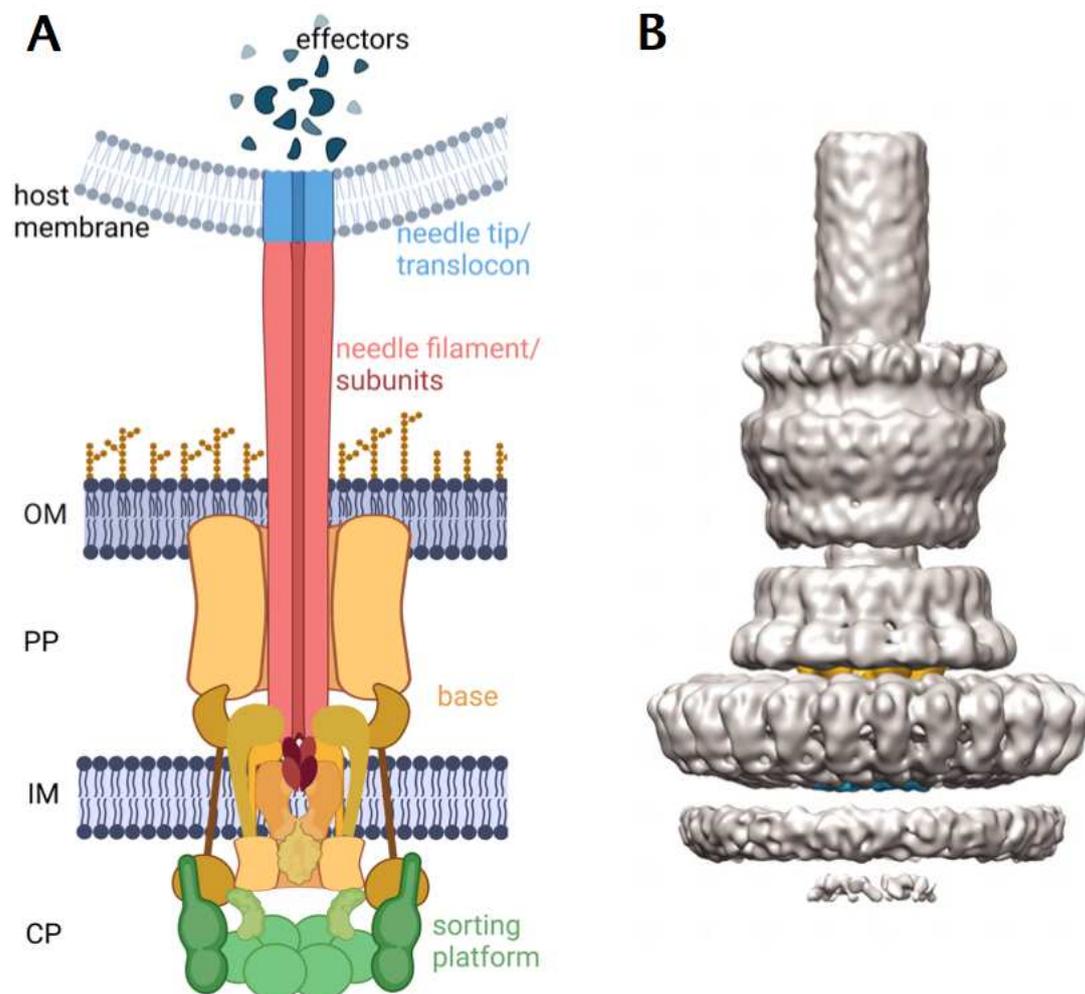
Figure 1.7 A) Schematic representation of a T3SS (created with BioRender.com). The cytoplasmic sorting platform is shown in green, the double membrane spanning base in yellow and orange, the needle filament in red and the needle tip, which is formed by intermediate substrates of T3SS, is shown in blue forming a translocation channel through a host membrane. This enables the secretion of a mixture of different effector proteins (blue random shapes). **B)** Surface representation of *Salmonella typhimurium's* needle complex adapted from (Schraidt and Marlovits, 2011).

The same holds true for the T4SS. In 2014 Low *et al.* published the first structure of a nearly complete T4SS in solution (Low et al., 2014). This was followed in 2018 and 2019 by electron tomography images of three more T4SSs of different subgroups (Figure 1.8 shows an F-plasmid encoded T4SS from *E. coli*) (Chetrit et al.,

2018, Hu et al., 2019a, Hu et al., 2019b). These four structures are very versatile in dimensions (diameters ranging from ~185 Å to ~400 Å in the OM) and underline the diversity found among T4SSs (Costa et al., 2020). These systems are unique in the sense that they are able to translocate protein-DNA complexes to the cytosol of a host in a process called conjugation, which contributed to their wide spread among Gram-positive and -negative bacteria as well as archaea (Souza et al., 2012). They are involved in pathogenesis, symbiosis and their ability to exchange whole plasmids contributes to the rising resistance against antibiotics among pathogens (Hubber et al., 2004, Huddleston, 2014).

Based on the secreted substrate(s) and substrate location, three subgroups of T4SSs have been defined. There are systems that secrete i) protein-DNA complexes into the cytosol of a host, ii) effector proteins into the cytosol of a host and iii) systems that secrete into or take up proteins and protein-DNA complexes from the surrounding environment (Zechner et al., 2012). Some bacteria express multiple T4SSs of different subgroups (Dehio, 2008). In addition to that, the structural organization of T4SSs is also very diverse but can be roughly divided in two groups: The 'minimized' and 'expanded' systems (Costa et al., 2020). The 'minimized systems' consist out of twelve different proteins while the 'expanded systems' utilize numerous accessory proteins in addition to these to build the translocation complex. However, all systems share common building blocks: A periplasmic stalk connects a complex in the OM (OMC), and a complex in the IM (IMC), while a pilus reaches out into the extracellular environment from the OMC (Figure 1.8) (Costa et al., 2020). In most cases the IMC contains a conserved membrane anchored ATPase (also called T4 coupling protein), which is involved in substrate recognition and sorting as well as energizing the assembly and transport via ATP hydrolysis together with two additional cytoplasmic ATPases of the IMC (Zechner et al., 2012, Llosa and Alkorta, 2017). Not much is known about the assembly of T4SSs, however, it is widely accepted that the double membrane spanning base has to assemble first in order to assemble the pilus. Pilus elongation is promoted by addition of post-translationally

modified pilin subunits at the cell-proximal end (Eisenbrandt et al., 1999, Clarke et al., 2008).

Substrates are believed to enter the translocation channel from the cytoplasm after being targeted to the machinery by specific chaperons (Pattis et al., 2007). This targeting involves at least the T4 coupling protein, the chaperon-substrate complex and translocation signals (TSs) on the substrate(s) (Zechner et al., 2012). Different TSs have been identified for different systems and their nature is extremely versatile: TSs located at the C-terminus can display clusters of positively charged residues (Vergunst et al., 2005), hydrophobic residues (Nagai et al., 2005) or short polar and negatively charged residues (Huang et al., 2011). TSs can also be distributed to both termini of a substrate (XVIPCD motif) (Alegria et al., 2005) or appear as a structural fold when the substrate is bound to its chaperon (Wagner et al., 2019). Identifying T4SS substrates is therefore a complex task, which is further complicated by the sheer amount of possible substrates. For example, the 'Dot/Icm' T4SS from *L. pneumophilia* secretes over 300 effector proteins during infection (Isaac and Isberg, 2014).

Substrate recognition for conjugative plasmids involves a multiprotein complex termed relaxosome, which assembles on the origin of transfer (oriT) of a respective plasmid (Rehman et al., 2019). These plasmids are usually self-transmissible, meaning that they encode all proteins necessary for transfer (F-plasmid) (Zechner et al., 2012). The relaxosome complex prepares and targets the DNA for transfer: A helicase unwinds the DNA allowing a relaxase to cut the strand destined for transfer (T-strand) at a defined position (*nic*) (Nelson et al., 1995). Through this nicking-reaction the relaxase is covalently bound to the 5' end of the T-strand forming a nucleoprotein that can be recognized by the T4 coupling protein (Byrd and Matson, 1997, Llosa et al., 2003). Transfer of this nucleoprotein is mostly dependent on host cell contact (Zechner et al., 2012, Costa et al., 2020); however, the recognition and signal transduction mechanisms are unknown so far as well as the actual pathway of the nucleoprotein through the translocation channel. Once the nucleoprotein reaches the cytosol of the host, the relaxase component initiates recircularization and

synthesis of the complementary strand (Draper et al., 2005). Complementary strand synthesis also happens in the donor, yielding two complete copies of the respective plasmid (Kingsman and Willetts, 1978).
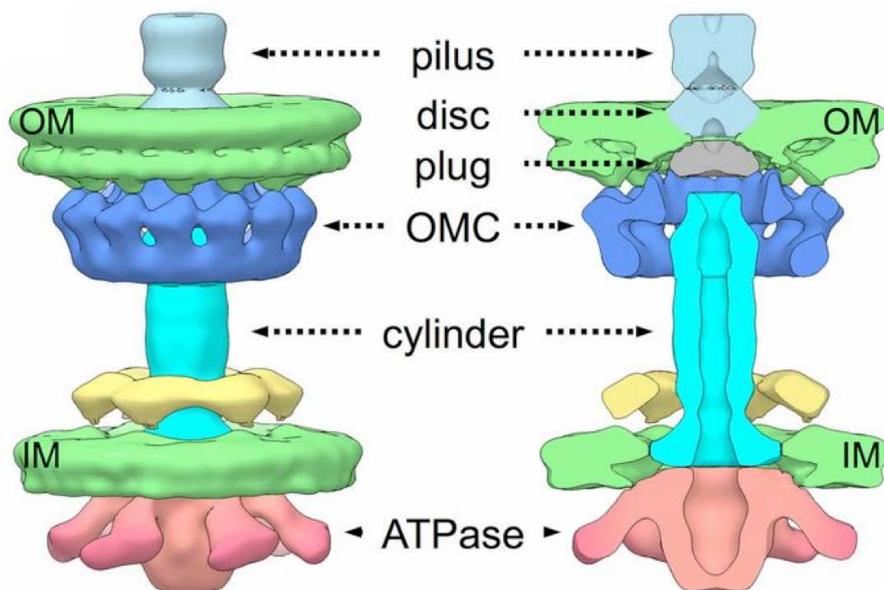
Figure 1.8 Three-dimensional surface rendering of cryoelectron tomography images of an F-plasmid encoded T4SS from *E. coli* adapted from (Hu et al., 2019a). A periplasmic stalk (cyan cylinder) connects a complex in the OM (dark blue) with a complex in the IM (yellow and red). The cutaway view on the right shows a channel that is closed to the cytoplasm by the ATPase and a plug (grey) that closes the channel towards the exterior. The pilus (light blue) emerges from the OM complex.

The last double membrane crossing system that will be introduced here is the T6SS (Figure 1.9). Although this system is widespread in Gram-negative and especially proteobacteria (Boyer et al., 2009), it was first defined as late as 2006 (Pukatzki et al., 2006). The system can deliver effector proteins to pro- and eukaryotic hosts and is sometimes compared to a crossbow as it assembles an arrow-like tube in the cytoplasm, whose tip is loaded with effector proteins, and 'shot' outwards into a neighboring cell (Sana et al., 2017). T6SS expressing cells commonly also express immunity proteins against these effectors (Ho et al., 2014). The resemblance to contractile tail machines of bacteriophages is apparent on both, functional and structural level (Ho et al., 2014).

There are at least 13 conserved proteins that account for the three building blocks of the system: The transmembrane complex (TMC), the cytoplasmic baseplate

and the cytoplasmic tail (Figure 1.9) (Boyer et al., 2009, Nguyen et al., 2018). The TMC consists out of four of these proteins. A complex of ~1.7 MDa corresponding to three of these TMC components was subjected to negative stain electron microscopy (ns-EM) and the structure was determined at ~12 Å (Durand et al., 2015). Surprisingly no pore was found in the OM and the internal channel only showed a diameter of 15 Å, which is too small for the arrow-like tube to pass through (Nguyen et al., 2018). The TMC, however, recruits components of the cytoplasmic baseplate to the IM and also the trimeric 'arrow-tip'-protein VgrG (Leiman et al., 2009, Brunet et al., 2015). The baseplate is the least understood building block of T6SSs but is thought to resemble those of myophages due to the similarity of the baseplate components (Nguyen et al., 2018). It connects the TMC and the tail structure (Zoued et al., 2013) and is believed to be involved in triggering the contraction event of the tail upon a so-far unknown signal (Stietz et al., 2018). Finally, the tail consists out of an outer sheath and an inner tube (Figure 1.9) (Wang et al., 2017a). In its expanded conformation the tail frequently spans the entire width and sometimes length of a cell (Basler et al., 2012). The contraction of the sheath reduces the tail length by roughly 50 % and pushes the tube up to 500 nm outwards of the cell (Basler et al., 2012, Ho et al., 2014). The tube is made up of hexameric Hcp subunits and the before-mentioned trimeric tip protein VgrG to which effector proteins can bind (Mougous et al., 2006, Flaugnatti et al., 2016). Furthermore, proteins of the PAAR-superfamily can also bind to VgrG, sharpening the tip and sometimes carrying additional effector domains (Shneider et al., 2013). The tube has an inner diameter between ~20 Å (*V. cholerae* (Wang et al., 2017a)) and ~40 Å (*P. aeruginosa* (Mougous et al., 2006)), which only allows unfolded effector proteins to bind to the inside of the tube (Silverman et al., 2013). In support of this, a chaperon function has been assigned to the main tube component Hcp (Silverman et al., 2013). Considering all the possible binding sites a cocktail of multiple effectors can be loaded to one 'arrow' and be delivered in one single translocation event to a host cell (Nguyen et al., 2018). The possible effectors are very diverse with some having cytoplasmic targets (e.g. nucleases) and some having periplasmic targets (e.g. Tme effectors), supporting

the respective bacteria in their specific ecological niche (Fridman et al., 2020). So far no active transport through the tube has been observed and it does not appear to be stable long enough to allow such transport (Nguyen et al., 2018). After one contraction event, which takes 5 ms or less (Basler et al., 2012), the sheath subunits expose a binding motif for a cytoplasmic ATPase, which disassembles the sheath thereby recycling the sheath subunits (Figure 1.9) (Kapitein et al., 2013, Wang et al., 2017a).

As already mentioned, the TMC shows a small inner diameter and no OM pore (Durand et al., 2015). Also lysozyme domains that would break down the peptidoglycan layer of a target cell and are commonly found on tube tips of T4 phages are mostly missing in T6SS VgrG proteins (Pukatzki et al., 2007). However, the force generated by the contraction of the sheath is enough to push the tube by brute force through all of these barriers. A calculation of this force is available for a T6SS from *V. cholera* based on cryo-EM structures of the extended and contracted tail (Wang et al., 2017a). The authors calculate that the contraction of a 1 µm long sheath would release ~44 000 kcal/mol of free energy, pushing the tube by 420 nm and rotating it at 40 000 – 50 000 rpm. The most surprising aspect of this model is the missing OM pore, which implies that T6SS expressing bacteria puncture their own OM in the secretion process. This is supported by the structure of the TMC in combination with biochemical experiments: In the closed-resting state (expanded sheath), the TMC protein TssM does not reach the exterior but only contacts the OM from the periplasmic side. However, during T6 secretion loops of TssM containing cysteine mutations can be labeled from the outside, showing that these loops reach the cell exterior during the secretion event (forced-open state, Figure 1.9) (Durand et al., 2015).
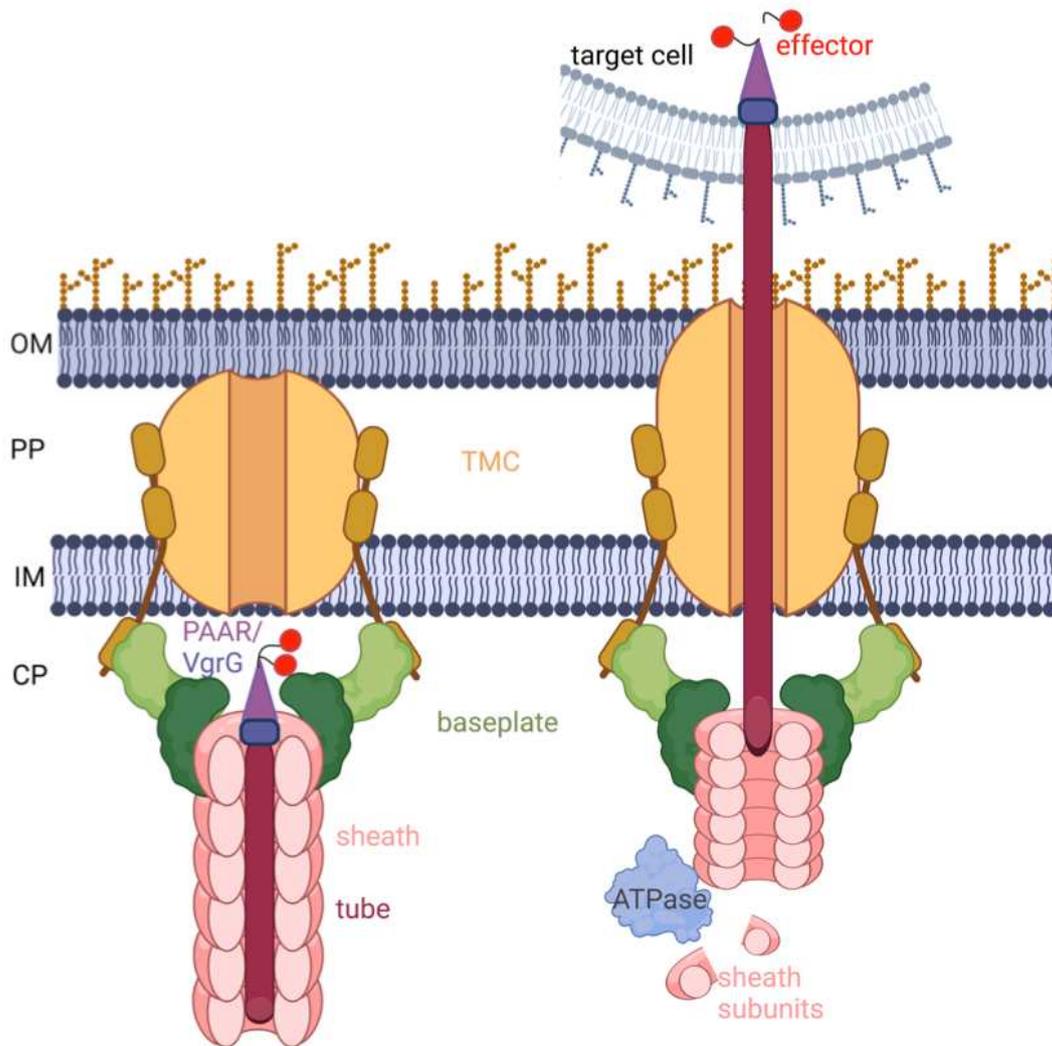
Figure 1.9 Schematic representation of a T6SS (created with BioRender.com). The TMC is shown in shades of yellow and the cytoplasmic baseplate in shades of green. During the resting state (left) the TMC does not cross the OM. After secretion (right) parts of the TMC become accessible from the outside. At the baseplate the specialized tip is assembled. It consists out of the trimeric tip protein VgrG (dark purple) and the tip is sharped by proteins of the PAAR family (lighter purple). Effector proteins (red) can bind to VgrG and the PAAR proteins (only one option is shown for simplicity). Below the tip, the tube (dark red) is assembled and surrounded by a sheath (salmon). Upon an unknown signal the sheath contracts pushing the tube out of the cell and through the membrane of a neighboring cell. A cytoplasmic ATPase (blue) disassembles the sheath thereby restoring the pool of sheath subunits. Note that the representation is not to scale. The sheath can span the entire width of the cell.

In conclusion, T3 and T4SSs both form tubes or pili, which cross the membrane of a respective host. Substrate delivery for both systems is achieved by active transport of unfolded substrates (or ssDNA for most T4SSs) through this pilus. In contrast to that, T2 and T6SSs attach substrates to the tip of a tube and then push this tube either by continuous assembly of the tube subunits (T2) or by a contractile

mechanism of a surrounding sheath structure (T6). All of these systems rely on single membrane crossing systems such as Sec, TAT or BAM to deliver and assemble at least some of their components (Costa et al., 2015).

1.2   Tripartite efflux pumps

Tripartite efflux pumps have a relatively simple architecture compared to other double membrane crossing secretion systems and are consequently widespread among Gram-negative bacteria (Linhartová et al., 2010, Weston et al., 2018). They consist out of one outer membrane protein (OMP) and two inner membrane (IM) proteins that together form a channel to transport a substrate across both membranes (Eswaran et al., 2004, Kanonenberg et al., 2018). One of these two IM proteins belongs to the family of membrane fusion proteins (MFPs), which connect the OMP and the second IM component. MFPs are often lipoproteins or bipartite membrane proteins and display homology to viral MFPs but they do not facilitate the actual fusion of membranes in the context of tripartite efflux pumps (Symmons et al., 2015). Therefore, they are also referred to as periplasmic adaptor proteins. Tripartite efflux pumps are further sub-grouped by their substrate spectrum (narrow or wide) and the identity of the second IM component. T1SSs, which belong to the family of tripartite efflux pumps, usually transport one or only a few related substrates and utilize an ABC (ATP binding cassette) transporter as a second IM component (Kanonenberg et al., 2018). Other efflux pumps generally display a wider substrate spectrum and use proteins from the ABC transporter family, RND (resistance nodulation cell division) family as well as MFS (major facilitator superfamily), SMR (small multidrug resistance), MATE (multidrug and toxic compound extrusion) or PACE (proteobacterial antimicrobial compound efflux) family as the second IM component (Anes et al., 2015). While both IM components are specific to each other and the respective substrate (spectrum), the OMP can be more promiscuous towards its interaction partners: For example, the OMP TolC of *E. coli* is used in the HlyA T1SS and the AcrAB RND-type efflux pump (Wandersman and Delepelaire, 1990, Fralick, 1996). Therefore, under certain experimental conditions these two tripartite

efflux pumps compete for the available TolC pool (Cescau et al., 2007). The following sections will give an overview of RND-type efflux pumps and T1SSs.

1.2.1 RND-type efflux pumps

RND-type efflux pumps are able to transport a variety of often unrelated toxic substrates outside of a cell and can thereby support multidrug resistance (MDR) of a respective pathogen (Nikaido and Pagès, 2012). These pumps are widely distributed among Gram-negative bacteria and their distribution is further enhanced by the frequent use of antibiotics in farming and healthcare (Anes et al., 2015). Understanding their efflux mechanism and finding potential inhibitors is therefore one strategy to tackle the problem of rinsing MDR among pathogens (Pagès et al., 2005).

In 2014 Du *et al.* published the first structure of an assembled RND-type efflux pump from *E. coli* at 16 Å (Du et al., 2014). This was followed in 2017 and 2019 by near-atomic resolution structures of fully assembled RND-type efflux pump of *E. coli* and *P. aeruginosa* respectively and both will be used as main examples for RND-type efflux pumps here (Figure 1.10) (Wang et al., 2017b, Tsutsumi et al., 2019). The RND-type efflux pump of *E. coli* is comprised of the MFP AcrA, the RND-type transporter AcrB and the OMP TolC (AcrAB-TolC). The pump of *P. aeruginosa* follows a similar nomenclature with the MFP MexA, the RND-type transporter MexB and the OMP OprM (MexAB-OprM). Both pumps are able to efflux organic solvents and many different antibiotics including but not limited to penicillins, tetracyclines, macrolides and chloramphenicol (Li and Poole, 1999, Masuda et al., 2000, Sulavik et al., 2001). Furthermore, at least AcrAB-TolC is also able to efflux compounds such as SDS (sodium dodecylsulfate), rhodamin 6G and ethidium bromide (Sulavik et al., 2001).

Both pumps display the same overall stoichiometry with a trimeric RND-type transporter and OMP, connected by a hexameric MFP (Wang et al., 2017b, Tsutsumi et al., 2019). The OMPs display the lowest sequence identity in this system (23 %) but their overall structure is very similar (Figure 1.10). They show a β-barrel in the OM and an α-helical extension that reaches about 100 Å deep into the periplasm thus

piercing the peptidoglycan layer. Each monomer has two helix-turn-helix (HTH) motifs at the periplasmic end thereby displaying six HTH motifs for interaction with the HTH motif of the hexameric MFPs (Wang et al., 2017b, Tsutsumi et al., 2019). AcrA and MexA both hold the conserved RLS motif (RxxxLxxxxxxS; x stands for any amino acid), which was proposed to facilitate the contact between MFP and OMP (Lee et al., 2012). However, the contact interface is more complex than that as neighboring residues are also necessary for OMP contact and R96 and S107, which are part of MexA's RLS motif, were shown to be non essential for complex formation by mutational studies (Tsutsumi et al., 2019).

Although *in vitro* structures of fully assembled pumps are available, there is still a heavy debate around the actual contact between OMP and MFP. Both pumps were only captured showing tip-to-tip interactions between MFP and OMP, hence undermining the earlier proposed 'wrapping' or 'deep interpenetration' model (Figure 1.10 A) (Symmons et al., 2009, Wang et al., 2017b, Tsutsumi et al., 2019). In this model MFP and OMP make extensive contacts along their helices and OMP and RND-type transporter make direct contact (Symmons et al., 2009). The model was proposed for AcrAB-TolC and supported by cross-linking experiments that showed a direct interaction between AcrB and TolC (Touzé et al., 2004, Tamura et al., 2005) and has recently gained more experimental support by mutational studies on TolC (Marshall and Bavro, 2020). Furthermore, docking studies with crystal structures of isolated AcrA and TolC showed how AcrA can interact with closed and open TolC in the 'wrapping model' (Lobedanz et al., 2007). Contrary to that, Jo *et al*. claim that the grooves on TolC, in which the helices of AcrA are fitted, disappear in the open state, rendering the 'wrapping model' "unreasonable" (Jo et al., 2019). Considering the limits of structure determination methods, it remains an open question if the contact is only mediated in a tip-to-tip manner as seen in cryo-EM structures of detergent solubilized complexes or is more dynamic during the whole secretion process. A switch between 'tip-to-tip' and 'wrapping' mode of interaction would compress the pump along the long axis by ~35 Å and would consequently compress the periplasm and induce membrane curvature. In this context it is noteworthy that the substrate-

bound AcrAB-TolC pump is ~10 Å shorter along the long axis (from IM to OM) than the resting AcrAB-TolC pump (Wang et al., 2017b). If the compression of the pump also takes place *in vivo*, remains to be determined.
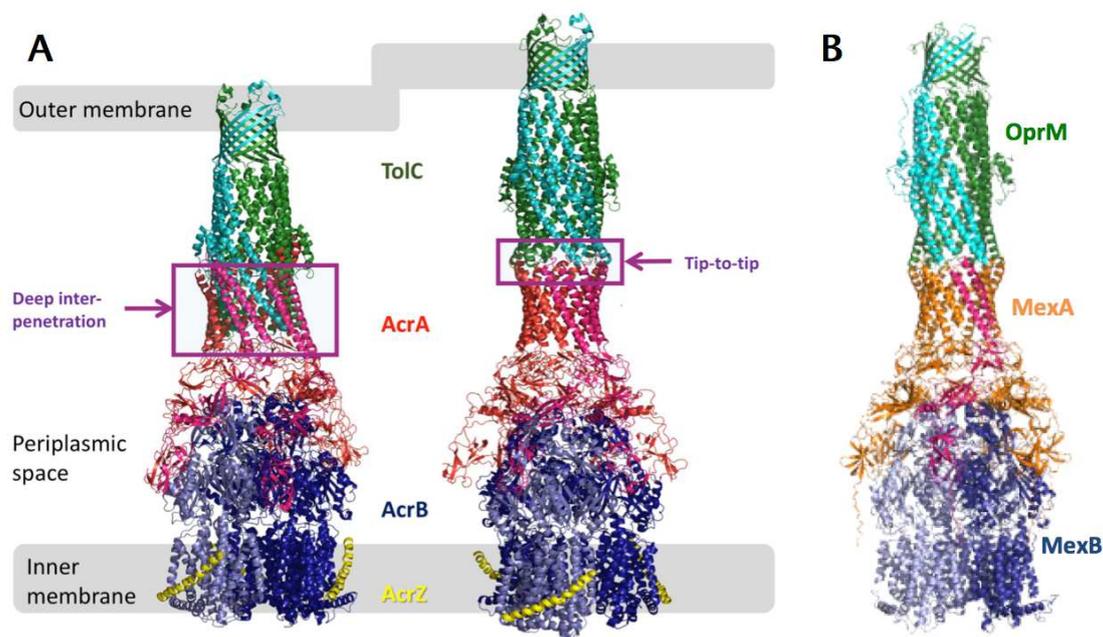
Figure 1.10 Structures of RND-type efflux pumps. A) Comparison of the 'deep interpenetration' model (also 'wrapping model') and the tip-to-tip connection of the AcrAB-TolC efflux pump of *E. coli* (adapted from (Marshall and Bavro, 2020)). TolC is shown in dark green with one monomer in cyan. AcrA is shown in pink and orange. AcrB is shown in shades of blue. AcrZ (yellow) is an accessory protein that modulates AcrB's activity but is only involved in the export of some of AcrB's substrates (Hobbs et al., 2012). **B)** Structure of the MexAB-OprM efflux pump from *P. aeruginosa* (PDB 6TA5, drawn with PyMOL). OprM is shown in dark green and one monomer in cyan. MexA is shown in orange and one monomer in pink. MexB is shown in dark blue and one monomer in light blue.

The before-mentioned HTH motifs of the MFPs are part of the 40-45 Å long α-helical domain, which is followed by a ring of six lipoyl domains (Wang et al., 2017b, Tsutsumi et al., 2019). These lipoyl domains neither contact the OMP nor the RND-type transporter but rather mediate the oligomerization of the MFP (Staron et al., 2014). Moving closer to the IM, two more rings with increasing diameter follow, formed by the β-barrel and membrane proximal (MP) domain respectively. Both of these make contact to the RND-type transporter displaying alternating larger and smaller contact areas, leading to the interpretation that the MFPs form a trimer of dimers (Wang et al., 2017b, Tsutsumi et al., 2019). Furthermore both MFPs, AcrA

and MexA, are lipoproteins that are anchored to the IM by a palmityl moiety but are also functional without it (Zgurskaya and Nikaido, 1999a, Yoneyama et al., 2000).

The trimeric RND-type transporter is the last component of the efflux pump and resides in the IM with a transmembrane (TM) domain containing 12 TM helices per monomer (Murakami et al., 2002, Sennhauser et al., 2009). Roughly the other half of the protein is located in the periplasm and can be separated into two domains: The porter domain and the funnel-like (MexB) or docking (AcrB) domain, which both contact the MFP (Wang et al., 2017b, Tsutsumi et al., 2019). The docking domain of AcrB has also been described as funnel-like but has been termed docking domain because it was believed to dock directly to TolC ('wrapping model') (Murakami et al., 2002). The porter domain contains the drug binding pocket and, in the case of AcrB, three drug entrance channels, which enable the transporter to take up drugs from the periplasm or the IM (Nakashima et al., 2011). During drug extrusion each monomer of the transporter adopts three distinct conformations: i) Access or 'loose' (L) conformation, in which the respective drug entrance channel is open but the actual binding pocket is still shrunken. ii) Binding or 'tight' (T) conformation, in which the drug binding pocket is expanded, the entrance gate is still open and the exit gate towards the funnel-like domain is still closed. iii) Extrusion or 'open' (O) conformation, in which the entrance gate is closed, the exit gate is open and the drug binding pocket is shrunken again, pushing the drug towards the exit gate (Murakami et al., 2006, Du et al., 2015). The voluminous drug binding pocket is able to accommodate for substrates of different size and displays multiple aromatic residues for multisite binding of different drugs (Nakashima et al., 2011). Both, AcrB and MexB, have been found as asymmetric trimers with each monomer displaying one of these three conformations (LTO) suggesting a functional rotation mechanism during which each monomer of the RND-type transporter cycles through these stages to promote drug efflux (Murakami et al., 2006, Sennhauser et al., 2009). RND-type transporters are pmf-dependent and the energy for cycling through these conformations is derived from proton antiport (Zgurskaya and Nikaido, 1999b, Ikonomidis et al., 2008). Symmetric trimers are mostly linked to resting or inactive

efflux pumps. For example, AcrB has been found mainly in the TTT conformation in the presence of an inhibitor and mainly in the LLL conformation in the absence of a substrate (resting-state) (Wang et al., 2017b).

Assembly of RND-type efflux pumps is a two-step process: The IM components assemble first into a subcomplex, AcrAB or MexAB, before recruiting the OMP (Shi et al., 2019, Tsutsumi et al., 2019). This recruitment does not seem to be substrate dependent (Touzé et al., 2004): For MexAB-OprM the pump can be assembled *in vitro* without substrates present (Tsutsumi et al., 2019). Localization studies on GFP (green fluorescent protein)-tagged AcrB showed that it is confined in the presence of TolC (and AcrA) and mobile in TolC-deletion strains (Yamamoto et al., 2016), which implies that AcrAB can bind to TolC in the absence of substrates. In the context of living cells this might be negligible as expression is tightly regulated and only induced upon certain environmental signals for example in the presence of an antibiotic (=substrate) (Weston et al., 2018). The substrate-free recruitment of the OMP might be transient and stabilized in the presence of substrates (Touzé et al., 2004).

A major difference between the introduced pumps is also linked to the recruitment of the OMP. In the structure of the assembled MexAB-OprM pump, OprM, was found in an open conformation (Figure 1.10 B) (Tsutsumi et al., 2019). Fitting the closed conformation of OprM from a crystal structure onto MexAB resulted in steric clashing of residues, among them L100 of the conserved RLS motif of MexA. The authors concluded that this clashing is circumvented *in vivo* by opening of OprM (Tsutsumi et al., 2019). This means that the assembly of the MexAB-OprM pump directly opens the OMP. This is in contrast to AcrAB-TolC, which was found assembled with TolC closed. Capturing the open conformation required the presence of a substrate or inhibitor (Wang et al., 2017b).

The extensive research on both pumps makes it possible to propose a secretion mechanism. In short, a stress signal induces the expression of the pump components (Weston et al., 2018). MFP and RND-type transporter form a complex in the IM, while the OMP is assembled (or already present) in the more rigid OM in a closed

conformation (Shi et al., 2019, Tsutsumi et al., 2019). The OMP pierces the peptidoglycan layer in the periplasm, which further hinders its lateral diffusion. The IM complex on the other hand can diffuse freely in the IM until it contacts the respective OMP (Yamamoto et al., 2016). Contact of MexAB to OprM induces an iris-like opening of OprM, while the complex is still sealed towards the periplasm (drug entrance channel is blocked by a subdomain of MexB) and cytoplasm as neither MexB nor AcrB form a channel though the IM (Murakami et al., 2002, Tsutsumi et al., 2019). Contact of AcrAB to TolC does not induce opening. For that conformational changes in AcrB are required, which are induced upon substrate binding. This conformational change is transferred to AcrA, which in turn induces the opening of TolC (Wang et al., 2017b). Both pumps extrude substrates in a pmf-dependent manner by the functional rotation mechanism (Zgurskaya and Nikaido, 1999b, Ikonomidis et al., 2008): Each monomer cycles through stages of drug access, binding and extrusion in order to increase the drug concentration in the OMP cavity until it exceeds the extracellular drug concentration. At this point the drug molecule will diffuse out of the channel (Du et al., 2015). When the initial stress signal subsides, transcriptional regulation returns to repressing the respective genes and membrane homeostasis pathways likely degrade the now 'useless' transporter components (Dalbey et al., 2012, Weston et al., 2018).

1.2.2 Type I secretion system (T1SS)

T1SSs represent a large subgroup of tripartite efflux pumps with a narrow substrate spectrum and an ABC transporter as a second IM component. As all other tripartite efflux pumps they also utilize an MFP and OMP (Holland et al., 2016). The ABC transporter and MFP are specific to the transported substrate and these three components (ABC, MFP, substrate) are often encoded adjacent in their operon. The OMP can be part of the same operon or be encoded elsewhere as it is not always specific to the system (Thomas et al., 2014). The substrates vary greatly in size from ~9 kDa (MccV) to ~1.5 MDa (IBA) and include proteases, lipases, toxins, iron-

scavenger proteins, bacteriocins and adhesins (Linhartová et al., 2010, Spitz et al., 2019).

T1SSs are assigned to one of three groups based on the identity of their ABC transporter (Kanonenberg et al., 2013). In general, ABC transporters contain a transmembrane domain (TMD) with multiple TM helices and a nucleotide binding domain (NBD), which can bind and hydrolyze ATP (Locher, 2016). The ABC transporters of group 1 and 2 contain an additional domain at their N-terminus, which can be an active peptidase (group 1) or an inactive peptidase (group 2), while ABC transporters of group 3 contain no additional domains (Kanonenberg et al., 2013). Recently a new subgroup has been identified with the help of phylogenetic analysis (Smith et al., 2018b). This subgroup secretes gigantic proteins, which is probably why the literature does not touch on the peptidase activity of the small N-terminal domain of the ABC transporter. The following sections will introduce each group and highlight their differences.

1.2.2.1 Group 1

ABC transporters of group 1 T1SSs hold an active cysteine peptidase at their N-terminus, which cleaves off the leader peptide during secretion of the substrate (Kanonenberg et al., 2013). One of the best investigated systems of this group is the MccV T1SS from *E. coli*, which consists out of the MFP CvaA, the ABC transporter CvaB and the OMP TolC (Fath et al., 1992). The substrate MccV belongs to the class of bacteriocins, which are small ribosomally synthesized peptides that display antibacterial activity. In Gram-negative bacteria bacteriocins are further subgrouped into the larger colicins (30-80 kDa) and the smaller microcins (1-10 kDa) (Rebuffat, 2011). It is important to note that MccV was first classified as a colicin and termed colicin V (ColV) before it was reclassified as a microcin and renamed by Duquesne *et al*. in 2007 (Duquesne et al., 2007). However, both names, ColV and MccV, are still commonly found in the literature. Microcins are even further subgrouped into class I (<5 kDa) and class II (>5 kDa), with class I microcins often undergoing extensive post-translational modifications before secretion (Duquesne et al., 2007). The class II

microcins, which MccV belongs to, can also undergo post-translational modifications but more importantly they share a common operon organization, which separates them from class I: In the minimal version the operon of class II contains the gene for the pro-microcin, a self-immunity protein, an ABC transporter and an MFP (class IIa). At times the operon is expanded by the presence of genes encoding for post-translational modification enzymes (class IIb) (Duquesne et al., 2007).

MccV in its mature form has a calculated size of 8.7 kDa and is active against related Gram-negative bacteria (Cohen et al., 2018). It is first synthesized from *cvaC* as a 103 amino acid long pro-microcin with the first 15 amino acids corresponding to the leader peptide. The mature MccV contains one disulfide bond between C78 and C89 (numbering corresponds to mature MccV without leader peptide), which is probably formed before secretion (Håvarstein et al., 1994). Not much is known about the actual transport process or spatial and temporal control of channel assembly. However, it is believed to be similar to other T1SSs and tripartite efflux pumps, like AcrAB-TolC: The MFP CvaA probably connects TolC and CvaB, which leads to the formation of one continuous channel from the cytoplasm to the extracellular space (Hwang et al., 1997). In order to protect the secreting bacterium from influx of toxic compounds and uncontrolled efflux through this channel, the opening and/or channel assembly is most likely connected to substrate recognition. So far it has been shown that CvaB cleaves the leader peptide of pro-MccV (Wu and Tai, 2004), showing that recognition could be mediated by the interaction of the leader peptide with the peptidase domain, but further interactions are possible. Once recognized, the CvaB peptidase domain cleaves the leader peptide after a conserved double-glycine motif in a Ca^{2+}-dependent manner (Wu and Tai, 2004). CvaB then translocates MccV from the cytoplasm to the proposed CvaA-TolC cavity in the periplasm likely by the 'alternating access' mechanism (Beis and Rebuffat, 2019). In the context of microcin secretion from Gram-negative bacteria this mechanism can best be explained with the ABC transporter McjD from *E. coli* since crystal structures of different conformations are available (Choudhury et al., 2014, Mehmood et al., 2016, Bountra et al., 2017).

McjD transports the class I microcin MccJ25 (Solbiati et al., 1999). Interestingly, the export of MccJ25 is dependent on TolC and an ABC transporter but does not seem to be facilitated by a T1SS, as the MFP is absent in the MccJ25 operon as well as in other class I microcin operons (Delgado et al., 1999, Duquesne et al., 2007). The four available crystal structures of McjD (Choudhury et al., 2014, Mehmood et al., 2016, Bountra et al., 2017) make it possible to postulate the transport mechanism as follows: In the absence of the substrate and ATP the transporter can adopt two different conformations: inward-occluded or inward-open. In both conformations the NBDs are disengaged and the transporter is sealed towards the periplasm (inward). In the inward-occluded conformation the transporter is also sealed towards the cytoplasm while in the inward-open conformation the substrate binding cavity is accessible for MccJ25. Substrate binding probably precedes ATP binding, because futile ATP binding and hydrolysis (in the absence of substrate) does not induce changes in the TMDs and thereby does not open the transporter towards the periplasm. Furthermore, ATP binding leads to the dimerization of the NBDs, which leads to an outward-occluded conformation that does not allow substrate entry into the cavity. Thus, after MccJ25 binds to the inward-open conformation of McjD, the NBDs bind ATP and dimerize. This induces first a transient outward-occluded conformation, which transitions into a short-lived outward-open conformation to release MccJ25. Substrate release transitions the outward-open conformation back to the outward-occluded conformation and ATP hydrolysis resets McjD to an inward conformation (inward-occluded or inward-open) (Figure 1.11) (Bountra et al., 2017). It should be noted that McjD does not contain a protease domain like CvaB and most likely does not interact with an MFP. The presence of the MFP CvaA, however, is required for CvaB activity (Zhong et al., 1996) and multiple residues apart from the catalytic triad in the protease domain of CvaB have been shown to influence MccV secretion activity (Wu et al., 2012). Therefore, the transport mechanism for CvaB might be more complex than for McjD, involving for example interactions with CvaA or signal transduction by the protease domain.
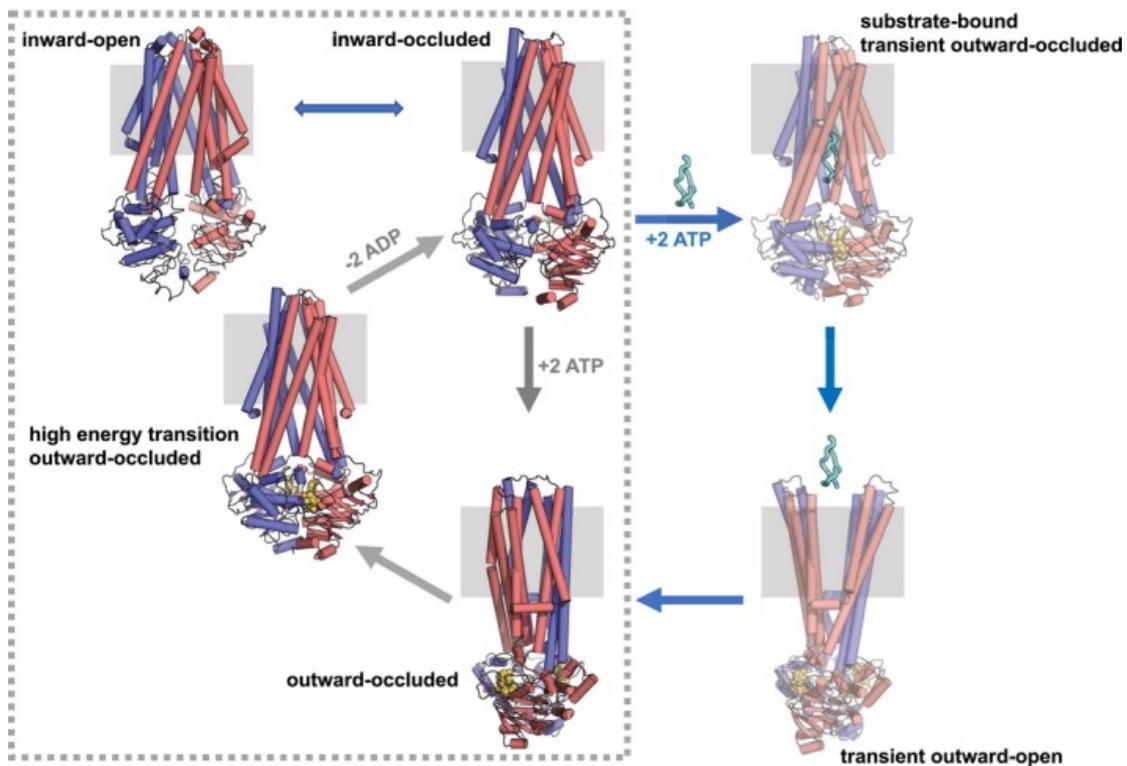
Figure 1.11 Mechanism of antibacterial peptide export by McjD (taken from (Bountra et al., 2017). McjD is shown in blue and red; the grey box represents the membrane. Bound nucleotides are shown in yellow and the substrate MccJ25 in cyan. McjD adopts either the inward-open or inward-occluded conformation. Futile ATP hydrolysis (grey dotted box) does not induce enough conformational change to open the TMDs towards the periplasm. Substrate and ATP binding changes McjD to a substrate-bound transient outward-occluded conformation. This is followed by a transient outward-open conformation that allows the substrate to leave the transporter. Without the substrate McjD adopts the outward-occluded conformation again and ATP hydrolysis and ADP and P_i release reset McjD to an inward conformation.

In summary, group 1 T1SSs transport the smallest substrates among T1SSs (bacteriocins) by the alternating access mechanism after cleavage of an N-terminal leader peptide by an N-terminal protease domain on the ABC transporter (Kanonenberg et al., 2013).

1.2.2.2 Group 2

Similar to group 1, ABC transporters of group 2 T1SSs also hold an N-terminal extension (Kanonenberg et al., 2013). This extension forms a domain, which resembles a C39-peptidase in sequence (~40 % identity) and fold (Figure 1.12 B) but shows no peptidase activity (Lecher et al., 2012). A famous example of this group is the HlyA T1SS from *E. coli*, with the ABC transporter HlyB, the MFP HlyD and the

OMP TolC, which together form a channel to transport the substrate HlyA in a one-step mechanism across both membranes into the extracellular environment (Holland et al., 2016). The structure of the N-terminal extension of HlyB was solved by nuclear magnetic resonance (NMR) spectroscopy in 2012 (Figure 1.12 A) and explained why no peptidase activity can be measured for this domain (Lecher et al., 2012). Firstly, the catalytically essential cysteine residue is replaced by a tyrosine (Tyr9). However, exchanging it with a cysteine residue does not recover proteolytic activity. Because secondly, the catalytically essential histidine residue (His83) is flipped by 180° (compared to an active C39-peptidase) and stabilized in this position by π-π interactions with a tryptophane (Trp77). In this orientation histidine cannot accept the proton from the nucleophilic residue and the stabilizing tryptophane is conserved among group 2 T1SS ABC transporters (Lecher et al., 2012). Therefore, this domain was termed C39-like domain (CLD). The exact function of this domain is still unclear but its presence is essential to the secretion process (Lecher et al., 2012, Reimann et al., 2016).

Substrates of group 2 T1SSs are transported without removal of a leader peptide. In fact, they do not display a classical N-terminal leader peptide at all but rather a C-terminal secretion signal, which is not cleaved and reaches the cell surface first (Lenders et al., 2015). The characteristics of this secretion signal are still heavily debated with implications for a structural motif, a sequence motif or a combination of both (Holland et al., 2016) (see chapter 3.3). Also the length of the secretion signal seems to differ among the substrates, with larger substrates having longer secretion signals: For example, HlyA has a size of 110 kDa (1024 amino acids) and a secretion signal of 50-60 residues (Gray et al., 1986, Hess et al., 1990, Jarchau et al., 1994). On the other hand, CyaA from *B. pertussis* has a size of 177 kDa (1706 amino acids) and a secretion signal of more than 70 residues (Sebo and Ladant, 1993).

As already seen in these two examples, the substrates of group 2 T1SSs are significantly larger than those of group 1. Furthermore, they are characterized by a repeated sequence motif, termed 'repeat in toxin' (RTX) motif or 'GG-repeat', with the consensus sequence GGxGxDxUx (where x can be any amino acid and U represents a

large hydrophobic amino acid) (Welch, 1991). Several of these motifs cluster in a C-terminal RTX domain, which precedes the C-terminal secretion signal. The number of RTX motifs varies between RTX proteins from less than 10 to more than 40 repeats and is loosely correlated to the size of the respective RTX protein (Linhartová et al., 2010, Bumba et al., 2016). This motif is not confined to the conserved C-terminal RTX domain as additional RTX motifs can be distributed throughout the protein (Linhartová et al., 2010). They are able to bind Ca^{2+} with a K_D of approximately 150 μM (HlyA) (Sánchez-Magraner et al., 2007). Since the intracellular Ca^{2+} concentration is tightly regulated in *E. coli* and other bacteria ($[Ca^{2+}]_{cytosol} \approx 300$ nM), Ca^{2+} binding takes place outside of the cell where the Ca^{2+} concentration is much higher, commonly exceeding 2 mM (Gangola and Rosen, 1987, Jones et al., 1999). Upon Ca^{2+} binding the RTX domain adopts a β-roll structure (Figure 1.12 C) (Baumann et al., 1993) where Ca^{2+} is coordinated by the side chain of aspartate and by the carbonyl groups of the backbones of residues 1 and 5 of the RTX motif (**G**GxG**xD**xUx). The β-roll is topped by a capping structure of five antiparallel β-sheets (Figure 1.12 C) and both structural features are conserved among RTX proteins (Bumba et al., 2016). Folding of the RTX domain is thought to induce folding of the rest of the protein, however, the N-terminal domain represents the functional domain and can therefore be very diverse (Linhartová et al., 2010). For example, CyaA from *B. pertussis* holds five consecutive RTX domains where folding of the most C-terminal RTX domain induces folding of the following RTX domains (in the presence of Ca^{2+}) (Bumba et al., 2016). The functional N-terminal adenylat cyclase domain, however, is able to refold in the presence of its substrate calmodulin but independently of the RTX domains (Karst et al., 2010).
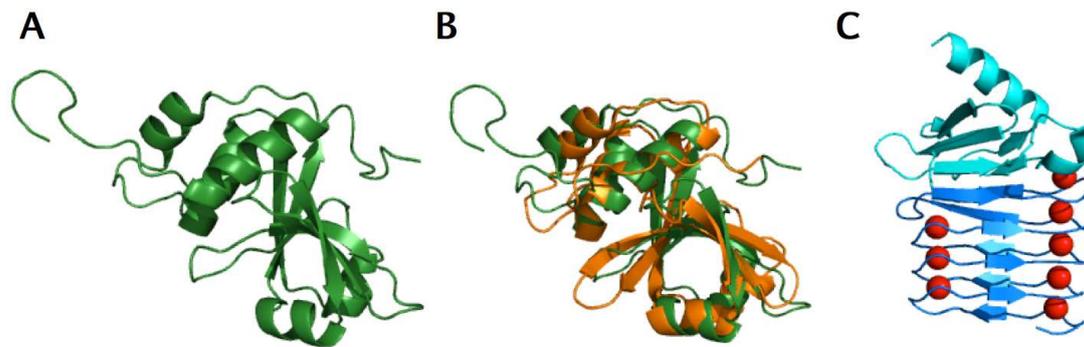
Figure 1.12 Structural features of group 2 T1SS. A) Cartoon representation of HlyB CLD from *E. coli* (PDB 3ZUA). **B)** Overlay of HlyB CLD (green) with active C39 peptidase ComA-PEP (orange) from *Streptococcus mutans* (PDB 3K8U). The structures are very similar with an RMSD of 1.8 Å over 137 C_α atoms. **C)** Cartoon representation of the most C-terminal RTX domain (RTX V; blue) of CyaA from *B. pertussis* (PDB 5CVW). Coordinated Ca^{2+} is shown as red spheres. The β-roll structure is topped by a capping structure (cyan). All structures were drawn with PyMOL.

The need for independent folding of domains is more apparent in multifunctional auto-processing RTX proteins (MARTX). They represent a subgroup of RTX proteins, which are secreted by group 2 T1SSs. They hold up to five different functional effector domains and a conserved auto-processing cysteine protease domain in their N-terminus. Furthermore, they also display the described C-terminal RTX domain and secretion signal and are therefore very large proteins with up to 5300 amino acids (Satchell, 2011). Upon secretion and folding of the RTX domain, MARTX proteins are able to form a pore in the membrane of a host cell, through which they can deliver their effector domains (Woida and Satchell, 2018). As most proteins of an OM, this pore is proposed to form a β-barrel with an estimated radius of 16.3 Å (Kim et al., 2008), which only allows for unfolded effectors to pass through (Kudryashova et al., 2014). Subsequently, these effector domains need to fold inside of the host cell where they modulate host cell signaling in various different ways (Woida and Satchell, 2018). The delivery of MARTX proteins to their host cell therefore resembles the T5SS where a single polypeptide chain can form a pore in the OM through which it releases effector domains by auto-processing (see section 1.1.1.).

There is no structure available for a fully assembled T1SS but the crystal structure of TolC in its open conformation suggests that HlyA and other T1SS substrates need to be transported unfolded (Koronakis et al., 2000). Furthermore,

specific interactions between ABC transporter and substrate were shown to depend on an unfolded substrate (Lecher et al., 2012). This represents a particular challenge for proteins of these sizes (commonly >100 kDa), as their C-terminus reaches the cell surface first (Lenders et al., 2015). The proteins have to be completely synthesized at the ribosome before the C-terminal secretion signal emerges and kept unfolded and undegraded in the cytosol of the producing bacteria. A chaperon usually performs such a task and several other secretion systems hold designated chaperons in their operon (Costa et al., 2015). However, no chaperon has been identified for any of the group 2 T1SS substrates (Kanonenberg et al., 2013). The absence of a chaperon and the experimentally prooven interaction of the CLD with the unfolded substrate led to the assumption that the CLD takes on the role of a chaperon for T1SS substrates (Lecher et al., 2012). This is supported by systems that lack a CLD or C39-peptidase at their N-terminus (group 3 T1SS) as they seem to depend on an additional chaperon (see next section) (Kanonenberg et al., 2013).

The interaction of the unfolded substrate with the components of the IM leads to the recruitment of the OMP (Thanabalu et al., 1998, Lecher et al., 2012). However, the already described 'alternating access' mechanism (see section 1.2.2.1) is unreasonable for substrates of group 2 T1SSs. They are simply too large to fit into any substrate binding cavity a respective ABC transporter could provide.

In 1992 Simon *et al.* postulated the 'Brownian ratchet' mechanism for which Bumba *et al.* provided experimental evidence in the context of CyaA (Simon et al., 1992, Bumba et al., 2016). The transport process is divided into two steps: i) the initial threading of a part of the protein into the transport machinery and ii) the actual subsequent transport (Simon et al., 1992). Surface plasmon resonance (SPR) studies with truncated version of HlyA and the isolated NBD of HlyB showed that they interact and that this interaction is disrupted in the presence of ATP (Benabdelhak et al., 2003). This points to a initial threading process (i), where HlyA binds to the nucleotide free form of HlyB NBD and subsequent ATP binding leads to the threading of the secretion signal into the transport channel. The actual transport process (ii) is thought to follow the 'Brownian ratchet' mechanism (Bumba et al.,

2016). Brownian motion at any given physiological temperature will lead to a back-and-forth movement of a protein chain though a translocation pore. In order to achieve directionality of transport the movement needs to be biased, which can be achieved by modifications of the protein chain on either the *cis* or *trans* side of transport (Simon et al., 1992). In the context of group 2 T1SSs the directionality of transport is achieved by Ca^{2+} induced folding outside of the cell, which prevents the protein from sliding back into the translocation pore (Bumba et al., 2016). The RTX motifs thereby represent an intramolecular 'ratchet' that biases the Brownian movement of the substrate towards the outside of the cell.

Following this model, secretion will be i) temperature sensitive as Brownian motion is temperature-dependent (Simon et al., 1992) and secretion will ii) still take place without the ratchet but will be drastically slower (~100 fold for CyaA; based on calculations from (Hepp and Maier, 2017)). The temperature dependency of secretion was shown for HlyA (Koronakis et al., 1991) and CyaA secretion was reduced by ~20 fold when the 'ratchet' was disabled by reducing the extracellular Ca^{2+} to a concentration that does not induce folding (Bumba et al., 2016). This reduction is less than expected from the model and points to additional mechanisms that support secretion. Furthermore, the secretion rate of HlyA was unaffected by disabling the 'ratchet' (reduction of extracellular Ca^{2+}) (Lenders et al., 2016). The difference in Ca^{2+}-dependency for secretion is mostly blamed on their different sizes and number of RTX motifs, as HlyA has six RTX motifs and a size of 1024 amino acids and CyaA has 17 RTX motifs and a size of 1706 amino acids (Hepp and Maier, 2017). Thus, CyaA is not only larger but holds more RTX motifs per residue. The aforementioned additional mechanism that supports secretion is most likely provided by the ABC transporter. A mutant of HlyB (G609D) that was able to bind but not hydrolyze ATP, was deficient in HlyA secretion. However, this mutant was still able to form a complex with HlyD, engage with HlyA and recruit TolC, pointing to an involvement of ATP hydrolysis in later transport steps (Thanabalu et al., 1998). Furthermore, there are implications that the pmf is also necessary for transport, as addition of the pmf-depleting agent CCCP (carbonyl cyanide m-chlorophenyl hydrazone) affected HlyA

secretion in early steps rather than late steps of transport (Koronakis et al., 1991). Taken together, these data infer that secretion of the large group 2 T1SS substrates depends on the pmf, ATP binding and hydrolysis and the extracellular folding induced by Ca^{2+} binding ('Brownian ratchet') (Koronakis et al., 1991, Bumba et al., 2016). The contribution of each process to the actual transport likely differs for different substrates.

In summary, group 2 T1SSs are characterized by the presence of a CLD on their ABC transporter, which could act as a chaperon for the respective substrate (Kanonenberg et al., 2013). These substrates share common features such as the presence of a C-terminal RTX domain, which is followed by an uncleaved C-terminal secretion signal (Linhartová et al., 2010). Transport proceeds from C- to N-terminus and is likely dependent on the pmf, the ability of the ABC transporter to bind and hydrolyze ATP as well as Ca^{2+}-induced extracellular folding (Koronakis et al., 1991, Thanabalu et al., 1998, Lenders et al., 2015, Bumba et al., 2016).

1.2.2.3 Group 3

The ABC transporters of group 3 T1SSs do not display an additional N-terminal domain and hence substrates of this group are also not cleaved (Kanonenberg et al., 2013). Similar to substrates of group 2, the substrates of group 3 have no N-terminal leader peptide but display a C-terminal secretion signal and are transported unfolded (Debarbieux and Wandersman, 2001, Kanonenberg et al., 2013). Most, but not all, also hold a C-terminal RTX domain. One famous example of this group without an RTX domain is the iron-scavenger protein HasA from *S. marcescens* (Létoffé et al., 1994a). It is a 19 kDa protein whose secretion by its dedicated T1SS is completely abolished by removal of the last 14 amino acids (Cescau et al., 2007). When heterologously expressed in *E. coli* TolC can replace the original OMP HasF to form a complex with the ABC transporter HasD and the MFP HasE (Létoffé et al., 1993). Furthermore, HasA secretion can also be mediated by the OMP PrtF: the OMP from a group 3 T1SS from *Dickeya dadantii* (former *Erwinia chrysanthemi*) (PrtDEF) (Létoffé et al., 1994b). This T1SS transports four related RTX domain containing proteases

(PrtA, PrtB, PrtC, PrtG) that are about 50 kDa in size with 3-4 RTX motifs each and a secretion signal of 40-55 residues (C-terminal) (Delepelaire and Wandersman, 1990, Ghigo and Wandersman, 1992). Importantly, the crystal structure of a homolog of the ABC component (PrtD) from *Aquifex aeolicus* (AaPrtD) was solved in an ADP-bound state showing unique features that may be conserved among the group (Figure 1.13) (Morgan et al., 2017).

A narrow channel through the TMDs was observed in the ADP-bound structure of AaPrtD (Figure 1.13 B) (Morgan et al., 2017). The channel is ~40 Å in length and closed on both sides: R307 blocks the channel towards the cytoplasm and a patch of conserved hydrophobic residues seals it towards the periplasm. The channel dimensions, although measured in a substrate-free state of the transporter, would only allow for an unfolded substrate to pass through (Morgan et al., 2017). The formation of a cavity that would allow binding of a completely folded 50 kDa protease (PrtA/B/C/G), as necessary for an 'alternating access' model, would require dramatic reorganization of the TM helices. Especially TM3 and TM6 constrict the channel since they display sharp kinks, thereby bending inward (Figure 1.13 C) (Morgan et al., 2017). In the bacteriocin exporting ABC transporter McjD (see 1.2.2.1), which acts according to the 'alternating access' model, these helices bend outward creating a substrate binding cavity (Bountra et al., 2017). In AaPrtD the inward bending of TM3 and TM6 creates a 'concave bowl' together with TM1, which is highly basic (Morgan et al., 2017). The authors note that this could enable the transporter to perform electrostatic interactions with acidic regions of the substrate. Notably, RTX proteins, including the substrates of AaPrtD, are often acidic with a theoretical pI value between 3.2 and 4.9 for 90 % of putative RTX proteins (Linhartová et al., 2010). Furthermore, residues that are involved in stabilizing the kinks in TM3 and TM6 are not found in peptide transporting ABC transporters (that function according to the 'alternating access model') but are conserved among group 2 and 3 T1SS ABC transporters including HlyB (group 2) and HasD (group 3) (Morgan et al., 2017). Although the structure of AaPrtD was solved in a substrate free state, it supports the theory of transport of unfolded substrates from the side of the

ABC transporter. This question is mostly addressed by measurements of the OMP pore or from the side of the substrate by analyzing its folding properties.
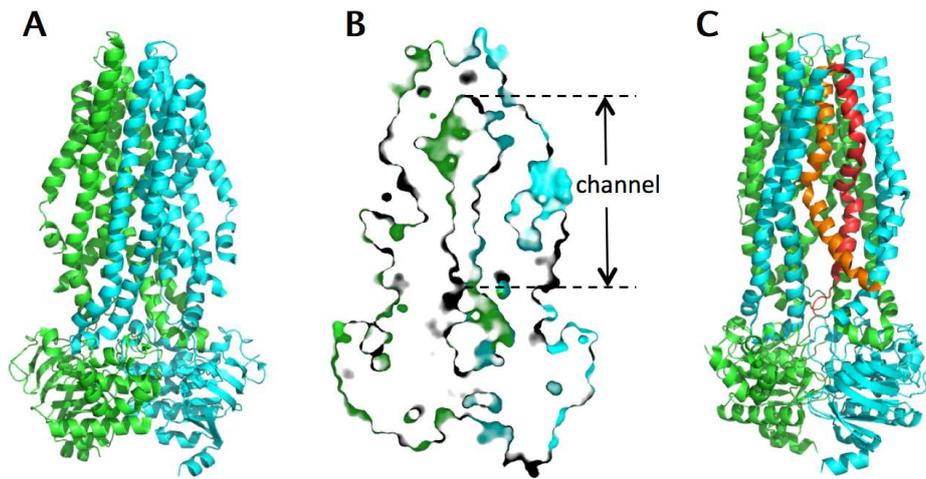
Figure 1.13 Structure of AaPrtD (PDB 5L22). A) Cartoon representation with one monomer in cyan and one monomer in green. **B)** Surface representation with same color coding and orientation as in (**A**) showing the channel through the TMDs. **C)** Same coloration as in (**A**) with TM3 highlighted in red and TM6 in orange of the cyan colored monomer. Both TM's display drastic kinks bending towards the dimer interface (inward).

The folding properties of HasA and the influence of substrate folding on assembly and secretion have been investigated in great detail. HasA rapidly folds in solution and *in vivo* (Wolff et al., 2003). By successively inducing HasA variants and the exporter components HasDE, Debarbieux and Wandersman could show that premature folding of HasA in the cytoplasm prevents secretion of newly synthesized HasA (Debarbieux and Wandersman, 2001). Under physiological conditions, SecB prevents premature folding of HasA in the cytoplasm and can also slow down HasA folding *in vitro* (Delepelaire and Wandersman, 1998, Wolff et al., 2003). The HasA T1SS is so far the only T1SS for which a SecB-dependency has been shown (Kanonenberg et al., 2013). The absence of a CLD or C39-peptidase domain in this whole group implies that also other group 3 T1SS substrates could depend on SecB or other chaperons for secretion. However, other members of this group mostly display at least one RTX domain with roughly four RTX motifs (Holland et al., 2016). These have been described as intramolecular chaperons, as they induce folding outside of the cell and *vice versa* prevent folding inside of the cell (Meier et al., 2007).

RTX domain containing substrates therefore might not need an additional chaperon. This does not undermine the suggested chaperon function of the CLD, as group 2 T1SS substrates are mostly larger than those of group 3 and might therefore depend on the additional chaperon function of the CLD in addition to the RTX domain(s) (Kanonenberg et al., 2013).

The secretion signal of HasA is the only one for which a specific function, apart from recognition and threading, has been shown (Holland et al., 2016). Cescau *et al.* demonstrated that the assembly of the tripartite complex is independent of this secretion signal and adding it *in trans* leads to disassembly of the system (Cescau et al., 2007). This raises the question about directionality of transport. For other systems it has been shown that the C-terminus of the substrate reaches the cell surface first (Lenders et al., 2015, Smith et al., 2018a). All of these systems display a C-terminal secretion signal, as does HasA. The interaction of the secretion signal that leads to disassembly of the Has systems is also dependent on ATP hydrolysis by HasD (Masi and Wandersman, 2010). How this could be mediated when the C-terminal secretion signal is the first part to leave the translocon is not easily explained. Therefore, Alva *et al.* depict the secretion process of HasA as N- to C-terminus in their recent and very comprehensive review (Alav et al., 2021). Furthermore, there are multiple regions in HasA, apart from the secretion signal, that are able to interact with HasD and induce OMP recruitment (Masi and Wandersman, 2010). These regions were identified by random insertions into a secretion incompetent variant of HasA in combination with an assembly assay that was applied to the HlyA T1SS in this work (see 3.4). The same study further validated that OMP recruitment is independent of ATP hydrolysis by HasD but that ATP hydrolysis is necessary for disassembly of the system (Masi and Wandersman, 2010). Secretion of PrtG, one of the four proteases of the Prt T1SS from *D. dadantii*, is also dependent on ATP hydrolysis by PrtD although it is unclear during which step (Morgan et al., 2017). Similar to group 2, multiple events, including but not limited to ATP binding and hydrolysis, might energize the transport. Based on the comparison of the larger CyaA with the smaller HlyA, where CyaA secretion efficiency is dependent on

$[Ca^{2+}]_{outside}$ and HlyA secretion rate is independent of $[Ca^{2+}]_{outside}$ (Bumba et al., 2016, Lenders et al., 2016), the effect of Ca^{2+} induced folding might be even smaller for the smaller group 3 T1SS substrates that contain an RTX domain. The 'Brownian ratchet' mechanism is still reasonable for the non-RTX protein HasA, as modifications on either side of the transport (including de-chaperoning on the *cis* side) can bias Brownian motion towards a certain direction (Simon et al., 1992). Therefore, removal of SecB inside of the cell and rapid folding outside of the cell could drive secretion. An impact of the pmf in group 3 T1SSs has so far not been discussed in literature.

In summary, group 3 T1SSs are characterized by the presence of a classical ABC transporter without additional N-terminal domains (Kanonenberg et al., 2013). Substrates are of intermediate size compared to group 1 and 2 and are also transported unfolded (Debarbieux and Wandersman, 2001). The confirmed chaperon-dependency of HasA as well as the postulated N-directional transport might correlate with the absence of RTX motifs but more examples need to be evaluated.

1.2.2.4 BTLCP-linked T1SSs

Independent of the group (1-3), secretion mediated by T1SSs is always described as a one-step process without any periplasmic intermediates (Costa et al., 2015, Holland et al., 2016). This view is challenged by the largest substrates of T1SS, adhesins, for which LapA from *Pseudomonas fluorescens* Pf0-1 represents the model system (Smith et al., 2018b).

Many already introduced features also apply to this group of T1SSs: The secretion apparatus is made of three proteins (ABC/MFP/OMP) that together form one continuous channel across both membranes (Hinsa et al., 2003). The substrates contain C-terminal secretion signals of similar sizes (e.g. LapA: 67 residues), which are preceded by one or several C-terminal RTX domains that fold in the presence of Ca^{2+} due to the conserved RTX motifs (Boyd et al., 2014, Smith et al., 2018b). Again, the functional domains are very diverse reflecting the specific task of the protein and often contain many repeated domains. One drastic example of this is the 1.5 MDa ice binding protein (IBP) or ice binding adhesin (IBA) from *Marinomonas primoryensis*,

which contains 120 ± 10 repeats of a Ca^{2+}-stabilized immunoglobulin-like domain (Guo et al., 2017). Secretion of these giant adhesins is also C-directional and no chaperons have been identified for any RTX-containing adhesin (Smith et al., 2018a, Smith et al., 2018b). Interestingly, the ABC transporters also contain an N-terminal domain that might interact with the substrate. Function and structure of this domain are still unclear but phylogenetic analysis showed that they neither group with the CLDs of group 2 T1SSs nor the C39-peptidases of group 1 T1SSs but rather form their own group (Smith et al., 2018a).

The most striking difference of this group compared to group 1-3, is that secretion happens as a two-step process. After assembly of the system and initial secretion of a C-terminal part of the substrate, a unique N-terminal domain folds presumably in the periplasm and definitely independent of the RTX domain(s) (Boyd et al., 2014). This domain clogs the transport channel, which prevents further secretion of the substrate thereby anchoring it to the OM though the OMP pore and forming a pseudo-periplasmic intermediate during secretion (Guo et al., 2017). This domain was termed retention module (RM) and is conserved among BTLCP-linked T1SS adhesins (Smith et al., 2018a). The acronym describes the protein family of bacterial transglutaminase-like cysteine proteinases (BTLCP) that are involved in the secretion process by removing the RM and subsequent release of the adhesin (Ginalski et al., 2004, Newell et al., 2011). BTLCPs localize to the periplasm and their activity is controlled by an IM-bound c-di-GMP receptor that is encoded adjacent to the BTLCP (Newell et al., 2009, Newell et al., 2011).

Therefore, these giant adhesins present a novel subgroup of T1SSs that function via a two-step secretion process (Smith et al., 2018b).

1.2.3 HlyA T1SS

The hemolysin A (HlyA) T1SS from *E. coli* UTI89 is the hallmark of group 2 T1SSs and the main focus of this thesis. Secretion of HlyA from *E. coli* and its Ca^{2+}-dependent hemolytic activity were first described as early as 1951 but the transport components, HlyB and HlyD, were not identified until the late 1970s and mid 1980s

(Robinson, 1951, Noegel et al., 1979, Mackman et al., 1985). The last transporter component, TolC, was identified as the OMP of this system in 1990 (Wandersman and Delepelaire, 1990). The following sections will introduce each component of the HlyA T1SS in more detail.

1.2.3.1 HlyA

The substrate of the system, HlyA, is a pore-forming RTX toxin that is active against many cell types including erythrocytes and macrophages (Menestrina et al., 1996). It is synthesized as a 1024 amino acid long inactive pro-HlyA and activation requires acylation at K564 and K690 by its specific acyltransferase HlyC prior to secretion (Stanley et al., 1994). The secretion process is unaffected by the acylation so many studies are performed with the non-toxic, non-acylated version pro-HlyA (Nicaud et al., 1985a). Other variants that are commonly used are truncated versions that contain different domains of which HlyA has three: The functional N-terminal domain, the RTX domain and the C-terminal secretion signal. The truncated version HlyA1 contains the secretion signal and roughly half of the RTX domain (3 of 6 RTX motifs). HlyA2 is derived from HlyA1 by deletion of the secretion signal (Figure 1.14 A). Although many other versions of HlyA have been used in different studies, these two are especially important as they were used to map interaction sites with the transporter components (Benabdelhak et al., 2003, Lecher et al., 2012).
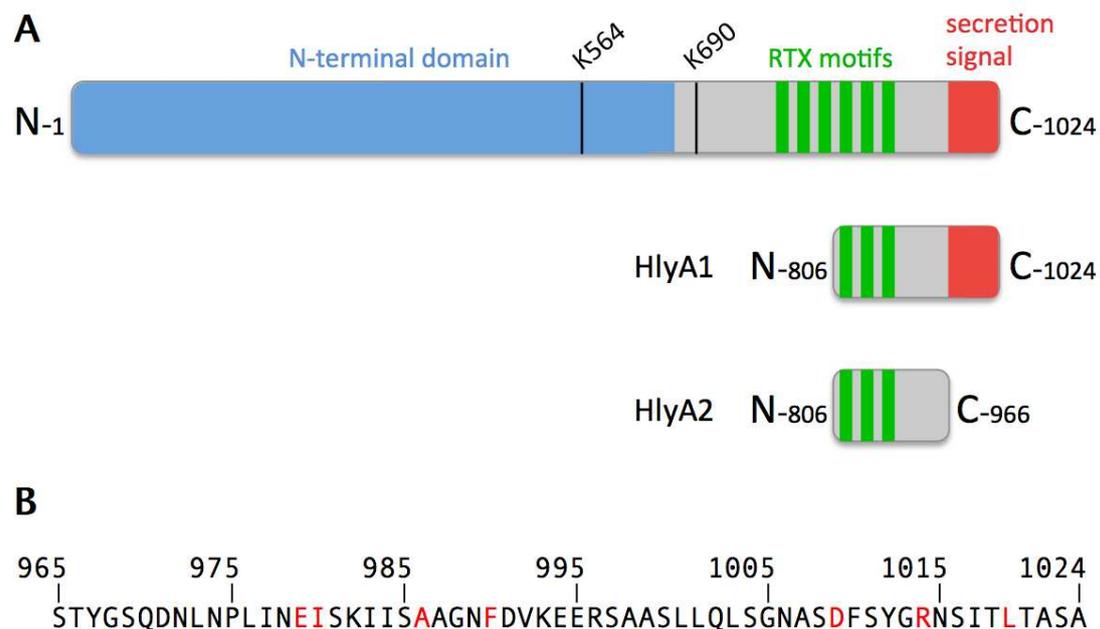
Figure 1.14 The substrate HlyA. A) Schematic representation of HlyA domains. The functional N-terminal domain is shown in blue, the RTX motifs in green and the secretion signal in red. N- and C-terminus are indicated. The two lysine residues that are acylated by HlyC are marked with black lines. The truncated construct HlyA1 contains three RTX motifs and the secretion signal. HlyA2 is derived from HlyA1 by deletion of the last 58 amino acids. **B)** Primary sequence of the secretion signal. Residues that reduced secretion in mutational studies are marked in red (Holland et al., 2016).

The functional domain of HlyA is able to form a pore with an estimated size of 30 Å that induces cell swelling and subsequent lysis (Bhakdi et al., 1986). While the N-terminal domain is required for pore formation, the RTX domain is also able to interact with a membrane (Sánchez-Magraner et al., 2007). However, there are also indications that HlyA can manipulate host cell signaling, for example by inhibition of the anti-apoptotic regulator Akt (Wiles et al., 2008). The interaction of a toxin with a host cell membrane is often dependent on specific receptors expressed by the host cell. For HlyA β2 integrin was shown to enhance HlyA activity but cells that do not present this receptor are still susceptible to HlyA (Ristow et al., 2019). Therefore, it is unclear which interactions initiate HlyA activity on a host cell.

The RTX domain holds six repeats of the name-giving RTX motif (GGxGxDxUx; x = any amino acid, U = large hydrophobic residue). They are able to bind Ca^{2+} with a K_D of 150 µM and Ca^{2+}-binding induces folding of the whole protein in which the RTX domain adopts a β-roll structure (Figure 1.12 C) (Baumann et al., 1993, Sánchez-Magraner et al., 2007). Furthermore, the RTX domain can interact with the CLD of

HlyB as shown by *in vitro* experiments with the isolated CLD and HlyA2. This interaction was not enhanced by the presence of the secretion signal (HlyA1) (Lecher et al., 2012).

The secretion signal was identified by N- and C-terminal deletions on HlyA. Deletion of the most C-terminal 27 residues completely abolished secretion (Gray et al., 1986). Deletions from the N-terminus identified the last 62 residues as the minimal secretory fragment (Jarchau et al., 1994). To promote efficient secretion of a fusion protein the secretion signal and part of the RTX domain are required (HlyA1) (Kenny et al., 1991). Nonetheless, the secretion signal of HlyA is mostly described consisting out of the last 50-60 residues (Figure 1.14) (Holland et al., 2016).

It is well accepted and has been shown *in vitro* that the secretion signal interacts with the transport components (Benabdelhak et al., 2003) and reaches the cell surface first (Lenders et al., 2015). However, it is unclear which features of the secretion signal promote this interaction. Sequence alignments with a wide variety of RTX toxins showed no apparent sequence motif and as shown by random mutagenesis many residues in the secretion signal are redundant (Kenny et al., 1994). There are, however, a few residues that significantly reduce HlyA secretion when mutated (Figure 1.14 B). Several other features of the linear sequence have been hypothesized to be important for HlyA secretion, namely a charged cluster (DVKEER) and an 'aspartate box' (residue 994-1009) but no functional role could be assigned to either of them (Zhang et al., 1993). There are also implications for a structural code and at least five different helices have been assigned to the secretion signal of HlyA, with some being amphipathic helices and some classical α-helices. Four of these five helices overlap with each other and with some of the residues that drastically reduced HlyA secretion when mutated (EISK) (Holland et al., 2016). Chapter 3 of this thesis focuses on this region and the presence of an amphipathic helix between residue 975 and 986.

1.2.3.2 HlyB

HlyB is the homodimeric ABC transporter of the HlyA T1SS from *E. coli* (Mackman et al., 1985). As already described in 1.2.2.2 HlyB has an additional N-terminal domain apart from the classical TMD and NBD. The N-terminal domain of HlyB resembles a C39-peptidase but is inactive and therefore termed C39-like domain (CLD) (Figure 1.12 B). The isolated CLD can be purified from *E. coli*, which allowed NMR and *in vitro* interaction studies. These studies showed that the CLD can interact with HlyA1 and HlyA2 (Figure 1.14 A) to the same extent, illustrating that the secretion signal is not needed for this interaction but rather the RTX domain (Lecher et al., 2012). Considering that an active C39-peptidase domain recognizes and cleaves after a double-glycine motif, the CLD might also interact with the glycine rich RTX motifs (Wu and Tai, 2004). Interestingly, the interaction site of HlyA is on the opposite site compared to a C39-peptidase (Lecher et al., 2012). The CLD is believed to take on the function of a chaperon for the large HlyA for multiple reasons: i) HlyA is transported unfolded and therefore cannot adopt a fold that would protect it from cytoplasmic proteases (Bakkes et al., 2010b). ii) Transport is C-directional and therefore cannot take place co-translational (Lenders et al., 2015). iii) There is no chaperon present in the *hly* operon (*hlyCABD*) (Thomas et al., 2014). iv) HlyA secretion does not depend on the general chaperon SecB (Bakkes et al., 2010a). v) The interaction of the CLD with HlyA depended on the unfolded state of the substrate (Lecher et al., 2012).

Similar to the CLD, the isolated NBD can also be purified from *E. coli*, which also allowed structural determination and *in vitro* interaction studies (Figure 1.15). Opposite to the CLD, the NBD can interact with HlyA1 but not with HlyA2, illustrating that the secretion signal interacts with the NBD (Benabdelhak et al., 2003). Interestingly, this interaction is disrupted in the presence of ATP (Benabdelhak et al., 2003) and ATP binding induces dimerization as seen from the crystal structures (Zaitseva et al., 2005). ATP is coordinated by motifs that are conserved among and define the NBD family (Schmitt and Tampé, 2002). The NBDs of HlyB are highly symmetric and therefore present two ATP binding sites. The

Walker A (residue G502-T510), Walker B motif (residue I626-E631) and the Q-loop (residue Q550-S558) of one monomer coordinate ATP together with the C-loop (residue G605-Q610) of the second monomer and *vice versa*. The D-loop (residue S634-D637) is involved in forming the dimer interface and the histidine (H662) of the H-loop (residue T656-R663), which was first believed to act as a γ-phosphate sensor, is actually involved in ATP hydrolysis and might be involved in monomer-monomer communication (Figure 1.15) (Zaitseva et al., 2005).
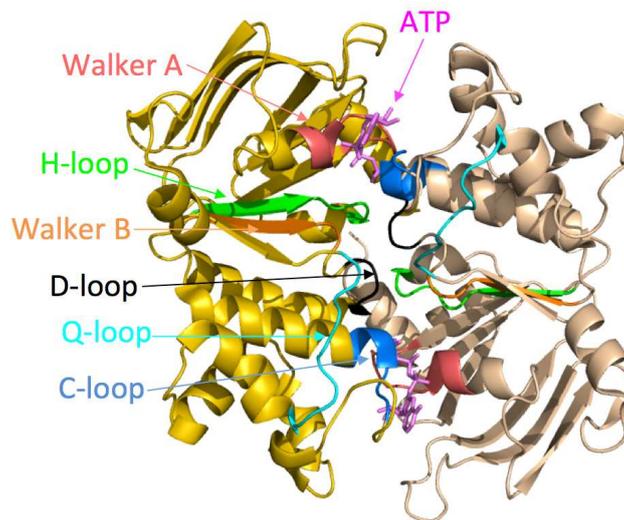
Figure 1.15 ATP bound dimer of HlyB NBD H662A from *E. coli* (PDB 2FGJ). One monomer is shown in yellow and the second monomer in beige. The conserved motifs are colored in both monomers but only labeled in one. The Walker A motif (red), the Walker B motif (orange) and the Q-loop (cyan) of one monomer coordinate ATP (pink sticks) together with the C-loop (blue) of the second monomer. The D-loop (black) is involved in forming the dimer interface. The H-loop (green) is necessary for ATP hydrolysis. The figure was drawn with PyMOL and inspiration taken from (Zaitseva et al., 2005).

In contrast to the CLD and NBD there is no structure available for the TMD. However, there are some indications that HlyA is transported through a hydrophilic cavity provided by the TMDs. In a cross-link study between detergent-solubilized HlyB and HlyA2, K322 of HlyB and N954 of HlyA were identified. The substitution to a bulky residue (K322W) drastically reduced secretion while K322A did not affect secretion (Reimann et al., 2016). When modeling the HlyB TMDs on the structure of PCAT-1, K322 points inward, suggesting that HlyA is transported directly through the TMDs (Reimann, 2017). Furthermore, the group 3 T1SS ABC transporter AaPrtD (see 1.2.2.3) shows a hydrophilic cavity in the TMDs (Figure 1.13 B) and residues that are

involved in forming and stabilizing this cavity are also present in HlyB (Morgan et al., 2017).

The CLD and NBD of HlyB can be isolated and interact with HlyA *in vitro* (Benabdelhak et al., 2003, Lecher et al., 2012). Also, HlyB can be purified with detergents or reconstituted in lipid-nanoparticles and ATPase activity is stimulated by addition of HlyA, showing that full-length HlyB and HlyA can interact independently of the other transporter components (Reimann et al., 2016, Kanonenberg et al., 2019). However, secretion only takes place when all transport components are present (Thanabalu et al., 1998). In addition to that, the complex of HlyB and HlyD cannot secrete from spheroblasts, which explains why no periplasmic intermediate of HlyA has been found (Koronakis et al., 1991).

1.2.3.3 HlyD

HlyD is the MFP of the HlyA T1SS from *E. coli* and therefore connects the ABC transporter HlyB with the OMP TolC (Mackman et al., 1985). Its N-terminus forms a small cytoplasmic domain (CD) of 60 residues, followed by a TM helix of 20 residues (60-80) and a large periplasmic domain (PPD) comprised of the remaining 398 residues (Schülein et al., 1994, Pimenta et al., 1999). Together these domains form a protein of 478 residues with a size of 55 kDa. The CD was shown to be essential for the secretion of HlyA and two features of it were identified based on sequence alignments: an amphipathic helix (AH) (residue 2-26) and a charged cluster (residue 34-38). Removal of the first 26 residues, which form the AH, drastically reduced pro-HlyA secretion to about 5-10 % compared to wild type HlyD. Removal of the first 45 residues and also removal of residue 26-45 completely abolished pro-HlyA secretion (Balakrishnan et al., 2001). The CD is able to directly interact with (pro-)HlyA and this interaction is required for TolC recruitment (Thanabalu et al., 1998, Balakrishnan et al., 2001). However, TolC recruitment by HlyD only takes place in the presence of HlyB and HlyA (Thanabalu et al., 1998).

In general, most MFPs display four domains in their PPD: an α-helical domain that contacts the OMP, a lipoyl domain that facilitates oligomerization, a β-barrel and

membrane proximal (MP) domain (Figure 1.16 A) that together contact the second IM component (Alav et al., 2021). In 2016 Kim *et al.* published a crystal structure of HlyD residue 96-372, which corresponds to the PPD without the first 16 and without the C-terminal ~100 residues (Figure 1.16 A). They identified a long α-helical domain (115 Å) displaying the typical helix-turn-helix (HTH) motif and a lipoyl domain (Kim et al., 2016). RND-type associated MFPs typically show shorter α-helical domains (e.g. AcrA: 60 Å) (Mikolosko et al., 2006). However, AaEmrA, an MFS-type associated MFP, also shows an elongated α-helical domain of 127 Å and a lipoyl and β-barrel domain, but no MP domain (Hinchliffe et al., 2014). Therefore Kim *et al.* suggest that HlyD also misses the MP domain and that the last 100 residues that are missing in the crystal structure form the β-barrel domain (Kim et al., 2016).

The tip-to-tip contact of AcrA and TolC is mediated by the RLS motif (RxxxLxxxxxS; x = any amino acid) (Figure 1.16 A) of which leucine is the most conserved residue (Lee et al., 2012). Alignments of the HTH region identified a DLA motif in HlyD instead of the RLS motif and D239 and L243 were subjected to mutational studies combined with cross-link experiments (Figure 1.16 A). This approach confirmed that these residues are involved in TolC contact suggesting that HlyD and AcrA interact with TolC in the same manner (Kim et al., 2016).

The oligomeric state of HlyD in the assembled T1SS is still under debate. So far only trimers have been isolated with the help of cross-linkers but a hexamer is often suggested, since it would be the smallest common denominator between the dimeric HlyB and trimeric TolC (Thanabalu et al., 1998, Balakrishnan et al., 2001). Furthermore, all structures of assembled tripartite efflux pumps show a hexameric organization of the MFP, which is facilitated by the lipoyl domains (Alav et al., 2021). The lipoyl domain of AcrA and HlyD align well with an RMSD (root mean square deviation) of 2.15 Å over 48 C_α atoms (Kim et al., 2016). This prompted multiple research groups to model a hexameric organization of HlyD (Figure 1.16 B), which can be docked to TolC in a tip-to-tip manner (Kim et al., 2016, Jo et al., 2019, Alav et al., 2021). Also the lipoyl domains arrange into a ring with an inner diameter of 20 Å, which would allow for an unfolded substrate to pass through (Kim et al., 2016).
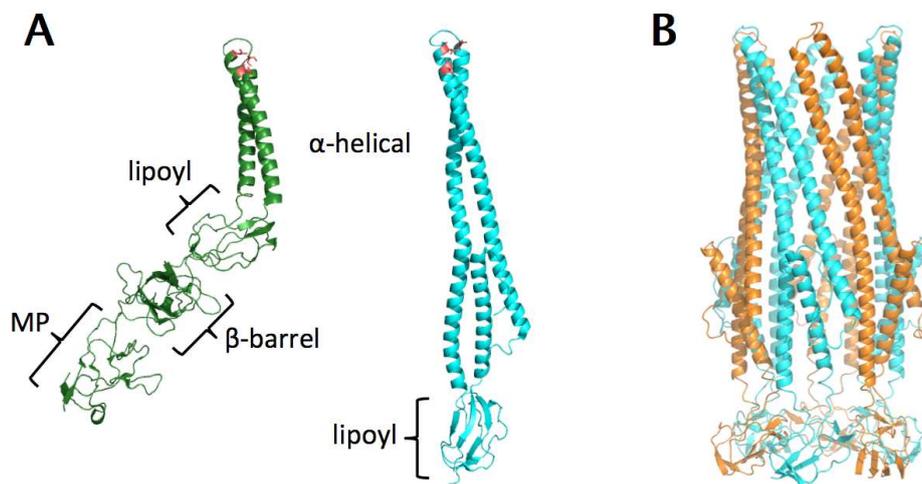
Figure 1.16 A) Cartoon representation of AcrA (green, PDB 5V78) and HlyD residue 96-372 (cyan, PDB 5C21). AcrA has four domains: an α-helical domain, a lipoyl domain, a β-barrel domain and a membrane proximal (MP) domain. In the partial structure of HlyD only an α-helical domain and a lipoyl domain were present. The RLS motif of AcrA and the DLA motif of HlyD that facilitate the contact to TolC are shown as red sticks. **B)** Hexameric model of HlyD (adapted from (Kim et al., 2016). The HlyD monomers are shown alternating in cyan and orange.

The components of the IM, HlyB and HlyD, form a complex in the IM even in the absence of the substrate (Thanabalu et al., 1998). In other tripartite efflux pumps periplasmic domains mediate this contact (Wang et al., 2017b, Tsutsumi et al., 2019). However, HlyB does not have such large periplasmic domains (Gentschev and Goebel, 1992) and HlyD most likely lacks the MP domain (Kim et al., 2016). Their contact could be mediated by TM helices in the IM or even in the cytoplasm, since HlyD and HlyB each display domains in the cytoplasm, which AcrA and AcrB do not.

1.2.3.4 TolC

TolC is an OMP of *E. coli*, which facilitates the efflux of many unrelated compounds. It can form an export pore for periplasmic substrates (e.g. MccJ25, see 1.2.2.1) or engage in a tripartite assembly (Wandersman and Delepelaire, 1990, Delgado et al., 1999). It is synthesized as a 493 amino acid precursor and transported to the periplasm by the Sec pathway (see 1.1.1) under removal of its leader peptide (residue 1-22) (Masi et al., 2009, Zgurskaya et al., 2011).

In 2000 the crystal structure of TolC was published and revealed a novel structure for an OM β-barrel protein (Koronakis et al., 2000). Other trimeric OMPs,

like OmpF, form one pore with each monomer (Yamashita et al., 2008). However, each monomer of TolC only contributes to a third of the single OM pore. TolC measures 140 Å along the long axis with 40 Å corresponding to the β-barrel in the OM with an inner diameter of 35 Å. An α-helical barrel extends 100 Å into the periplasm with each monomer displaying two HTH motifs. The equatorial domain of TolC consists of a mixture of α-helices and β-strands that wrap around the mid-section of the α-helical barrel (Figure 1.10 A). The α-helices are coiled and twisted in a way that closes the TolC channel towards the periplasm and an iris-like opening is induced by interaction with another protein (Koronakis et al., 2000).

Most other OMPs are assembled into the OM by the BAM complex after being targeted to it by a chaperon like SurA or Skp (see 1.1.1) (Kim et al., 2012). However, TolC assembly into the OM neither requires SurA nor Skp and deep-rough LPS (lipopolysaccharide) mutations also had no influence on TolC assembly (Werner et al., 2003). After TolC is transported to the periplasm by the Sec translocon it adopts an insertion-competent form that is proteinase K sensitive (Werner et al., 2003) but likely already a trimer for two reasons: TolC without its leader peptide trimerizes and folds in the cytoplasm (Masi et al., 2009) and the monomer interfaces display large hydrophilic surfaces that would by energetically unfavoured when exposed to the hydrophobic lipid bilayer during insertion of one monomer (Koronakis et al., 2000, Werner et al., 2003). Thus, after trimerization in the periplasm, TolC is inserted into the OM with the help of BamA and YaeT (Werner and Misra, 2005, Bennion et al., 2010); both are members of the Omp85 family that is involved in OMP biogenesis (Gentle et al., 2005, Kim et al., 2012). Once in the OM, TolC is mostly protected from proteinase K digestion as only a carboxy-terminal fragment of ~5 kDa can be cleaved from it, leaving a 46 kDa fragment in the OM (Koronakis et al., 1997, Werner et al., 2003). Interestingly, assembly into the HasA T1SS converts TolC into a proteinase K sensitive form again, that can be fully degraded (Masi and Wandersman, 2010).

1.2.3.5 Secretion process

Secretion of HlyA from *E. coli* is coupled to the bacterial growth cycle taking place during early and mid logarithmic phase and is switched off before the stationary phase (Nicaud et al., 1985b, Koronakis et al., 1989). In genetically modified *E. coli* strains, such as *E. coli* BL21 (DE3), expressing HlyBD and HlyA from inducible plasmids can circumvent this confinement. The OMP TolC is constitutively expressed at 4472 ± 1231 trimers per cell (in *E. coli* BL21 (DE3) as well as UTI89) and overexpression of the HlyA T1SS components does not influence the TolC copy number (Beer, 2020).

The IM proteins HlyB and HlyD are most likely inserted into the IM by the co-translational SRP dependent pathway of the Sec translocon (see 1.1.1) (Denks et al., 2014), where they form a IM complex even in the absence of HlyA (Thanabalu et al., 1998). Engagement of the IM complex with unfolded HlyA triggers TolC recruitment (Figure 1.17) (Thanabalu et al., 1998). The exact interactions are unknown but involve all cytoplasmic domains of the IM complex (CLD and NBD of HlyB, CD of HlyD), as removal or inactivation of only one of them abolishes secretion (Thanabalu et al., 1998, Balakrishnan et al., 2001, Lecher et al., 2012). One the side of HlyA, at least the last 62 residues are necessary but the additional presence of at least part of the RTX domain (HlyA1) enhances secretion compared to the 62 amino acid fragment (Jarchau et al., 1994). *In vitro* experiments showed that the secretion signal interacts with HlyB NBD and the RTX domain with HlyB CLD (Benabdelhak et al., 2003, Lecher et al., 2012). HlyD and HlyA were also shown to interact with each other *in vivo* in the absence of HlyB (Thanabalu et al., 1998). So far it is unclear if this direct interaction also takes place in the presence of HlyB during the secretion process.

After TolC is recruited by HlyD into the transport channel, secretion takes place at a rate of 16 amino acids per second per transporter with the C-terminus of HlyA reaching the cell surface first (Lenders et al., 2015, Lenders et al., 2016). As soon as the RTX domain reaches the outside, it will bind Ca^{2+} ions, which induce folding of the protein thereby preventing backsliding into the TolC pore (Bakkes et al., 2010b).

Secretion of full length HlyA takes about 1 min and is ATP dependent at least in later steps of secretion (Koronakis et al., 1991, Zaitseva et al., 2005, Lenders et al., 2016). The IM complex and TolC disengage after the secretion event, which finishes this one-step transport mechanism (Figure 1.17) (Thanabalu et al., 1998).
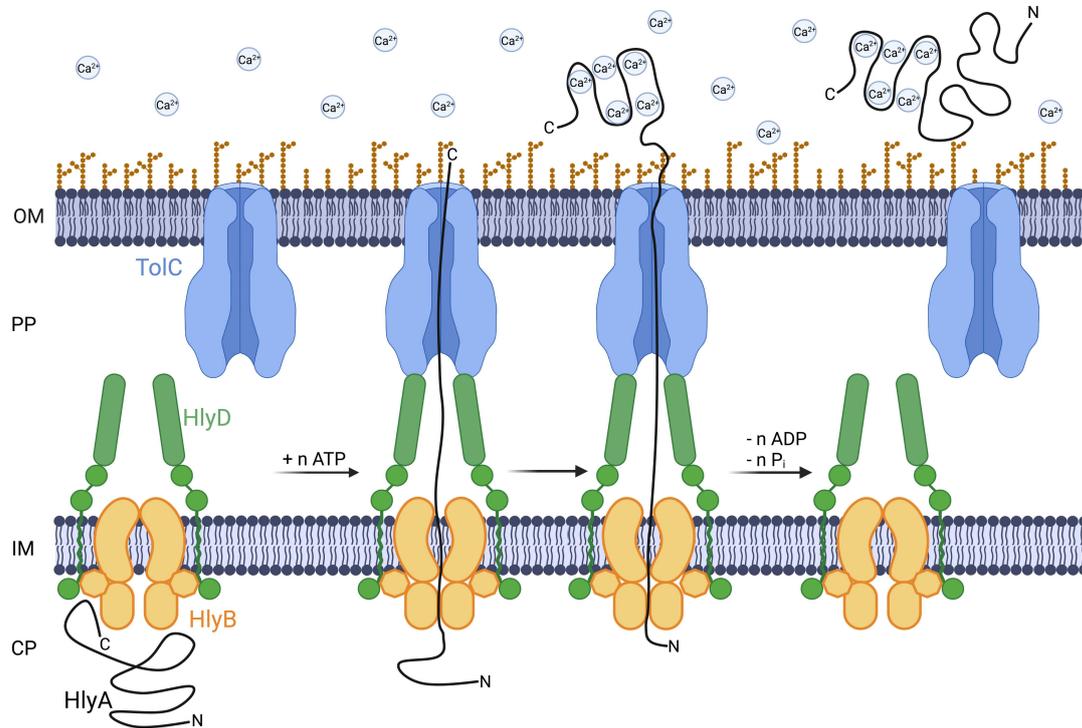
Figure 1.17 Secretion of HlyA by its dedicated T1SS. HlyB (orange) and HlyD (green) form a stable complex in the IM (HlyBD). Only two HlyD are depicted for simplicity. Engagement of the unfolded HlyA (black line) with the cytoplasmic domains of HlyBD triggers TolC (blue) recruitment. Secretion is C-directional and ATP dependent. The RTX domain of HlyA will bind Ca^{2+} (grey spheres) once it reaches the extracellular space and start folding. After secretion of the complete HlyA, it will adopt its mature form and HlyBD and TolC disengage. (Created with BioRender.com).

2.   Aim

Gram-negative bacteria have evolved many strategies to communicate with their environment as reflected by the variety of secretion systems. Depending on the context it can be desirable to manipulate these systems either negatively, for example during an infection, or positively for biotechnological approaches. Both aims profit from a deep functional understanding of the respective system.

The main focus of this work was placed on the HlyA T1SS from *E. coli*, as T1SSs are wide spread, hold a simple architecture and the HlyA T1SS represents the most investigated member of this group. The overall aim of this work was to gain a deeper understanding of the secretion mechanism; to achieve this two structural aims and one functional aim were set.

The structure of the OMP TolC was already solved in the open and closed state (Koronakis et al., 2000, Bavro et al., 2008) and the NBDs of HlyB have been captured in many different intermediate steps of ATP hydrolysis (Schmitt et al., 2003, Zaitseva et al., 2005, Zaitseva et al., 2006). The CLD of HlyB has also been investigated as an isolated domain (Lecher et al., 2012), leaving the TMDs of HlyB as the last missing puzzle piece of the ABC transporter. However, solving only the structure of the TMDs would not uncover the organization of the domains towards each other. For this the structure of the whole protein is needed. Unfortunately, HlyB shows a tendency to aggregation during crystallization trials, which is why a homology approach was initiated during this work. The aim was to identify at least one homolog of HlyB that can be expressed, purified and crystallized (chapter 3.5). The second structural aim was directed towards HlyD. Data on the cytosolic domain (CD) of HlyD is scarce but this domain is essential to the secretion process (Balakrishnan et al., 2001). Its small size of 60 amino acids makes it a perfect candidate for NMR studies so that purification trials of this domain were performed during this work (see section 4).

The functional aim was directed towards the assembly of the HlyA T1SS. As described in 1.2.3.5 it is unclear which domain interactions lead to the recruitment of TolC. For the HasA T1SS this question was addressed with a proteinase K susceptibility assay, which was transferred to the HlyA T1SS during this work (chapter 3.4) (Masi and Wandersman, 2010).

3. Publications

3.1 Chapter 1 – Type I secretion system – it takes three and a substrate

Title: Type I secretion system – it takes three and a substrate

Authors: Kerstin Kanonenberg, **Olivia Spitz**, Isabelle N. Erenburg, Tobias Beer and Lutz Schmitt

Published in: *FEMS Microbiology Letters* (2018)

Impact Factor 1.987 (2019)

Own Work: 15 %

Writing of the manuscript

OXFORD

FEMS JOURNALS investing in science

MINIREVIEW – Physiology & Biochemistry

Type I secretion system—it takes three and a substrate

Kerstin Kanonenberg, Olivia Spitz, Isabelle N. Erenburg, Tobias Beer and Lutz Schmitt*

Institute of Biochemistry, Heinrich Heine University, 40225 Düsseldorf, Germany

*Corresponding author: Institute of Biochemistry, Heinrich hein University, Universitätsstr 1,40225 Düsseldorf, Germany.
Tel: +49-211-81-10773; Fax: +49-211-81-15310; E-mail: lutz.schmitt@hhu.de
One sentence summary: An overview of type I secretion systems of Gram-negative bacteria and a summary of the recent developments is provided.
Editor: Lily Karamanou

ABSTRACT

Type I secretion systems are widespread in Gram-negative bacteria and mediate the one-step translocation of a large variety of proteins serving for diverse purposes, including nutrient acquisition or bacterial virulence. Common to most substrates of type I secretion systems is the presence of a C-terminal secretion sequence that is not cleaved during or after translocation. Furthermore, these protein secretion nanomachineries are always composed of an ABC transporter, a membrane fusion protein, both located in the inner bacterial membrane, and a protein of the outer membrane. These three membrane proteins transiently form a 'tunnel channel' across the periplasmic space in the presence of the substrate. Here we summarize the recent findings with respect to structure, function and application of type I secretion systems.

Keywords: protein secretion; ABC transporter; secretion sequence; RTX toxin

INTRODUCTION

Bacteria have a need for secreting a variety of proteins and other molecules to the extracellular space, for nutrient acquisition (e.g. iron-scavenger proteins), biofilm formation (adhesins) or host invasion (virulence factors, e.g. exotoxins).

Secretory pathways have been of major research interest over the past decades and depending on the definition applied, a minimum of 15 different secretion systems has been identified so far in Gram-negative bacteria (reviewed in Costa et al. 2015). Here, the outer membrane imposes an additional problem as secreted macromolecules have to cross a second, the outer membrane. These secretion systems are capable of exporting a diverse range of small molecules, DNA and proteins to the extracellular space or even directly into the cytosol of a target cell. They vary greatly in composition and molecular mechanism, but can be easily divided into two major subgroups based on the presence or absence of a periplasmic transport intermediate during the secretion process.

Type I, III and IV secretion systems are double-membrane-spanning export machineries where the substrate is secreted in one step from the cytosol to the extracellular space (type I). The latter two are even capable of delivering their substrate directly into the cytosol of the target cell, thus traversing three membranes (Fig. 1). Obviously, all these secretion systems require a 'tunnel channel'-like architecture, composed of a minimum of 3 but up to more than 10 membrane-localized proteins (Fig. 1). For further information, the reader is referred to Economou and Dalbey (2014) and Costa et al. (2015) for a review series covering the details of most bacterial secretion systems.

This review highlights the recent advances in research concerning specifically type I secretion systems (T1SS), setting the focus mainly on new structural insights that have been obtained over the last years. T1SS are often referred to as the most 'simple' representative considering that they are composed of only three membrane proteins (also see Delepelaire 2004; Kanonenberg, Schwarz and Schmitt 2013; Thomas, Holland and Schmitt 2014; Holland et al. 2016 for various aspects of T1SS).

Many Gram-negative pathogens make use of T1SS to secrete a great variety of virulence factors. The discovery of the first T1SS substrate dates back to as far as 1979 when the Goebel
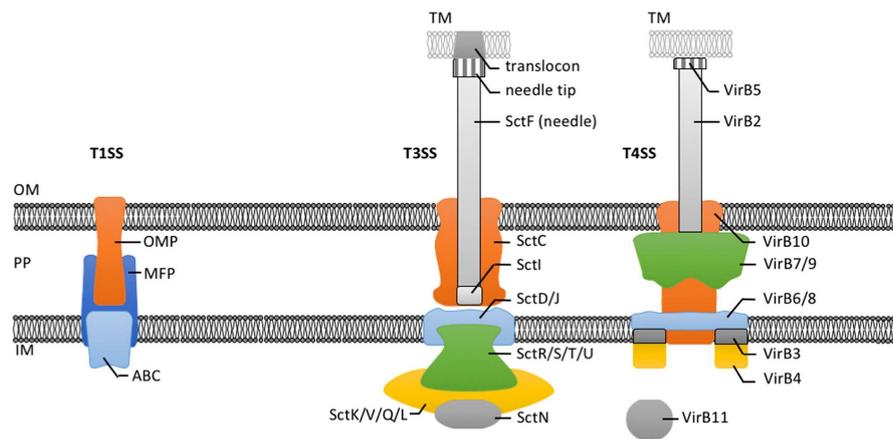
1

Figure 1. Cartoon of secretion systems from Gram-negative bacteria that translocate their substrates in one step across two (T1SS) or three membranes (T3SS and T4SS). OM: outer membrane, PP: periplasm, IM: inner membrane, OMP: outer membrane protein, MFP: membrane fusion protein, ABC: ABC transporter. Proteins forming the T3SS and T4SS and their putative location are indicated.

laboratory identified hemolysin A (HlyA), named after its ability to lyse erythrocytes from uropathogenic *Escherichia coli* strains (Noegel *et al.* 1979). Subsequently, the nucleotide sequence of HlyA was determined (Felmlee, Pellett and Welch 1985). Additional studies from the laboratories of Koronakis, Holland and Goebel (Mackman and Holland 1984; Mackman *et al.* 1985a,b, 1987; Gray *et al.* 1986, 1989; Felmlee and Welch 1988; Koronakis, Koronakis and Hughes 1989; Gentschev, Hess and Goebel 1990; Chervaux *et al.* 1995) demonstrated that secretion of HlyA occurred without any periplasmic intermediate and was Sec independent. Moreover, HlyA carried a C-terminal secretion signal indicating an unknown secretion mechanism.

The analysis of the sequence of the *hly* operon (Felmlee, Pellett and Welch 1985) revealed the presence of two additional membrane proteins and a third component in addition to the substrate HlyA. The third component, HlyC, turned out to be essential for the activation of HlyA, but not for secretion per se (Nicaud *et al.* 1985). HlyC was shown to act as a cytosolic acyltransferase acylating two internal lysine residues of the unfolded HlyA prior to secretion. This required equimolar amounts of the acyl carrier protein (Issartel, Koronakis and Hughes 1991; Stanley, Koronakis and Hughes 1991; Stanley *et al.* 1994, 1999; Thomas, Smits and Schmitt 2014). Only recently, the crystal structure of an HlyC homolog was reported (Greene *et al.* 2015) that will open up new approaches to understand the function of this unusual acyltransferase at the molecular level.

HlyD, one of the membrane proteins encoded by the *hly* operon, belongs to the family of bacterial membrane fusion proteins (MFPs) (Symmons, Marshall and Bavro 2015) that is unique to Gram-negative bacteria. The second membrane protein, HlyB, is a member of the ABC transporter family (Davidson *et al.* 2008), which is found in all kingdoms of life. Since HlyB and HlyD were localized to the inner membrane (Mackman *et al.* 1985a,b; Wang *et al.* 1991; Pimenta *et al.* 1999), the lack of periplasmic intermediates raised an obvious question—How does HlyA reach the extracellular space? This issue was addressed by the Wandersman group, who identified TolC, a ubiquitous and polyvalent outer membrane protein, as the missing, third component of the HlyA-T1SS (Wandersman and Delepelaire 1990). A complex of the two inner membrane proteins and TolC form the 'tunnel

channel' that allows the one-step secretion of HlyA. Apart from T1SS, TolC is involved in the extrusion of toxic components (Koronakis, Eswaran and Hughes 2004), for example by being part of tripartite drug efflux systems such as the AcrA-AcrB/TolC complex (Du *et al.* 2014).

In other T1SS, more than one transport substrate (Letoffe, Delepelaire and Wandersman 1990) or the TolC homolog (Letoffe, Ghigo and Wandersman 1994) can be encoded in the operon. Thus, there are no strict requirements on the genetic level for the operon organization of T1SS, but several lines of evidence suggest that a minimal unit composed of the gene coding for the transport substrate, the ABC transporter and the MFP is present in all operons. In addition, some degree of promiscuity with respect to the transported substrate exists, since the Hly system of *E. coli* was successfully used to secrete, for example, CyaA from *Bordatella pertussis* (Masure *et al.* 1990; Sebo and Ladant 1993), FrpA from *Neisseria meningitis* (Thompson and Sparling 1993) or PaxA from *Pasteurella aerogenes* (Kuhnert *et al.* 2000).

T1SS SUBSTRATES

For the vast majority of T1SS substrates, all the information necessary and sufficient for secretion is encoded in the extreme C-terminus, which is not cleaved during or after translocation. This was recognized early on (Gray *et al.* 1986) and was one of the first indications that HlyA was secreted independently of the Sec system. However, a small group of substrates (class II microcins) contain an N-terminal propeptide, which is cleaved by a C39 peptidase domain on the ABC transporter prior to secretion (Hwang, Zhong and Tai 1997).

The actual secretion signal of the Hly system was shown to be confined to the last 50 to 60 most extreme C-terminal amino acids (Koronakis, Koronakis and Hughes 1989) but its size and nature varies from system to system. The reader is referred to a review (Holland *et al.* 2016), which summarizes our current knowledge on T1SS secretion signals, still leaving many unanswered questions that need to be addressed in future research.

Upstream to the secretion sequence of HlyA, aspartate and glycine-rich nonapeptide repeats were identified (Welch 1991).
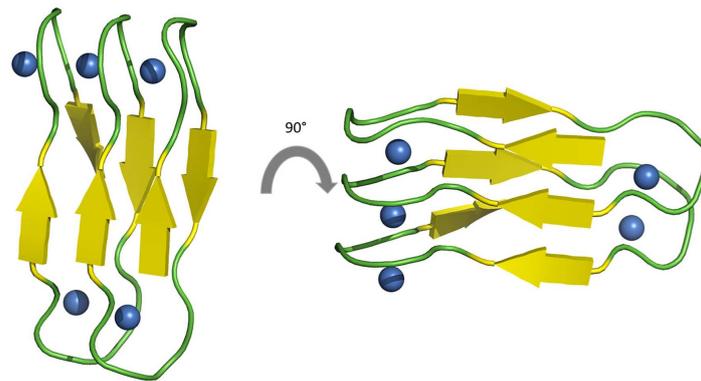
Figure 2. A zoom into the β-roll motif of an alkaline protease (Baumann *et al.* 1993). Ca^{2+} ions and the Ca^{2+} binding region are shown as blue spheres and in cartoon representation, respectively.

These have the consensus sequence GGXGXDXUX (where X can be any amino acid and U is a large hydrophobic amino acid) and the term 'GG repeats' or 'repeats-in-toxins' (RTX) was coined by Rod Welch. These repeats form the hallmark of an entire family of proteins including lipases, proteases, adhesins, S-layer proteins or toxins (reviewed by Linhartova *et al.* 2010).

Structural studies on the T1SS-secreted alkaline protease from *Pseudomonas aeruginosa* (Baumann *et al.* 1993) and other substrates (Baumann *et al.* 1993; Izadi-Pruneyre *et al.* 1999; Meier *et al.* 2007; Griessl *et al.* 2013) revealed that the coordination of one Ca^{2+} by two RTX motifs via the side chains of the aspartate residues and the backbone of the first two glycine residues creates a so-called β-roll or β-helix motif (Fig. 2).

RTX substrates of T1SS bind Ca^{2+} ions in the high micromolar range, for example ~500 μM for CyaA from *Bordetella pertussis* (Chenal *et al.* 2009; Sotomayor Perez *et al.* 2015) or ~150 μM for HlyA from *E. coli* (Ostolaza, Soloaga and Goni 1995; Sanchez-Magraner *et al.* 2007; Thomas *et al.* 2014). This binding induces folding of the entire RTX protein. As the concentration of free Ca^{2+} ions in the bacterial cytosol is in the high nanomolar range (Jones *et al.* 1999), RTX proteins remain unfolded until they reach the extracellular space, where Ca^{2+} concentrations of up to 10 mM result in immediate binding and protein folding.

The N-terminal moiety of T1SS substrates encodes for functionality, i.e. lipolytic, hemolytic, proteolytic, adhesive or any other activity. A recent data mining approach of 840 bacterial genome sequences (Linhartova *et al.* 2010) identified ~1000 RTX proteins, being extremely variable in size (up to 900 kDa, Hinsa *et al.* 2003) and function, but conforming to the general arrangement of T1SS substrates, functional domain/RTX domain/secretion sequence. The number of RTX domains in an individual RTX-protein scales to some extent with the molecular weight (Linhartova *et al.* 2010), but the presence of these characteristic motifs is ubiquitous and highlights their functional importance.

The iron siderophore HasA from *Serratia marcescens* represents an exception (Letoffe, Ghigo and Wanderman 1994). With a size of 19 kDa it is the smallest substrate of a T1SS identified so far and interestingly lacks the entire RTX domain, but contains a C-terminal secretion sequence (Izadi-Pruneyre *et al.* 1999). Interestingly, it is the only T1SS substrate known which requires a chaperone, SecB, for secretion (Sapriel, Wanderman and Delepelaire 2002; Bakkes *et al.* 2010).

FUNCTIONAL INSIGHTS

Early on, Koronakis and coworkers (Thanabalu *et al.* 1998; Balakrishnan, Hughes and Koronakis 2001) demonstrated for the HlyA T1SS that upon interaction of the substrate with the inner membrane proteins (HlyB and HlyD), TolC is recruited and a 'channel tunnel' is formed through which HlyA is secreted at the cost of ATP hydrolysis. After substrate translocation is completed, TolC disassembles from the complex, leaving HlyB and HlyD as a stable complex in the inner membrane, ready to start another round of substrate secretion.

Deletion studies proved that the cytosolic domain of HlyD (residues 1–60) forms the hub from which assembly of the secretion complex is initiated (Balakrishnan, Hughes and Koronakis 2001). Complementary data were provided by *in vitro* surface plasmon resonance experiments demonstrating that the isolated nucleotide-binding domain (NBD) of the ABC transporter also interacts with the substrate (Benabdelhak *et al.* 2003). Interestingly, this interaction was strictly limited to the C-terminal secretion signal.

Indirect (Kenny, Haigh and Holland 1991; Debarbieux and Wanderman 2001) and direct evidence (Bakkes *et al.* 2010) demonstrated that substrates of T1SS are translocated in an unfolded state. In an elegant set of experiments, Wanderman and colleagues observed that the presence of folded HasA actually inhibited the secretion of newly synthesized HasA (Debarbieux and Wanderman 2001). Subsequently, they addressed the underlying principles of this cis inhibition (Cescau, Debarbieux and Wanderman 2007) and surprisingly, the results of this study demonstrated that the interaction of unfolded HasA with the inner membrane complex also occurs outside the region encoding the secretion sequence, identifying a second, non-overlapping binding site. This interaction resulted in stable recruitment of the outer membrane protein TolC, which could be reversed by adding in cis the isolated secretion sequence. This pointed toward an intermolecular activity that triggered complex dissociation (Cescau, Debarbieux and Wanderman 2007).

All structural and functional data obtained for ABC transporters so far indicate that the transport mechanism used by these primary active transporters to shuttle their substrates from one side of the membrane to the other follows the 'alternating two site access model' for membrane transporters (Jardetzky 1966). However, the unfolded state and the mere

physical length (up to 9000 amino acids) of substrates of T1SS make it impossible to apply this generally accepted mechanism also for ABC transporters involved in T1SS. Although some research has been carried out on this issue, there is still very little understanding of the mechanism of secretion through the transenvelope channel.

In contrast to the canonical organization of ABC transporters, HlyB harbors an additional N-terminal domain, a cytosolic appendix (Kanonenberg, Schwarz and Schmitt 2013). Based on the primary structure of HlyB, the first ~130 amino acids belong to the family of C39 peptidases, a subfamily of the papain superfamily of cysteine proteases (Havarstein, Diep and Nes 1995; Wu and Tai 2004). These peptidases are unique to ABC transporters and are only found in bacteriocin exporters (Havarstein, Diep and Nes 1995). In principle, the protein family of bacteriocins is limited to Gram-positive bacteria, but a few members can also be found in Gram-negative strains, e.g. Colicin V in *E. coli*. Using a type I secretion apparatus but retaining the typical N-terminally cleaved propeptide, these peptides form a small yet unique group amongst type I substrates (Hwang, Zhong and Tai 1997).

However, in many T1SS ABC transporters such as HlyB the catalytically active cysteine residue is replaced by a tyrosine, resulting in a corrupted catalytic triad. Lecher *et al.* (2012) therefore established the term 'C39 peptidase-like domain' (CLD). NMR studies revealed an identical tertiary structure compared to C39 peptidase domains (Ishii *et al.* 2010). A conserved interaction of the histidine residue in the corrupted active center with a tryptophan residue was discovered, which is now commonly used to distinguish C39 peptidase domains from CLDs (Lecher *et al.* 2012; Kanonenberg, Schwarz and Schmitt 2013). In a set of *in vitro* functional and structural studies, Lecher *et al.* (2012) confirmed binding of unfolded substrate to the isolated CLD that was independent of the secretion signal. The substrate-binding site was mapped by chemical perturbation experiments and results were subsequently confirmed by mutational studies. These results suggest that the CLD acts as a receptor that grabs unfolded HlyA and positions it for subsequent translocation. It remains speculative whether the CLD also plays a role in preventing folding and degradation of the substrate in the cytosol and further studies are needed to establish its precise function and mechanism of action.

Within the field of study, the question of directionality of type I secretion has long been under debate. The concept of stalling the HlyA T1SS by using substrate N-terminally fused to fast folding enhanced green fluorescent protein (eGFP) (Evdokimov *et al.* 2006) finally answered this question (Lenders *et al.* 2015). Folded eGFP in the cytosol served as a 'plug' while the C-terminal moiety inserted into the channel tunnel and protruded partially into the extracellular space, where it prevented backsliding by adopting its tertiary structure. A combination of fluorescence and super-resolution microscopy exploiting the autofluorescence of eGFP in the cytosol and immunofluorescence-based methods to detect the secreted C-terminus of the substrate demonstrated that the secretion sequence appears first on the external surface of the cell envelope.

QUANTITATIVE ANALYSIS

The concept of stalling a T1SS (as described in section 'Functional insights') not only offered the possibility to address the directionality of transport but was also exploited to determine the secretion rate of the HlyA T1SS (Lenders *et al.* 2016). Importantly,

the fluorescence of Cy3-labeled antibody was first employed to quantify the total number of active HlyA T1SS translocons per cell. Interestingly, the derived number was in good agreement with the absolute number of HlyB dimers present in the membrane, as determined by quantitative western blot analysis, using a standard of purified HlyB (Reimann *et al.* 2016) of known concentrations. By experimentally quantifying the amount of secreted substrate, the secretion rate of the HlyA T1SS was determined to be 16 amino acids per transporter per second. Thus, it requires 90 s to secrete one complete HlyA molecule. This rate is roughly 10-fold lower than the rate of SecA-dependent protein translocation across the inner membrane, which operates at a calculated rate of ~152–228 amino acids per second per transporter (Schiebel *et al.* 1991; Uchida, Mori and Mizushima 1995). Intriguingly, the rate of HlyA secretion is very similar to the rate of bacterial protein synthesis at the ribosome, which was calculated to be 10–20 amino acids per second (Young and Bremer 1976). Whether this similarity is of any relevance and results from some sort of connection still has to be addressed experimentally.

In earlier studies, the proton motive force (pmf) was identified as being essential for substrate secretion (Koronakis, Hughes and Koronakis 1991). In our hands, however, an influence of the pmf on the secretion rate was not observed (unpublished data), supporting recent results on CyaA, an exotoxin from *B. pertussis* (Bumba *et al.* 2016), whose secretion is also independent from the pmf. This seminal study also demonstrated a clear influence of the extracellular Ca^{2+} concentration on the secretion efficiency. Thus, the presence of Ca^{2+} accelerated CyaA secretion by generating intramolecular Brownian ratchets. In other words, this process is passive but involves ratcheted translocation events. Nevertheless, these data do not support the hypothesis that the binding of Ca^{2+} ions to the RTX domains represents a driving force of prime importance for secretion (Chenal *et al.* 2009; Thomas *et al.* 2014). This is in striking contrast to the secretion of HlyA (Lenders *et al.* 2016), where changes in the Ca^{2+} concentration in the medium or even the complete absence of Ca^{2+} did not influence the secretion rate, which remained at 16 amino acids per transporter per second. However, one has to stress that Ca^{2+} is crucial for the functionality of HlyA and that in the absence of Ca^{2+} the pore-forming activity was abolished (Lenders *et al.* 2016). These findings suggest a certain variety in the molecular mechanism of secretion amongst different T1SS, which may be influenced by the size of the substrate or the arrangement of the RTX domains. Thus, it is suggested that a Brownian ratchet mechanism combined with a pulling force is operational in CyaA (Bumba *et al.* 2016), but absent in HlyA (Lenders *et al.* 2016) and further experiments especially involving other T1SS are required to settle this issue.

STRUCTURAL INSIGHTS

Structural elucidation, together with functional characterization, is a powerful tool to investigate the transport mechanisms of membrane proteins. The very first structural information of a T1SS component was derived from two-dimensional crystals of the outer membrane protein TolC from *E. coli*. Even at a resolution of 12 Å, apart from the trimeric β-barrel nature, the presence of a novel periplasmic domain became evident (Koronakis *et al.* 1997). Only a few years later, solving the crystal structure of TolC at 2.1 Å revealed the novel fold of this funnel-like domain (Koronakis *et al.* 2000). This is composed out of 12 α-helices that protrude 100 Å into the periplasmic space. Including the
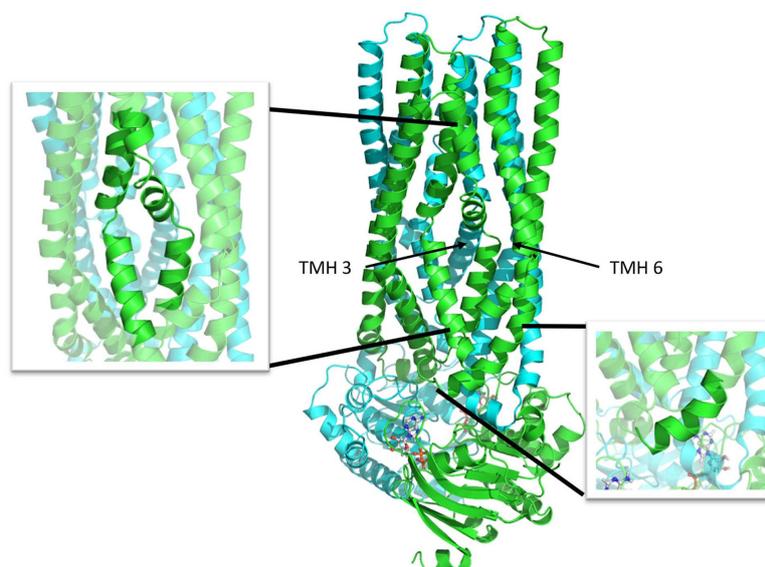
Figure 3. Cartoon representation of AaPrtD (Morgan, Acheson and Zimmer 2017). Monomers are shown in green and cyan. The bound nucleotides are shown in stick representation. The kinked helices 3 and 6 (TMH3 and TMH6) are highlighted for one monomer (left zoom-in). The right zoom-in highlights the coupling helix 1.
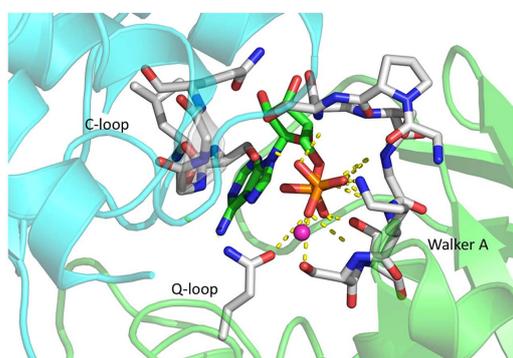

Figure 4. Zoom into the nucleotide-binding site of AaPrtD (Morgan, Acheson and Zimmer 2017) formed by both NBDs colored in green and cyan. ADP is shown in stick representation, the co-factor Mg^{2+} as magenta sphere. The conserved motifs interacting with the bound ADP molecule are labeled and highlighted in stick representation. Interactions are visualized by dashed, yellow lines.

12-stranded β-barrel (3 per monomer) the total length of the protein adds up to 140 Å. The structure was interpreted to represent a closed state as the inner diameter of the water-filled β-barrel of 20 Å narrows to only 3.9 Å at the periplasmic gate of TolC.

Based on this structural information, alanine mutations were placed within the region of the periplasmic gate in order to disrupt the closed state of TolC. Conductivity measurements in black lipid membranes led to a model, in which an 'iris-like' opening of the inner helices opens the periplasmic gate and therefore allows substrate translocation (Andersen *et al.* 2002a,b). Later on, this model was confirmed by structural information (Bavro *et al.* 2008; Pei *et al.* 2011).

Starting in 2003, a series of crystal structures paired with functional analysis on the NBD of the ABC transporter HlyB was

published (Benabdelhak *et al.* 2003, 2005; Schmitt *et al.* 2003; Zaitseva *et al.* 2005a, 2006). These studies offered valuable insights into the motor domain of a T1SS and its detailed molecular mechanism of ATP hydrolysis (Zaitseva *et al.* 2005b,c; Hanekop *et al.* 2006; Oswald, Holland and Schmitt 2006).

Additional insights into the structure of T1SS were obtained only recently, when the crystal structure of the ABC transporter of a putative T1SS from the hyperthermophilic Gram-negative bacterium *Aquifex aeolicus* (AaPrtD) was published. It shares a sequence identity of 40% with PrtD from *Dickeya dadantii*, but neither the substrate nor the MFP homolog of AaPrtD was identified (Morgan, Acheson and Zimmer 2017). The structure of the homodimer was determined at a resolution of 3.15 Å and reflected the ADP/Mg^{2+}-bound state (Fig. 3). The presence of six transmembrane helices (TMH) per monomer and the canonical fold of the NBDs is typical for ABC transporters. Interestingly, the arrangement of the NBDs in the ADP-bound state seems to represent the occluded state (Morgan, Acheson and Zimmer 2017), which contradicts the accepted view that ATP binding induces dimerization of the two NBDs (Locher 2004, 2016; Oswald, Holland and Schmitt 2006).

The ADP molecule is coordinated by residues of the Walker A motif and the glutamine residue of the Q-loop, and also by the serine residue of the C-loop of the opposing NBD resulting in dimerization (Fig. 4). However, a similar architecture has been observed in the crystal structure of Sav1866 from *Staphylococcus aureus* (Dawson and Locher 2006) raising the question of how this architecture fits into the well-established view that only ATP induces formation of the NBD dimer.

The architecture of the TMHs of AaPrtD is distinct from that of other ABC export systems. Generally, ABC transporters contain two coupling helices (CH1 and CH2), which interact with the NBDs. In AaPrtD, the interaction conferred by CH1 is taken over by TMH2, which extends into a loop region that continues without any secondary structure into TMH3.

Figure 5. Crystal structure of a soluble fragment of HlyD from *E. coli* (Kim *et al.* 2016) that highlights the coiled-coil interaction of the helices. The lipoyl domain is colored in red and blue. N- and C-termini are indicated.

Furthermore, TMH3 and TMH6 are kinked at the approximate position of the lipid head groups, splitting TMH6 into two and TMH3 into three separate helices—a novel architecture—not been observed before (Morgan, Acheson and Zimmer 2017), neither in the bacteriocin transporter McjD (Choudhury *et al.* 2014) nor in the peptide transporter TAP 1/2 (Oldham, Grigorieff and Chen 2016). The kinks create a restriction within the putative substrate channel. The fact that residues lining this restriction are stabilized by interactions with conserved amino acids emphasizes their functional importance. Although the functionality of AaPrtD has not been demonstrated yet and the identity of the MFP and the transport substrate remain elusive, the structure provided the first glimpse of a T1SS ABC transporter and presents a platform to design new and exciting experiments.

Proteins from the family of MFPs are not only an indispensible part of T1SS but have also been intensively studied in the context of bacterial tripartite drug efflux pumps, such as the AcrB-AcrA-TolC system (Du *et al.* 2014) or the MacA-MacB-TolC system (Fitzpatrick *et al.* 2017). Until 2016, structural information was limited to MFPs involved in drug efflux (for a recent review, see Symmons, Marshall and Bavro 2015), while for T1SS MFPs only recently has some limited structural information became available (Kim *et al.* 2016). In 2016, the crystal structure of a soluble fragment of HlyD, comprising the α-helical domain and lipoyl domain, was published (Kim *et al.* 2016). However, the construct used for structure determination lacked not only the first 95 N-terminal amino acids, including the cytoplasmic domain (residues 1–59) and the single TMH of HlyD (residues 60–80), but also the last 106 C-terminal residues (residues 373–478), corresponding to the entire membrane proximal domain (Fig. 5).

In contrast to the majority of structurally described MFPs that contain two α-helices, the unusually long (115 Å) α-helical domain of HlyD is built out of three helices, of which helix 3 in-

teracts in an anti-parallel coiled-coil fashion with helix 1 and 2. Based on structural comparison with AcrA (Kim *et al.* 2010) and sequence conservation analysis, a model was proposed, where the α-helical tip located between helix 2 and 3 forms the interaction site with TolC (Kim *et al.* 2016). Only recently, in 2017, the crystal structure of the α-helical domain and the lipoyl domain of LipC, the MFP of the lipase secretion system from *Serratia marcescens*, was reported (Murata *et al.* 2017). Interestingly, the α-helical domain also contained three helices, which might be a common feature of T1SS MFPs.

The AcrA-AcrB-TolC (Du *et al.* 2014) and the MacA-MacB-TolC (Fitzpatrick *et al.* 2017) structures revealed a hexameric arrangement of the corresponding MFPs. Contradictory findings provided by cross-linking studies in *E. coli* suggest a trimeric state of HlyD as the functional unit (Thanabalu *et al.* 1998). Even though in principle a hexameric state seems more likely and is also supported by a model (Kim *et al.* 2016) based on the available crystal structure of MacA (Yum *et al.* 2009), the oligomeric state of T1SS MFPs is still under debate and the subject is still in need of further investigation.

BIOTECHNOLOGICAL APPLICATIONS

By achieving the secretion of fusion proteins in high amounts into the extracellular medium, large-scale purification can be significantly simplified, which reflects an attractive approach for biotechnological applications.

The relatively simple nature of T1SS and the C-terminal secretion signal raised interest in their use in biotechnological applications. Two main areas became the focus of intensive research: heterologous protein secretion, in general (Blight and Holland 1994) and antigen production for vaccination (Sebo *et al.* 1999; Spreng *et al.* 1999).

Here, we will only focus on the applicability of T1SS for the secretion of heterologous proteins. Based on the T1SS of TliA, a thermostable lipase from *P. fluorescence* (Park *et al.* 2012), a versatile secretion system was engineered (Ryu *et al.* 2015). The C-terminal secretion signal of TliA was fused either to GFP or alkaline phosphatase, and both proteins were secreted into the medium. The hydrophobic nature of the secretion signal even allowed subsequent purification via hydrophobic interaction chromatography. Secretion of fusion proteins was further enhanced by engineering a negative net charge by introducing aspartate clusters in the fusion proteins of interest (Byun *et al.* 2017). These constructs go hand in hand with the direction of the membrane potential, which is positive on the surface of bacteria. This favors the translocation of negatively charged proteins since the directionality of the potential acts electrophoretically (Cao, Kuhn and Dalbey 1995).

The other T1SS that has been extensively studied for the purpose of heterologous protein secretion is the HlyA T1SS from *E. coli*. With the identification of HlyA1 (Nicaud *et al.* 1986), a 23 kDa, C-terminal fragment of HlyA encouraging the secretion sequence and three of six RTX domains, research efforts were intensified to exploit the system for the secretion of heterologous proteins. The studies of Debarbieux and Wandersman (2001) and Bakkes *et al.* (2010) stressed the importance of folding rates in the successful secretion of fusion proteins. Remarkably, the natural folding rate of a protein fused to HlyA1 decreased dramatically, which increased the range of possible fusion partners and increased the yields of the proteins of interest (Bakkes *et al.* 2010).

Only recently, a new expression vector was established that impressively improved secretion efficiencies of various

fusion proteins (Khosa *et al.* 2018). Here, a 5′-untranslated region (UTR) was identified that increased the amount of secreted HlyA1 several-fold and that, most surprisingly, does contain nucleotides belonging to the coding sequence of HlyC. The region was mapped to an area enriched in adenosine- and uracil nucleotides, located ∼36 base pairs upstream from the start codon of HlyA1. Nevertheless, the most striking result to emerge from this data in terms of biotechnological application is that besides boosting the secretion efficiency of HlyA1, the vector enabled the secretion of fast-folding fusion proteins that could not be secreted until then (Khosa *et al.* 2018). This is likely due to the fact that this 5′-UTR is recognized by ribosomal protein S1 and targeted to the ribosome for faster and more efficient translation. Certainly, these findings could open up new avenues for the exploitation of T1SS and highlights their potential biotechnological and pharmaceutical value. Furthermore, they emphasize the necessity not to limit biotechnological engineering to the mere coding sequence of a protein.

OUTLOOK

Since its discovery nearly 40 years ago, T1SS have been the subject of numerous fruitful studies and the basic outline of the secretion process is by now well established. However, much uncertainty still exists about the detailed mechanism of transport. Ongoing, exciting research should address, for example, additional structural information, the stoichiometry of the T1SS complexes, the nature of the secretion signal and biochemical insights into recognition and processing of the substrate.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

Andersen C, Koronakis E, Bokma E *et al.* Transition to the open state of the TolC periplasmic tunnel entrance. *P Natl Acad Sci USA* 2002a;**99**:11103–8.

Andersen C, Koronakis E, Hughes C *et al.* An aspartate ring at the TolC tunnel entrance determines ion selectivity and presents a target for blocking by large cations. *Mol Microbiol* 2002b;**44**:1131–9.

Bakkes PJ, Jenewein S, Smits SH *et al.* The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin Type 1 secretion system. *J Biol Chem* 2010;**285**:40573–80.

Balakrishnan L, Hughes C, Koronakis V. Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *J Mol Biol* 2001;**313**:501–10.

Baumann U, Wu S, Flaherty KM *et al.* Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif. *EMBO J* 1993;**12**:3357–64.

Bavro VN, Pietras Z, Furnham N *et al.* Assembly and channel opening in a bacterial drug efflux machine. *Mol Cell* 2008;**30**:114–21.

Benabdelhak H, Kiontke S, Horn C *et al.* A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A. *J Mol Biol* 2003;**327**:1169–79.

Benabdelhak H, Schmitt L, Horn C *et al.* Positive cooperative activity and dimerization of the isolated ABC-ATPase domain of HlyB from *E. coli. Biochem J* 2005;**368**:1–7.

Blight MA, Holland IB. Heterologous protein secretion and the versatile *Escherichia coli* haemolysin translocator. *Trends Biotechnol* 1994;**12**:450–5.

Bumba L, Masin J, Macek P *et al.* Calcium-driven folding of RTX domain beta-Rolls ratchets translocation of RTX proteins through type I secretion ducts. *Mol Cell* 2016;**62**:47–62.

Byun H, Park J, Kim SC *et al.* A lower isoelectric point increases signal sequence-mediated secretion of recombinant proteins through a bacterial ABC transporter. *J Biol Chem* 2017;**292**:19782–91.

Cao G, Kuhn A, Dalbey RE. The translocation of negatively charged residues across the membrane is driven by the electrochemical potential: evidence for an electrophoresis-like membrane transfer mechanism. *EMBO J* 1995;**14**:866–75.

Cescau S, Debarbieux L, Wandersman C. Probing the in vivo dynamics of type I protein secretion complex association through sensitivity to detergents. *J Bacteriol* 2007;**189**:1496–504.

Chenal A, Guijarro JI, Raynal B *et al.* RTX calcium binding motifs are intrinsically disordered in the absence of calcium. *J Biol Chem* 2009;**284**:1781–9.

Chervaux C, Sauvonnet N, Le Clainche A *et al.* Secretion of active beta-lactamase to the medium mediated by the Escherichia coli haemolysin transport pathway. *Mol Gen Genet* 1995;**249**:237–45.

Choudhury HG, Tong Z, Mathavan I *et al.* Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. *P Natl Acad Sci USA* 2014;**111**:9145–50.

Costa TR, Felisberto-Rodrigues C, Meir A *et al.* Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* 2015;**13**:343–59.

Davidson AL, Dassa E, Orelle C *et al.* Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol R* 2008;**72**:317–64, table of contents.

Dawson RJ, Locher KP. Structure of a bacterial multidrug ABC transporter. *Nature* 2006;**443**:180–5.

Debarbieux L, Wandersman C. Folded HasA inhibits its own secretion through its ABC exporter. *EMBO J* 2001;**20**:4657–63.

Delepelaire P. Type I secretion in gram-negative bacteria. *BBA-Mol Cell Res* 2004;**1694**:149–61.

Du D, Wang Z, James NR *et al.* Structure of the AcrAB-TolC multidrug efflux pump. *Nature* 2014;**509**:512–5.

Economou A, Dalbey RE. Preface to special issue on protein trafficking and secretion in bacteria. *Biochim Biophys Acta* 2014;**1843**:1427.

Evdokimov AG, Pokross ME, Egorov NS *et al.* Structural basis for the fast maturation of Arthropoda green fluorescent protein. *EMBO Rep* 2006;**7**:1006–12.

Felmlee T, Pellett S, Welch RA. Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. *J Bacteriol* 1985;**163**:94–105.

Felmlee T, Welch RA. Alterations of amino acid repeats in the *Escherichia coli* hemolysin affect cytolytic activity and secretion. *P Natl Acad Sci USA* 1988;**85**:5269–73.

Fitzpatrick AWP, Llabres S, Neuberger A *et al.* Structure of the MacAB-TolC ABC-type tripartite multidrug efflux pump. *Nat Microbiol* 2017;**2**:17070.

Gentschev I, Hess J, Goebel W. Change in the cellular localization of alkaline phosphatase by alteration of its carboxy-terminal sequence. *Mol Gen Genet* 1990;**222**:211–6.

Gray L, Baker K, Kenny B *et al.* A novel C-terminal signal sequence targets *Escherichia coli* haemolysin directly to the medium. *J Cell Sci* 1989;**11**:45–57.

Gray L, Mackman N, Nicaud JM *et al.* The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*. *Mol Gen Genet* 1986;**205**:127–33.

Greene NP, Crow A, Hughes C *et al.* Structure of a bacterial toxin-activating acyltransferase. *P Natl Acad Sci USA* 2015;**112**:E3058–66.

Griessl MH, Schmid B, Kassler K *et al.* Structural insight into the giant Ca^{2+}-binding adhesin SiiE: implications for the adhesion of salmonella enterica to polarized epithelial cells. *Structure* 2013;**21**:741–52.

Hanekop N, Zaitseva J, Jenewein S *et al.* Molecular insights into the mechanism of ATP-hydrolysis by the NBD of the ABC-transporter HlyB. *FEBS Lett* 2006;**580**:1036–41.

Havarstein LS, Diep DB, Nes IF. A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol Microbiol* 1995;**16**:229–40.

Hinsa SM, Espinosa-Urgel M, Ramos JL *et al.* Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. *Mol Microbiol* 2003;**49**:905–18.

Holland IB, Peherstorfer S, Kanonenberg K *et al.* Type I protein secretion-deceptively simple yet with a wide range of mechanistic variability across the family. *EcoSal Plus* 2016;**7**:1–46.

Hwang J, Zhong X, Tai PC. Interactions of dedicated export membrane proteins of the colicin V secretion system: CvaA, a member of the membrane fusion protein family, interacts with CvaB and TolC. *J Bacteriol* 1997;**179**:6264–70.

Ishii S, Yano T, Ebihara A *et al.* Crystal structure of the peptidase domain of *Streptococcus* ComA, a Bifunctional ATP-binding cassette transporter involved in the quorum-sensing pathway. *J Biol Chem* 2010;**285**:10777–85.

Issartel JP, Koronakis V, Hughes C. Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. *Nature* 1991;**351**:759–61.

Izadi-Pruneyre N, Wolff N, Redeker V *et al.* NMR studies of the C-Terminal secretion signal of the haem-binding protein, HasA. *Eur J Biochem* 1999;**261**:562–8.

Jardetzky O. Simple allosteric model for membrane pumps. *Nature* 1966;**211**:969–70.

Jones HE, Holland IB, Baker HL *et al.* Slow changes in cytosolic free Ca2+ in *Escherichia coli* highlight two putative influx mechanisms in response to changes in extracellular calcium. *Cell Calcium* 1999;**25**:265–74.

Kanonenberg K, Schwarz CK, Schmitt L. Type I secretion systems - a story of appendices. *Res Microbiol* 2013;**164**:596–604.

Kenny B, Haigh R, Holland IB. Analysis of the haemolysin transport process through the secretion from *Escherichia coli* of PCM, CAT or beta-galactosidase fused to the Hly C-terminal signal domain. *Mol Microbiol* 1991;**5**:2557–68.

Khosa S, Scholz R, Schwarz C *et al.* An A/U-Rich enhancer region is required for high-level protein secretion through the HlyA Type I secretion system. *Appl Environ Microb* 2018;**84**:e01163–17.

Kim HM, Xu Y, Lee M *et al.* Functional relationships between the AcrA hairpin tip region and the TolC aperture tip region for the formation of the bacterial tripartite efflux pump AcrAB-TolC. *J Bacteriol* 2010;**192**:4498–503.

Kim JS, Song S, Lee M *et al.* Crystal structure of a soluble fragment of the membrane fusion protein HlyD in a Type I secretion system of gram-negative bacteria. *Structure* 2016;**24**:477–85.

Koronakis V, Eswaran J, Hughes C. Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annu Rev Biochem* 2004;**73**:467–89.

Koronakis V, Hughes C, Koronakis E. Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. *EMBO J* 1991;**10**:3263–72.

Koronakis V, Koronakis E, Hughes C. Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes. *EMBO J* 1989;**8**:595–605.

Koronakis V, Li J, Koronakis E *et al.* Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol Microbiol* 1997;**23**:617–26.

Koronakis V, Sharff A, Koronakis E *et al.* Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 2000;**405**:914–9.

Kuhnert P, Heyberger-Meyer B, Nicolet J *et al.* Characterization of PaxA and its operon: a cohemolytic RTX toxin determinant from pathogenic *Pasteurella aerogenes*. *Infect Immun* 2000;**68**:6–12.

Lecher J, Schwarz CK, Stoldt M *et al.* An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. *Structure* 2012;**20**:1778–87.

Lenders MH, Beer T, Smits SH *et al.* In vivo quantification of the secretion rates of the hemolysin A Type I secretion system. *Sci Rep* 2016;**6**:33275.

Lenders MH, Weidtkamp-Peters S, Kleinschrodt D *et al.* Directionality of substrate translocation of the hemolysin A Type I secretion system. *Sci Rep* 2015;**5**:12470.

Letoffe S, Delepelaire P, Wandersman C. Protease secretion by *Erwinia chrysanthemi*: the specific secretion functions are analogous to those of *Escherichia coli* alpha-haemolysin. *EMBO J* 1990;**9**:1375–82.

Letoffe S, Ghigo JM, Wandersman C. Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *J Bacteriol* 1994;**176**:5372–7.

Linhartova I, Bumba L, Masin J *et al.* RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol Rev* 2010;**34**:1076–112.

Locher KP. Structure and mechanism of ABC transporters. *Curr Opin Struct Biol* 2004;**14**:426–31.

Locher KP. Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nat Struct Mol Biol* 2016;**23**:487–93.

Mackman N, Baker K, Gray L *et al.* Release of a chimeric protein into the medium from *Escherichia coli* using the C-terminal secretion signal of haemolysin. *EMBO J* 1987;**6**:2835–41.

Mackman N, Holland IB. Functional characterization of a cloned haemolysin determinant from *E. coli* of human origin, encoding information for the secretion of a 107K polypeptide. *Mol Gen Genet* 1984;**196**:129–34.

Mackman N, Nicaud JM, Gray L *et al.* Genetical and functional organisation of the *Escherichia coli* haemolysin determinant 2001. *Mol Gen Genet* 1985a;**201**:282–8.

Mackman N, Nicaud JM, Gray L *et al.* Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*. *Mol Gen Genet* 1985b;**201**:529–36.

Masure HR, Au DC, Gross MK *et al.* Secretion of the *Bordetella pertussis* adenylate cyclase from *Escherichia coli* containing the hemolysin operon. *Biochemistry* 1990;**29**:140–5.

Meier R, Drepper T, Svensson V *et al.* A Calcium-gated lid and a large beta-Roll sandwich are revealed by the crystal structure of extracellular lipase from *Serratia marcescens*. *J Biol Chem* 2007;**282**:31477–83.

Morgan JL, Acheson JF, Zimmer J. Structure of a Type-1 secretion system ABC transporter. *Structure* 2017;**25**:522–9.

Murata D, Okano H, Angkawidjaja C *et al.* Structural basis for the *Serratia marcescens* lipase secretion system: Crystal structures of the membrane fusion protein and Nucleotide-Binding domain. *Biochemistry* 2017;**56**:6281–91.

Nicaud JM, Mackman N, Gray L *et al.* Regulation of haemolysin synthesis in *E. coli* determined by HLY genes of human origin. *Mol Gen Genet* 1985;**199**:111–6.

Nicaud JM, Mackman N, Gray L *et al.* The C-terminal, 23 kDa peptide of *E. coli* haemolysin 2001 contains all the information necessary for its secretion by the haemolysin (Hly) export machinery. *FEBS Lett* 1986;**204**:331–5.

Noegel A, Rdest U, Springer W *et al.* Plasmid cistrons controlling synthesis and excretion of the exotoxin alpha-haemolysin of Escherichia coli. *Mol Gen Genet* 1979;**175**:343–50.

Oldham ML, Grigorieff N, Chen J. Structure of the transporter associated with antigen processing trapped by herpes simplex virus. *Elife* 2016;**5**:e21289.

Ostolaza H, Soloaga A, Goni FM. The binding of divalent cations to *Escherichia coli* alpha-haemolysin. *Eur J Biochem* 1995;**228**:39–44.

Oswald C, Holland IB, Schmitt L. The motor domains of ABC-transporters. *N-S Arch Pharmacol* 2006;**372**:385–99.

Park Y, Moon Y, Ryoo J *et al.* Identification of the minimal region in lipase ABC transporter recognition domain of *Pseudomonas fluorescens* for secretion and fluorescence of green fluorescent protein. *Microb Cell Fact* 2012;**11**:60.

Pei XY, Hinchliffe P, Symmons MF *et al.* Structures of sequential open states in a symmetrical opening transition of the TolC exit duct. *P Natl Acad Sci USA* 2011;**108**:2112–7.

Pimenta AL, Young J, Holland IB *et al.* Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator. *Mol Gen Genet* 1999;**261**:122–32.

Reimann S, Poschmann G, Kanonenberg K *et al.* Interdomain regulation of the ATPase activity of the ABC transporter haemolysin B from *Escherichia coli*. *Biochem J* 2016;**473**:2471–83.

Ryu J, Lee U, Park J *et al.* A vector system for ABC transporter-mediated secretion and purification of recombinant proteins in pseudomonas species. *Appl Environ Microb* 2015;**81**:1744–53.

Sanchez-Magraner L, Viguera AR, Garcia-Pacios M *et al.* The Calcium-binding C-terminal domain of *Escherichia coli* alpha-hemolysin is a major determinant in the surface-active properties of the protein. *J Biol Chem* 2007;**282**:11827–35.

Sapriel G, Wandersman C, Delepelaire P. The N terminus of the HasA protein and the SecB chaperone cooperate in the efficient targeting and secretion of HasA via the ATP-binding cassette transporter. *J Biol Chem* 2002;**277**:6726–32.

Schiebel E, Driessen AJ, Hartl FU *et al.* Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* 1991;**64**:927–39.

Schmitt L, Benabdelhak H, Blight MA *et al.* Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. *J Mol Biol* 2003;**330**:333–42.

Sebo P, Ladant D. Repeat sequences in the *Bordetella pertussis* adenylate cyclase toxin can be recognized as alternative carboxy-proximal secretion signals by the *Escherichia coli* alpha-haemolysin translocator. *Mol Microbiol* 1993;**9**:999–1009.

Sebo P, Moukrim Z, Kalhous M *et al.* In vivo induction of CTL responses by recombinant adenylate cyclase of *Bordetella Pertussis* carrying multiple copies of a viral CD8(+) T-cell epitope. *FEMS Immunol Med Mic* 1999;**26**:167–73.

Sotomayor Perez AC, Karst JC, Davi M *et al.* Characterization of the regions involved in the calcium-induced folding of the intrinsically disordered RTX motifs from the *Bordetella pertussis* adenylate cyclase toxin. *J Mol Biol* 2015;**397**:534–49.

Spreng S, Dietrich G, Goebel W *et al.* The *Escherichia coli* haemolysin secretion apparatus: a potential universal antigen delivery system in gram-negative bacterial vaccine carriers. *Mol Microbiol* 1999;**31**:1596–8.

Stanley P, Hyland C, Koronakis V *et al.* An ordered reaction mechanism for bacterial toxin acylation by the specialized acyltransferase HlyC: formation of a ternary complex with acylACP and protoxin substrates. *Mol Microbiol* 1999;**34**:887–901.

Stanley P, Koronakis V, Hughes C. Mutational analysis supports a role for multiple structural features in the C-terminal secretion signal of *Escherichia coli* haemolysin. *Mol Microbiol* 1991;**5**:2391–403.

Stanley P, Packman LC, Koronakis V *et al.* Fatty acylation of two internal lysine residues required for the toxic activity of *Escherichia coli* hemolysin. *Science* 1994;**266**:1992–6.

Symmons MF, Marshall RL, Bavro VN. Architecture and roles of periplasmic adaptor proteins in tripartite e fflux assemblies. *Front Microbiol* 2015;**6**:513.

Thanabalu T, Koronakis E, Hughes C *et al.* Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *EMBO J* 1998;**17**:6487–96.

Thomas S, Bakkes PJ, Smits SH *et al.* Equilibrium folding of pro-HlyA from *Escherichia coli* reveals a stable calcium ion dependent folding intermediate. *BBA-Proteins Proteom* 2014;**1844**:1500–10.

Thomas S, Holland IB, Schmitt L. The Type 1 secretion pathway - The hemolysin system and beyond. *Biochim Biophys Acta* 2014;**1843**:1621–41.

Thomas S, Smits SH, Schmitt L. A simple in vitro acylation assay based on optimized HlyA and HlyC purification. *Anal Biochem* 2014;**464**:17–23.

Thompson SA, Sparling PF. The RTX cytotoxin-related FrpA protein of *Neisseria meningitidis* is secreted extracellularly by meningococci and by HlyBD+ *Escherichia coli*. *Infect Immun* 1993;**61**:2906–11.

Uchida K, Mori H, Mizushima S. Stepwise movement of preproteins in the process of translocation across the cytoplasmic membrane of *Escherichia coli*. *J Biol Chem* 1995;**270**:30862–8.

Wandersman C, Delepelaire P. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *P Natl Acad Sci USA* 1990;**87**:4776–80.

Wang RC, Seror SJ, Blight M *et al.* Analysis of the membrane organization of an *Escherichia coli* protein translocator, HlyB, a

member of a large family of prokaryote and eukaryote surface transport proteins. *J Mol Biol* 1991;**217**:441–54.

Welch RA. Pore-forming cytolysins of gram-negative bacteria. *Mol Microbiol* 1991;**5**:521–8.

Wu KH, Tai PC. Cys 32 and His 105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter. *J Biol Chem* 2004;**279**:901–9.

Young R, Bremer H. Polypeptide-chain-elongation rate in *Escherichia coli* B/r as a function of growth rate. *Biochem J* 1976;**160**:185–94.

Yum S, Xu Y, Piao S *et al.* Crystal structure of the periplasmic component of a tripartite macrolide-specific efflux pump. *J Mol Biol* 2009;**387**:1286–97.

Zaitseva J, Jenewein S, Jumpertz T *et al.* H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *EMBO J* 2005a;**24**:1901–10.

Zaitseva J, Jenewein S, Oswald C *et al.* A molecular understanding of the catalytic cycle of the nucleotide-binding domain of the ABC transporter HlyB. *Biochm Soc Trans* 2005b:990–5.

Zaitseva J, Jenewein S, Wiedenmann A *et al.* Functional characterization and ATP-induced dimerization of the isolated ABC-domain of the haemolysin B transporter. *Biochemistry* 2005c;**44**:9680–90.

Zaitseva J, Oswald C, Jumpertz T *et al.* A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer. *EMBO J* 2006;**25**:3432–43.

3.2   Chapter 2 - Type I Secretion Systems – One Mechanism for All?

Title: Type I Secretion Systems – One Mechanism for All?

Authors: **Olivia Spitz**, Isabelle N. Erenburg, Tobias Beer, Kerstin Kanonenberg,

I. Barry Holland and Lutz Schmitt

Published in: *Microbiology Spectrum* (2019)

Impact Factor 5.465 (2019)

Own Work: 40 %

Writing of the manuscript

Type I Secretion Systems— One Mechanism for All?

OLIVIA SPITZ,[1] ISABELLE N. ERENBURG,[1] TOBIAS BEER,[1] KERSTIN KANONENBERG,[1] I. BARRY HOLLAND,[2] and LUTZ SCHMITT[1]

[1]Institute of Biochemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany; [2]Institute of Genetics and Microbiology, University of Paris-Sud, Orsay, France

ABSTRACT Type I secretion systems (T1SS) are widespread in Gram-negative bacteria, especially in pathogenic bacteria, and they secrete adhesins, iron-scavenger proteins, lipases, proteases, or pore-forming toxins in the unfolded state in one step across two membranes without any periplasmic intermediate into the extracellular space. The substrates of T1SS are in general characterized by a C-terminal secretion sequence and nonapeptide repeats, so-called GG repeats, located N terminal to the secretion sequence. These GG repeats bind Ca^{2+} ions in the extracellular space, which triggers folding of the entire protein. Here we summarize our current knowledge of how Gram-negative bacteria secrete these substrates, which can possess a molecular mass of up to 1,500 kDa. We also describe recent findings that demonstrate that the absence of periplasmic intermediates, the "classic" mode of action, does not hold true for all T1SS and that we are beginning to realize modifications of a common theme.

INTRODUCTION

Gram-negative bacteria are equipped with at least seven dedicated secretion systems that mediate the export of proteins beyond the outer membrane (1, 2). These are called type 1 to 6 and type 9 secretion systems (T1SS to T6SS and T9SS). Among those, T3SS, T4SS, and T6SS are even capable of delivering their cargo directly into the cytosol of the host cell. In this minireview, we place the major emphasis on the hemolysin A (HlyA) secretion system in *Escherichia coli*. This is by far the most studied and illustrates very well the largely conserved, essential features of T1SS. Interestingly, however, an important mechanistic variation in the translocation of some of the unusually extended giant RTX proteins—adhesins—was discovered recently (3) and is also discussed.

T1SS substrates are usually defined by the presence of several blocks of nonapeptide-binding sequences with the consensus GGxGxDxUx (4, 5), where x can be any amino acid and U is a large hydrophobic amino acid. The exceptions are the SiiE-like adhesins (Fig. 1) (6). These nonapeptides gave rise to the abbreviation RTX (repeats in toxins), the name for the family. These motifs, also called GG repeats, specifically bind Ca^{2+} (see below) and are implicated in posttranslocation folding. The RTX domain (Fig. 1) is located N terminal to the secretion signal at the extreme C terminus.

Like T1SS substrates, the very large and widespread group of peptide bacteriocins in Gram-negative bacteria (7–9) also require an ABC transporter, a membrane fusion protein (MFP), and an outer membrane (OM) protein for secretion. However, these antimicrobials lack RTX repeats, have a cleavable N-terminal secretion sequence instead of the "classical" C-terminal signal, and
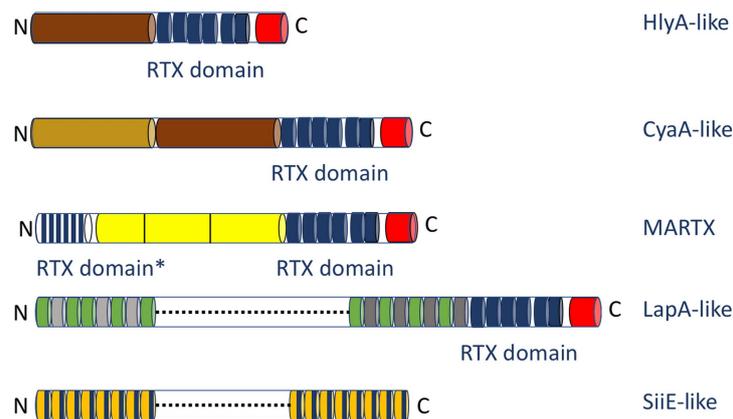
Spitz et al.

FIGURE 1 Architecture of substrates of T1SS. The primary structure of a canonical substrate of a T1SS is shown as white cylinder with the N and C termini labeled by "N" and "C," respectively. The secretion sequence (approximately 50 to 100 amino acids depending on the substrate) at the C terminus is in red, the GG repeats forming the classic RTX domain are in blue (six GG repeats as in the case of HlyA have been chosen as an example), and the functional, N-terminal domain is in brown. However, the number and types of architectures of this functional domain have increased in recent years. HlyA-like proteins contain only one domain with dedicated activity (pore-forming activity in the case of HlyA), while, for example, CyaA-like proteins contain two domains, which possess an adenylate cyclase (light brown) and a pore-forming (brown) activity in the case of CyaA. A third class are MARTX proteins (exemplified here by a MARTX protein from *V. cholerae*). The effector domains (yellow and separated by black vertical lines) that are autocatalytically excised after secretion are flanked by an N-terminal RTX-like domain (marked as RTX domain*) and a C-terminal RTX domain. The C-terminal domain corresponds to the canonical sequence, while the conserved aspartate is missing in the N-terminal one. Another architecture is present in LapA-like adhesins (or bacterial transglutaminase-like cysteine proteinases) that contain multiple, different domains. In the case of LapA, two different colors indicate two different domains. However, the number of different domains is not restricted to two. Additionally, the double-alanine motif in the N termini of LapA-like RTX adhesins is not shown. Finally, SiiE-like adhesins contain multiple identical domains, such as the 53 copies of the BIg domain in the case of SiiE ([6](#), [71](#)). The vertical blue line indicates that the GG repeats are integrated within the Ig-like domains and do not form a separate RTX domain. Please note that the drawing of the functional domains is not to scale.

have a quite distinctive translocation mechanism ([10](#), [11](#)): alternating access rather than extrusion through an OM "tunnel." In view of these properties, we decided not to include them in this minireview. The interested reader is directed to references [7](#) to [11](#).

The first molecular identification of a T1SS, the secretion machinery of the pore-forming toxin HlyA from *E. coli* ([12](#)), was made in the 1980s and 1990s ([13](#)–[15](#)) with the demonstration that two inner membrane proteins, an ABC transporter and an MFP, encoded together with the toxin in the same operon were required for secretion. A fourth gene in the *hly* operon encoded an acyltransferase, HlyC ([16](#)), catalyzing the posttranslational modification of two internal lysine residues ([17](#)–[19](#)). This

modification, with fatty acids ranging from C_{14} to C_{17} in length, requires acyl carrier protein ([20](#)) and is essential for HlyA to form a pore in the host membrane. Thus, only the acylated, toxic form of hemolysin should be called HlyA, while the nonacylated form should correctly be called pro-HlyA. Such acylation conferring toxicity is observed not only in HlyA but also in other hemolysins, leukotoxins, and cytolysins that are members of the RTX family ([4](#)). The recently published crystal structure of a homologue of the *E. coli* HlyC ([21](#)) allows a more detailed understanding of how acylation is installed. Notably, however, acylation is not required for secretion into the extracellular space. On the other hand, the proteins encoded in the *hly* operon are not sufficient for

secretion of (pro-)HlyA. The OM component of the translocon is TolC. The TolC protein is encoded elsewhere in the chromosome and was first described for a related T1SS by Wandersman and Delepelaire (22).

Equally important, the work by the laboratories of Koronakis, Holland, and Goebel demonstrated that substrate secretion by the T1SS is a one-step process, i.e., directly from the cytosol into the extracellular space without any periplasmic intermediate. Furthermore, these data established that the entire process was Sec independent, relying on a novel C-terminal secretion signal (5, 23–31).

However, the view that T1SS is mediated by the one-step translocation of proteins has been challenged. Recently, so-called periplasmic intermediates for a proposed two-step secretion process were described for the adhesins LapA and IBA (3). The exciting results identified a "retention module" (RM) at the N terminus that anchors the adhesion to the cell surface by stalling further translocation. This leaves a stalled short stub in the periplasm, apparently stuck in TolC, and a fully translocated, functional adhesin in the extracellular space. When conditions change, for example, in the case of LapA, to conditions unfavorable for biofilm formation, proteolysis removes the RM and releases the adhesin. Therefore, the secretion of the adhesin, as the authors described it, occurs in two steps. Our interpretation is that this is an exciting and important variation of the T1SS but that translocation is still effectively one step, and therefore, we prefer to call the adhesin-TolC-RM complex a pseudoperiplasmic intermediate.

Here we summarize our current knowledge of the molecular processes that underlie the T1SS and focus on the molecular events that result in secretion of substrates that harbor a C-terminal secretion sequence.

THE SUBSTRATES OF THE T1SS

The N-terminal domain of an RTX protein like HlyA contains one functional domain, the HlyA pore-forming toxin. CyaA from *Bordetella pertussis* (32) harbors an HlyA-like toxin but also an adenylate cyclase that, following translocation into the cytosol of a host cell, manipulates cAMP levels. More complex architectures are present in MARTX (multifunctional autoprocessing repeats in toxins), LapA, and SiiE-like proteins. MARTX proteins are of enormous size (approximately 500 to 900 kDa). This protein family is encoded in a chromosomal island in human pathogens such as *Vibrio cholerae* (33, 34). The extreme N-terminal part of these proteins is composed of an RTX-like domain that, however, lacks

the conserved aspartate residue that normally coordinates the Ca^{2+} ion. Spaced between this domain and the RTX domain near the C terminus are effector proteins that are autoprocessed, posttranslocationally, to release a cocktail of different effectors into a host cell (35). However, little is still known about the mechanism of secretion of MARTX proteins. The RTX adhesins, LapA from *Pseudomonas fluorescens*, and SiiE, the RTX-like protein from *Salmonella*, are even larger, reaching up to 1.5 MDa (3, 36, 37). In LapA (3, 38), the functional domain contains a varying number of domains that mediate adhesin functions. Strikingly, the SiiE-like adhesins deviate from the canonical architecture of RTX proteins, particularly with respect to calcium binding (39). Ca^{2+} binding sites are distributed virtually throughout the entire molecule, which is composed of 53 bacterial immunoglobulin-like (BIg) domains constituting the functional domain—the adhesin (Fig. 1). Ca^{2+} type I sites (three aspartate residues) fulfill the role of RTX repeats in secretion and are positioned at all the interfaces between two BIg domains (6). On the other hand, the translocon is composed of the familiar tripartite complex; translocation depends on a C-terminal secretion sequence (40) inferred to be extruded first (39).

A C-terminal secretion signal remains as a signature characteristic of RTX substrates. Signals appear to be conserved only within groups of related proteins, with no evidence of widespread conservation as far as we are aware. For the hemolysin group, competitive hypotheses have postulated a specific linear code, a structural code, or a combination of the two, but the question remains unresolved (see the extensive discussion in reference 41).

RTX Motifs and Ca^{2+} Promote Extracellular Folding of Substrates

A bioinformational approach based on the presence of the GG repeats revealed more than 1,000 putative RTX proteins in approximately 250 bacterial species (4). Since that study was published in 2010, the number of putative RTX proteins is necessarily much larger today, given the enormous number of genomes now sequenced. However, only the compilation of Linhartová et al. (4) is currently available. The number of identified RTX repeats ranged from below 10 to more than 40, with a slight tendency of the number of repeats to correlate with molecular weight. Additionally, more than 90% of the putative RTX proteins displayed an isoelectric point below 5.0, suggesting that electrophoretic mobility (42) might be important for the secretion process.

Structural studies of the alkaline protease from *Pseudomonas aeruginosa* (43) and other substrates of the

Spitz et al.

T1SS ([6](), [44](), [45]()) confirmed that the nonapeptide repeats bind Ca^{2+} ions. Two GG repeats coordinate one Ca^{2+} ion by interaction of the side chain of the aspartate residue and the carbonyl oxygens of the amino acids forming the repeat. This architecture creates a right-handed, so-called β-roll motif ([Fig. 2]()). Functional *in vitro* studies demonstrated that Ca^{2+} ions are a strict requirement for folding. In other words, in the absence of Ca^{2+} ions, substrates, such as HlyA from *E. coli* ([46–49]()) or CyaA from *Bordetella pertussis* ([50–52]()), remain unfolded or in a molten globular state ([53]()). Subsequent studies revealed that the dissociation constant of Ca^{2+} ions from the RTX domain is in the high micromolar range. The concentration of free Ca^{2+} ions in the bacterial cytosol is strictly regulated and remains in the high nanomolar range (100 to 300 nM in *E. coli*) ([54]()). Secretion of RTX proteins therefore presumably occurs in the unfolded state. This hypothesis was indeed experimentally verified by the fusion of maltose binding protein to a C-terminal fragment of HlyA that only harbored the secretion signal and three of the six GG repeats ([55]()). Given that the extracellular con-

centration of Ca^{2+} ions is normally in the millimolar range ([54]()), this suggests that RTX repeats immediately bind Ca^{2+} upon exit from the bacterium. Elegant *in vitro* studies with CyaA have also demonstrated that binding of Ca^{2+} ions to the RTX domain induces immediate folding of the entire protein ([46–48]()), suggesting that Ca^{2+} ions act as a chemical foldase.

CURRENT WORKING MODEL FOR CLASSIC T1SS

The process of secreting a substrate by a T1SS starts at the ribosome. However, only after the extreme C terminus of the substrate containing the secretion signal ([Fig. 1]()), around 50 to 100 amino acids, has been synthesized will secretion be initiated, since all information necessary and sufficient for secretion is encoded in the secretion signal. Bearing in mind that the sizes of T1SS proteins range from 20 kDa up to 1,500 kDa, two obvious questions arise: why do substrates of T1SS not aggregate and precipitate prior to secretion, and why are these proteins not immediately degraded by cytosolic proteases? Unfortunately, we do not yet have answers to these important questions.

In the second step, the unfolded substrate interacts with both of the two membrane proteins of the inner membrane, the ABC transporter and the MFP ([56](), [57]()). Based on cross-linking studies with the HlyA system, these two proteins were shown to form a stable complex in the inner membrane, a dimer of the ABC transporter and a trimer of the MFP ([57]()). However, the remarkable similarity of the T1SS translocon to tripartite drug efflux pumps, such as the AcrB-AcrA-TolC system from *E. coli*, in which there is a 2:6 stoichiometry (ABC:MFP) ([58]()), makes it most likely that the T1SS MFP is also a hexamer. Nevertheless, further research should be undertaken to resolve this obvious discrepancy. Deletion studies by the Koronakis laboratory showed that the cytoplasmic domain of the MFP is required to recruit the OM component, TolC in the case of the HlyA machinery ([56]()). However, the engagement occurred only in the presence of the substrate, indicating that docking of HlyA with the inner membrane complex transmits a signal to the periplasmic domain of HlyD that results in the formation of a transient HlyB-HlyD-TolC complex, a so-called "channel-tunnel" bridging the entire distance from the cytosol to the extracellular space across the periplasm and two membranes. The timing of these events also explains why deletion or inactivation of one of the three translocon components completely abolishes secretion without the appearance of a periplasmic intermediate.

FIGURE 2 Structure of GG repeats of alkaline protease (PDB entry 1KAP) from *P. aeruginosa* in its Ca^{2+}-bound state, resulting in the classic β-roll motif. **(A)** The five Ca^{2+} ions are shown as blue spheres. For simplicity, only the first three GG repeats are shown in ball-and-stick representation. The carbon atoms of GG repeat one are in gray, the carbon atoms of the second GG repeat in green, and the ones of the third repeat in yellow. The interactions of repeat one with the bound Ca^{2+} ion are indicated by gray dashed lines, and the interaction of the third repeat with the bound Ca^{2+} ions is in yellow. As it is evident, one Ca^{2+} ion is coordinated by repeat *n* and repeat *n* + 2. **(B)** RTX domain of alkaline protease from *P. aeruginosa* in cartoon representation. The orientation is identical to that in panel A, and the gray and yellow dashed lines indicate the interactions.
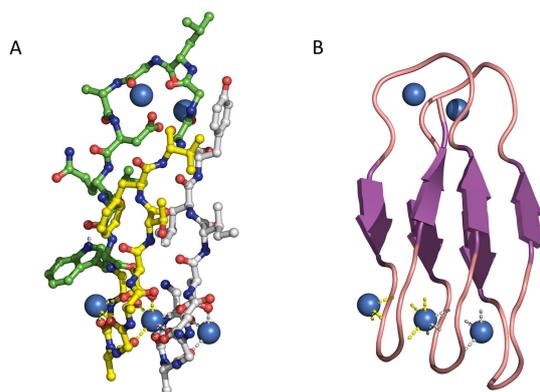
Finally, biophysical studies with the isolated nucleotide binding domain of HlyB, the ABC transporter of the HlyA secretion machinery, demonstrated an interaction with the substrate in the low micromolar range that required the secretion signal ([59]).

As soon as the outer membrane protein is engaged and a continuous channel tunnel has formed, the substrate enters the translocation pathway ([Fig. 3A]). For HlyA, it was experimentally demonstrated that secretion is directional, with the C terminus extended first onto the cell surface ([60]). Furthermore, the entire process proceeds with a secretion rate of 16 amino acids per transporter per second ([61]). At this stage, Ca^{2+} ions must bind to the RTX motifs and induce folding as soon as the substrate appears at the cell surface ([Fig. 3A]). This should prevent backsliding of the entire protein. Interestingly, reducing the external Ca^{2+} concentration below the dissociation constant of the ion from the RTX motif did not reduce the secretion rate in the HlyA system ([61]). This clearly demonstrates that the secretion rate is independent of Ca^{2+} and that Ca^{2+}-induced folding does not represent a driving force for secretion. A seemingly different scenario was observed for the much larger adenylate cyclase toxin (CyaA) from *B. pertussis* ([62]). A Ca^{2+} concentration of 2 mM in the media accelerated the efficiency of secretion. However, even when the Ca^{2+} concentration was reduced

FIGURE 3 Schematic summary of the classic T1SS-mediated substrate secretion **(A)** and the recently discovered secretion mechanism for some RTX adhesins in which secretion stalls just before completion, creating a so-called two-step process with a pseudoperiplasmic intermediate **(B)**. The ABC transporter and the MFP are shown in blue and green, respectively, and the OM protein is in maroon. **(A)** The unfolded substrate is secreted with its C terminus first. At the cell surface, Ca^{2+} ions (blue spheres) bind to the GG repeats and induce folding, which results in formation of the β-roll (indicated in cartoon representation). **(B)** In the case of adhesins such as IBA or LapA, the N-terminal domain starts folding prior to or during secretion, which plugs the translocon (indicated by the light brown polygon) and tethers the entire substrate at the cell surface within the OM component of the translocon of the T1SS. The brown cubes and distorted ellipse represent folded domains of the substrate. This scheme clearly demonstrates that the classic T1SS disassembles only after the entire substrate is translocated, while in two-step T1SS disassembly earlier, e.g., when the N-terminal plug domain has not passed the OM. For further details, see the text. IM, inner membrane; NBD, nucleotide binding domain; TMD, transmembrane domain.

Spitz et al.
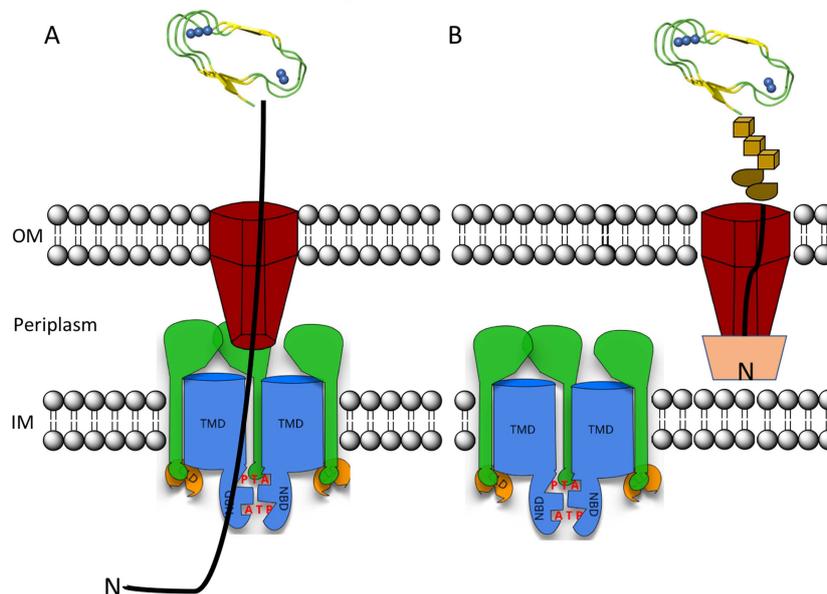
to 0.1 mM (which does not allow folding of CyaA), 50% of produced CyaA still reached the cell surface. These differences might be due to the diverging sizes of the two RTX proteins or details of the architecture of the RTX domain. Thus, further experimental approaches and analysis of additional T1SS substrates are required to completely understand the molecular mechanisms of secretion, the influence of Ca^{2+} ion folding and secretion (if any), and the molecular signals that regulate substrate translocation across two membranes in one step for this group of classic T1SS substrates.

RTX ADHESINS—NEW KIDS ON THE BLOCK

RTX proteins include not only toxins but also lipases, S-layer proteins, MARTX, and adhesins, which are extremely large in size. Recently, the structure of an ice-binding adhesin (IBA) of the marine Gram-negative bacterium *Marinomonas primoryensis* (molecular mass, 1.5 MDa) was determined and the putative mechanism of translocation modeled (37). IBA contains the hallmarks of substrates of a T1SS, an RTX domain and a C-terminal secretion sequence. N terminal to the RTX domain, three additional domains are located, namely, peptide-, sugar-, and ice-binding domains. While interactions of the peptide- and sugar-binding domains with surface receptors of other microorganisms allow formation of mixed aggregates of microorganisms, the ice-binding domain anchors *M. primoryensis* to ice in seas, lakes, or rivers. Surprisingly, and in contrast to the classical RTX proteins described earlier that are directly secreted into the extracellular space, IBA is translocated but then retained on the cell surface (Fig. 3B). Guo et al. (37) identified a conserved region (homologous to RM in LapA described above) at the extreme N terminus of IBA that they proposed could plug the channel-tunnel of the T1SS. Based on their structural analysis, Guo et al. proposed that this N-terminal region forms two domains: a proximal sequence that folds and a distal region that is sufficiently unfolded to traverse the TolC homologue into the outer membrane. This prevents further translocation and retains the adhesin at the cell surface. In 2018, exciting new data concerning the LapA adhesin from *Pseudomonas fluorescens* (3; for a summary, see reference 38) provided direct experimental evidence for this model in a comprehensive multidisciplinary study. The 160-residue RM was shown to be essential to tether the adhesin to the translocator and thus to the cell surface (Fig. 3B). Moreover, the RM consists of two domains, folded and unfolded, with the former specifically cleaved by a dedicated protease, LapG, to release the

adhesin. In another exciting twist, LapG is normally inactivated by binding to its membrane receptor, LapD. Binding is controlled by c-di-GMP to favor binding under conditions suitable for biofilm formation (63). Finally, we note that both of these studies (for IBA and LapA) confirm the directionality of translocation, C terminal first, for T1SS secretion.

Finally, it must be stressed that while IBA and LapA are anchored to the surface by stalling translocation, other strategies are used to retain adhesins at the cell surface. SiiE from *Salmonella enterica* contains a putative coiled-coil motif that facilitates immobilization of the entire protein on the surface of the cell envelope, which is controlled by SiiA and SiiB (36). Thus, continued efforts are needed to see whether additional mechanisms and modifications of the classic T1SS exist that are used by Gram-negative bacteria to cope with the demands of their ecological niches.

SUMMARY AND OUTLOOK

An enormous amount of data on T1SS has been gathered since the discovery of the first system. The amount of available structural information on the components of the translocon machinery is increasing constantly. These components include the OM protein TolC (64), a closely related homologue of HlyC (21), isolated domains of the ABC transporter HlyB (65–67) and other ABC transporters (68), a soluble fragment of the MFP HlyD (69), and an entire structure of an ABC transporter (70) of a putative T1SS with unknown substrate from *Aquifex aeolicus*. This article is unable to cover all aspects of type I secretion; however, it provides a broad summary of the accumulating data on functional aspects of the secretion process. The review does not engage with the possibilities of the T1SS in biotechnological applications (for a recent summary, see reference 41) that go well beyond basic research and would allow large-scale protein production and isolation via protein secretion.

However, we are still some distance from a systematic understanding of the T1SS since the nature of molecular signals and intramolecular communication within this nanomachinery remains unclear. In summary, there are still many open questions that have to be addressed and many more fascinating variations and novel insights to be discovered for the T1SS in Gram-negative bacteria.

REFERENCES

1. Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* 13:343–359. http://dx.doi.org/10.1038/nrmicro3456.

2. Veith PD, Glew MD, Gorasia DG, Reynolds EC. 2017. Type IX secretion: the generation of bacterial cell surface coatings involved in virulence, gliding motility and the degradation of complex biopolymers. *Mol Microbiol* 106:35–53. http://dx.doi.org/10.1111/mmi.13752.

3. Smith TJ, Font ME, Kelly CM, Sondermann H, O'Toole GA. 2018. An N-terminal retention module anchors the giant adhesin LapA of *Pseudomonas fluorescens* at the cell surface: a novel sub-family of type I secretion systems. *J Bacteriol* 200:e00734-17. http://dx.doi.org/10.1128/JB.00734-17.

4. Linhartová I, Bumba L, Mašín J, Basler M, Osička R, Kamanová J, Procházková K, Adkins I, Hejnová-Holubová J, Sadílková L, Morová J, Sebo P. 2010. RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol Rev* 34:1076–1112. http://dx.doi.org/10.1111/j.1574-6976.2010.00231.x.

5. Felmlee T, Welch RA. 1988. Alterations of amino acid repeats in the *Escherichia coli* hemolysin affect cytolytic activity and secretion. *Proc Natl Acad Sci U S A* 85:5269–5273. http://dx.doi.org/10.1073/pnas.85.14.5269.

6. Griessl MH, Schmid B, Kassler K, Braunsmann C, Ritter R, Barlag B, Stierhof YD, Sturm KU, Danzer C, Wagner C, Schäffer TE, Sticht H, Hensel M, Muller YA. 2013. Structural insight into the giant Ca^{2+}-binding adhesin SiiE: implications for the adhesion of *Salmonella enterica* to polarized epithelial cells. *Structure* 21:741–752. http://dx.doi.org/10.1016/j.str.2013.02.020.

7. Håvarstein LS, Diep DB, Nes IF. 1995. A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol Microbiol* 16:229–240. http://dx.doi.org/10.1111/j.1365-2958.1995.tb02295.x.

8. Håvarstein LS, Holo H, Nes IF. 1994. The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria. *Microbiology* 140:2383–2389. http://dx.doi.org/10.1099/13500872-140-9-2383.

9. Michiels J, Dirix G, Vanderleyden J, Xi C. 2001. Processing and export of peptide pheromones and bacteriocins in Gram-negative bacteria. *Trends Microbiol* 9:164–168. http://dx.doi.org/10.1016/S0966-842X(01)01979-5.

10. Choudhury HG, Tong Z, Mathavan I, Li Y, Iwata S, Zirah S, Rebuffat S, van Veen HW, Beis K. 2014. Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. *Proc Natl Acad Sci U S A* 111:9145–9150. http://dx.doi.org/10.1073/pnas.1320506111.

11. Husada F, Bountra K, Tassis K, de Boer M, Romano M, Rebuffat S, Beis K, Cordes T. 2018. Conformational dynamics of the ABC transporter McjD seen by single-molecule FRET. *EMBO J* 37:e100056. http://dx.doi.org/10.15252/embj.2018100056.

12. Felmlee T, Pellett S, Welch RA. 1985. Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. *J Bacteriol* 163:94–105.

13. Härtlein M, Schiessl S, Wagner W, Rdest U, Kreft J, Goebel W. 1983. Transport of hemolysin by *Escherichia coli*. *J Cell Biochem* 22:87–97. http://dx.doi.org/10.1002/jcb.240220203.

14. Noegel A, Rdest U, Springer W, Goebel W. 1979. Plasmid cistrons controlling synthesis and excretion of the exotoxin alpha-haemolysin of *Escherichia coli*. *Mol Gen Genet* 175:343–350. http://dx.doi.org/10.1007/BF00397234.

15. Springer W, Goebel W. 1980. Synthesis and secretion of hemolysin by *Escherichia coli*. *J Bacteriol* 144:53–59.

16. Nicaud JM, Mackman N, Gray L, Holland IB. 1985. Characterisation of HlyC and mechanism of activation and secretion of haemolysin from *E. coli* 2001. *FEBS Lett* 187:339–344. http://dx.doi.org/10.1016/0014-5793(85)81272-2.

17. Stanley P, Hyland C, Koronakis V, Hughes C. 1999. An ordered reaction mechanism for bacterial toxin acylation by the specialized acyltransferase HlyC: formation of a ternary complex with acylACP and protoxin substrates. *Mol Microbiol* 34:887–901. http://dx.doi.org/10.1046/j.1365-2958.1999.01648.x.

18. Stanley P, Packman LC, Koronakis V, Hughes C. 1994. Fatty acylation of two internal lysine residues required for the toxic activity of *Escherichia coli* hemolysin. *Science* 266:1992–1996. http://dx.doi.org/10.1126/science.7801126.

19. Trent MS, Worsham LM, Ernst-Fonberg ML. 1998. The biochemistry of hemolysin toxin activation: characterization of HlyC, an internal protein acyltransferase. *Biochemistry* 37:4644–4652. http://dx.doi.org/10.1021/bi971588y.

20. Issartel JP, Koronakis V, Hughes C. 1991. Activation of *Escherichia coli* prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. *Nature* 351:759–761. http://dx.doi.org/10.1038/351759a0.

21. Greene NP, Crow A, Hughes C, Koronakis V. 2015. Structure of a bacterial toxin-activating acyltransferase. *Proc Natl Acad Sci U S A* 112:E3058–E3066. http://dx.doi.org/10.1073/pnas.1503832112.

22. Wandersman C, Delepelaire P. 1990. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc Natl Acad Sci U S A* 87:4776–4780. http://dx.doi.org/10.1073/pnas.87.12.4776.

23. Chervaux C, Sauvonnet N, Le Clainche A, Kenny B, Hung AL, Broome-Smith JK, Holland IB. 1995. Secretion of active beta-lactamase to the medium mediated by the *Escherichia coli* haemolysin transport pathway. *Mol Gen Genet* 249:237–245. http://dx.doi.org/10.1007/BF00290371.

24. Gentschev I, Hess J, Goebel W. 1990. Change in the cellular localization of alkaline phosphatase by alteration of its carboxy-terminal sequence. *Mol Gen Genet* 222:211–216. http://dx.doi.org/10.1007/BF00633820.

25. Gray L, Baker K, Kenny B, Mackman N, Haigh R, Holland IB. 1989. A novel C-terminal signal sequence targets *Escherichia coli* haemolysin directly to the medium. *J Cell Sci Suppl* 11:45–57. http://dx.doi.org/10.1242/jcs.1989.Supplement_11.4.

26. Gray L, Mackman N, Nicaud JM, Holland IB. 1986. The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*. *Mol Gen Genet* 205:127–133. http://dx.doi.org/10.1007/BF02428042.

27. Koronakis V, Koronakis E, Hughes C. 1989. Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes. *EMBO J* 8:595–605. http://dx.doi.org/10.1002/j.1460-2075.1989.tb03414.x.

28. Mackman N, Baker K, Gray L, Haigh R, Nicaud JM, Holland IB. 1987. Release of a chimeric protein into the medium from *Escherichia coli* using the C-terminal secretion signal of haemolysin. *EMBO J* 6:2835–2841. http://dx.doi.org/10.1002/j.1460-2075.1987.tb02580.x.

29. Mackman N, Holland IB. 1984. Functional characterization of a cloned haemolysin determinant from *E. coli* of human origin, encoding information for the secretion of a 107K polypeptide. *Mol Gen Genet* 196:129–134. http://dx.doi.org/10.1007/BF00334104.

30. Mackman N, Nicaud JM, Gray L, Holland IB. 1985. Genetical and functional organisation of the *Escherichia coli* haemolysin determinant 2001. *Mol Gen Genet* 201:282–288. http://dx.doi.org/10.1007/BF00425672.

31. Mackman N, Nicaud JM, Gray L, Holland IB. 1985. Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*. *Mol Gen Genet* 201:529–536. http://dx.doi.org/10.1007/BF00331351.

32. Goyard S, Sebo P, D'Andria O, Ladant D, Ullmann A. 1993. *Bordetella pertussis* adenylate cyclase: a toxin with multiple talents. *Zentralbl Bakteriol* 278:326–333. http://dx.doi.org/10.1016/S0934-8840(11)80849-2.

33. Satchell KJ. 2011. Structure and function of MARTX toxins and other large repetitive RTX proteins. *Annu Rev Microbiol* 65:71–90. http://dx .doi.org/10.1146/annurev-micro-090110-102943.

34. Satchell KJF. 2015. Multifunctional-autoprocessing repeats-in-toxin (MARTX) toxins of vibrios. *Microbiol Spectr* 3(3):VE-0002-2014. http:// dx.doi.org/10.1128/microbiolspec.VE-0002-2014.

35. Kim BS, Gavin HE, Satchell KJ. 2015. Distinct roles of the repeat-containing regions and effector domains of the *Vibrio vulnificus* multifunctional-autoprocessing repeats-in-toxin (MARTX) toxin. *mBio* 6: e00324-15. http://dx.doi.org/10.1128/mBio.00324-15.

36. Barlag B, Hensel M. 2015. The giant adhesin SiiE of *Salmonella enterica*. *Molecules* 20:1134–1150. http://dx.doi.org/10.3390/molecules20011134.

37. Guo S, Stevens CA, Vance TDR, Olijve LLC, Graham LA, Campbell RL, Yazdi SR, Escobedo C, Bar-Dolev M, Yashunsky V, Braslavsky I, Langelaan DN, Smith SP, Allingham JS, Voets IK, Davies PL. 2017. Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. *Sci Adv* 3:e1701440. http://dx.doi.org/10.1126/sciadv .1701440.

38. Smith TJ, Sondermann H, O'Toole GA. 2018. Type 1 does the two-step: type 1 secretion substrates with a functional periplasmic intermediate. *J Bacteriol* 200:e00168-18. http://dx.doi.org/10.1128/JB.00168-18.

39. Peters B, Stein J, Klingl S, Sander N, Sandmann A, Taccardi N, Sticht H, Gerlach RG, Muller YA, Hensel M. 2017. Structural and functional dissection reveals distinct roles of Ca2+-binding sites in the giant adhesin SiiE of *Salmonella enterica*. *PLoS Pathog* 13:e1006418. http://dx.doi.org /10.1371/journal.ppat.1006418.

40. Wagner C, Polke M, Gerlach RG, Linke D, Stierhof YD, Schwarz H, Hensel M. 2011. Functional dissection of SiiE, a giant non-fimbrial adhesin of *Salmonella enterica*. *Cell Microbiol* 13:1286–1301. http://dx.doi .org/10.1111/j.1462-5822.2011.01621.x.

41. Holland IB, Peherstorfer S, Kanonenberg K, Lenders M, Reimann S, Schmitt L. 2016. Type I protein secretion—deceptively simple yet with a wide range of mechanistic variability across the family. *EcoSal Plus* 7:ESP-0019-2015. http://dx.doi.org/10.1128/ecosalplus.ESP-0019-2015.

42. Cao G, Kuhn A, Dalbey RE. 1995. The translocation of negatively charged residues across the membrane is driven by the electrochemical potential: evidence for an electrophoresis-like membrane transfer mechanism. *EMBO J* 14:866–875. http://dx.doi.org/10.1002/j.1460-2075.1995 .tb07068.x.

43. Baumann U, Wu S, Flaherty KM, McKay DB. 1993. Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif. *EMBO J* 12:3357–3364. http://dx.doi.org/10.1002/j.1460-2075.1993.tb06009.x.

44. Baumann U, Bauer M, Létoffé S, Delepelaire P, Wandersman C. 1995. Crystal structure of a complex between *Serratia marcescens* metallo-protease and an inhibitor from *Erwinia chrysanthemi*. *J Mol Biol* 248:653–661. http://dx.doi.org/10.1006/jmbi.1995.0249.

45. Meier R, Drepper T, Svensson V, Jaeger KE, Baumann U. 2007. A calcium-gated lid and a large beta-roll sandwich are revealed by the crystal structure of extracellular lipase from *Serratia marcescens*. *J Biol Chem* 282:31477–31483. http://dx.doi.org/10.1074/jbc.M704942200.

46. Ostolaza H, Soloaga A, Goñi FM. 1995. The binding of divalent cations to *Escherichia coli* alpha-haemolysin. *Eur J Biochem* 228:39–44.

47. Sánchez-Magraner L, Viguera AR, García-Pacios M, Garcillán MP, Arrondo JL, de la Cruz F, Goñi FM, Ostolaza H. 2007. The calcium-binding C-terminal domain of *Escherichia coli* alpha-hemolysin is a major determinant in the surface-active properties of the protein. *J Biol Chem* 282:11827–11835. http://dx.doi.org/10.1074/jbc.M700547200.

48. Soloaga A, Ramírez JM, Goñi FM. 1998. Reversible denaturation, self-aggregation, and membrane activity of *Escherichia coli* alpha-hemolysin, a protein stable in 6 M urea. *Biochemistry* 37:6387–6393. http://dx.doi.org /10.1021/bi9730994.

49. Thomas S, Bakkes PJ, Smits SH, Schmitt L. 2014. Equilibrium folding of pro-HlyA from *Escherichia coli* reveals a stable calcium ion dependent

folding intermediate. *Biochim Biophys Acta* 1844:1500–1510. http://dx .doi.org/10.1016/j.bbapap.2014.05.006.

50. Blenner MA, Shur O, Szilvay GR, Cropek DM, Banta S. 2010. Calcium-induced folding of a beta roll motif requires C-terminal entropic stabilization. *J Mol Biol* 400:244–256. http://dx.doi.org/10.1016/j.jmb.2010.04 .056.

51. Chenal A, Guijarro JI, Raynal B, Delepierre M, Ladant D. 2009. RTX calcium binding motifs are intrinsically disordered in the absence of calcium: implication for protein secretion. *J Biol Chem* 284:1781–1789. http://dx.doi.org/10.1074/jbc.M807312200.

52. Sotomayor Pérez AC, Karst JC, Davi M, Guijarro JI, Ladant D, Chenal A. 2010. Characterization of the regions involved in the calcium-induced folding of the intrinsically disordered RTX motifs from the *Bordetella pertussis* adenylate cyclase toxin. *J Mol Biol* 397:534–549. http://dx.doi .org/10.1016/j.jmb.2010.01.031.

53. Zhang L, Conway JF, Thibodeau PH. 2012. Calcium-induced folding and stabilization of the *Pseudomonas aeruginosa* alkaline protease. *J Biol Chem* 287:4311–4322. http://dx.doi.org/10.1074/jbc.M111.310300.

54. Jones HE, Holland IB, Baker HL, Campbell AK. 1999. Slow changes in cytosolic free Ca2+ in *Escherichia coli* highlight two putative influx mechanisms in response to changes in extracellular calcium. *Cell Calcium* 25:265–274. http://dx.doi.org/10.1054/ceca.1999.0028.

55. Bakkes PJ, Jenewein S, Smits SH, Holland IB, Schmitt L. 2010. The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin type 1 secretion system. *J Biol Chem* 285:40573–40580. http://dx .doi.org/10.1074/jbc.M110.173658.

56. Balakrishnan L, Hughes C, Koronakis V. 2001. Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *J Mol Biol* 313:501–510. http://dx.doi.org/10.1006 /jmbi.2001.5038.

57. Thanabalu T, Koronakis E, Hughes C, Koronakis V. 1998. Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *EMBO J* 17:6487–6496. http://dx.doi.org/10.1093/emboj/17.22 .6487.

58. Du D, Wang Z, James NR, Voss JE, Klimont E, Ohene-Agyei T, Venter H, Chiu W, Luisi BF. 2014. Structure of the AcrAB-TolC multidrug efflux pump. *Nature* 509:512–515. http://dx.doi.org/10.1038/nature13205.

59. Benabdelhak H, Kiontke S, Horn C, Ernst R, Blight MA, Holland IB, Schmitt L. 2003. A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A. *J Mol Biol* 327:1169–1179. http://dx.doi.org/10.1016/S0022 -2836(03)00204-3.

60. Lenders MHH, Weidtkamp-Peters S, Kleinschrodt D, Jaeger K-E, Smits SHJ, Schmitt L. 2015. Directionality of substrate translocation of the hemolysin A type I secretion system. *Sci Rep* 5:12470. http://dx.doi .org/10.1038/srep12470.

61. Lenders MH, Beer T, Smits SH, Schmitt L. 2016. In vivo quantification of the secretion rates of the hemolysin A type I secretion system. *Sci Rep* 6:33275. http://dx.doi.org/10.1038/srep33275.

62. Bumba L, Masin J, Macek P, Wald T, Motlova L, Bibova I, Klimova N, Bednarova L, Veverka V, Kachala M, Svergun DI, Barinka C, Sebo P. 2016. Calcium-driven folding of RTX domain β-rolls ratchets translocation of RTX proteins through type I secretion ducts. *Mol Cell* 62:47–62. http://dx.doi.org/10.1016/j.molcel.2016.03.018.

63. Monds RD, Newell PD, Gross RH, O'Toole GA. 2007. Phosphate-dependent modulation of c-di-GMP levels regulates *Pseudomonas fluorescens* Pf0-1 biofilm formation by controlling secretion of the adhesin LapA. *Mol Microbiol* 63:656–679. http://dx.doi.org/10.1111/j.1365-2958 .2006.05539.x.

64. Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C. 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405:914–919. http://dx.doi.org/10.1038 /35016007.

65. Lecher J, Schwarz CK, Stoldt M, Smits SH, Willbold D, Schmitt L. 2012. An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. *Structure* 20:1778–1787. http://dx.doi .org/10.1016/j.str.2012.08.005.

66. Zaitseva J, Jenewein S, Jumpertz T, Holland IB, Schmitt L. 2005. H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *EMBO J* 24:1901–1910. http://dx.doi.org /10.1038/sj.emboj.7600657.

67. Zaitseva J, Oswald C, Jumpertz T, Jenewein S, Wiedenmann A, Holland IB, Schmitt L. 2006. A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer. *EMBO J* 25:3432–3443. http://dx.doi.org/10.1038/sj.emboj.7601208.

68. Murata D, Okano H, Angkawidjaja C, Akutsu M, Tanaka SI, Kitahara K, Yoshizawa T, Matsumura H, Kado Y, Mizohata E, Inoue T, Sano S, Koga Y, Kanaya S, Takano K. 2017. Structural basis for the *Serratia marcescens* lipase secretion system: crystal structures of the membrane fusion protein and nucleotide-binding domain. *Biochemistry* 56:6281–6291. http://dx.doi.org/10.1021/acs.biochem.7b00985.

69. Kim JS, Song S, Lee M, Lee S, Lee K, Ha NC. 2016. Crystal structure of a soluble fragment of the membrane fusion protein HlyD in a type i secretion system of Gram-negative bacteria. *Structure* 24:477–485. http:// dx.doi.org/10.1016/j.str.2015.12.012.

70. Morgan JLW, Acheson JF, Zimmer J. 2017. Structure of a type-1 secretion system ABC transporter. *Structure* 25:522–529. http://dx.doi .org/10.1016/j.str.2017.01.010.

71. Gerlach RG, Jäckel D, Stecher B, Wagner C, Lupas A, Hardt WD, Hensel M. 2007. *Salmonella* pathogenicity island 4 encodes a giant non-fimbrial adhesin and the cognate type 1 secretion system. *Cell Microbiol* 9:1834–1850. http://dx.doi.org/10.1111/j.1462-5822.2007 .00919.x.

3.3 Chapter 3 – *In silico* analyses of HlyA T1SS components

Title: *In silico* analyses of HlyA T1SS components

Authors: **Olivia Spitz** and Lutz Schmitt

Published in: *to be submitted*

Own Work: 90 %

In silico analyses (alignments, predictions, modeling)

Expression and localization of HlyD CD

Preparation of the figures

Writing of the manuscript

In silico analyses of HlyA T1SS components

Olivia Spitz and Lutz Schmitt

Institute of Biochemistry, Heinrich-Heine-University Düsseldorf

Corresponding author: Lutz Schmitt, Institute of Biochemistry, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, E-Mail: Lutz.Schmitt@hhu.de

Key words: HlyA T1SS, amphipathic helix, EISKIIS motif

Abstract

Protein structures are a powerful basis to investigate protein function and interaction. However, structure determination still requires large amounts of purified protein or protein complexes, which can be hard to obtain. *In silico* tools circumvent this problem by comparing protein sequences and assigning already identified protein features to similar proteins. With constantly growing numbers of protein structures in databases the accuracy of these predictions is steadily increasing, but information drawn from these tools still requires critical evaluation.

We used a combination of secondary structure prediction tools, 3D modeling and alignments to analyze data on the hemolysin A (HlyA) Type 1 Secretion System (T1SS) of *Escherichia coli* with a main focus on the initial interactions between the substrate HlyA and the transport components of the inner membrane, HlyB and HlyD. Our findings strongly support the presence of an amphipathic helix (AH) in the extreme C-terminus of HlyA as well as in the cytoplasmic domain (CD) of HlyD. Two possible binding pockets for the AH of HlyA were identified in the nucleotide binding domain (NBD) of the ABC transporter HlyB and mapped on a 3D model of the transporter. Furthermore a conserved motif in the membrane fusion protein HlyD was identified and the interactions of HlyA with the different domains of HlyB and HlyD are discussed separately.

Introduction

Type I secretion systems (T1SSs) are common in Gram-negative bacteria and transport a variety of substrates among those repeats-in-toxin (RTX) proteins (Holland et al., 2016). Although our knowledge about this secretion process is constantly increasing a lot of questions regarding the initiation of this process remain unanswered. The hemolysin A (HlyA) T1SS from *Escherichia coli* is one of the best-studied T1SSs and also the focus of this study.

The substrate of this system, HlyA, belongs to the family of RTX proteins and has a size of 110 kDa or 1024 amino acids (aa) (Felmlee et al., 1985, Welch, 1991). It is able to lyse several cell types, among them erythrocytes, by inserting its N-terminal domain into the membrane of target cells (Ludwig et al., 1991). For this activity acylation of two internal lysine residues prior to secretion by the acyltransferase HlyC is necessary (Stanley et al., 1994). This acylation only affects the activity of HlyA but not its secretion, since pro-HlyA, the non-acylated version, is secreted with the same efficiency as the wild type (Nicaud et al., 1985).

Downstream of this functional, N-terminal domain the RTX domain is located, which gave rise to the name of this protein family. A motif of nine conserved residues is repeated several times in this domain and was termed "GG repeat" due to its consensus sequence: GGxGxDxUx (with x standing for any aa and U for a large hydrophobic aa) (Linhartová et al., 2010). These GG repeats bind Ca^{2+} and induce the folding of the whole protein (Baumann et al., 1993, Thomas et al., 2014). In HlyA the GG repeats have an approximal K_D of ~150 µM Ca^{2+}, which suggests that Ca^{2+} binding and subsequent folding of the protein only happens after secretion, since Ca^{2+} concentration in the cytosol of *E. coli* is only around 90-270 nM and strictly controlled (Gangola and Rosen, 1987, Jones et al., 1999, Sánchez-Magraner et al., 2007).

The secretion signal of HlyA has been localized to the C-terminal 50 to 60 amino acids (Gray et al., 1986, Hess et al., 1990, Jarchau et al., 1994), which reach the cell surface first in the secretion process (Lenders et al., 2015). Neither the

secretion signal nor any other part of HlyA is cleaved during the secretion process (Felmlee et al., 1985).

The translocation channel of T1SSs generally consists out of three membrane proteins: An outer membrane protein (OMP) and two proteins residing in the inner membrane (IM), which belong to the family of ABC transporters and membrane fusion proteins (MFPs). For the HlyA T1SS these proteins are TolC (OMP), HlyB (ABC transporter) and HlyD (MFP) (Mackman et al., 1985, Wandersman and Delepelaire, 1990). The structure of TolC has been solved by crystallography in 2000 and supports the notion that HlyA is transported unfolded (Koronakis et al., 2000).

The two proteins of the IM, HlyB and HlyD, form a stable complex in the absence of the substrate and only recruit TolC upon substrate engagement (Thanabalu et al., 1998). To this date it remains unknown which interactions lead to the recruitment of TolC and subsequent assembly of the whole transport channel. However, the importance of several domains has been shown by deletion of these, which resulted in a lack of HlyA secretion. For example, the deletion of the N-terminal extension (CLD) of HlyB as well as the deletion of the 40 N-terminal residues of HlyD results in a lack of secretion and in intracellular accumulation of HlyA (Pimenta et al., 1999, Lecher et al., 2012).

The interaction of the CLD of HlyB with HlyA has been shown by pull-down experiments with isolated proteins and the region of the CLD that interacts with the substrate has been mapped by chemical shift experiments (Lecher et al., 2012). Furthermore the nucleotide binding domain (NBD) of HlyB has been shown to interact with the HlyA secretion signal by surface plasmon resonance (SPR) studies (Benabdelhak et al., 2003). Both of these interactions only take place with the unfolded substrate. Additionally crosslinking experiments showed that HlyA is also able to interact with HlyD in the absence of HlyB (Thanabalu et al., 1998).

While the structure of the isolated CLD and NBD of HlyB have been solved by NMR (CLD) and crystallography (NBD) (Schmitt et al., 2003, Lecher et al., 2012), the structure of the cytoplasmic domain (CD) of HlyD as well as the structure of HlyA remain not solved and can only be predicted using *in silico* methods. Amphipathic

helices (AHs) have been predicted in both, the CD of HlyD and in the secretion signal of HlyA (Koronakis et al., 1989, Balakrishnan et al., 2001). While only a few studies exist on the CD of HlyD (Pimenta et al., 1999, Balakrishnan et al., 2001), the secretion signal of HlyA has been studied in great detail (Stanley et al., 1991, Kenny et al., 1992, Zhang et al., 1993, Chervaux and Holland, 1996). Still, no consensus sequence could be identified for the secretion signal and the recognition of the substrate was therefore contributed to the secondary structure of this region. In the most recent study a consensus motif in the predicted AH of HlyA was identified by aligning the last 100 aa of 40 RTX toxins and this motif was termed 'EISKIIS-motif' due to its primary sequence (Kanonenberg, 2018). Proline mutations and mutations identified in other studies (Chervaux and Holland, 1996) were introduced to this region to disrupt the AH. All mutants displayed the same hemolytic activity and their secretion rates were determined and compared to wild type HlyA (Lenders et al., 2016, Kanonenberg, 2018).

We analyzed these mutants in regards to their secretion rate phenotype with a combination of *in silico* tools, including secondary structure prediction and peptide modeling, and our findings strongly support the presence of an AH at this position. Furthermore we analyzed the available structures of the NBD of HlyB and found two possible interaction sites for the AH of HlyA. These interaction sites as well as the previously identified HlyA interaction site of the CLD were mapped onto a model of HlyB. The CD of HlyD was also subjected to *in silico* analyses and the possible interactions with HlyA are discussed.

Results and Discussion

Evaluation of HlyA amphipathic helix (AH)

The amphipathic helix (AH) in the C-terminal secretion signal of HlyA was proposed to be located between residue L973 and F990 (Koronakis et al., 1989). The secretion rate of 21 mutants in this region has been previously determined with eleven mutants displaying a strongly reduced secretion rate (Figure 1 A) (Kanonenberg, 2018). Ten of these eleven mutants are proline substitutions,

however, there are single proline mutations that display the same secretion rate as the wild type, for example E979P. We analyzed the wild type and all 21 mutants with a combination of *in silico* tools, which included secondary structure prediction by two different tools (AmphipaSeeK and Quick2D) as well as peptide modeling (PEP-FOLD3). For HlyA wild type AmphipaSeeK, all five models provided by PEP-FOLD3 and all four prediction methods used by Quick2D predicted the same AH starting at residue P975 ± 1 and ending at residue A986 ± 1 (Figure 1 B and C).
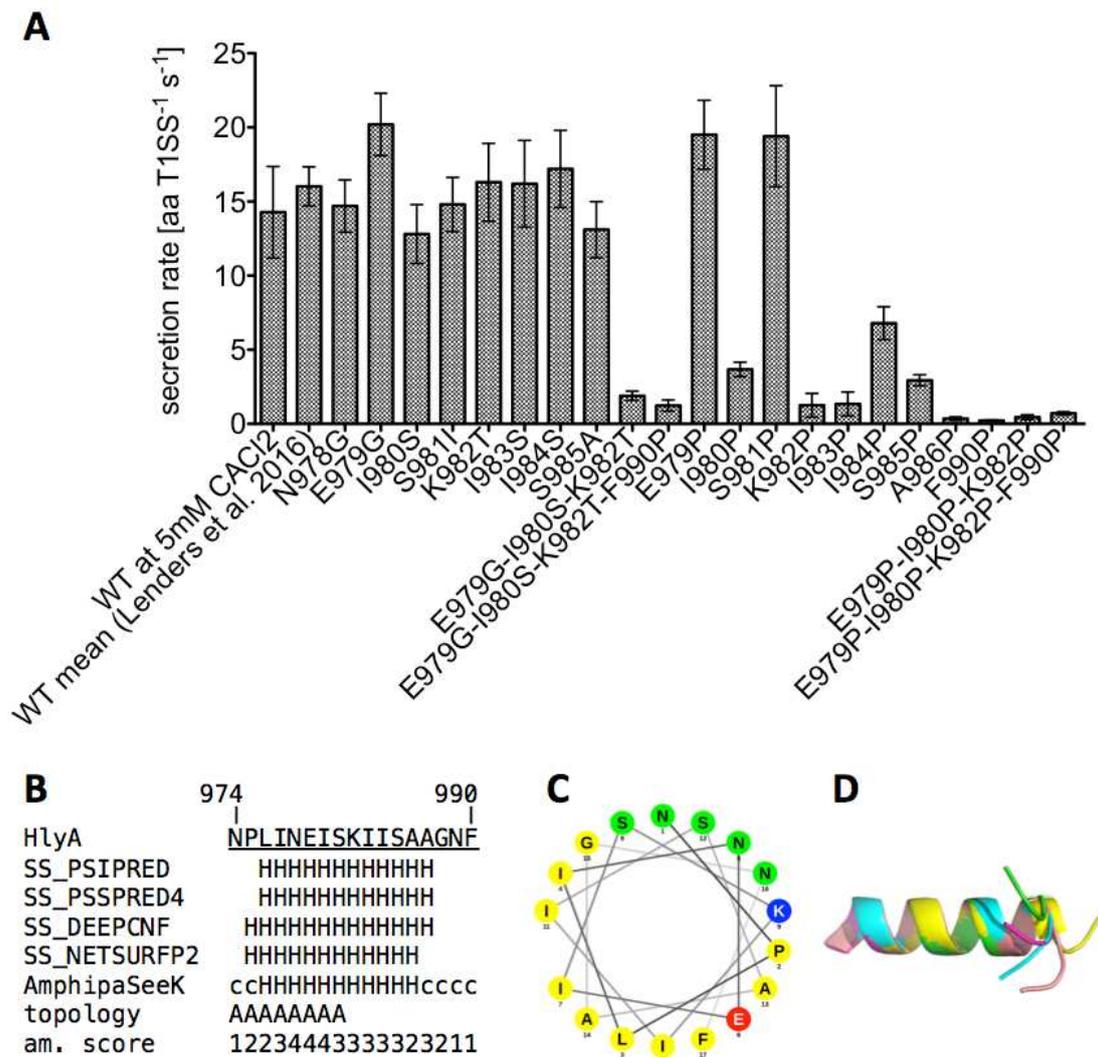
Figure 1: A) Secretion rates of HlyA mutants (Kanonenberg, 2018). B-D) *In silico* **analyses of HlyA amphipathic helix (AH). B)** Results of secondary structure prediction with different tools. H = α-helix, c = coiled coil. Predictions labeled with "SS" are from Quick2D, that utilizes multiple prediction algorithms. Topology = predicted by AmphipaSeeK: (A) marks residues that are predicted to be inserted parallel into the membrane. Am. score = amphipathy score predicted by AmphipaSeeK with 1 = low amphipathy and 5 = high amphipathy. **C)** Helical wheel projection of residue 974 – 990 of HlyA. Non-polar residues are colored yellow, lysine blue, glutamine red and polar residues green. **D)** Overlay

of five PEP-FOLD3 models of residue 974 – 990 of HlyA. The helix is similar in all five models with the C-terminal tail sticking out in different directions.

The single mutants that were secreted with the same secretion rate as the wild type displayed almost identical helix predictions, which are shown for the single proline mutants E979P and S981P (Figure 2 A and B). Proline is known as a so-called helix breaker, due to its unique conformation and rigid rotation. Its preferred position in helices is at the N-terminus (Richardson and Richardson, 1988) but helices with proline in or close at the center, as in the mutants E979P and S981P, are still possible (Kim and Kang, 1999).

The single-proline mutants I980P and I984P exhibited a reduced secretion rate of 3.7 ± 0.5 and 6.8 ± 1.1 aa T1SS^{-1}*s^{-1}, respectively, and the prediction programs predicted strong impairments of the AH in both cases (Figure 2 C and D).

The triple mutant E979G I980S K982T is of particular interest, since each single mutant displayed the same secretion rate (and helix prediction) as the wild type (Figure 1 A), while the triple mutant shows a strongly reduced rate of 1.9 ± 0.3 aa T1SS^{-1}*s^{-1}. The prediction tools are in agreement that a helix is unlikely with this primary amino acid sequence (Figure 2 E).

A E979P SR = 19.5 ± 2.3 aa T1SS^{-1} * s^{-1}

```
            974              990
             |                |
HlyA        NPLINPISKIISAAGNF
SS_PSIPRED      HHHHHHHHH
SS_PSSPRED4     HHHHHHHH
SS_DEEPCNF      HHHHHHHH
SS_NETSURFP2    HHHHHHH
AmphipaSeeK   ccHHHHHHHHHHHHcccc
topology      AAAAAAAA
am. score     12134443333323211
```

B S981P SR = 19.4 ± 3.4 aa T1SS^{-1} * s^{-1}

```
            974              990
             |                |
HlyA        NPLINEIPKIISAAGNF
SS_PSIPRED      HHHHHHHHHHH
SS_PSSPRED4     HHHHHHHHHHH
SS_DEEPCNF      HHHHHHHHHHH
SS_NETSURFP2    HHHHHHHHHH
AmphipaSeeK   ccHHHHHHHHHHHHcccc
topology      AAAAAAAA
am. score     12234443333323211
```

C I980P SR = 3.7 ± 0.5 aa T1SS^{-1} * s^{-1}

```
            974              990
             |                |
HlyA        NPLINEPSKIISAAGNF
SS_PSIPRED      HHHHH
SS_PSSPRED4     HHHHHH
SS_DEEPCNF      HHHHHHH
SS_NETSURFP2    HHHHHHH
AmphipaSeeK   cccccccccH??EHcccc
topology      AAAA
am. score     21134443322334311
```

D I984P SR = 6.8 ± 1.1 aa T1SS^{-1} * s^{-1}

```
            974              990
             |                |
HlyA        NPLINEISKIPSAAGNF
SS_PSIPRED      HHHHH
SS_PSSPRED4     HHHHHHH
SS_DEEPCNF      HHHHHH
SS_NETSURFP2    HHHHHHH   H
AmphipaSeeK   ccHHHHHcccccccccc
topology      AAAAAA
am. score     12245433322322211
```

E E979G I980S K982T SR = 1.9 ± 0.3 aa T1SS^{-1} * s^{-1}

```
            974              990
             |                |
HlyA        NPLINGSSTIISAAGNF
SS_PSIPRED       EEEE
SS_PSSPRED4      EEEEE
SS_DEEPCNF        EE
AmphipaSeeK   ccccccccEEEEEcccc
topology
am. score     21124333211234533
```

Figure 2: Secondary structure predictions of five different mutants of HlyA. Predictions labeled with "SS" are from Quick2D, that utilizes multiple prediction algorithms. Topology = predicted by AmphipaSeeK: (A) marks residues that are predicted to be inserted parallel into the membrane. Am. score = amphipathy score predicted by AmphipaSeeK with 1 = low amphipathy and 5 = high amphipathy. H = α-helix, c = coiled coil, E = β-sheet, ? = no prediction. Mutated residues are marked in red. Secretion rate (SR) is given for each mutant as mean ± SD of three independent measurements (Kanonenberg, 2018). **A** and **B**) Single proline mutations with SR like wild type. **C** and **D**) Single proline mutations with reduced SR. **E**) Triple mutant without proline with reduced SR.

The secondary structure prediction tools predict impairments of the AH for almost all mutants, that exhibit a reduced secretion rate. There are four mutants, whose secretion rate is strongly reduced, but a helix is still predicted: F990P, which is not part of the AH, K982P, S985P and A986P. However, their phenotype can be explained with the help of additional *in silico* tools.

The latter two, S985P and A986P, show a slightly shortened AHs in the predictions while A986 marks the end of the AH in the wild type (Figure 3 A and C). Another Ala residue and a Gly residue follow. This region ((S)AAG) is therefore flexible, which is reflected by the five different models from PEP-FOLD3 whose tails stick out in different directions (Figure 1 D). This flexibility is impaired if a Pro is introduced at this position, which likely explains the reduced secretion rate. In addition to the reduced flexibility, the polarity of the polar side of the AH is lessened for S985P which is evident in the helical wheel projection (Figure 3 D).

The mutant K982P also results in a change of polarity and displays a reduced secretion rate of 1.3 ± 0.8 aa $T1SS^{-1}*s^{-1}$. However, the mutant K982T, which eliminates the positive charge at this position as well, shows a secretion rate of 16.3 ± 2.6 aa $T1SS^{-1}*s^{-1}$, which is like the wild type (16.0 ± 1.3 aa $T1SS^{-1}*s^{-1}$ (Lenders et al., 2016)) (Figure 1 A). Therefore K982T shows that a positive charge at this position is not needed for the secretion process.

However, a proline at this position introduces a bend to the AH as seen in the PEP-FOLD3 models (Thévenet et al., 2012). The proline substitution at n-1 (S981P) also shows a bend of the AH but no impairment of the secretion rate within experimental error (19.4 ± 3.4 aa $T1SS^{-1}*s^{-1}$) (Figure 3 F). However, these two mutants bend in the opposite direction with S981P resembling the wild type more than K982P. This further supports the theory that the precise secondary structure of this motif is essential for secretion.
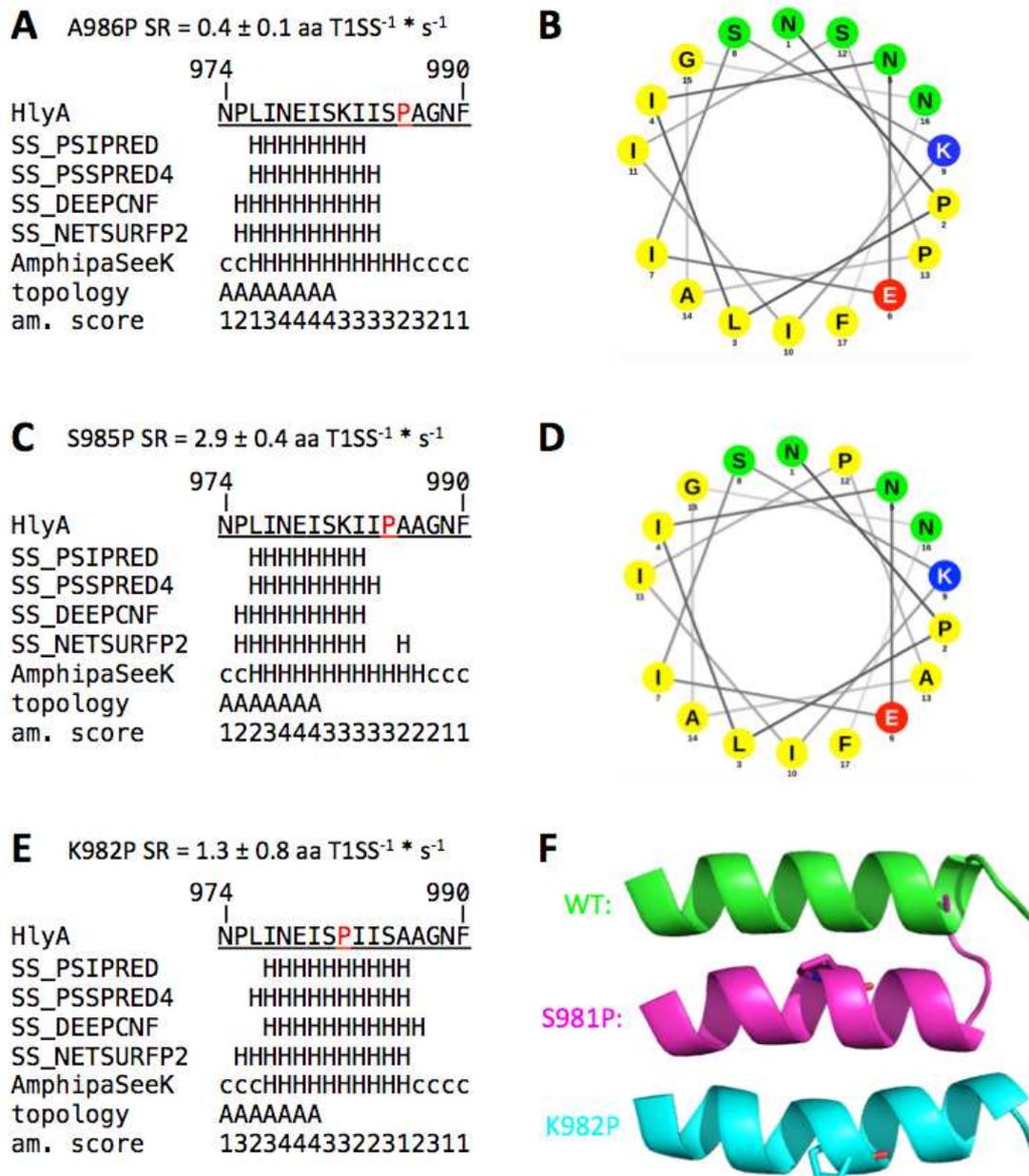
A A986P SR = 0.4 ± 0.1 aa T1SS⁻¹ * s⁻¹

```
                974                990
                 |                  |
HlyA            NPLINEISKIISPAGNF
SS_PSIPRED       HHHHHHHH
SS_PSSPRED4      HHHHHHHHH
SS_DEEPCNF       HHHHHHHHH
SS_NETSURFP2     HHHHHHHHH
AmphipaSeeK     ccHHHHHHHHHHHHHcccc
topology        AAAAAAAA
am. score       12134444333323211
```

C S985P SR = 2.9 ± 0.4 aa T1SS⁻¹ * s⁻¹

```
                974                990
                 |                  |
HlyA            NPLINEISKIIPAAGNF
SS_PSIPRED       HHHHHHHH
SS_PSSPRED4      HHHHHHHHH
SS_DEEPCNF       HHHHHHHHH
SS_NETSURFP2     HHHHHHHHH   H
AmphipaSeeK     ccHHHHHHHHHHHHHccc
topology        AAAAAAA
am. score       12234443333322211
```

E K982P SR = 1.3 ± 0.8 aa T1SS⁻¹ * s⁻¹

```
                974                990
                 |                  |
HlyA            NPLINEISPIISAAGNF
SS_PSIPRED        HHHHHHHHHH
SS_PSSPRED4       HHHHHHHHHHH
SS_DEEPCNF         HHHHHHHHHHH
SS_NETSURFP2      HHHHHHHHHHHH
AmphipaSeeK     cccHHHHHHHHHHcccc
topology        AAAAAAA
am. score       13234443322312311
```

Figure 3: *In silico* **analysis of three different mutants of HlyA. A, C, E)** Predictions labeled with "SS" are from Quick2D, that utilizes multiple prediction algorithms. Topology = predicted by AmphipaSeeK: (A) marks residues that are predicted to be inserted parallel into the membrane. Am. score = amphipathy score predicted by AmphipaSeeK with 1 = low amphipathy and 5 = high amphipathy. H = α-helix, c = coiled coil. Mutated residues are marked in red. Secretion rate (SR) is given for each mutant as mean ± SD of three independent measurements (Kanonenberg, 2018). **B** and **D)** Helical wheel projection of A986P (**B**) and S985P (**D**). Non-polar residues are colored yellow, lysine blue, glutamine red and polar residues green. Proline at position 985 reduces polarity on the polar site of the AH compared to wild type HlyA (Figure 1 C). **F)** Cartoon representation of PEP-FOLD3 models of HlyA wild type (WT) in green, S981P in pink and K982P in cyan. Mutated proline residues are shown as sticks. The models were superimposed and then distributed vertically. K982P and S981P bend in the opposite direction.

F990 is not part of the predicted AH but highly conserved and essential for secretion (Kanonenberg, 2018). In previous studies it has been shown, that a

substitution of this residue to His, Cys, Ala, Ser, Ile, Asn or Pro strongly reduced the secretion of HlyA to <20% compared to wild type (Chervaux and Holland, 1996). The substitution to Tyr allowed a secretion of ~35% compared to wild type (Chervaux and Holland, 1996) and Tyr can be found in some other RTX proteins at this position as indicated by alignments (Kanonenberg, 2018).

Since the most efficient secretion can be achieved by a substitution with another aromatic residue, π-π interactions can be assumed, that are disrupted in F990P. Vernon *et al.* recently provided an extensive study on protein crystal structures (N=5718) analyzing π-π interactions. Amongst other findings they show that Phe and Tyr have very similar preferences for the nature of their contacts and that π-π stacking with non-aromatic residues is actually more common than aromatic-aromatic stacking by roughly 13 to 1. Furthermore they identified Arg as the first or second most likely interaction partner for any given aromatic side chain (Vernon et al., 2018). Conserved Arg residues can be found for example in the cytosolic domain (CD) of HlyD and could present an interaction partner to F990.

Taken together these phenotypes and their *in silico* analysis strongly support the presence of an AH at this position and underline the importance of the secondary and primary structure for secretion.

Evaluation of HlyB NBD

Benabdelhak *et al.* demonstrated that the secretion signal of HlyA, which harbors the discussed AH, and the NBD of HlyB likely interact with each other and that this interaction is disrupted in the presence of ATP. Assuming that this interaction takes place via the AH of HlyA a hydrophobic binding pocket on or in the NBD is to be expected. Fortunately, many structures of the isolated NBD are available covering the full catalytic circle of ATP hydrolysis: The wild type is available as a nucleotide free monomer (1MT0 (Schmitt et al., 2003)), as an ADP bound monomer (2FF7 (Zaitseva et al., 2006)) and as a TNP-ADP bound monomer (2PMK (Oswald et al., 2008)). The dimeric structure is available for the hydrolysis-impaired mutant E631Q with two ATP bound (2FGK (Zaitseva et al., 2006)), for the hydrolytic inactive

mutant H662A (activity <0.1% (Zaitseva et al., 2005)) with two ATP bound (2FGJ) (Zaitseva et al., 2006) and for the mutant H662A with two ATP and two Mg^{2+} bound (1XEF (Zaitseva et al., 2005)).

To identify hydrophobic binding pockets we used the yrb-script (Hagemans et al., 2015). Since Benabdelhak *et al.* only observed the interaction without ATP present, the interaction of HlyA with HlyB NBD likely takes place with the nucleotide free form of the NBD. Therefore we searched for a hydrophobic binding pocket in the nucleotide free monomer (1MT0). The interaction seems to be disrupted in the presence of ATP (Benabdelhak et al., 2003) so the identified possible binding pockets (pbp's) were analyzed in the dimeric ATP-bound structure (1XEF) as well to see if they are disrupted by ATP binding and/or dimerization. The positional change of the residues can be analyzed by aligning both structures and calculating the RMSD (root mean square deviation) value only for the residues involved in forming the pbp's.

We identified two regions of high hydrophobicity, which we called possible binding pocket inside (pbp-in) and outside (pbp-out).

The pbp-in is located closely to the dimer interface and the ATP binding site (Figure 4 A). While some residues that from this pocket are considered polar amino acids, they contribute to the unpolar binding pocket with the carbon atoms of their side chain. Furthermore, charges at the side of a binding pocket for an AH may interact with the polar side of the AH and help its orientation. The residues that form pbp-in are: F475, Y477, K478, I484, T510, K513, Q516, F518, Y519 (Figure 4 C).

The pbp-in shares at least three residues with the ATP binding site (Y477, I484, T510): Y477 interacts with the adenine base, I484 with the ribose moiety and T510 with the Pα of ATP (Zaitseva et al., 2005). When aligning the ATP free monomer (1MT0) with the ATP bound dimer (1XEF) the residues of the pbp-in show a change in position of approximately 2 Å as reflected by the RMSD value. Both Tyr (Y477 and Y519) and both Lys (K487 and K513) display the largest change in position (Figure 4 C). The RMSD value is 1.3 Å when compared to the ADP-bound state (2FF7,(Zaitseva et al., 2006)).

The pbp-out is located opposite to the membrane and is thus exposed to the cytosol (Figure 4 B and F). It is made up by the following residues: V675, E677, K680, V682, E683, L697, Y700, L701 and L704. Y700, L701 and L704 point towards the dimer interface and are involved in monomer-monomer contact (Zaitseva et al., 2005). The pbp-out changes less than the pbp-in upon ATP binding and dimerization as shown by the RMSD value of 1.1 Å (Figure 4 D). The difference to the ADP-bound monomer is only 0.9 Å.

A secretion process of a toxin like HlyA with a size of 1024 amino acids likely involves multiple transient contacts between the transporter components and the toxin. Both identified pbp's hold the potential to interact with the C-terminal part of HlyA: i) They display a hydrophobic area that matches the length of the AH in HlyA. ii) They hold residues that are able to form π bonds with their side chain and could therefore act as an interaction partner to F990 of HlyA. iii) They change upon ATP binding and dimerization, offering an explanation for the observations of Benabdelhak *et al.*.

Experimental investigation of both pbp's is needed to draw further conclusions. So far both are possible and do not exclude each other.
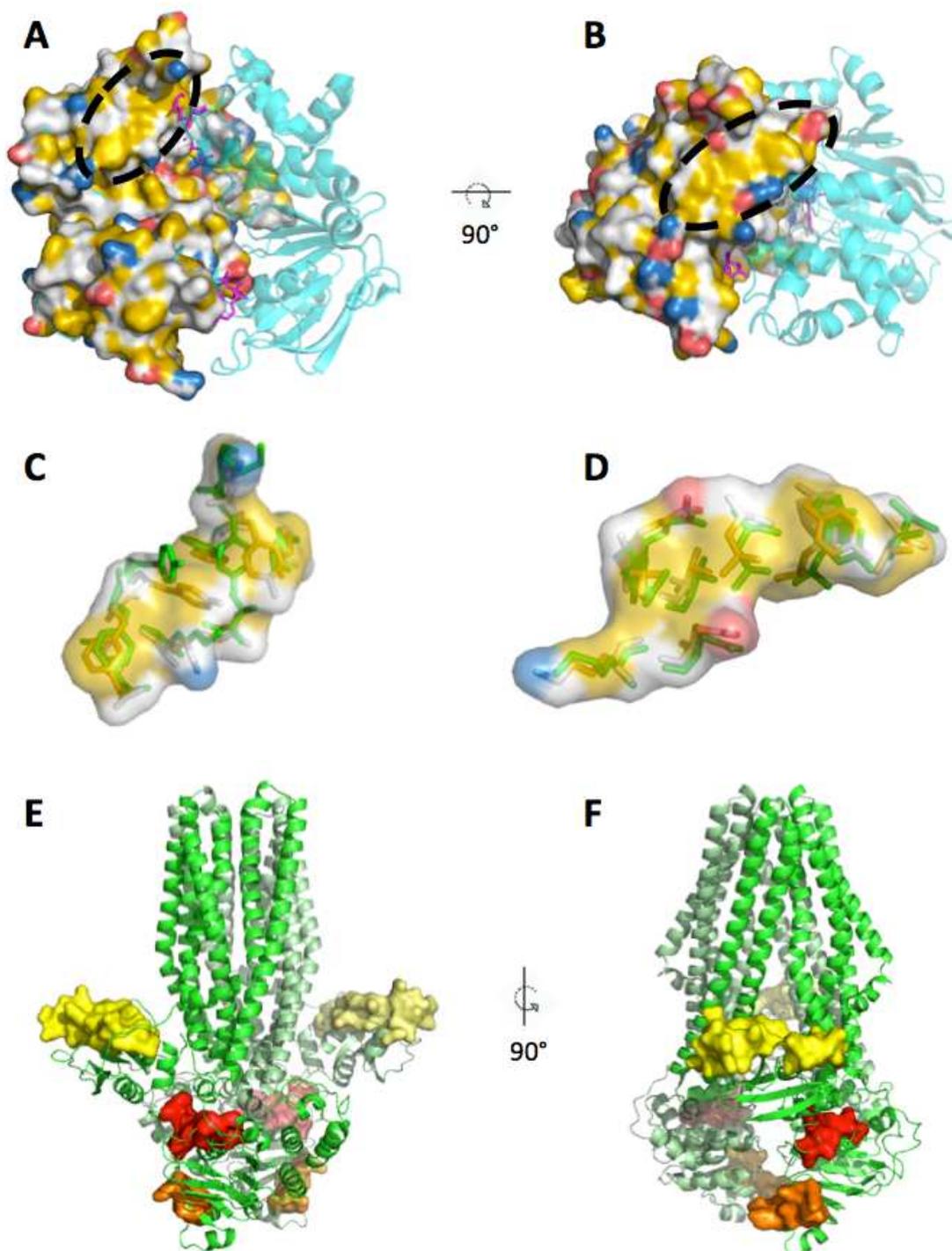
Figure 4: Possible binding pockets (pbp) in HlyB for HlyA AH. A and **B)** The HlyB NBD monomer (1MT0) is shown as a surface representation. The surface has been colored with the yrb-script (Hagemans et al., 2015), which highlights carbon atoms that are not bound to oxygen or nitrogen in yellow, the charged oxygens of Glu and Asp in red, the charged nitrogens of Lys and Arg in blue while all other atoms are white. The monomer has been superimposed on the dimeric NBD structure (1XEF) for reference, which is shown as transparent cyan cartoon representation. The pbp's are circled with a black dashed line. From **A**, which shows the pbp-in, to **B**, which shows the pbp-out, the molecules have been rotated towards the reader. ATP molecules from the dimeric structure (1XEF) are shown as pink sticks. In (**A**) ATP overlaps with the pbp-in. **C** and **D)** The pbp-in (**C**) and pbp-out (**D**) from the monomeric structure (1MT0) are shown as stick representations with transparent surface colored with the yrb-script as well. Color code is as in **A** and **B**. Green sticks represent the same residues but in the dimeric structure (1XEF). The RMSD for the residues of the pbp-in is 1.981 Å and 1.125 Å for pbp-out.

E and **F)** Cartoon representation of a model of HlyB based on PCAT1 in green (Lin et al., 2015). The identified pbp's are shown as surface representation with pbp-in in red and pbp-out in orange. Yellow surface maps the HlyA-interaction region on the CLD (Lecher et al., 2012). Pale colors correspond to the same regions in the second monomer.

Evaluation of HlyB CLD

Secretion of HlyA is not possible without the CLD of HlyB (Lecher et al., 2012). Furthermore, Lecher *et al*. performed pull-down assays with the isolated CLD and unfolded or folded HlyA as well as truncated versions of HlyA, which they termed HlyA1 (C-terminal 217 residues) and HlyA2 (HlyA1 lacking the C-terminal 60 residues that make up the secretion signal). The absence of the discussed AH of HlyA in the truncated version HlyA2 had no impact on the observed interaction. Instead, the interaction was disrupted when the experiment was performed with folded versions of the substrate (Lecher et al., 2012). Therefore the authors contributed the interaction to the conserved GG repeats in the RTX domain of HlyA, which induce folding of RTX proteins upon Ca^{2+} binding (Baumann et al., 1993).

By a combination of chemical shift experiments and mutational studies Lecher *et al.* were able to map the region in the CLD that interacts with HlyA (Figure 4 E and F). Since HlyB seems to interact with different regions of the substrate, we were interested if simultaneous binding would be possible.

The distance between the C-terminus of HlyA2 (smallest fragment to interact with CLD) and the beginning of the AH of HlyA is only 14 amino acids. The last GG repeat of HlyA2 is located 122 amino acids upstream from this AH. Assuming 3.8 Å per amino acid in an unfolded protein (Carrion-Vazquez et al., 1999, Crecca and Roitberg, 2008) the distance between CLD-interaction site and NBD-interaction site of HlyA is between 53.2 Å and 463.6 Å.

In order to measure the distances between the pbp's of the NBD and the interaction site on the CLD the whole structure of HlyB is needed. Unfortunately, this structure is not available and can only be modeled based on other ABC transporters such as PCAT1 (Lin et al., 2015) (Figure 4 E and F). The model presented here was calculated with the newly introduced modeling tool TopModel (Mulnaes et al., 2020).

In this model the outmost residue of the HlyA interaction site in the CLD is E92. The minimal distance between E92 and pbp-in measures ~40 Å and the maximal distance from E92 to pbp-in is ~64 Å. The pbp-out shows a distance to E92 between 54 Å and 68 Å. HlyA can easily bridge these distances and simultaneous binding of the substrate HlyA to the CLD and NBD of HlyB is possible.

If multiple interaction sites on HlyB are occupied simultaneously by one substrate molecule this would result in a strictly ordered substrate arrangement, which could confer specificity between substrate and transport components. The transport components HlyB and HlyD can also secrete heterologously expressed RTX toxins such as FrpA from *N. meningitidis* and HlyIA from *A. pleuropneumoniae* serotype 1 (Gygi et al., 1990, Thompson and Sparling, 1993). Both are predicted to have an AH in their C-terminus and a GG repeat can be found 122 residues (FrpA) and 121 residues (HlyIA) upstream of this AH.

Evaluation of HlyD CD

The MFP HlyD is critical for HlyA export and able to interact with the substrate even in the absence of the ABC transporter (Thanabalu et al., 1998). Although interaction of HlyA with the periplasmic domain of HlyD have already been shown (Pimenta et al., 2005), the experiments of Thanabalu *et al.* suggest additional interaction(s) in the cytosol, which will be further discussed here.

HlyD has a small cytoplasmic domain (CD), made up of the first 60 N-terminal residues (Schülein et al., 1992). An AH has been identified in this region (Balakrishnan et al., 2001) and we performed the secondary structure prediction of the CD with the above introduced tools.

In silico *analysis of HlyD CD - AH*

The tools are in agreement with each other with the AH starting at T3 or W4 and ending at Q28. The two tools in Figure 5 A, that do not show a helix at this position are tools that predict disordered regions (Hanson et al., 2016, Klausen et al., 2019). AmphipaSeeK also predicts a small portion of this AH to be in-plane, however,

the prediction score for this topology is quite low compared to known in-plane membrane anchors (Sapay et al., 2006a). The prediction score takes on values between -1 and 1 and values >0 lead to the prediction of in-plane topology. The score is given for each residue and an average can be calculated for the residues predicted to be in-plane. With this method the AH of for example FtsA from *E. coli* (Pichoff and Lutkenhaus, 2005) reaches an average prediction score of 0.106 and the AH of NS5A from bovine viral diarrhea virus (BVDV) (Sapay et al., 2006b) reaches 0.213. For HlyD the value is only 0.045 (Figure 5 A). However, when overexpressing the CD of HlyD in *E. coli*, a portion of the protein localizes to the membrane fraction even though the TM-helix is not present (residue 61-80) (Figure 5 C). The affinity to the membrane is likely conferred by the AH. Purification trials of this fragment, which would allow *in vitro* interaction studies, failed so far.

Sapay *et al*. demonstrated that Lys, Phe and Trp are the most common residues in in-plane helices, with Trp being the only hydrophobic residue more abundant in in-plane membrane anchors than in TM-helices (Sapay et al., 2006a). The AH of HlyD shows three Trp residues and two Phe residues on the hydrophobic side, which further supports the hypothesis that the AH is inserted into the membrane (Figure 6 C and D).

Since HlyD is fixed to the inner membrane by the TM-helix the AH does not need to function as an anchor even if inserted into the membrane. The possible insertion of the AH might therefore present a signal to the inner membrane complex for example to recruit TolC. If the AH is not constantly inserted into the membrane but the insertion functions as a signal, this might explain the low prediction score for the in-plane topology as well.

In silico *analysis of HlyD CD – second helical region*

The prediction tools also detected a second helical region, approximately from K36 to E50. Here the tools are not in agreement with each other, with two presenting one continuous helix, three showing two helix fragments, one predicting no helix at all and one predicting a disordered region from H45 to V60 (Figure 5 A). The

disagreement among the tools may reflect a certain flexibility of this region, which will be further discussed later in the context of the 3D models.

The second helical region overlaps with a region that has been shown to be essential for HlyA secretion by deletion of residue 26 to 45 (Balakrishnan et al., 2001). It contains a highly conserved motif of eight amino acids (FLPAHLEL) (Figure 5 B). When this motif is used in a protein-BLAST (basic local alignment search tool) search, in- or excluding *E. coli*, only HlyD-like proteins are among the first 100 results.

Balakrishnan *et al.* assigned the importance of region 26 to 45 to a charged cluster of five amino acids (R34 to E38), which are less conserved. However, all ten analyzed homologs show charged or polar amino acids in this region, which are flanked by a highly conserved Arg and a Glu (Figure 5 B).

Figure 5 *In silico* analysis of HlyD CD. **A)** Secondary structure predictions. Predictions labeled with "SS" and "DO" are from Quick2D, which utilizes multiple prediction algorithms. Predictions labeled with "DO" only predict disordered regions (D). Topology = predicted by AmphipaSeeK: (A) marks residues that are predicted to be inserted parallel into the membrane. Am. score = amphipathy score predicted by AmphipaSeeK with 1 = low amphipathy and 5 = high amphipathy. H = α-helix, c = coiled coil, E = β-sheet. Below the am. score the prediction score for the topology is given. Values >0 indicate membrane insertion. **B)** Alignment of ten MFPs to HlyD performed with Clustal Omega (Madeira et al., 2019). The MFPs are labeled with their UniProt-ID (except WP_049080580.1, which is from NCBI). (*) marks the same residue, while (:) and (.) mark similar residues. Residues of the discussed charged cluster (R34 – E38) are colored with positively charged residues in blue and negatively charged residues in red. The green box highlights a conserved motif of eight residues. **C)** Western Blot of initial expression of HlyD CD (9.2 kDa). 2 h, 3 h and 4 h after induction of protein expression, samples were taken, cells disrupted and centrifuged twice to collect the membrane fraction. 1: Supernatant of first centrifugation (cytosolic fraction and membrane fragments). 2: Pellet of first centrifugation (undisrupted cells, cell debris, inclusion bodies). 3: Supernatant of second centrifugation (cytosolic fraction). 4: Pellet of second centrifugation (membrane fraction). At all time points HlyD CD can be found in the membrane fraction (4) as well, with the strongest signal after 4 h. M = molecular weight marker. White spots in black bands indicate saturation of the detector.

PEP-FOLD3 models of HlyD CD

In order to gain a deeper understanding of the HlyD CD, we modeled the first 50 aa (which include the AH and the second helical region) with PEP-FOLD3 (Thévenet et al., 2012, Shen et al., 2014, Lamiable et al., 2016). Interestingly, four of five models presented a break in the AH, while the fifth model showed a strong bend at a similar position (Figure 6 A). This break is inconsistent with all six secondary structure prediction tools used above (Figure 5 A) and earlier predictions (Balakrishnan et al., 2001).

When shortening the input sequence to the first 42 aa all five models showed one continuous AH with one model showing a similar bend to that observed in the 50 aa model (Figure 6 B).

Only one out of the ten models did not display a helix in the second helical region. In the modeling approach with 42 aa of HlyD four models show only one helix, all starting at K36 and ending at position 40-42. In the 50 aa approach two models display one helix and three models display two helices. Four out of this five also start at K36 (one at D37) and four end at L48 (Figure 6 A and B).
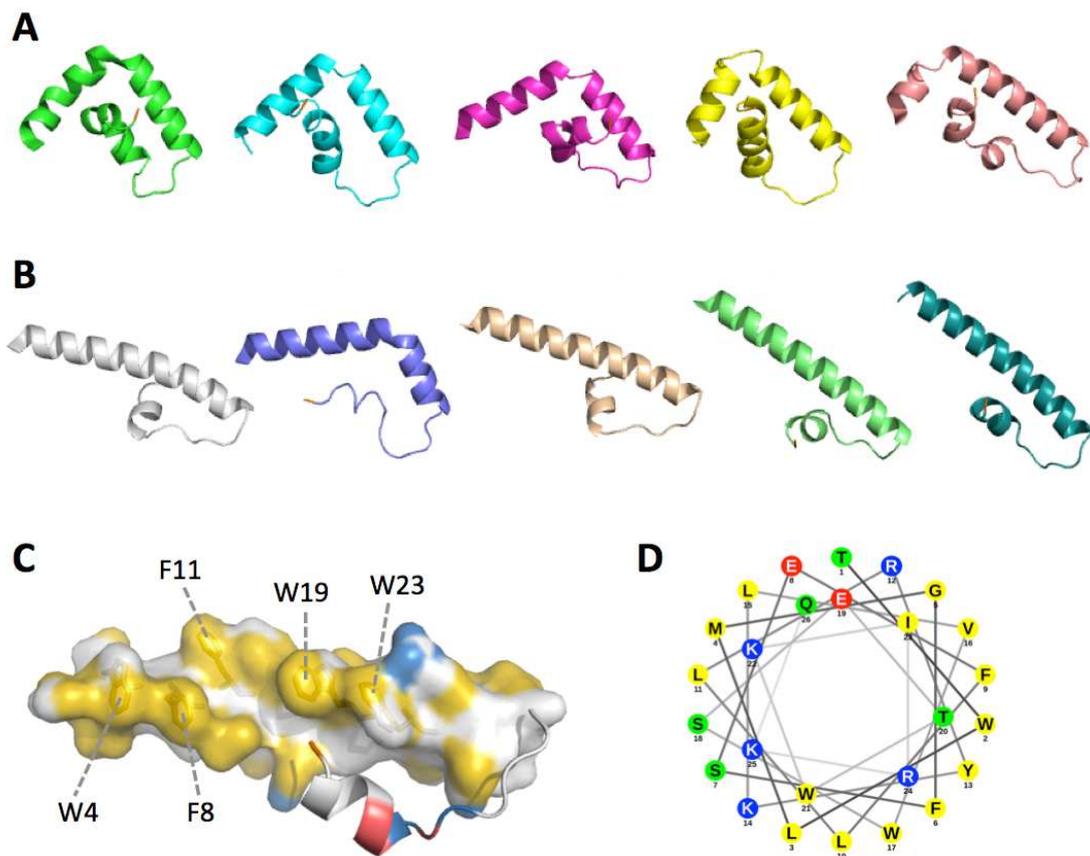
Figure 6: PEP-FOLD3 models of HlyD CD (Lamiable et al., 2016). A) Modeling approach with N-terminal 50 aa. All models show a break or bend (green) in the AH. The second helical region either shows one helix (cyan and yellow) or two helices (green, pink, salmon). **B)** Modeling approach with N-terminal 42 aa. Only one model shows a bend in the AH (purple), which is also the only model that does not show a helix in the second helical region. **C)** Shown is model 5 from the 42 aa modeling approach (shown in **B** as petrol cartoon). The AH is shown as a surface representation colored with the yrb script (Hagemans et al., 2015), which highlights carbon atoms that are not bound to oxygen or nitrogen in yellow, the charged oxygens of Glu and Asp in red, the charged nitrogens of Lys and Arg in blue while all other atoms are white. The second helical region is shown as a cartoon representation with the discussed charged cluster (R34 to E38) colored in red (Glu and Asp) and blue (Arg and Lys). All C-termini in **A**, **B**, and **C** are colored orange. **D)** Helical wheel projection of HlyD CD AH (T3 to Q28), with Glu in red, Arg and Lys in blue, polar residues in green and non-polar residues in yellow (Mól et al., 2018).

In both modeling approaches and also the secondary structure predictions, the region between the AH and the second helical region is not annotated with a secondary structure and therefore flexible. This allows multiple orientations of the AH and the second helical region towards each other. In all models they perform some kind of contact with the 50 aa models bending around the second helical region (Figure 6 A).

Furthermore all five of the 50 aa models display the highly conserved motif (FLPAHLEL) to be at least part of a helix.

As already mentioned, the inconsistency in prediction and modeling for this region might reflect its flexibility, which could be important for HlyD to sense an interaction partner, which would (de-)stabilize the secondary or even tertiary structure. A possible scenario could be that the second helical region interacts with the AH (similar to the modeled structures) keeping the AH from inserting into the membrane. An interaction partner such as HlyA could bind to the second helical region, disrupting the interaction to the AH allowing insertion into the membrane. However, there are multiple ways in which the substrate HlyA could interact with the cytoplasmic domain of HlyD.

As mentioned earlier Arg is the most favored interaction partner for π-π interactions with any aromatic residue (Vernon et al., 2018) and HlyA secretion is highly dependent on F990 (in HlyA). Both residues, R34 (HlyD) and F990 (HlyA) are highly conserved and might mediate the interaction by π-π stacking.

The AH of HlyD shows a high frequency of aromatic residues on the non-polar side and although aromatic-to-aromatic side chain π-π interactions have been found to be less common (Vernon et al., 2018) an interaction between F990 and the aromatic patch of the AH of HlyD cannot be excluded.

The highly conserved motif (FLPAHLEL) also contains residues able to perform π-π interactions (F, H and E) and could therefore act as an interaction site.

The cluster of charged amino acids is also highly conserved among the membrane fusion proteins and a charged cluster can also be found in the extreme C-terminus of HlyA made up of KEER (residue 993 to 996). Electrostatic interactions between both charged clusters are also possible.

An interaction between both AHs is another possibility how HlyA and HlyD could interact in the cytoplasm, since helices are known to facilitate protein-protein interactions (Jones and Thornton, 1995, Guharoy and Chakrabarti, 2007).

Summary of findings

An amphipathic helix (AH) in the secretion signal of HlyA has long been proposed and could be validated by mutational studies with a combination of *in silico* approaches, including secondary structure prediction, alignments and modeling.

Two possible binding sites in the NBD of HlyB for this AH were identified and require further investigations.

Modeling of the whole ABC transporter HlyB allowed measuring of distances between the domains known to interact with HlyA C-terminal fragments showing that simultaneous binding of the substrate to two domains of the transporter is possible.

An AH has also long been proposed for the cytoplasmic domain (CD) of HlyD (Balakrishnan et al., 2001). The prediction could be reproduced with multiple tools utilizing larger databases then ever before. The orientation of the CD towards the membrane has been shown and can be attributed to the AH. Alignments of this fragment to other membrane fusion proteins showed two additional regions of interest: a charged cluster and an eight amino acid long motif, which seems to be unique for this class of proteins.

Material and Methods

Secondary structure prediction

Two different tools were used: Quick2D (Zimmermann et al., 2018) and AmphipaSeeK (Combet et al., 2000, Sapay et al., 2006a).

Quick2D is a toolkit, which utilizes a variety of published prediction methods, therefore giving the user multiple outputs for one sequence. It is able to predict α-, π- and TM-helices, β-strands, coiled coils, as well as disordered regions (Zimmermann et al., 2018).

AmphipaSeeK on the other hand is specifically designed to identify amphipathic helices that insert themselves into the membrane in a parallel manner; so-called in-plane membrane anchors (Sapay et al., 2006a). It provides the user with a secondary structure prediction, a predicted membrane topology (in-plane or not-in-

plane), a prediction score for the proposed membrane topology and an amphipathy score for each residue in dependence to the neighboring residues.

Modeling

PEP-FOLD3 was used to model peptides of HlyD and HlyA. It allows quick modeling of peptides ranging from 5 to 50 amino acids in length and their downstream analysis for example in PyMOL (Thévenet et al., 2012, Shen et al., 2014, Lamiable et al., 2016). The structure of HlyB was modeled based on the structure of PCAT1 with the tool TopModel (Lin et al., 2015, Mulnaes et al., 2020).

Illustration of different structural models

The amphipathic characteristic of a helix is best visualized by a helical wheel projection. We choose NetWheel for its clear layout and easy handling (Mól et al., 2018). Protein and peptide structures were processed in PyMOL (Schrodinger, LLC. 2010. The PyMOL molecular Graphics System, Version 1.8.6.0 Enhanced for Mac OS X). In order to illustrate and identify hydrophobic surfaces the yrb-script was applied in PyMOL, which highlights carbon atoms that are not bound to oxygen or nitrogen in yellow, the charged oxygens of Glu and Asp in red, the charged nitrogens of Lys and Arg in blue while all other atoms are white (Hagemans et al., 2015).

Alignments

Alignments were performed using Clustal Omega (Madeira et al., 2019). Sequences of the homologous proteins were taken from UniProt (https://www.uniprot.org/) or NCBI (https://www.ncbi.nlm.nih.gov/) as indicated in Figure 5 B.

Expression of HlyD CD

Chemically competent *E. coli* BL21(DE3) cells were transformed by the heat shock method with pET22b-HlyD60, which encodes for the cytoplasmic domain (CD) of HlyD (residue 1-60) with a C-terminal 6xhis-tag (Froger and Hall, 2007) and were

selected on LB-agar plates containing 100 μg/mL ampicillin. Clones from this plate were used to inoculate 5 mL LB-media cultures supplemented with the same antibiotic concentration, which were grown for 16 h at 37°C and 180 rpm. They were used to inoculate 50 mL cultures (same media and antibiotic concentration) to an OD_{600} (optical density at 600 nm) of 0.1. At OD_{600} 0.5-0.7 protein expression was induced by adding 1 mM IPTG. Samples of 1 mL were taken 2 h, 3 h and 4 h after induction, cells collected by centrifugation (11,000xg, 1 min, RT) and mechanically disrupted by vortexing with glass beads. The cell suspension was fractionated by centrifugation: First a centrifugation at 11,000xg for 5 min to remove glass beads and cell debris. The supernatant of this centrifugation was subjected to a high spin centrifugation at 120,000xg for 30 min at 4°C to collect the membrane. A sample of both supernatants was mixed with SDS-sample buffer (100 mM Tris pH 6.8, 3.3% (w/v) SDS, 0.02% (w/v) bromophenol blue, 40% (v/v) glycerol) and the pellets were resuspended in resuspension buffer (50 mM Na_2HPO_4 pH 8, 300 mM NaCl) before also being mixed with SDS-sample buffer. Samples of both supernatants and both pellets were applied to a 15% SDS-PAGE. Proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane with a Trans-Blot SD Semi-Dry Transfer Cell by BioRad according to manufactures instructions. The membrane was blocked with 10% (w/v) milk powder overnight at 4°C, incubated with Penta His Antibody (Qiagen) for 1 h at room temperature (RT), then incubated with a secondary antibody coupled to horse radish peroxidase (HRP) for 1 h at RT, washed three times in between each step with TBS-T buffer (20 mM Tris pH 8, 250 mM NaCl, 0.1% (v/v) Tween-20) and analyzed using Chem Genius[2] bio imaging system by Syngene.

Acknowledgments

References

Balakrishnan, L., Hughes, C. and Koronakis, V. (2001) 'Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*', *Journal of Molecular Biology,* 313(3), pp. 501-510.

Baumann, U., Wu, S., Flaherty, K. M. and McKay, D. B. (1993) 'Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif', *The EMBO Journal,* 12(9), pp. 3357-3364.

Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B. and Schmitt, L. (2003) 'A Specific Interaction Between the NBD of the ABC-transporter HlyB and a C-Terminal Fragment of its Transport Substrate Haemolysin A', *Journal of Molecular Biology,* 327(5), pp. 1169-1179.

Carrion-Vazquez, M., Marszalek, P. E., Oberhauser, A. F. and Fernandez, J. M. (1999) 'Atomic force microscopy captures length phenotypes in single proteins', *Proceedings of the National Academy of Sciences,* 96(20), pp. 11288-11292.

Chervaux, C. and Holland, I. B. (1996) 'Random and directed mutagenesis to elucidate the functional importance of helix II and F-989 in the C-terminal secretion signal of *Escherichia coli* hemolysin', *J Bacteriol,* 178(4), pp. 1232-6.

Combet, C., Blanchet, C., Geourjon, C. and Deléage, G. (2000) 'NPS@: network protein sequence analysis', *Trends Biochem Sci,* 25(3), pp. 147-50.

Crecca, C. R. and Roitberg, A. E. (2008) 'Using distances between α-carbons to predict protein structure', *International Journal of Quantum Chemistry,* 108(15), pp. 2782-2792.

Felmlee, T., Pellett, S. and Welch, R. A. (1985) 'Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin', *Journal of Bacteriology,* 163(1), pp. 94-105.

Froger, A. and Hall, J. E. (2007) 'Transformation of plasmid DNA into *E. coli* using the heat shock method', *J Vis Exp,* (6), pp. 253.

Gangola, P. and Rosen, B. P. (1987) 'Maintenance of intracellular calcium in *Escherichia coli*', *J Biol Chem,* 262(26), pp. 12570-4.

Gray, L., Mackman, N., Nicaud, J. M. and Holland, I. B. (1986) 'The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*', *Mol Gen Genet,* 205(1), pp. 127-33.

Guharoy, M. and Chakrabarti, P. (2007) 'Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions', *Bioinformatics,* 23(15), pp. 1909-1918.

Gygi, D., Nicolet, J., Frey, J., Cross, M., Koronakis, V. and Hughes, C. (1990) 'Isolation of the *Actinobacillus pleuropneumoniae* haemolysin gene and the activation and secretion of the prohaemolysin by the HlyC, HlyB and HlyD proteins of Escherichia coli', *Mol Microbiol,* 4(1), pp. 123-8.

Hagemans, D., van Belzen, I. A., Morán Luengo, T. and Rüdiger, S. G. (2015) 'A script to highlight hydrophobicity and charge on protein surfaces', *Front Mol Biosci,* 2, pp. 56.

Hanson, J., Yang, Y., Paliwal, K. and Zhou, Y. (2016) 'Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks', *Bioinformatics,* 33(5), pp. 685-692.

Hess, J., Gentschev, I., Goebel, W. and Jarchau, T. (1990) 'Analysis of the haemolysin secretion system by PhoA-HlyA fusion proteins', *Mol Gen Genet,* 224(2), pp. 201-8.

Holland, I. B., Peherstorfer, S., Kanonenberg, K., Lenders, M., Reimann, S. and Schmitt, L. (2016) 'Type I Protein Secretion-Deceptively Simple yet with a Wide Range of Mechanistic Variability across the Family', *EcoSal Plus,* 7(1).

Jarchau, T., Chakraborty, T., Garcia, F. and Goebel, W. (1994) 'Selection for transport competence of C-terminal polypeptides derived from *Escherichia coli* hemolysin: the shortest peptide capable of autonomous HlyB/HlyD-dependent secretion comprises the C-terminal 62 amino acids of HlyA', *Molecular and General Genetics MGG,* 245(1), pp. 53-60.

Jones, H. E., Holland, I. B., Baker, H. L. and Campbell, A. K. (1999) 'Slow changes in cytosolic free Ca2+ in *Escherichia coli* highlight two putative influx mechanisms in response to changes in extracellular calcium', *Cell Calcium,* 25(3), pp. 265-274.

Jones, S. and Thornton, J. M. (1995) 'Protein-protein interactions: A review of protein dimer structures', *Progress in Biophysics and Molecular Biology,* 63(1), pp. 31-65.

Kanonenberg, K. (2018) *Biochemical Characterisation of the Escherichia coli Haemolysin A Type I Secretion System.* PhD Doctoral Dissertation, Heinrich-Heine-University Düsseldorf.

Kenny, B., Taylor, S. and Holland, I. B. (1992) 'Identification of individual amino acids required for secretion within the haemolysin (HlyA) C-terminal targeting region', *Mol Microbiol,* 6(11), pp. 1477-89.

Kim, M. K. and Kang, Y. K. (1999) 'Positional preference of proline in alpha-helices', *Protein Sci,* 8(7), pp. 1492-9.

Klausen, M. S., Jespersen, M. C., Nielsen, H., Jensen, K. K., Jurtz, V. I., Sønderby, C. K., Sommer, M. O. A., Winther, O., Nielsen, M., Petersen, B. and Marcatili, P. (2019) 'NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning', *Proteins,* 87(6), pp. 520-527.

Koronakis, V., Koronakis, E. and Hughes, C. (1989) 'Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes', *Embo j,* 8(2), pp. 595-605.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B. and Hughes, C. (2000) 'Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export', *Nature,* 405(6789), pp. 914-9.

Lamiable, A., Thévenet, P., Rey, J., Vavrusa, M., Derreumaux, P. and Tufféry, P. (2016) 'PEP-FOLD3: faster *de novo* structure prediction for linear peptides in solution and in complex', *Nucleic Acids Res,* 44(W1), pp. W449-54.

Lecher, J., Schwarz, Christian K. W., Stoldt, M., Smits, Sander H. J., Willbold, D. and Schmitt, L. (2012) 'An RTX Transporter Tethers Its Unfolded Substrate during Secretion via a Unique N-Terminal Domain', *Structure,* 20(10), pp. 1778-1787.

Lenders, M. H. H., Beer, T., Smits, S. H. J. and Schmitt, L. (2016) '*In vivo* quantification of the secretion rates of the hemolysin A Type I secretion system', *Scientific Reports,* 6(1), pp. 33275.

Lenders, M. H. H., Weidtkamp-Peters, S., Kleinschrodt, D., Jaeger, K.-E., Smits, S. H. J. and Schmitt, L. (2015) 'Directionality of substrate translocation of the hemolysin A Type I secretion system', *Scientific Reports,* 5(1), pp. 12470.

Lin, D. Y.-w., Huang, S. and Chen, J. (2015) 'Crystal structures of a polypeptide processing and secretion transporter', *Nature,* 523(7561), pp. 425-430.

Linhartová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., Kamanová, J., Procházková, K., Adkins, I., Hejnová-Holubová, J., Sadílková, L., Morová, J. and Šebo, P. (2010) 'RTX proteins: a highly diverse family secreted by a common mechanism', *FEMS Microbiology Reviews,* 34(6), pp. 1076-1112.

Ludwig, A., Schmid, A., Benz, R. and Goebel, W. (1991) 'Mutations affecting pore formation by haemolysin from *Escherichia coli*', *Mol Gen Genet,* 226(1-2), pp. 198-208.

Mackman, N., Nicaud, J. M., Gray, L. and Holland, I. B. (1985) 'Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*', *Mol Gen Genet,* 201(3), pp. 529-36.

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D. and Lopez, R. (2019) 'The EMBL-EBI search and sequence analysis tools APIs in 2019', *Nucleic acids research,* 47(W1), pp. W636-W641.

Mól, A. R., Castro, M. S. and Fontes, W. (2018) 'NetWheels: A web application to create high quality peptide helical wheel and net projections', *bioRxiv,* pp. 416347.

Mulnaes, D., Porta, N., Clemens, R., Apanasenko, I., Reiners, J., Gremer, L., Neudecker, P., Smits, S. H. J. and Gohlke, H. (2020) 'TopModel: Template-Based Protein Structure Prediction at Low Sequence Identity Using Top-Down Consensus and Deep Neural Networks', *Journal of Chemical Theory and Computation,* 16(3), pp. 1953-1967.

Nicaud, J. M., Mackman, N., Gray, L. and Holland, I. B. (1985) 'Characterisation of HlyC and mechanism of activation and secretion of haemolysin from *E. coli* 2001', *FEBS Lett,* 187(2), pp. 339-44.

Oswald, C., Jenewein, S., Smits, S. H., Holland, I. B. and Schmitt, L. (2008) 'Water-mediated protein-fluorophore interactions modulate the affinity of an ABC-ATPase/TNP-ADP complex', *J Struct Biol,* 162(1), pp. 85-93.

Pichoff, S. and Lutkenhaus, J. (2005) 'Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA', *Mol Microbiol,* 55(6), pp. 1722-34.

Pimenta, A. L., Racher, K., Jamieson, L., Blight, M. A. and Holland, I. B. (2005) 'Mutations in HlyD, Part of the Type 1 Translocator for Hemolysin Secretion, Affect the Folding of the Secreted Toxin', *Journal of Bacteriology,* 187(21), pp. 7471-7480.

Pimenta, A. L., Young, J., Holland, I. B. and Blight, M. A. (1999) 'Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator', *Molecular and General Genetics MGG,* 261(1), pp. 122-132.

Richardson, J. S. and Richardson, D. C. (1988) 'Amino acid preferences for specific locations at the ends of alpha helices', *Science,* 240(4859), pp. 1648-52.

Sánchez-Magraner, L., Viguera, A. R., García-Pacios, M., Garcillán, M. P., Arrondo, J. L., de la Cruz, F., Goñi, F. M. and Ostolaza, H. (2007) 'The calcium-binding C-terminal domain of *Escherichia coli* alpha-hemolysin is a major determinant in the surface-active properties of the protein', *J Biol Chem,* 282(16), pp. 11827-35.

Sapay, N., Guermeur, Y. and Deléage, G. (2006a) 'Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier', *BMC Bioinformatics,* 7(1), pp. 255.

Sapay, N., Montserret, R., Chipot, C., Brass, V., Moradpour, D., Deléage, G. and Penin, F. (2006b) 'NMR structure and molecular dynamics of the in-plane membrane anchor of nonstructural protein 5A from bovine viral diarrhea virus', *Biochemistry,* 45(7), pp. 2221-33.

Schmitt, L., Benabdelhak, H., Blight, M. A., Holland, I. B. and Stubbs, M. T. (2003) 'Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains', *J Mol Biol,* 330(2), pp. 333-42.

Schülein, R., Gentschev, I., Mollenkopf, H.-J. and Goebel, W. (1992) 'A topological model for the haemolysin translocator protein HlyD', *Molecular and General Genetics MGG,* 234(1), pp. 155-163.

Shen, Y., Maupetit, J., Derreumaux, P. and Tufféry, P. (2014) 'Improved PEP-FOLD Approach for Peptide and Miniprotein Structure Prediction', *Journal of Chemical Theory and Computation,* 10(10), pp. 4745-4758.

Stanley, P., Koronakis, V. and Hughes, C. (1991) 'Mutational analysis supports a role for multiple structural features in the C-terminal secretion signal of *Escherichia coli* haemolysin', *Mol Microbiol,* 5(10), pp. 2391-403.

Stanley, P., Packman, L. C., Koronakis, V. and Hughes, C. (1994) 'Fatty acylation of two internal lysine residues required for the toxic activity of *Escherichia coli* hemolysin', *Science,* 266(5193), pp. 1992-6.

Thanabalu, T., Koronakis, E., Hughes, C. and Koronakis, V. (1998) 'Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore', *Embo j,* 17(22), pp. 6487-96.

Thévenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P. and Tufféry, P. (2012) 'PEP-FOLD: an updated *de novo* structure prediction server for both linear and disulfide bonded cyclic peptides', *Nucleic Acids Res,* 40(Web Server issue), pp. W288-93.

Thomas, S., Bakkes, P. J., Smits, S. H. and Schmitt, L. (2014) 'Equilibrium folding of pro-HlyA from *Escherichia coli* reveals a stable calcium ion dependent folding intermediate', *Biochim Biophys Acta,* 1844(9), pp. 1500-10.

Thompson, S. A. and Sparling, P. F. (1993) 'The RTX cytotoxin-related FrpA protein of *Neisseria meningitidis* is secreted extracellularly by *meningococci* and by HlyBD+ *Escherichia coli*', *Infect Immun,* 61(7), pp. 2906-11.

Vernon, R. M., Chong, P. A., Tsang, B., Kim, T. H., Bah, A., Farber, P., Lin, H. and Forman-Kay, J. D. (2018) 'Pi-Pi contacts are an overlooked protein feature relevant to phase separation', *Elife,* 7.

Wandersman, C. and Delepelaire, P. (1990) 'TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion', *Proc Natl Acad Sci U S A,* 87(12), pp. 4776-80.

Welch, R. A. (1991) 'Pore-forming cytolysins of gram-negative bacteria', *Mol Microbiol,* 5(3), pp. 521-8.

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B. and Schmitt, L. (2005) 'H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB', *Embo j,* 24(11), pp. 1901-10.

Zaitseva, J., Oswald, C., Jumpertz, T., Jenewein, S., Wiedenmann, A., Holland, I. B. and Schmitt, L. (2006) 'A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer', *Embo j,* 25(14), pp. 3432-43.

Zhang, F., Greig, D. I. and Ling, V. (1993) 'Functional replacement of the hemolysin A transport signal by a different primary sequence', *Proceedings of the National Academy of Sciences,* 90(9), pp. 4211-4215.

Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A. N. and Alva, V. (2018) 'A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core', *Journal of Molecular Biology,* 430(15), pp. 2237-2243.

3.4 Chapter 4 – Proteinase K susceptibility of TolC in a stalled HlyA T1SS

Title: Proteinase K susceptibility of TolC in a stalled HlyA T1SS

Authors: **Olivia Spitz** and Lutz Schmitt

Published in: *to be submitted*

Own Work: 90 %

Cloning of most plasmids (80 %)

PK susceptibility assay

SDS-sensitivity assay

Growth curves

Data analysis

Preparation of the figures

Writing of the manuscript

Chapter 4 – Proteinase K susceptibility of TolC in a stalled HlyA T1SS

Proteinase K susceptibility of TolC in a stalled HlyA T1SS

Olivia Spitz and Lutz Schmitt

Institute of Biochemistry, Heinrich-Heine-University Düsseldorf

Corresponding author: Lutz Schmitt, Institute of Biochemistry, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, E-Mail: Lutz.Schmitt@hhu.de

Key words: HlyA T1SS, assembly, limited proteolysis, TolC

Abstract

Gram-negative bacteria have evolved several specialized systems to secrete a variety of substrates from the cytosol to the extracellular space. The type 1 secretion system (T1SS) seems to be one of the more simple systems since it is only comprised out of three membrane proteins that together form a continuous channel to secrete the designated substrate. In the case of the 110 kDa substrate hemolysin A (HlyA) these membrane components are the ABC transporter HlyB, the membrane fusion protein (MFP) HlyD and the outer membrane protein (OMP) TolC. The ABC transporter and MFP form a stable complex in the inner membrane and only recruit the OMP upon substrate engagement. This recruitment process, which leads to the assembly of the T1SS, has been analyzed for the HasA T1SS by proteolytic digestion of the OMP by an unspecific protease and by monitoring the SDS-sensitivity of the host cells. We aimed to transfer both of these assays from the HasA T1SS to the HlyA T1SS and encountered several problems that are discussed here.

Introduction

Type 1 secretion systems (T1SSs) represent one of the several specialized secretion systems by which Gram-negative bacteria secrete substrates to their environment (Costa et al., 2015). They are comprised of three membrane bound components that together form a continuous channel from the cytosol to the extracellular space in order to secrete the substrate without any periplasmic intermediate across both membranes (Gray et al., 1986, Felmlee and Welch, 1988, Thanabalu et al., 1998). The hemolysin A (HlyA) T1SS from *Escherichia coli* is one of the best studied T1SS. It is made up of the outer membrane protein TolC, the ABC transporter HlyB and the membrane fusion protein (MFP) HlyD (Mackman et al., 1985, Wandersman and Delepelaire, 1990). The MFP and ABC transporter reside in the inner membrane and form a stable complex in the absence of the substrate HlyA (Thanabalu et al., 1998). Upon substrate engagement TolC is recruited and HlyA is secreted unfolded with its C-terminus reaching the cell surface first (Thanabalu et al., 1998, Bakkes et al., 2010, Lenders et al., 2015). Prior to secretion two internal lysine residues are acylated by the acyltransferase HlyC (Stanley et al., 1994). This acylation does not affect secretion, since the non-acylated version, pro-HlyA, is secreted to the same extent, but is necessary for HlyA activity (Nicaud et al., 1985, Stanley et al., 1994). The secretion can be blocked by fusing enhanced green fluorescent protein (eGFP) to the N-terminus of HlyA. The C-terminus still reaches the cell surface but the fast-folding eGFP is too large to pass the translocation channel and is therefore retained in the cytoplasm (Lenders et al., 2015). It has been shown that different regions in HlyA can interact with HlyB but to date it is not known which of these interactions leads to the recruitment of TolC and therefore assembly of the T1SS (Benabdelhak et al., 2003, Lecher et al., 2012).

The HasA T1SS from *S. marcescens* shares several features with the HlyA T1SS and can be heterologously expressed in *E. coli* in which it also uses TolC as the outer membrane component (Létoffé et al., 1994). Masi *et al.* applied two different assays to this system in order to investigate the assembly process and were able to identify multiple regions in the substrate HasA that interact with the transport components

(Masi and Wandersman, 2010). In one of these assays they permeabilized the outer membrane of *E. coli* to allow an unspecific protease to enter the periplasm. The digestion of TolC by this protease was different for the assembled and non-assembled system (Masi and Wandersman, 2010). In a second assay they exploited the fact that TolC is also used in the tripartite efflux pump AcrAB-TolC (Fralick, 1996). This pump is able to efflux many toxic substrates among them SDS (Cescau et al., 2007). Stalling the T1SS depletes the TolC pool of the cells leaving them more sensitive to SDS since they are not able to form a functional AcrAB-TolC complex. This sensitivity leads to reduced growth on SDS containing agar plates (Masi and Wandersman, 2010).

We aimed to apply both assays to a library of secretion deficient mutants of the HlyA T1SS in order to identify domains and/or regions important for TolC recruitment.

Results

PK susceptibility assay

When investigating the assembly of a transient complex, a permanently assembled complex and a permanently non-assembled complex are needed as positive and negative controls, respectively. A non-assembled T1SS can be achieved by removal of the transport components of the inner membrane or by removal of the substrate (Thanabalu et al., 1998). Therefore, the transport components for HlyA were expressed from one vector under the control of the same promotor, while the substrate was expressed from a second vector. Consequently, in experiments with non-assembled complex, the protein expressing plasmids (transporter components or substrate) were substituted with empty vectors so that all cells used in the assay had the same antibiotic resistances and were therefore treated equally. This lead to two possible plasmid combinations as a control for a non-assembled system: pK184-HlyBD + pUC19 (BD-empty) or pK184 + pBAD-eGFP-HlyA (empty-(eGFP)A).

In order to ensure permanent assembly of the T1SS, a fusion variant of HlyA was used with eGFP (enhanced green fluorescent protein) fused to its N-terminus.

While the C-terminus of HlyA can still be detected on the cell surface, the N-terminus is retained in the cytoplasm where fast folding eGFP acts as a plug stalling the T1SS (Lenders et al., 2015). Therefore all HlyA variants that are secreted to the same extent without fusion to eGFP can be used as controls for a permanently assembled system when eGFP is fused to their N-terminus. We choose full-length wild type HlyA instead of the shortened version HlyA1 leading to the plasmid combination pK184-HlyBD + pBAD-eGFP-HlyA (BD-(eGFP)A) as a control for a permanently assembled system.

When an assay is applied to a similar but different system, minor adjustments in experimental parameters are to be expected. In general, *E. coli* cells harboring different plasmids were grown, protein expression was induced, cells were harvested by centrifugation and the outer membrane was permeabilized by resuspension in a buffer containing 20% sucrose, 20 mM Tris-HCl pH 8.0, 10 mM EDTA and 0.5 % Triton X-100 to allow entry of proteinase K (PK) to the periplasm. The assay was stopped after 10 min by adding the protease inhibitor phenylmethylsulfonyl fluoride (PMSF) and samples were analyzed via SDS-PAGE and immunoblotting with an antibody directed against TolC (Werner et al., 2003, Masi and Wandersman, 2010).

Establishment of assay conditions and analysis method

In a first attempt to apply this assay to the HlyA T1SS two different buffers and an increased PK concentration (100 µg/mL final concentration) were used in comparison to Werner *et al.* and Masi *et al.* (Werner et al., 2003, Masi and Wandersman, 2010). As described above BD-(eGFP)A was used as a control for an assembled system and BD-empty as a control for a non-assembled system. Similar to Masi *et al.* an ATPase deficient mutant of HlyB (HlyB H662A (Zaitseva et al., 2005)) in combination with HlyD and eGFP-HlyA was also used (Masi and Wandersman, 2010). Boiling the samples at 95°C for 5 min led to an inability to detect TolC by immunoblotting (Supplementary Figure 1 A). When the samples were not heated only the trimeric state of TolC could be detected. In a buffer containing $CaCl_2$ instead of EDTA, the controls for assembled and non-assembled T1SS showed the same band pattern. In the buffer of Masi *et al.*, no PK dependent digestion of TolC was observed

in the non-assembled control (Figure 1 A) (Masi and Wandersman, 2010). This is in contrast to the findings of Werner *et al.*, who described the maturation pathway of TolC and found a 46 kDa digestion product of TolC when it is assembled into the outer membrane and digested with PK (Werner et al., 2003). A defined band for TolC was only visible in PK treated samples, while a blurry smear was detected in the non-treated samples (Figure 1 A).

Detergent solubilized TolC was digested with PK in the same buffers for 10 and 30 min in order to gain a better understanding of the possible band patterns (Figure 1 B). Samples were either boiled for 5 min at 90°C or not. Monomeric TolC was only observed in boiled samples while PK activity was observed in all samples. After 10 min a digestion product of >40 kDa was visible in the boiled samples, which disappeared in the 30 min samples. Two bands were visible between 35 and 25 kDa, which do not disappear in the 30 min samples, with the upper band likely corresponding to PK itself (28.9 kDa). Almost no time dependent difference was observed in the unboiled samples showing the importance of analyzing monomeric TolC instead of trimeric TolC. Also PK was not detected in the unboiled samples.

Therefore the assay buffer and the SDS-sample buffer were slightly changed: 6 M urea and 0.3% Coomassie were added to the SDS-sample buffer and Triton X-100 was removed from the assay buffer, while the sucrose concentration was increased to 30%. This allowed boiling of the samples and subsequent analysis by SDS-PAGE and immunoblotting (Supplementary Figure 1 B and C). The assay was performed for 1 and 5 min at 10°C, 20°C and 30°C with three plasmid combinations. Two of which would not allow assembly (BD-empty and empty-(eGFP)A) and one with permanent assembly (BD-(eGFP)A). The antibody used for this analysis displayed a band at 70 kDa, which is likely caused by unspecific binding. However digestion bands could be observed between 35 and 25 kDa in the PK-treated samples of the non-assembled controls, which were stronger in the 5 min samples than in the 1 min samples and correspond to the digestion product observed in the *in vitro* digestion (Figure 1 C). The difference between temperatures was not strong so that 20°C was chosen for further assays (Supplementary Figure 2 A and B). The 46 kDa

fragment described by Werner *et al.* and Masi *et al.* was not observed (Werner et al., 2003, Masi and Wandersman, 2010).

It is possible that PK further digests this fragment and therefore the PK-concentration was lowered to 10 µg/mL for further experiments. Also samples were treated with TCA (final concentration 15% (v/v)) to precipitate all digestion products, which were then resuspended in the same SDS-sample buffer. Since the 46 kDa fragment was still not visible (Supplementary Figure 2 C), we focused on the digestion products between 35 and 25 kDa. The identification of the correct digestion band was complicated by the cross-reactivity of the antibody. The digestion products showed weak signals in comparison to the unspecific band at 70 kDa and the band for non-digested TolC (55 kDa), which both gave a signal that saturated the detector (indicated by pink pixels).

A time-dependent assay was therefore performed in which samples were taken every two minutes to determine whether an increase in at least one of the bands between 35 and 25 kDa occurred (Figure 1 D). The increase was not visible by eye and quantification of the different bands was attempted.
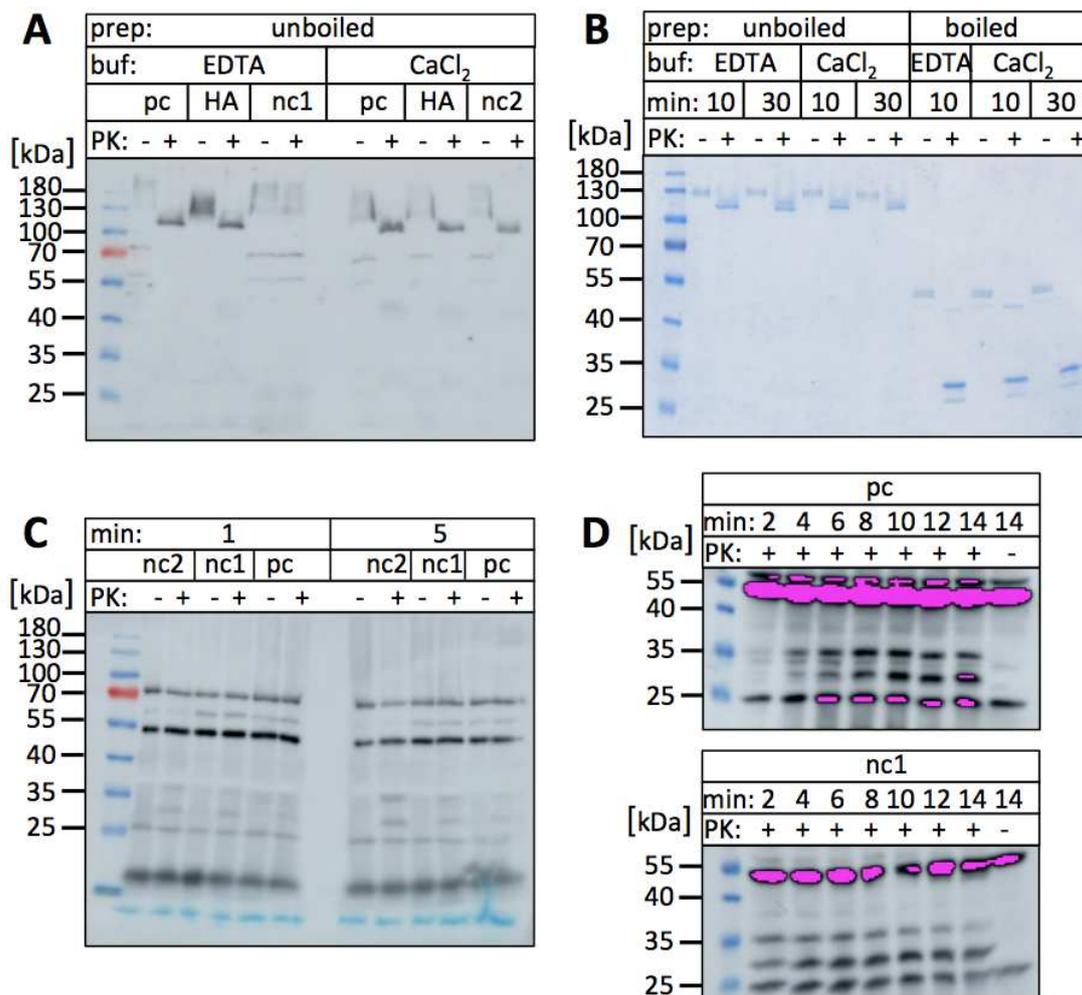
Figure 1: Initial PK susceptibility assay with TolC antibody that displays cross-reactivity. A, C, D) *E. coli* cells harboring either pK184-HlyBD and pBAD-eGFP-HlyA (pc) as a control for permanently assembled T1SS, pK184-HlyBD and pUC19 (nc1) or pK184 and pBAD-eGFP-HlyA (nc2) as controls for a non-assembled system were either treated with PK (+) or with buffer (-) as described in the material and methods section. Shown are Western Blots with an antibody against TolC (monomer: 55 kDa). Pink pixels indicate saturation of the detector. **A)** The assay was performed in two different buffers (buf) for 10 min. HA: pK184-HlyB*H662A-HlyD and pBAD-eGFP-HlyA. This HlyB mutant is deficient in ATPase activity (Zaitseva et al., 2005). **B)** Coomassie stained SDS-PAGE of *in vitro* digestion of detergent solubilized TolC with (+) or without (-) PK (28.9 kDa) for 10 and 30 min in two different buffers (buf). Monomeric TolC is only visible in boiled samples. **C)** The SDS-sample buffer was altered to contain Urea and Coomassie, which allowed boiling of the samples and detection of monomeric TolC. Signals above 55 kDa are caused by cross-reactivity of the antibody. The assay was performed in the presence of 30% sucrose for 1 and 5 min at 10°C. **D)** The assay was performed as in **C)** but at 20°C and samples were taken at more time points, which are indicated. All samples shown in this figure were quenched with PMSF.

Without a suitable standard, quantifications are still relative and should at least be normalized to a standard protein. We wanted to normalize the amount of TolC digestion product to the amount of HlyB for two reasons: i) the antibody against HlyB is directed against the NBD, which is located in the cytoplasm. A constant signal for

HlyB-NBD would therefore show that PK does not reach the cytoplasm and that the inner membrane is still intact after permeabilization and during the assay. ii) The amount of possible assembled T1SSs depends on the amount of HlyB (and HlyD). However, no HlyB was detected at all (Figure 2 A). Careful analysis of the buffer composition and preparation procedures revealed that handling the sample prior to SDS-PAGE analysis was not the reason for the inability to detect HlyB. Rather it appeared that PMSF precipitated under the conditions used. Thus, the use of aminoethylbenzolsulfonyl fluoride (AEBSF), as a water-soluble alternative to PMSF, was investigated. However, the reagent did not quench the reaction. No bands for TolC or HlyB were visible in the PK-treated samples but only in the samples not treated with PK (Figure 2 B and C). A suitable alternative was the addition of trifluoric acid (TFA) since TolC and HlyB could both be detected in samples quenched by this method. From this experiment on all further assays were quenched by addition of TFA and samples were neutralized with NaOH after 30 min of incubation to allow analysis by SDS-PAGE. Interestingly samples no longer required boiling to visualize monomeric TolC.

With this new quenching method normalization to HlyB was possible. The digestion product and HlyB migrate at different levels on a SDS-PAGE so that after transfer to a PVDF membrane this membrane can be cut and incubated with different antibodies. However, the signal for HlyB resolved by a 15 % SDS-PAGE shows a pattern of roughly three bands, which is unfavorable for quantification (Figure 2 D). Therefore the type of SDS-PAGE was changed to gradient gels (4-20%), which allowed analysis of HlyB and the digestion product of TolC at the same time (Figure 2 E). However, the cross-reactivity of the antibody was still a problem for two reasons: i) the number of bands between 35 and 25 kDa was not consistent and ii) the bands were not always resolved well enough to allow quantification.
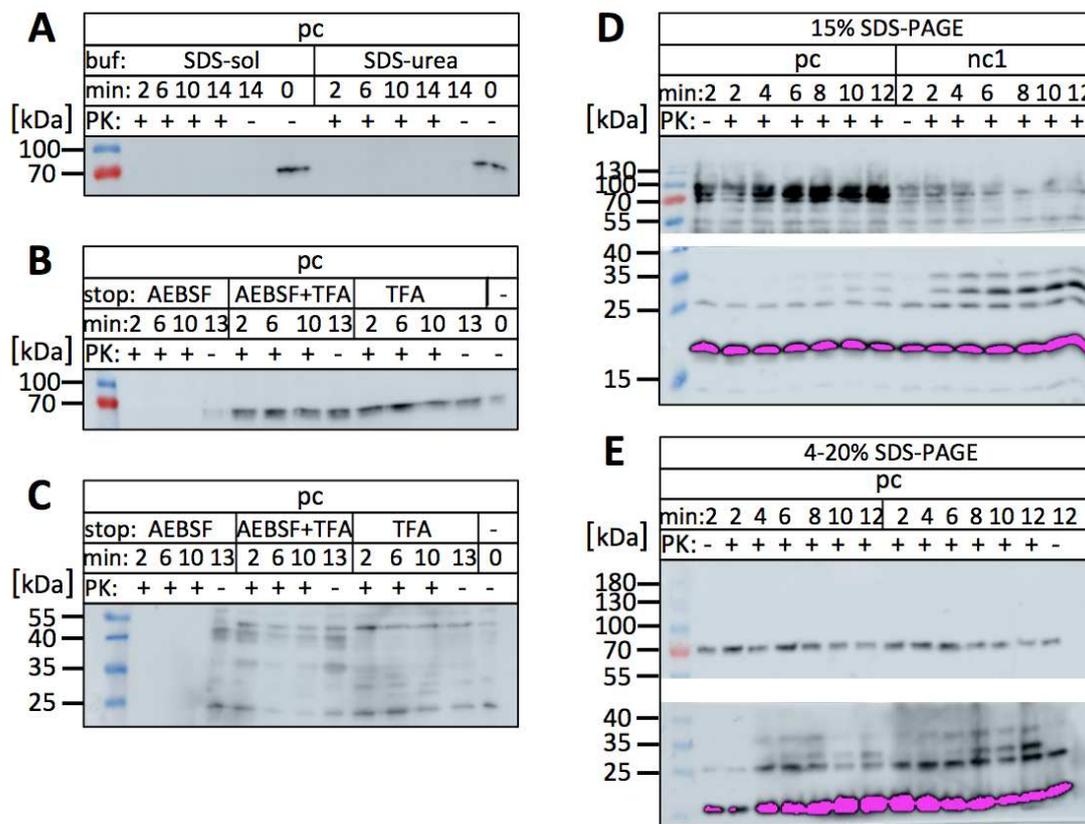
Figure 2: Western Blots of whole cell samples treated with PK. *E. coli* cells harboring either pK184-HlyBD and pBAD-eGFP-HlyA (pc) as a control for a permanently assembled T1SS or pK184-HlyBD and pUC19 (nc1) as a control for a non-assembled T1SS were treated with PK (+) or buffer (-) as described in material and methods. **A)** Western Blot with an antibody directed against HlyB-NBD. Samples were mixed with two different SDS-sample buffers (buf). Incubation time with PK is indicated in min. 0 min incubation means that the sample was not treated with PK and also not with quenching solution. The permeabilizied cells were directly mixed with SDS-sample buffer without any incubation. **B** and **C)** Different quenching methods (stop) were tested. Same samples were immunoblotted against HlyB-NBD (**B**) and TolC (**C**). Incubation time with PK is indicated in min. **D** and **E)** Different types of SDS-PAGEs were used for Western Blot. The membranes were cut and the upper half incubated with an antibody against HlyB-NBD and the lower half with an antibody against TolC. Incubation time is indicated in min. Pink pixels indicate saturation of the detector.

Finally, employing a newly produced TolC antibody allowed detection of the 46 kDa fragment observed by Werner *et al.* and Masi *et al.* (Figure 3 A, lowest panel) (Werner et al., 2003, Masi and Wandersman, 2010).

Introduction of a new TolC antibody

In an initial assay with the new TolC antibody the signal for the digestion product was fairly weak so that PK concentration was increased from 10 to 50 μg/mL (as final concentration) and assay time was elongated to 1 h. Samples were taken after 5, 10, 30 and 60 min and quenched with TFA. A sample not treated with PK

was taken after 5 and 60 min. The assay was performed with two different plasmid combinations: BD-(eGFP)A as a control for permanently assembled T1SS and BD-empty as a control for non-assembled T1SS. The samples were blotted with antibodies against TolC, HlyB-NBD and HlyD (Figure 3 A). Contrary to Masi *et al.* TolC was not sensitive to PK in the assembled system, while the non-assembled system showed a strong decrease of TolC (~55 kDa) and increase of a digestion product (~45 kDa) (Masi and Wandersman, 2010). Blotting against a cytoplasmic domain of HlyB showed that PK does not reach into the cytosol since the signal of HlyB did not decrease over time. However, different expression levels of HlyB in the two controls were visible, although the same plasmid and growth conditions were used in these controls. Interestingly, the levels of HlyD expression in both controls were similar. This points towards a mechanism in which HlyB is stabilized by the presence of the substrate, but HlyD is not affected. This also raised the question, if the observed sensitivity of TolC is caused by a lower expression level of HlyB.

Nonetheless the difference in TolC digestion by PK in BD-(eGFP)A and BD-empty was obvious. Therefore, the assay was performed with multiple constructs, which carried a mutation or domain deletion in either HlyA, HlyB or HlyD. They were analyzed in the same way and comparative samples of BD-(eGFP)A and BD-empty were always used as controls (Figure 3 B). It was expected that the digestion pattern of TolC in combination with the mutated constructs would either match with BD-(eGFP)A or BD-empty. However, a lot of constructs displayed the digestion bands while intact TolC did not seem to decrease therefore displaying an intermediate phenotype in regards to the used controls. Since total digestion of TolC was never reached (also not after over night incubation with 100 µg/mL PK, Figure 3 C) statements regarding the assembly would have to base on amount of digestion for which a quantification was required. Normalization to HlyB was deemed unsuitable for two reasons: On the one hand the expression of HlyB was very different and varied from cell to cell and on the other hand the same Blot could no longer be used for analysis with two antibodies as the signal separation was too little. Therefore, the amount of digestion fragment was normalized to the amount of total TolC.

Several problems were encountered. Samples not treated with PK were always needed as negative controls but contained the highest amount of TolC. An illumination time had to be determined in which this band does not show saturation to allow reliable quantification, but the digestion fragments are still visible. Furthermore intact TolC and digestion product have to be well separated to be quantified individually. This can either be achieved with longer migration times during the SDS-PAGE or with higher acrylamide/bisacrylamid content. The latter led to unfocused bands, which are unfavorable for quantification, but longer migration times increased the 'smile effect', which is also unfavorable for quantification. Gradient gels were also tested but showed different behavior during transfer to a PVDF membrane leading to Western Blots unsuitable for quantification. But even after systematic errors in quantification were removed by changing the quantification software and compromises for illumination and running parameters of SDS-PAGEs were found still no construct showed as much digestion of TolC as BD-empty implying at least partial assembly in all 16 constructs.
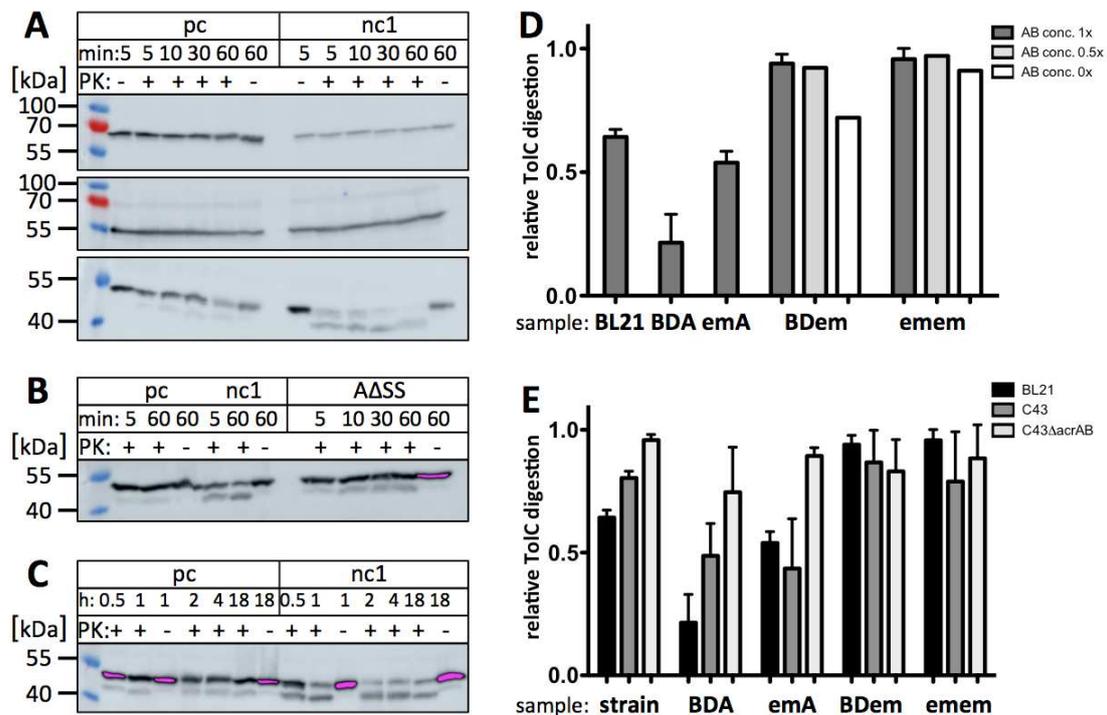
Figure 3: Western bolts and quantification of PK susceptibility assay with specific TolC antibody. A, B and C) Western Blots of whole cell samples that were incubated with PK (+) or buffer (-) for different times as indicated in min and h. Pc = positive control for permanently assembled system with plasmid combination: pK184-HlyBD and pBAD-eGFP-HlyA. Nc1 = negative control 1 for non-assembled system with plasmid combination: pK184-HlyBD and pUC19. Pink pixels indicate saturation of the detector. **A)** Same samples were blotted with different antibodies: Upper panel shows HlyB (~80 kDa), which shows more expression in pc than nc1, middle panel shows HlyD (~55 kDa), which is expressed equally in all samples and lowest panel shows TolC (~55 kDa), which is digested more in nc1 than pc. **B)** Shown is the digestion pattern of TolC by PK for pc and nc1 in comparison to AΔSS (plasmid combination: pK184-HlyBD and pBAD-eGFP-HlyAΔSS → the secretion signal (SS) of HlyA (C-terminal 60 residues) was deleted). The construct AΔSS shows more digestion than pc but less than nc1. **C)** Digestion of TolC with 100 µg/mL PK in constructs pc and nc1 for up to 18 h. A band for monomeric TolC is visible in all samples. Complete digestion was not possible under these conditions. **D** and **E)** Quantification of TolC digestion in 60 min samples as described in material and methods. BD = HlyBD expressed from pK184-HlyBD. A = eGFP-HlyA expressed from pBAD-eGFP-HlyA. Em = empty vector (pK184 when replacing HlyBD, pUC19 when replacing eGFP-HlyA). **D)** BL21 = *E. coli* BL21(DE3) cells without any plasmid. The assay was performed with the usual antibiotic concentration (AB conc. 1x, grey). The assay was also performed at 0.5x antibiotic concentration (light grey) and without antibiotics (white) for BDem and emem with n =1. Grey bars represent mean ± SD of at least two independent experiments. **E)** Strain = *E. coli* strain without any plasmid with *E. coli* BL21(DE3) in black, *E. coli* C43(DE3) in grey and *E. coli* C43(DE3)Δ*acrAB* in white. All bars represent mean ± SD of at least two independent experiments.

Introduction of multiple controls for non-assembled system

We therefore decided to analyze the digestion of TolC by PK using additional controls and also assayed *E. coli* cells carrying both empty plasmids (empty-empty) and no plasmids at all (BL21) leading to five constructs: BD-(eGFP)A, BD-empty, empty-(eGFP)A, empty-empty and BL21.

It should be noted that cells not carrying any plasmid grew faster and were not treated with inducer (IPTG and arabinose). All other constructs were induced with 1 mM IPTG and 1 mM arabinose at OD_{600} 0.5-0.7 and grown for 2 h before harvesting. *E. coli* cells without any plasmid were harvested and used for the assay when they reached a comparable OD_{600} to the other constructs, which was usually between 2.0 and 3.0.

By applying students t-test to the quantification results the five control constructs could be divided in three groups. The lowest digestion was observed for BD-(eGFP)A, an intermediate digestion for empty-(eGFP)A and BL21 and highest digestion for BD-empty and empty-empty (Figure 3 D). Intermediate and high digestion was therefore present in constructs that would not allow assembly of the T1SS but was significantly different between the two groups. The common denominator in the constructs that displayed high digestion was the presence of pUC19 as an empty vector to confer ampicillin resistance. Interestingly, such high digestion was not observed in experiments in which only pUC19 was present and therefore only ampicillin was used.

In an attempt to reduce stress of the cells we performed the assay in the absence of antibiotics for the plasmid combinations BD-empty and empty-empty, but TolC was still highly sensitive to PK treatment (Figure 3 D).

As mentioned in the introduction TolC is also used in other T1SSs and tripartite efflux pumps such as AcrAB-TolC, which confers resistance to some antibiotics, among them ampicillin (Sulavik et al., 2001). Ampicillin resistance is conferred by pBAD-eGFP-HlyA and therefore present in the culture medium of each construct. An influence of AcrAB and competition for the available TolC pool are therefore possible. An *acrAB* knock out strain in the *E. coli* C43(DE3) background was available in our laboratory and we transformed all four plasmid combinations in *E. coli* C43(DE3) and *E. coli* C43(DE3)Δ*acrAB* and repeated the assay twice including each strain without any plasmids (as for *E. coli* BL21(DE3)). Both strains are able to secrete pro-HlyA and are therefore able to assemble the HlyA T1SS.

The digestion of TolC in regards to the different plasmid combinations was similar in *E. coli* C43(DE3) compared to *E. coli* BL21(DE3) although digestion was overall higher (Figure 3 E). In *E. coli* C43(DE3)Δ*acrAB* the digestion of TolC was overall very high and no difference was visible among the plasmid combinations. Since the digestion was very different for the three strains even without any plasmid present no definite conclusions regarding the involvement of the AcrAB-TolC complex could be drawn. Also the protection of TolC by the permanently assembled system (BD-(eGFP)A) was gradually lost from *E. coli* BL21(DE3) to *E. coli* C43(DE3) to *E. coli* C43(DE3)Δ*acrAB.*

Quantification of HlyD digestion

The membrane fusion protein HlyD has a large periplasmic domain and likely changes its conformation when recruiting TolC (Pimenta et al., 1999). This change in conformation could result in different sensitivity towards PK and the digestion of HlyD by PK was therefore also investigated.

Digestion was observed for the permanently assembled system (BD-(eGFP)A) as well as for the non-assembled system (BD-empty) and was more severe for the non-assembled system (Figure 4 A). However even without PK treatment HlyD displays degradation signals (Supplementary Figure 2 D) and the used antibody displays cross-reactivity. While unspecific binding at 70 kDa could be ignored, the antibody also binds a protein, which migrates at the same height as HlyD, which is best visible in samples where a truncated version of HlyD was used (Figure 4 B).

This renders quantification of HlyD digestion bands and intact HlyD impossible and in fact no time-dependent increase in digestion product or time-dependent decrease of intact HlyD could be observed.
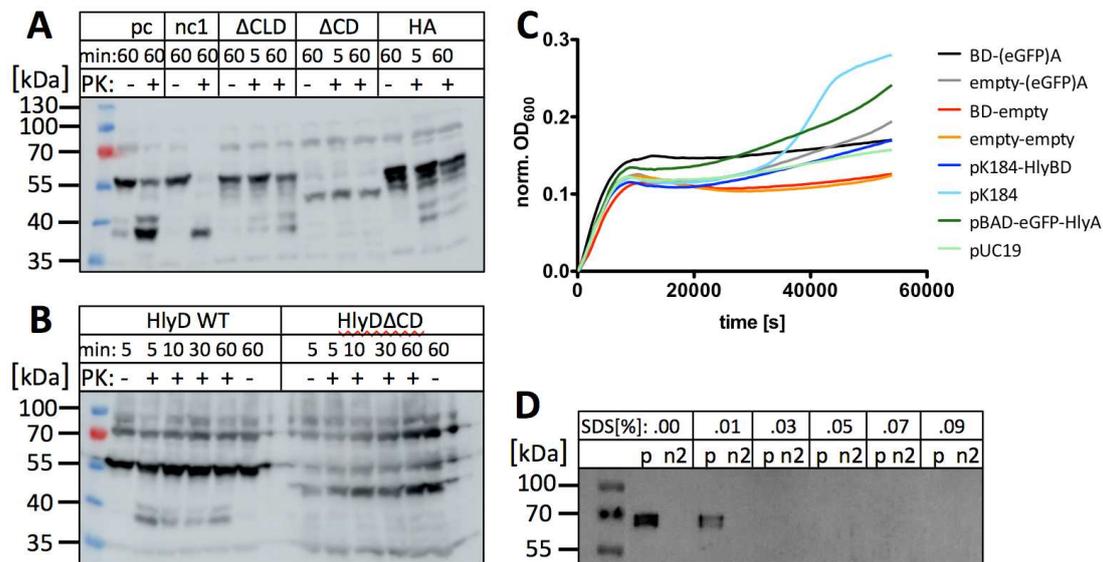
Figure 4: Explored alternatives to TolC quantification and PK susceptibility assay. A and **B)** Western Blots of whole cell samples that were incubated with PK (+) or buffer (-) for different times as indicated. Western Blots were developed with an antibody directed against HlyD (55 kDa). The antibody displays cross-reactivity with a protein of ~70 kDa. **A)** Pc = positive control for permanently assembled system with plasmid combination pK184-HlyBD and pBAD-eGFP-HlyA. Nc1 = negative control 1 for non-assembled system with plasmid combination pK184-HlyBD and pUC19. The constructs ΔCLD, ΔCD and HA all hold pBAD-eGFP-HlyA and pK184-HlyBD with different mutations. ΔCLD = C39-like domain of HlyB (residue 2 – 145) has been deleted. ΔCD = cytoplasmic domain of HlyD (residue 1-60) has been deleted (HlyDΔCD = 47 kDa). HA = point mutation H662A in HlyB, which leaves HlyB ATPase inactive (Zaitseva et al., 2005). Degradation bands of HlyD are visible at ~40 kDa but also present in samples without PK (-). In the 60 min sample nc1 shows almost no undigested HlyD. In **B)** the cross-reactivity of the HlyD antibody is shown. Samples of HlyD wild type (WT) show a clear band at 55 kDa. HlyDΔCD is smaller with 47 kDa in size but a clear band is visible at 55 kDa as well. Also the signals above 70 kDa are likely caused by cross-reactivity. **C)** Growth curves of *E. coli* BL21(DE3) with different plasmid combinations. BD is always expressed from pK184-HlyBD. (eGFP)A is always expressed from pBAD-eGFP-HlyA. Empty refers to empty vectors pK184 or pUC19. Cells were grown to OD_{600} of 0.5-0.7, protein expression was induced and OD_{600} was measured over 16 h. OD_{600} was normalized (norm.) by subtracting initial OD_{600} from all samples. **D)** Western Blot of whole cell samples with an antibody against HlyB (~80 kDa). Cells were grown in the presence of different SDS-concentrations (% (w/v)). P = pc (as in **A**). n2 = negative control 2 for non-assembled system with plasmid combination pK184 and pBAD-eGFP-HlyA. The signal for HlyB in p decreases with increasing SDS-concentrations.

Growing behavior of PK susceptibility assay controls

In addition to analyzing the PK susceptibility of TolC and HlyD the growth behavior of several plasmid combinations was analyzed. They were grown in a 96-well plate over night in the presence and absence of inducers of protein expression and their OD_{600} was monitored. The plasmid combinations that showed the highest TolC digestion (BD-empty, empty-empty) also showed the least growth independent of induction of protein expression (Figure 4 C and Supplementary

Figure 3). This reduced growth could not be contributed to any of the plasmids themselves (pK184-HlyBD, pK184-empty, pUC19), since cells with only one of these plasmids grew to higher ODs. Also the phenotype is not simply explained by the presence of two plasmids, since BD-(eGFP)A and empty-(eGFP)A also hold two plasmids and grew to higher ODs. In all four two-plasmid systems the origin of replication of the plasmids as well as the sensitivity markers are compatible. However, in the two two-plasmid systems that grew normally, different promotors control protein expression (pK184(-HlyBD) holds the lac promotor, pBAD-eGFP-HlyA the arabinose promotor). In the two two-plasmid systems that showed reduced growth the same promotor controls protein expression (lac promotor). Without further analysis it is not clear if the growth phenotype is caused by the combination of two plasmids with the same promotor. Also the combination of pK184-HlyBD with pSU2726-HlyA, which shares most of its backbone with pUC19 and also holds the lac promotor, is frequently used for HlyA secretion (Reimann et al., 2016, Khosa et al., 2018). Still the reduced growth of BD-empty and empty-empty shows that the cells are more stressed than in BD-(eGFP)A and empty-(eGFP)A and this might explain why only these two plasmid combinations showed such a high TolC digestion by PK. The outer membrane might be easier to permeabilize allowing more PK to enter the periplasm leading to more digestion.

TolC digestion in comparison to other controls for non-assembled system

The different growth behavior of BD-empty and empty-empty shows that they are not suitable as controls for a non-assembled T1SS. Therefore samples from the PK susceptibility assay were compared to empty-(eGFP)A or *E. coli* BL21(DE3). For a reliable comparison by quantification the samples should be immunoblotted together on the same membrane and samples were therefore re-analyzed. Interestingly, BD-(eGFP)A showed more digestion product of TolC than before and the assay was repeated twice for several constructs including BD-(eGFP)A, empty-(eGFP)A and *E. coli* BL21(DE3). However the protection conferred by BD-(eGFP)A was reduced although all cultures of the PK susceptibility assay had been induced from the same

cryo-conserved culture, which was stored at -80°C. Due to the increased digestion of TolC in BD-(eGFP)A the difference between assembled and non-assembled system was little and according to statistical analysis not significant anymore (Supplementary Figure 4). The former observed significance might have been regained with more experiments or freshly transformed cultures.

<u>SDS-sensitivity assay</u>

Since the PK susceptibility assay did not yield any clear results regarding the assembly of the HlyA T1SS, an SDS-sensitivity assay was also applied. If the T1SS is permanently assembled less or no TolC is available for the AcrAB-TolC efflux pump leaving cells more sensitive to SDS (Cescau et al., 2007). This sensitivity would lead to fewer colonies on SDS containing agar plates or to reduced growth in SDS containing liquid medium. In a first attempt liquid cultures of BD-(eGFP)A and empty-(eGFP)A were grown in the presence of different SDS concentrations. However, Western Blot analysis revealed that the expression of HlyB is drastically reduced in the presence of 0.01% (w/v) SDS and non detectable in the presence of 0.03% (w/v) SDS. This effect was also observed when cells were pre-incubated without SDS to allow protein expression and then transferred to SDS-containing medium (Figure 4 D). The SDS sensitivity assay is therefore not suited for the HlyA T1SS under the tested conditions.

Discussion

Secretion of the 110 kDa toxin HlyA is a complex process that can be divided into multiple steps. Since TolC recruitment was so far only observed in the presence of the substrate, substrate recognition precedes TolC recruitment (Thanabalu et al., 1998).

Substrate recognition can be investigated by *in vitro* pull down assays and has successfully been performed with isolated parts of HlyB and truncated versions of HlyA (Benabdelhak et al., 2003, Lecher et al., 2012). Also the stimulation of HlyB

ATPase activity by HlyA can be assayed and different mutants of these proteins could be applied to this assay (Reimann et al., 2016).

Interactions of different parts of HlyA with different parts of HlyB have been determined and it is not clear which of these interactions or a combination of them triggers the recruitment of TolC. With a combination of secretion experiments and the here described assembly assays Masi *et al*. were able to identify multiple regions distributed along HasA that interact with HasD and called them primary recognition sites (Masi and Wandersman, 2010).

We attempted to transfer both assembly assays to the HlyA T1SS but encountered a number of problems. The SDS-sensitivity assay is based on a depletion of available TolC by a permanently assembled T1SS, which leaves the cells more sensitive to SDS. However, HlyB expression was not detectable at SDS concentrations that would have an impact on cell growth and depletion of the available TolC pool was therefore not possible. This might point to a protection mechanism of the cells that might be circumvented by using a different *E. coli* strain, different expression parameters for HlyB or another substrate for the AcrAB-TolC system instead of SDS.

Proteins usually change their conformation when interacting with each other, which can lead to different digestion patterns upon protease treatment. Therefore permeabilizing the outer membrane to allow entry of an unspecific protease to the periplasm is a possibility to investigate the assembly of a complex.

A specific antibody is required to draw reliable conclusions, which was a reoccurring problem during the establishment of the PK susceptibility assay. Also *in vivo* experiments are influenced by a number of cellular processes and the variability among samples should be reduced as much as possible. Although the same plasmid and growth conditions for HlyBD expression were used, different expression levels of HlyB were observed in the presence or absence of the substrate HlyA. Furthermore different growing behaviors were observed among the positive and negative controls for the assay indicating different levels of stress on the cells. In this context stress is used as a loose term and the cellular responses to that were not further investigated. Differences in the proteome as well as the lipidome of the cells are possible. Both

could impact the stability of TolC itself as well as the permeabilization efficiency, which would increase the variability between samples.

Apart from determining the optimal assay conditions (temperature, protease concentration, time, assay buffer, quench solution, etc.) and sample preparation methods (SDS-sample buffer, (no) heating, PAGE style, etc.), evaluation is also a critical step. A simple black or white answer regarding the digestion of TolC was not possible so that quantification was required. Differences in signal intensities of PK treated and not PK treated samples as well as poor vertically and horizontally separation of the signals complicated the analysis.

In conclusion the variability between the samples in combination with the sensitive evaluation method made definite statements regarding the assembly of the T1SS impossible. Therefore, other methods should be investigated.

Thanabalu *et al.* used an unspecific crosslinker and subsequent purification of the crosslinked product to investigate if TolC is recruited in the presence or absence of different T1SS components (Thanabalu et al., 1998). This assay could be repeated with mutated versions of the proteins to gain more detailed information.

The AcrAB-TolC complex is able to efflux a variety of substrates with SDS being only one of them (White et al., 1997, Tsukagoshi and Aono, 2000, Yu et al., 2003). The reduction of HlyB expression, which made the SDS-sensitivity assay not applicable, might not be present when another substrate of AcrAB-TolC is used.

Another *in vivo* approach could be the addition of FRET sensors to HlyD and TolC. This would likely need an extensive establishment period in which the correct localization of both tagged proteins has to be verified and whether secretion is still possible in the presence of the sensors.

The variability between constructs could be reduced the most if the HlyA T1SS could be reconstituted *in vitro*. This approach raises its one set of issues: The inner membrane complex as well as TolC has to be reconstituted separately. Furthermore the substrate is transported in an unfolded state and has to be kept unfolded in order to interact with the inner membrane complex. Therefore, a method or buffer system has to be found, which keeps HlyA unfolded while HlyBD and TolC are unaffected.

Also it is not know if other periplasmic components such as peptidoglycan are required for correct assembly of the T1SS.

Material and Methods

Plasmids, transformation and *E. coli* cultures

If not stated otherwise, *E. coli* BL21(DE3) was used for all experiments. As described in the text, the PK susceptibility assay was also performed in *E. coli* C43(DE3) and *E. coli* C43(DE3)Δ*acrAB*. All *E. coli* cells were made competent by incubation in 0.1 M CaCl$_2$ and transformed with the heat shock method and clones were selected on antibiotic containing agar plates (100 µg/mL ampicillin and/or 30 µg/mL kanamycin) (Mandel and Higa, 1970, Froger and Hall, 2007). If two plasmids were used cells were transformed sequentially. HlyB, HlyD and variants of them (e. g. HlyB H662A or HlyDΔCD) were expressed from pK184-HlyBD (Jenewein, 2008). In some experiments this vector was supplemented with the empty vector pK184. The fusion protein eGFP-HlyA and variants of it were expressed from pBAD-eGFP-HlyA, which was supplemented with the empty vector pUC19 in some experiments as indicated in the text. All liquid cultures were incubated at 37°C and 180 rpm. In general, a mixture of transformed clones from the agar plate was used to inoculate 5 mL cultures of 2xYT media with the corresponding antibiotics, which were grown for 16 h. These cultures were harvested by centrifugation, resuspended in 500 µL of fresh media, mixed with glycerol (final concentration: 25% (v/v)) and stored at -80°C. These cryo-conserved cultures were used to inoculate 5 mL precultures with the same media and antibiotics, which were grown for 16 h. Their OD$_{600}$ was measured and they were used to inoculate 50 mL main cultures (same media and antibiotic concentration) to OD$_{600}$ 0.1. At an OD$_{600}$ of 0.5-0.7, protein expression was induced with 1 mM IPTG and 1 mM arabinose and cultures were grown for 2 h. After this 5 to 12 mL of the culture were harvested by centrifugation at 4000xg for 10 min at 4°C. As mentioned in the text *E. coli* cells that did not contain any plasmid grew faster, were not treated with antibiotics or inducers and were harvested at an OD$_{600}$ comparable to the *E. coli* cultures containing plasmids.

PK susceptibility assay

The general protocol for this assay was taken from Masi *et al.* and was adapted several times (Masi and Wandersman, 2010). Cells treated as described above were resuspended to OD_{600} of 15 in permeabilization buffer to allow entry of PK to the periplasm. Three different permeabilization buffers were used: i) $CaCl_2$-buffer: 20 mM Tris pH 8.0, 10 mM $CaCl_2$, 20% (w/v) sucrose, 0.5% (v/v) Triton X-100 ii) EDTA-buffer: 20 mM Tris pH 8.0, 10 mM EDTA, 20% (w/v) sucrose, 0.5% (v/v) Triton X-100 iii) final buffer: 20 mM Tris pH 8.0, 10 mM EDTA, 30% (w/v) sucrose.

The assay was started by mixing 90 µL permeabilized cell suspension with 10 µL PK-solution. Different final PK concentrations were used as indicated in the text (10-100 µg/mL). PK from *Tritirachium album* (Sigma Aldrich, Germany) was solved and stored according to manufactures instructions. Negative controls contained 10 µL of 20 mM Tris pH 8.0 instead of PK. The assay was performed at 20°C and 650 rpm in a thermocycler if not stated otherwise. Different digestion times were tested ranging from 1 min to 18 h.

Three different quenching methods were tested. At first 50 µL PK digested cell suspension were mixed with 50 µL 10 mM phenylmethylsulfonyl fluoride (PMSF) in ethanol. Then PMSF was exchanged for 10 mM aminoethylbenzolsulfonyl fluoride (AEBSF) in water. The third and final quenching method was mixing 50 µL PK digested cell solution with 4 µL 25% (v/v) trifluoric acid (TFA). The samples were incubated with TFA for 30 min before being neutralized with ~2.5 µL 5 M NaOH. 20 µL SDS-sample buffer were added to this solution. Two different sample buffers were used: SDS-soluble (100 mM Tris pH 6.8, 3.3% (w/v) SDS, 0.02% (w/v) bromophenol blue, 40% (v/v) glycerol) and SDS-urea (same as SDS-soluble + 6 M urea and 0.2% (w/v) Coomassie). At first SDS-soluble was used which did not allow boiling (95°C for 5-10 min) of the SDS-samples as described in the text. Then the sample buffer was changed to SDS-urea. When the quenching method was changed to TFA, SDS-soluble was used again.

Analysis of PK susceptibility assay

Samples were analyzed on different types of SDS-PAGEs (10%, 15% and 4-20%) as described in the text. Mostly 10% SDS-PAGES were used. For Western Blot analysis the proteins were transferred to a polyvinylidene difluoride (PVDF) membrane with a Trans-Blot SD Semi-Dry Transfer Cell by Bio Rad according to manufactures instructions (1 h, 0.1 A per membrane, adjustable voltage). The membranes were blocked with 10% (w/v) milk powder overnight at 4°C, incubated with the respective primary antibody for 1.5-2 h at room temperature (RT), then incubated with a secondary antibody coupled to horse radish peroxidase (HRP) for 1 h at RT, washed three times in between each step with TBS-T buffer (20 mM Tris pH 8, 250 mM NaCl, 0.1% (v/v) Tween-20) and analyzed using Amersham™ Imager 600 by GE Healthcare. Pictures were taken every 10 s until saturation of the detector was reached. The image 10 s prior to saturation was used for quantification.

At first the open-source software FIJI (Schindelin et al., 2012) was used for quantification and later changed to GelAnalyzer (GelAnalyzer 19.1 (www.gelanalyzer.com) by Istvan Lazar Jr., PhD and Istvan Lazar Sr., PhD, CSc). In general, a region of interest is defined and the software creates an intensity plot for this region. Baselines are set either manually (FIJI) or automatically (GelAnalyzer) and the integral of the peaks, corresponding to the intensity of the bands on the Western Blot, is used for further evaluation. The scaling of the y-axis in the intensity plot needs to be constant when the amount of two different bands should be compared. The free software FIJI does not allow this, which is why we changed to GelAnalyzer after detecting this problem. In both approaches the digestion product of TolC at ~45 kDa and the remaining TolC at 55 kDa were quantified individually and the amount of digestion product was divided by the amount of total TolC (signal at 46 kDa + signal at 55 kDa).

TCA precipitation

Trichloroacetic acid (TCA) was added to the samples to a final concentration of ~13% (v/v). Samples were incubated at 4°C for >16 h to allow precipitation of

proteins and peptides. Samples were centrifuged at 20,000xg for 20 min at 4°C and the pellets were washed with 100% acetone three times before being resuspended in SDS-sample buffer (soluble). They were analyzed by SDS-PAGE and Western Blot as described above.

Growth curves

Pre- and main cultures were prepared from cryo-conserved cultures as described above. 1 h after induction of protein expression, the OD_{600} was measured and cultures were used to inoculate 96 well plates (same media, antibiotics and inducer concentration) to an OD_{600} of 0.15. The growth was monitored for 16 h in a microplate reader (Tecan). Each plasmid combination was measured as triplicates. As mentioned in the text the growth was also monitored in the absence of inducers. These cultures were treated the same way but were never treated with IPTG and/or arabinose (Supplementary Figure 3).

SDS-sensitivity assay

Pre- and main cultures were prepared from cryo-conserved cultures as described above. 1 h after induction of protein expression, the OD_{600} was measured and cultures were used to inoculate cultures with different SDS-concentrations (from 0 to 0.09% (w/v) SDS). Their OD_{600} was monitored for several hours. Also samples were taken after 5 h, 6 h and 7 h and analyzed with SDS-PAGE and Western Blot for the presence of HlyB, HlyD and HlyA of which only the Western Blot of HlyB 6 h after protein induction is shown (Figure 4 D).

Acknowledgments

project name Identity and Dynamics of Membrane Systems – From molecules to Cellular Functions (project A01 to L.S.).

References

Bakkes, P. J., Jenewein, S., Smits, S. H., Holland, I. B. and Schmitt, L. (2010) 'The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin type 1 secretion system', *J Biol Chem,* 285(52), pp. 40573-80.

Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B. and Schmitt, L. (2003) 'A Specific Interaction Between the NBD of the ABC-transporter HlyB and a C-Terminal Fragment of its Transport Substrate Haemolysin A', *Journal of Molecular Biology,* 327(5), pp. 1169-1179.

Cescau, S., Debarbieux, L. and Wandersman, C. (2007) 'Probing the in vivo dynamics of type I protein secretion complex association through sensitivity to detergents', *J Bacteriol,* 189(5), pp. 1496-504.

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M. and Waksman, G. (2015) 'Secretion systems in Gram-negative bacteria: structural and mechanistic insights', *Nature Reviews Microbiology,* 13(6), pp. 343-359.

Felmlee, T. and Welch, R. A. (1988) 'Alterations of amino acid repeats in the *Escherichia coli* hemolysin affect cytolytic activity and secretion', *Proceedings of the National Academy of Sciences,* 85(14), pp. 5269.

Fralick, J. A. (1996) 'Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of Escherichia coli', *J Bacteriol,* 178(19), pp. 5803-5.

Froger, A. and Hall, J. E. (2007) 'Transformation of plasmid DNA into *E. coli* using the heat shock method', *J Vis Exp,* (6), pp. 253.

Gray, L., Mackman, N., Nicaud, J. M. and Holland, I. B. (1986) 'The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*', *Mol Gen Genet,* 205(1), pp. 127-33.

Jenewein, S. (2008) *The Escherichia coli haemolysin transporter - A paradigm for Type I secretion.* PhD Doctoral Dissertation, Heinrich-Heine-University Duesseldorf.

Khosa, S., Scholz, R., Schwarz, C., Trilling, M., Hengel, H., Jaeger, K. E., Smits, S. H. J. and Schmitt, L. (2018) 'An A/U-Rich Enhancer Region Is Required for

High-Level Protein Secretion through the HlyA Type I Secretion System', *Appl Environ Microbiol,* 84(1).

Lecher, J., Schwarz, Christian K. W., Stoldt, M., Smits, Sander H. J., Willbold, D. and Schmitt, L. (2012) 'An RTX Transporter Tethers Its Unfolded Substrate during Secretion via a Unique N-Terminal Domain', *Structure,* 20(10), pp. 1778-1787.

Lenders, M. H. H., Weidtkamp-Peters, S., Kleinschrodt, D., Jaeger, K.-E., Smits, S. H. J. and Schmitt, L. (2015) 'Directionality of substrate translocation of the hemolysin A Type I secretion system', *Scientific Reports,* 5(1), pp. 12470.

Létoffé, S., Ghigo, J. M. and Wandersman, C. (1994) 'Secretion of the *Serratia marcescens* HasA protein by an ABC transporter', *J Bacteriol,* 176(17), pp. 5372-7.

Mackman, N., Nicaud, J. M., Gray, L. and Holland, I. B. (1985) 'Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*', *Mol Gen Genet,* 201(3), pp. 529-36.

Mandel, M. and Higa, A. (1970) 'Calcium-dependent bacteriophage DNA infection', *J Mol Biol,* 53(1), pp. 159-62.

Masi, M. and Wandersman, C. (2010) 'Multiple signals direct the assembly and function of a type 1 secretion system', *J Bacteriol,* 192(15), pp. 3861-9.

Nicaud, J. M., Mackman, N., Gray, L. and Holland, I. B. (1985) 'Characterisation of HlyC and mechanism of activation and secretion of haemolysin from *E. coli* 2001', *FEBS Lett,* 187(2), pp. 339-44.

Pimenta, A. L., Young, J., Holland, I. B. and Blight, M. A. (1999) 'Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator', *Molecular and General Genetics MGG,* 261(1), pp. 122-132.

Reimann, S., Poschmann, G., Kanonenberg, K., Stühler, K., Smits, Sander H. J. and Schmitt, L. (2016) 'Interdomain regulation of the ATPase activity of the ABC transporter haemolysin B from Escherichia coli', *Biochemical Journal,* 473(16), pp. 2471-2483.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. (2012) 'Fiji: an open-source platform for biological-image analysis', *Nature Methods,* 9(7), pp. 676-682.

Stanley, P., Packman, L. C., Koronakis, V. and Hughes, C. (1994) 'Fatty acylation of two internal lysine residues required for the toxic activity of *Escherichia coli* hemolysin', *Science,* 266(5193), pp. 1992-6.

Sulavik, M. C., Houseweart, C., Cramer, C., Jiwani, N., Murgolo, N., Greene, J., DiDomenico, B., Shaw, K. J., Miller, G. H., Hare, R. and Shimer, G. (2001) 'Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes', *Antimicrob Agents Chemother,* 45(4), pp. 1126-36.

Thanabalu, T., Koronakis, E., Hughes, C. and Koronakis, V. (1998) 'Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore', *Embo j,* 17(22), pp. 6487-96.

Tsukagoshi, N. and Aono, R. (2000) 'Entry into and release of solvents by *Escherichia coli* in an organic-aqueous two-liquid-phase system and substrate specificity of the AcrAB-TolC solvent-extruding pump', *J Bacteriol,* 182(17), pp. 4803-10.

Wandersman, C. and Delepelaire, P. (1990) 'TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion', *Proc Natl Acad Sci U S A,* 87(12), pp. 4776-80.

Werner, J., Augustus, A. M. and Misra, R. (2003) 'Assembly of TolC, a structurally unique and multifunctional outer membrane protein of *Escherichia coli* K-12', *J Bacteriol,* 185(22), pp. 6540-7.

White, D. G., Goldman, J. D., Demple, B. and Levy, S. B. (1997) 'Role of the acrAB locus in organic solvent tolerance mediated by expression of marA, soxS, or robA in *Escherichia coli*', *J Bacteriol,* 179(19), pp. 6122-6.

Yu, E. W., Aires, J. R. and Nikaido, H. (2003) 'AcrB multidrug efflux pump of *Escherichia coli*: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity', *J Bacteriol,* 185(19), pp. 5657-64.

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B. and Schmitt, L. (2005) 'H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB', *Embo j,* 24(11), pp. 1901-10.
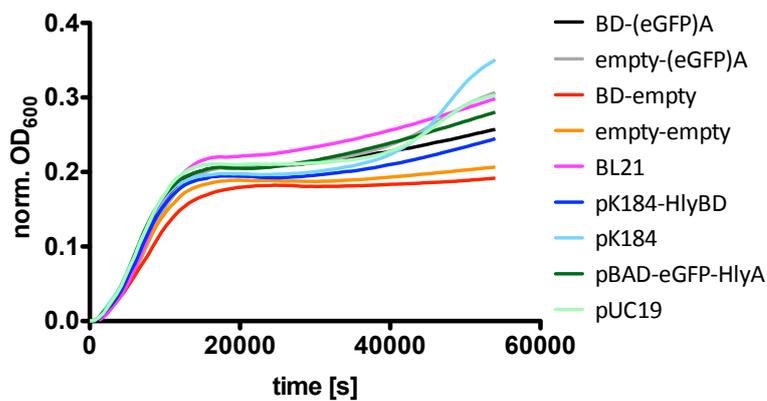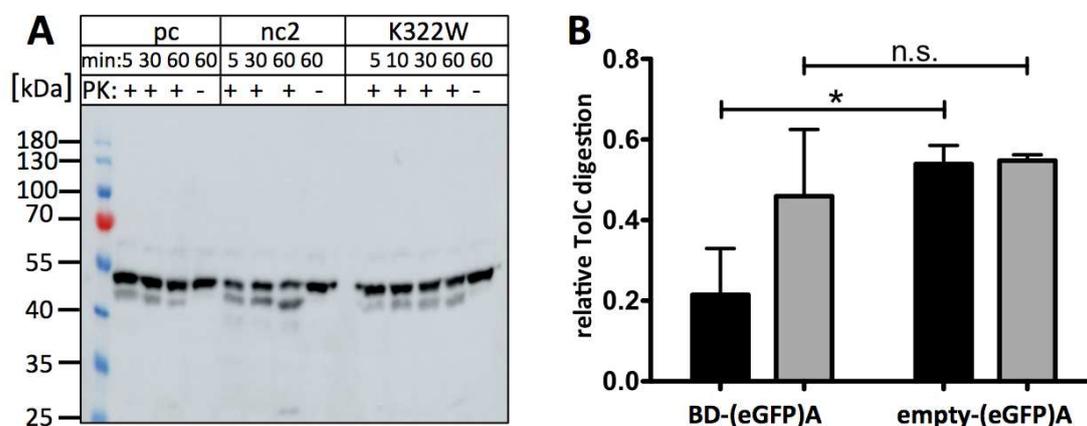
Supplement
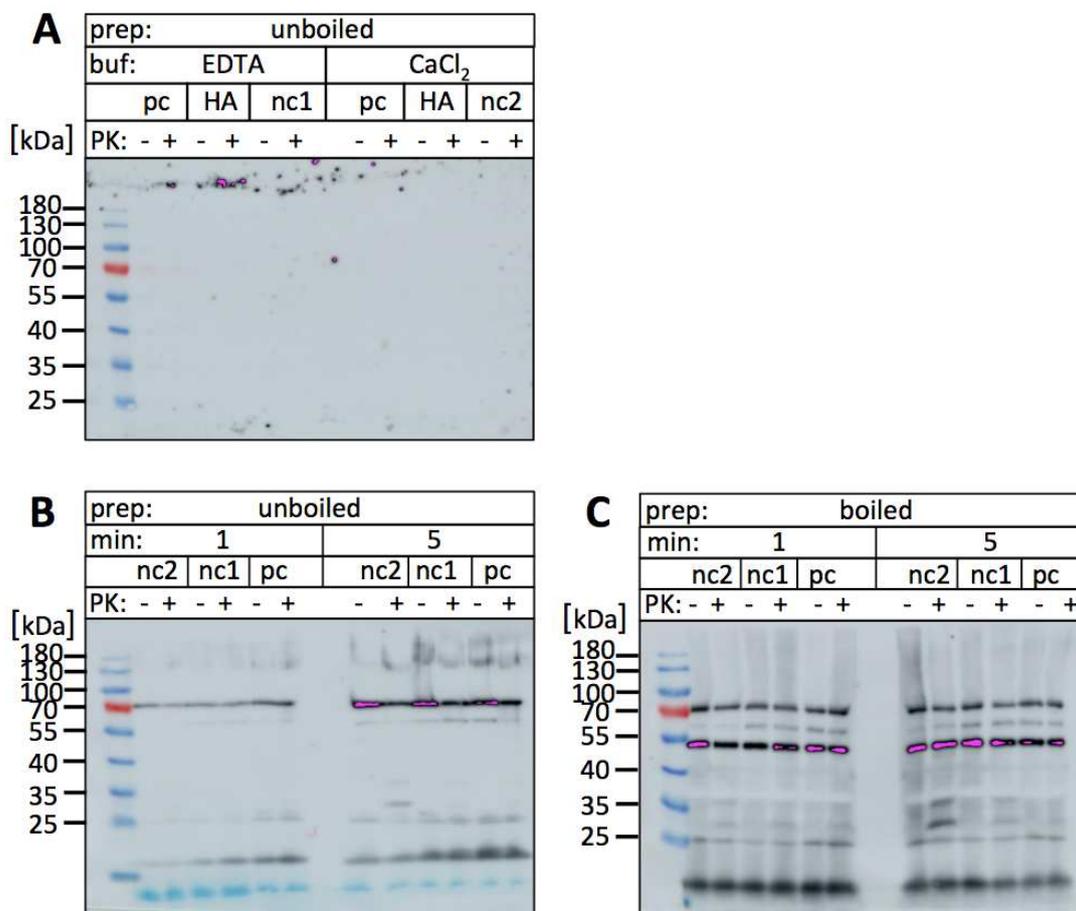
Supplementary Figure 1: Western Blots of whole cell samples analyzed with an antibody directed against TolC that displays cross-reactivity. Samples were either incubated with PK (+) or buffer (-) for 10 min in **(A)** and 1 and 5 min in **(B)** and **(C)**. Pc = positive control for a permanently assembled system with plasmid combinations pK184-HlyBD and pBAD-eGFP-HlyA. Nc = negative control for non-assembled system. Nc1 = pK184-HlyBD and pUC19. Nc2 = pK184 and pBAD-eGFP-HlyA. **A)** HA = ATPase inactive mutant of HlyB H662A. The assay was performed in two different buffers (buf) but samples were boiled for 5 min at 95°C. No TolC could be detected. **B** and **C)** show the same samples unboiled **(B)** and boiled **(C)**. In comparison to **(A)** the SDS-sample buffer and assay buffer has been changed, which allowed boiling of the samples and detection of monomeric TolC (55 kDa) in **(C)**. Pink pixels show saturation of the detector. The signal at 70 kDa is caused by cross-reactivity of the antibody.
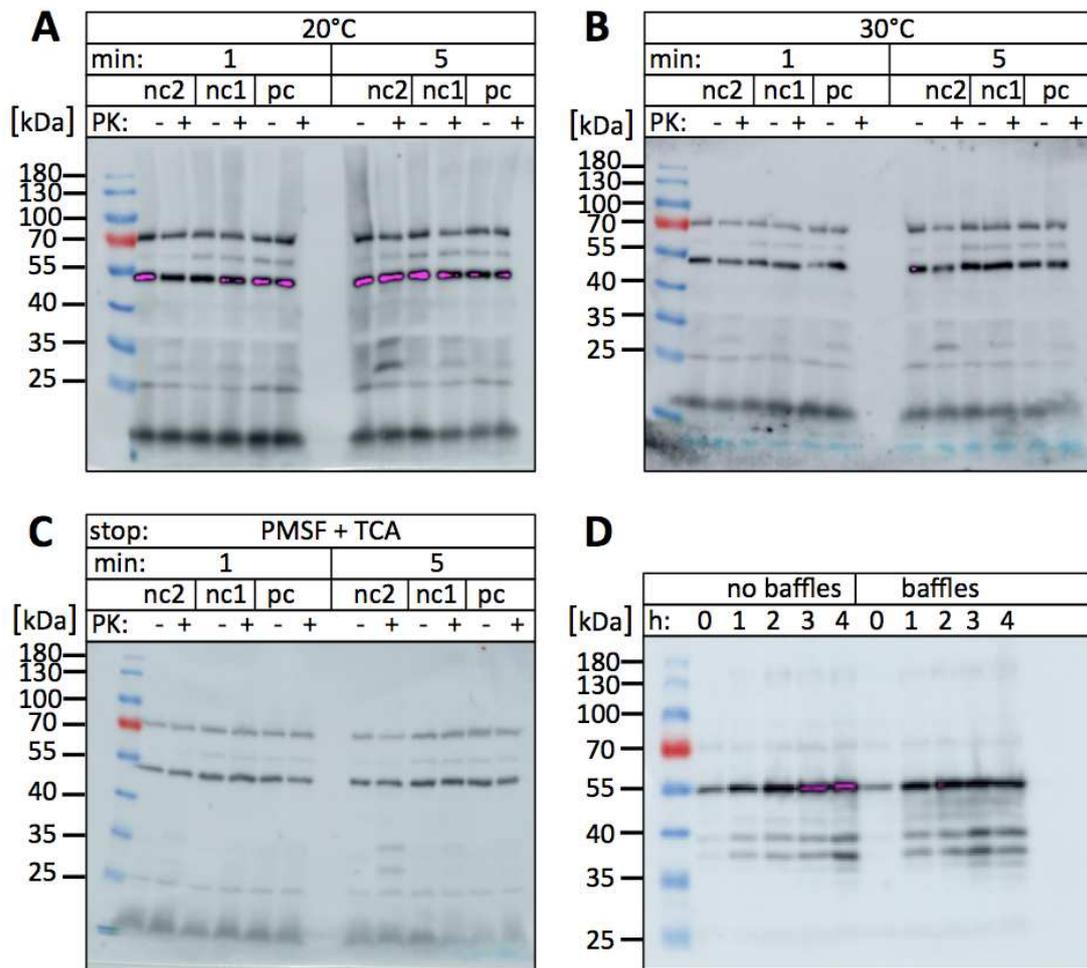
Supplementary Figure 2: Western Blots of whole cell samples. Samples were either incubated with PK (+) or with buffer (-) for 1 or 5 min. Pc = positive control for a permanently assembled system with plasmid combinations pK184-HlyBD and pBAD-eGFP-HlyA. Nc = negative control for non-assembled system. Nc1 = pK184-HlyBD and pUC19. Nc2 = pK184 and pBAD-eGFP-HlyA. **A** and **B)** The assay was performed at 20°C **(A)** and 30°C **(B)** and quenched with PMSF. In **C)** the samples were also treated with trichloroacetic acid (TCA) to precipitate all digestion products. **D)** An unrelated secretion experiment in which pK184-HlyBD was used for HlyBD expression in two different flask types (with and without baffles). Samples were taken before (0 h) and at several time points after induction of protein expression. The Western Blot was developed using an antibody directed against HlyD (55 kDa). Samples were not treated with any protease, but degradation bands of HlyD are still visible.

Supplementary Figure 3: Growth curves of non-induced _E. coli_ BL21(DE3) cultures. Cells were treated as described in material and methods. Although protein expression was never induced, the constructs are named after the plasmids they hold. BD refers to the presence of pK184-HlyBD and (eGFP)A stands for the presence of pBAD-eGFP-HlyA. Empty refers to empty vectors pK184 or pUC19. BL21 refers to _E. coli_ BL21(DE3) cells without any plasmid. OD$_{600}$ was normalized (norm.) by subtracting initial OD$_{600}$ from all samples. Color coding is as in Figure 4 C.

Supplementary Figure 4: A) Western Blot directed against TolC (55 kDa) of whole cell samples treated with PK (+) or with buffer (-) for different times as indicated. Pc = positive control for a permanently assembled system with plasmid combinations pK184-HlyBD and pBAD-eGFP-HlyA. Nc2 = negative control 2 for non-assembled system with plasmid combination pK184 and pBAD-eGFP-HlyA. K322W = mutant of HlyB with plasmid combination pK184-HlyB K322W-HlyD and pBAD-eGFP-HlyA. Pc shows more digestion than before (Figure 3 A). **B)** Quantification of TolC digestion for pc (BD-(eGFP)A; n=4) and nc2 (empty-(eGFP)A; n=2). Shown is mean ± SD. Black bars represent older data and grey bars represent repetition of the assay roughly a year later as described in the text. Significance was calculated using students t-test. * = p<0.05 and n.s. = not significant (p>0.05).

137

3.5 Chapter 5 – Identification and initial characterization of HlyB homologs from Gram-negative bacteria

Title: Identification and initial characterization of HlyB homologs from Gram-negative bacteria

Authors: **Olivia Spitz**, Manuel T. Anlauf, Sander H. J. Smits and Lutz Schmitt

Published in: *to be submitted*

Own Work: 60%

In silico analyses

Cloning

Expression analysis of HlyB*

Solubilization analysis

Purification of HlyB* from Cv, Mh and Kk

Secretion of HlyA and HlyA1

Preparation of the figures

Writing of the manuscript

Identification and initial characterization of HlyB homologs from Gram-negative bacteria

Olivia Spitz, Manuel T. Anlauf, Sander H. J. Smits and Lutz Schmitt

Institute of Biochemistry, Heinrich-Heine-University Düsseldorf

Corresponding author: Lutz Schmitt, Institute of Biochemistry, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, E-Mail: Lutz.Schmitt@hhu.de

Abstract

Type I secretion systems (T1SSs) represent a large group of export systems that share a common protein composition. An ABC transporter and membrane fusion protein (MFP) reside in the inner membrane and recruit an outer membrane protein (OMP) upon substrate engagement to form a continuous channel from the cytoplasm to the extracellular space. Although similar in architecture, T1SSs can be further subdivided into three groups based on the N-terminal extension of the ABC transporter, which correlates to the properties of the substrate such as size and the presence of a N- or C-terminal secretion signal. In this study we focused on group 2 T1SSs, which is defined by an ABC transporter with an inactive peptidase domain, and identified 25 organisms that hold homologous systems. The ABC transporters of five of these organisms were successfully expressed in *E. coli* so that initial purifications and secretion experiments could be performed as well as *in silico* analyses. Secretion approaches with chimeric ABC transporters as well as domain-specific alignments, which were compared to other groups of T1SSs, underlined the importance of the nucleotide binding domain (NBD) and transmembrane domain (TMD) in the transport process. Additionally, two conserved motifs were identified in the MFP that are specific to MFPs of group 2 T1SSs.

Introduction

Type 1 secretion systems (T1SSs) are found in almost all Gram-negative bacteria and transport a variety of substrates from the cytosol across both membranes to the extracellular space (Linhartová et al., 2010). These substrates include toxins, proteases, lipases, adhesins, bacteriocins and hemophores and vary in size from ~9 kDa to 900 kDa (Springer and Goebel, 1980, Gilson et al., 1990, Duong et al., 1992, Létoffé et al., 1994, Akatsuka et al., 1995, Satchell, 2011). The transport machinery is made up of an outer membrane protein (OMP), an ABC transporter and a membrane fusion protein (MFP), both in the inner membrane (IM), which together form a continuous channel from the cytoplasm across the periplasm to the extracellular space (Mackman et al., 1985, Nicaud et al., 1985b, Wandersman and Delepelaire, 1990). While the OMP is often used in multiple systems, the IM components are specific to the transported substrate (Wandersman and Delepelaire, 1990, Koronakis et al., 2004).

The different T1SSs can be further subdivided into three groups (Kanonenberg et al., 2013): Group 1 T1SSs are classified by the presence of an N-terminal extension on the ABC transporter, that resembles a C39 peptidase and cleaves the N-terminal signal sequence of the substrate. A member of this group is the colicin V or MccV T1SS from *E. coli* with CvaB as the ABC transporter, CvaA as the MFP and TolC as the OMP (Gilson et al., 1987, Gilson et al., 1990). ABC transporters of group 2 T1SSs also have an N-terminal extension, that resembles a C39-peptidase but is not active and is therefore termed C39-like domain (CLD) (Lecher et al., 2012). The substrate is not cleaved before, during or after transport (Gray et al., 1986). Members of this group include the HlyA T1SS from *E. coli* or the CyaA T1SS from *B. pertussis* (Glaser et al., 1988). Group 3 includes T1SSs whose ABC transporters have no N-terminal extension but are instead made up of the classical nucleotide binding domain (NBD) and transmembrane domain (TMD). The iron scavenger protein HasA from *S. marcescens* is secreted by such a T1SS with HasD as the ABC transporter and HasE as the MFP. The system can be heterologously expressed in *E. coli* utilizing TolC

as an OMP (Létoffé et al., 1994). Comparing these systems can help to identify key players in the secretion process and features unique for the groups or systems.

One of the most extensively studied T1SSs is the HlyA T1SS from *E. coli*, which belongs to group 2. The ABC transporter of this system, HlyB, carries a peptidase-inactive N-terminal extension, called CLD, whose structure has been solved by NMR (Lecher et al., 2012). Like all ABC transporters it has a TMD and NBD and the structure of the NBD has been solved by X-ray crystallography in many different states (Schmitt et al., 2003, Zaitseva et al., 2005a, Zaitseva et al., 2006, Oswald et al., 2008). Still the structure of the TMDs is missing and can only be modeled based on other T1SS ABC transporters such as PCAT1. However, PCAT1 has no N-terminal extension and is derived from the Gram-positive organism *Clostridium thermocellum* (Lin et al., 2015). No complete structure of a group 2 T1SS ABC transporter has been solved so far.

HlyB forms a stable IM complex with the MFP HlyD even in the absence of the substrate (Thanabalu et al., 1998). HlyD has a small cytoplasmic domain (CD) of 60 amino acids followed by a transmembrane (TM) helix and a large periplasmic domain (PPD) (Schülein et al., 1994, Pimenta et al., 1999). The structure of a part of this PPD has also been solved by X-ray crystallography (Kim et al., 2016) and is often compared to MFPs of tripartite efflux pumps, since there is no structure available for a T1SS MFP.

Upon substrate engagement the IM complex forms a continuous channel across the periplasm with the OMP TolC (Thanabalu et al., 1998), whose structure has been solved by X-ray crystallography as well (Koronakis et al., 2000).

The substrate, HlyA, belongs to the repeats-in-toxin (RTX) family, which is classified by the presence of a C-terminal secretion signal, that is not cleaved, and a repeating motif of nine amino acids, which is responsible for Ca^{2+}-binding and subsequent folding of the protein, called RTX motif or GG-repeat (Felmlee and Welch, 1988, Linhartová et al., 2010). Prior to secretion two internal lysine residues of HlyA are acylated by the acyltransferase HlyC (Stanley et al., 1994). This acylation is necessary for HlyA function but not for secretion, as the non-acylated version, pro-

HlyA, is secreted with the same efficiency (Nicaud et al., 1985a). The 110 kDa toxin HlyA is transported unfolded with the C-terminus reaching the cell surface first (Lenders et al., 2015). Outside of the cell, the conserved motif binds Ca^{2+} and HlyA folds into its active form, which is able to lyse several cell types among them erythrocytes (Goebel and Hedgpeth, 1982, Baumann et al., 1993, Thomas et al., 2014).

HlyB can be purified from *E. coli* membranes and also be reconstituted into lipid nano particles (Kanonenberg et al., 2019b) but structural approaches towards the TMD or the whole protein have failed so far. We therefore aimed to find a more suitable homolog of HlyB and performed an extensive database search, which yielded 25 group 2 T1SSs from other Gram-negative organisms. Eight of these were analyzed in more detail and five were heterologously expressed in *E. coli*. Apart from structural insight also functional studies together with *in silico* analyses were performed with these homologs to deepen the understanding of transport mechanism.

Results

Identifying homologous T1SSs

The first step was the identification of T1SSs of group 2 in other organisms. Therefore the Protein Basic Local Alignment Search Tool (pBLAST) of the National Center for Biotechnology Information (NCBI) was used. In a first approach, the primary sequence of HlyB (UniProt-ID: Q1R2T6) was employed and the search was limited to archaea, which are known to live in extreme environments with therefore stable proteins (Cowan, 1992, Danson and Hough, 1998). At the time of the first search, 2016, no homologs of HlyB were identified, since none of the transporters had an N-terminal extension. Since then more genomes have been added to the database and now (January 2021) at least six transporters with an N-terminal extension can be found by this method.

However, in 2016 the search for homologs was then continued by pBLAST, this time excluding *Escherichia coli*, the original host of HlyB. This yielded hundreds of results reflecting the huge size of sequenced genomes.

To allow identification of a group 2 T1SS ABC transporters, the first 100 aa (amino acid(s)) of a respective transporter were searched for the presence of a Cys residue. The absence of such a Cys indicates, that the N-terminal extension is not able to function as a C39 peptidase, which then classifies this extension as a C39 peptidase-like domain (CLD) (Lecher et al., 2012, Kanonenberg et al., 2013).

After identifying such an ABC transporter, the genome of the respective organism was searched for the presence of a HlyD-like MFP by pBLAST with the HlyD sequence (UniProt-ID: Q1R2T7) and for the presence of an RTX toxin by pBLAST with the HlyA sequence (UniProt-ID: Q1R2T5). In most of the organisms multiple RTX containing proteins (often hypothetical proteins) were identified. Their size and number of conserved RTX motifs, so-called GG repeats, was noted (Table 1) (Felmlee and Welch, 1988, Linhartová et al., 2010). Organisms whose genomes did not contain an MFP and RTX protein were excluded from the list of interesting homologs.

In a last step, the genome of the respective organisms was analyzed for the presence of a HlyC-like protein by pBLAST with the HlyC sequence (UniProt-ID: Q1R2T4). Since RTX toxins can have various functions that are usually mediated by the N-terminal domain (Linhartová et al., 2010), which seems to play a subordinate role in secretion (Gray et al., 1989, Jarchau et al., 1994), the absence of an HlyC-like protein was not used to exclude organisms but rather as a factor to distinguish between the identified T1SSs. It should be noted that the sole presence of a HlyC-like protein does not mean, that the RTX toxin interacts with this protein. For that functional studies are required.

The sequences of 25 HlyB homologs (from here on referred to as HlyB*, listed in Supplementary Table 1) were used in an alignment performed with Clustal Omega, which also provides a phylogenetic tree (Madeira et al., 2019). They could be subdivided into four groups of different size (Supplementary Figure 1). Three organisms from the largest group (which included *E. coli*), two organisms from both medium sized groups and one organism from the smallest group were picked for further analysis, resulting in eight homologs in total. The organisms and the identity of the putative T1SS components, as given by the alignments from the pBLAST

search, are listed in Table 1. They will be abbreviated by using the first letter of their genus and species. For *Escherichia coli* this results in Ec and for *Xylella fastidiosa* this results in Xf and so on.

Table 1: Identity and characteristics of putative T1SS components of different organisms. Proteins were identified by pBLAST search. The number (#) of GG repeats sometimes shows a range. This depends on how strict the motif (GGxGxDxUx, where U stands for a large hydrophobic residue and x stands for any amino acid) is applied (Linhartová et al., 2010). For Cv four RTX proteins were identified. aa = amino acids.

organism	identity compared to [%]				putative RTX toxin	
	HlyB	HlyD	HlyA	HlyC	size [aa]	# of GG repeats
Xylella fastidiosa	61	41	37	32	1814	16-18
Aeromonas diversa CDC 2478-85	65	38	43	-	351	9-10
Avibacterium paragallinarum	69	48	32	-	2286	15
Mannheimia haemolytica	82	61	43	55	953	5
Bibersteinia trehalosi	82	59	42	50	955	6
Cardiobacterium valvarum	70	41	34, 35, 41, 49	-	217, 569, 665, 558	4, 8-9, 2, 5-6
Moraxella bovis	69	41	43	56	927	5
Kingella kingae	72	40	44	62	956	6-7

Cloning of HlyB*

For cloning procedures the FX cloning kit was used which allows high-throughput and only leaves one artificial amino acid at either end of the cloned sequence (Geertsma, 2013). A pBAD-derived plasmid was chosen as an expression vector with an N-terminal 10xhis-tag, which would allow purification by affinity chromatography and detection by Western Blot. With this method seven of the eight picked HlyB* were successfully cloned into the chosen expression vector and the results were verified by sequencing (Microsynth SeqLab, Göttingen).

<u>Expression of HlyB* in *E. coli*</u>

Expression of membrane proteins can be complicated and lead for example to cell death. Therefore, three different expression strains were chosen that are derived from the already well-established *E. coli* C41(DE3) and C43(DE3) strains (Miroux and Walker, 1996). The first two strains tested were *E. coli* C43(DE3)Δ(*acrAB*) and C41(DE3)Δ(*ompF-acrAB*) (Kanonenberg et al., 2019a). The expression was tested at 18°C, 25°C and 37°C for multiple hours and overnight and verified by whole cell analysis on Western Blots with an antibody directed against the N-terminal his-tag.

Similar results were observed for both strains: Five of the seven HlyB* expressed sufficiently while HlyB* from Bt and Mb showed no expression in these strains under all tested conditions (Figure 1 A-D). Their expression was therefore also tested in *E. coli* C43(DE3)Δ(*ompF-acrAB*). While HlyB* from Mb still showed no detectable expression, HlyB* from Bt showed an unexpected band pattern on the Western Blot (Supplementary Figure 2 D): The signal migrated higher than expected: at ~90 kDa for an ~80 kDa transporter, while HlyB migrates at ~70 kDa though it has a size of ~80 kDa. Furthermore the signal increased from 1 h to 2 h but dissipated soon after. HlyB* from Bt and Mb were therefore excluded from further analysis.

<u>Solubilization of HlyB*</u>

In order to stabilize a membrane protein in solution a membrane mimetic is required. By now there are several options available, with detergent based solubilization still being the most frequently used method (Lin and Guidotti, 2009, Gulati et al., 2014, Frauenfeld et al., 2016). Ionic detergents such as Fos-Cholines and SDS are very efficient in solubilizing the membrane, but often interfere with the protein structure. Non-ionic detergents, such as DDM, are believed to be less harsh and maintain protein structure and therefore function, while often being less efficient leading to decreased protein yields (Helenius and Simons, 1975, Helenius et al., 1979).

It is not possible to predict the best detergent for a membrane protein based on its primary sequence alone and even homologous proteins can display very different

solubilization efficiencies with the same detergent. Therefore we applied 88 different detergents and analyzed the soluble membrane fractions via dot blot technique (Ellinger et al., 2013) (data not shown). To gain deeper understanding of the state of the respective protein and the solubilization efficiency, the solubilization was repeated in a smaller scale (with seven detergents) and soluble and unsoluble fractions were analyzed via SDS-PAGE and Western Blot (Figure 1 E-I). Three detergents were tested, that were not part of the dot blot analysis, and three non-ionic detergents from dot blot analysis. They were compared to Fos-Choline 14, which showed high solubilization efficiency for all tested homologs.

This analysis confirmed two points: i) highest solubilization efficiency is achieved with ionic detergents and ii) all five HlyB* localized to the membrane fraction.
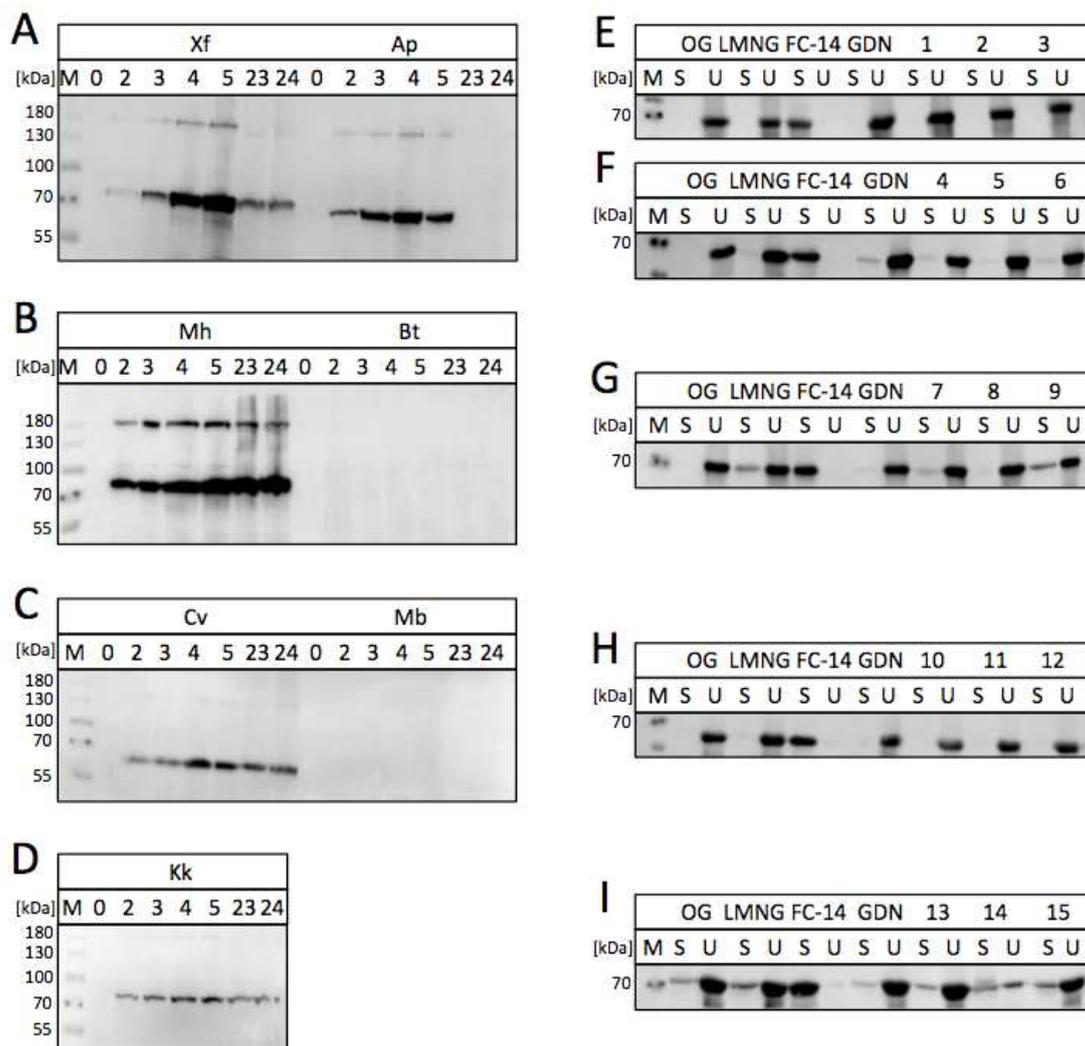
Figure 1: Expression (A-D) and solubilization (E-I) of HlyB* from different organisms. A-D) Cultures were inoculated from overnight cultures containing *E. coli* C41(DE3)Δ(*ompF-acrAB*), which had been transformed with an expression plasmid for each HlyB*. At late logarithmic growth phase protein expression was induced (0 h) at 25°C and samples were taken after 2 h, 3 h, 4 h, 5 h, 23 h and 24 h and analyzed with SDS-PAGE and Western Blot with an antibody directed against the his-tag. Two letter acronyms stand for genus and species of the respective organism (see Table 1). **E-I)** Membranes from the shown expressions were prepared, adjusted to the same concentration and solubilized with different detergents for 1 h at 8°C. Soluble (S) and unsoluble (U) fractions were separated by high speed centrifugation and analyzed via SDS-PAGE and Western Blot. Membranes containing HlyB* from Xf (**E**), Ap (**F**), Mh (**G**), Cv (**H**) and Kk (**I**) were analyzed. OG: n-octyl-β-D-glucopyranoside, LMNG: lauryl-maltose-neopentyl-glycol, FC-14: Fos-Choline 14, GDN: glyco-diosgenin, 1: n-decyl-β-D-thiomaltopyranoside, 2: C12E10, 3: Triton-X-114, 4: n-dodecyl-β-D-maltopyranoside (DDM), 5: n-dodecyl-α-D-maltopyranoside, 6: C12E10, 7: C12E9, 8: C10E9, 9: n-undecyl-β-D-thiomaltopyranoside, 10: N,N'-bis-(3-D-glucoamidopropyl)cholamide, 11: n-undecyl-α-D-maltopyranoside, 12: C12E8, 13: C10E8, 14: n-Dodecyl-N,N-dimethylglycine, 15: n-nonyl-β-D-maltopyranoside. M: molecular weight marker. The 70 kDa band of the molecular weight marker was marked with two dots during image acquisition.

Purification and ATPase activity

Affinity chromatography is one of the most frequently used methods to purify proteins (Lin and Guidotti, 2009, Urh et al., 2009). As already mentioned, all HlyB* were expressed with an N-terminal his-tag whose presence was confirmed by Western Blot and would allow purification via immobilized metal ion affinity chromatography (IMAC). In general two different strategies were applied: i) Solubilization with an ionic detergent and subsequent purification via IMAC during which the detergent was changed to a non-ionic detergent as well as ii) solubilization with a non-ionic detergent followed by IMAC.

After each IMAC the fractions containing the respective HlyB* were combined and concentrated before applying them to a size exclusion chromatography (SEC). This allowed removal of imidazole and aggregates, buffer exchange for downstream applications and estimations about oligomeric state and stability of each HlyB* in different buffer systems.

The purification attempts for HlyB* from Cv and Kk were least successful. HlyB* from Cv, solubilized with n-undecyl-α-D-maltosid, showed very weak binding to the IMAC matrix and eluted together with all nonspecifically bound proteins, indicating inaccessibility of the his-tag. HlyB* of Kk showed better binding to the IMAC matrix eluting at ~100 mM imidazole and could therefore be separated from most other proteins. However, the yields were consistently extremely low (below 90 µg protein from one liter of cell culture after SEC) even when the membrane was solubilized with Fos-Choline 14, leaving no pellet after 1 h of centrifugation at 120,000xg. The low yield could be contributed to a low expression (compared to other HlyB*) as well as to low stability of HlyB* from Kk. Precipitation was often observed, especially when concentrating the protein for SEC analysis or storing at 4°C overnight. HlyB* from Kk was therefore determined unsuitable for crystallization attempts.

Purification for HlyB* from Xf was more successful reaching yields of ~4 mg protein from one liter of cell culture after SEC. The best protein expression was achieved in *E. coli* C41(DE3)Δ(*ompF-acrAB*) (Kanonenberg et al., 2019a) after 5 h

(after induction) at 25°C. The whole membrane was solubilized with 1% (w/v) Fos-Choline 14 overnight (~16 h) minimizing the loss of protein in this step. IMACs were consistent in that HlyB* from Xf eluted at imidazole concentrations >120 mM independent of the detergent used (Fos-Choline 14, DDM or GDN). However, SEC analysis revealed that the protein was least stable with DDM as a large portion of the protein eluted in the void volume of the column indicating aggregation, which was not observed with Fos-Choline 14 or GDN. Each purification yielded sufficient amounts of protein to allow activity analysis with an ATPase assay, but no activity could be measured under the tested conditions.

The best protein expression was observed for HlyB* from Mh with increasing band intensities on Western Blots even after 24 h of expression at 25°C. The protein could be solubilized with the non-ionic detergent C12E9 but increasing the solubilization time from 1 h to 16 h had no significant effect on protein yield, which was between 300 and 400 µg protein per liter of cell culture after SEC. The yield could be increased to 650 µg protein per liter of cell culture when Fos-Choline 14 was used for solubilization and changed to C12E9 during IMAC. The affinity to the IMAC matrix was even higher than for HlyB* from Xf with HlyB* from Mh eluting at imidazole concentrations >150 mM. However, SEC analysis always showed a significant portion of the protein eluting in the void volume of the column indicating aggregation. This likely happened during concentration of the protein before SEC analysis, since precipitation could be observed in the concentrator when protein concentrations of 8-10 mg/mL were exceeded. The highest yield (1.5 mg protein per liter of cell culture) was therefore achieved when SEC was not performed but imidazole was instead removed with a PD-10 column, which required less volume reduction. It is still possible that a significant amount of protein purified by this approach was aggregated. While HlyB* from Xf showed no activity in the ATPase assay, HlyB* from Mh was either too active or contaminated with another ATPase. In the assay the free phosphate is stained with malachite green and measured photometrically. Reliable photometric values (OD <1) could only be achieved when protein concentration was diluted to 0.1 mg/mL and staining time was reduced below

10 min. However, the standard was not sufficiently stained under these conditions. Since the yield was still fairly low, HlyB* from Mh was rather used for crystallization attempts than for activity assays.

The most promising HlyB* was the one from Ap. The first purification attempt initially yielded 2.5 mg active protein (activity measured by ATPase assay) from one liter of cell culture and purification was therefore only slightly altered. The protein was expressed in *E. coli* C41(DE3)Δ(*ompF-acrAB*) for 4 h at 25°C, solubilized with Fos-Choline 14 for 16 h, which was changed to GDN during IMAC in which the protein showed a strong affinity to the matrix (elution with >150 mM imidazole in broad peak). In later purifications the IMAC column was washed with increasing imidazole concentrations in a stepwise manner before the protein was eluted in a sharp peak with 400 mM imidazole (Figure 2 A). This increased the yield (after SEC) from 2.5 mg to 3.5 – 4.5 mg active protein from one liter of cell culture. The buffer system was changed from phosphate to HEPES during SEC in which the protein showed two different elution volumes, which were not baseline separated and showed no difference on SDS-PAGE (Figure 2 C and D). Both peaks exhibited ATPase activity which showed no significant decrease after storage at 4°C over seven days (Figure 2 E). To ensure that ATPase activity was due to HlyB* from Ap a mutant was cloned in which two residues important for ATPase activity (Zaitseva et al., 2005b, Zaitseva et al., 2006) were exchanged: The glutamate of the Walker B motif (E633Q) and the histidine of the H-loop (H664A). This double mutant was expressed and purified in the same manner as the wild type protein and showed no ATPase activity (Figure 2 E).

In summary HlyB* from Cv did not bind to the IMAC column and could not be purified. HlyB* from Kk could be purified, but only low yields were obtained and the protein was not stable. HlyB* from Xf could be purified to high yields, but was not active. HlyB* from Mh was purified to mediocre yields, but its activity could not be prooven. HlyB* from Ap was sufficiently purified in an active state and sufficiently stable. HlyB* from Ap and Mh were both used in crystallization attempts, which to

date did not (yet) yield any crystals. However, HlyB* from Ap showed less tendency to aggregation in crystallization trials than HlyB from *E. coli*.
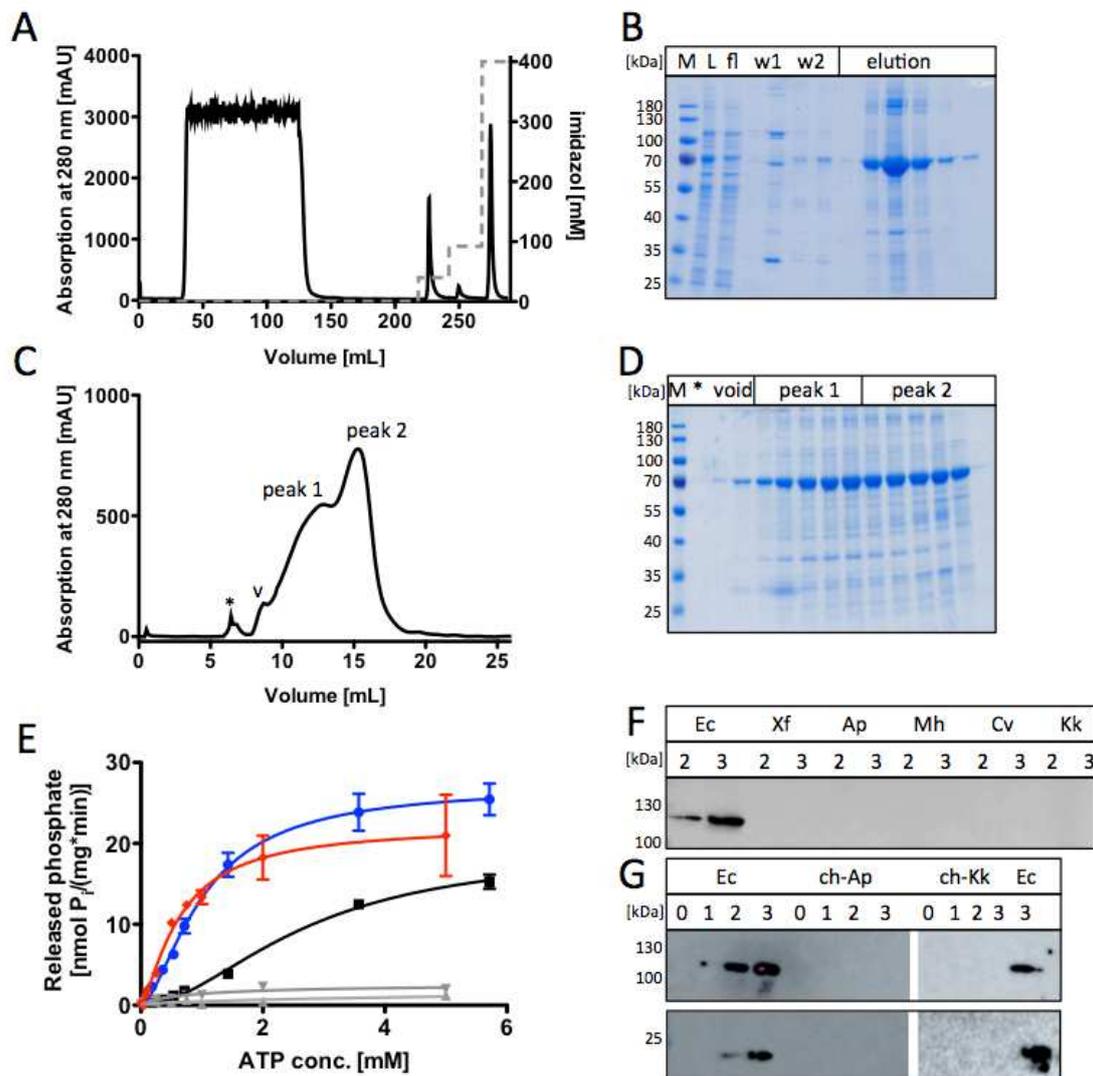

Figure 2: Purification and ATPase activity of HlyB* from Ap. HlyB* from Ap was purified via IMAC (**A**). The concentration of imidazole is shown as a gray dotted line. Load (L), flow through (fl), wash 1 (w1, 225-245 mL), wash 2 (w2, 245-270 mL) and elution fractions (270-290 mL) were analyzed with SDS-PAGE (**B**). Elution fractions were concentrated and applied to a SEC (**C**). A small peak appeared before the void volume (v) of the column and is marked with * in (**C**) and (**D**). (**D**) SDS-PAGE of SEC fractions. Peak 1 and 2 are marked in **C** and **D**. Fractions from both peaks were concentrated and ATPase activity was measured (**E**). Lines in (**E**) show an allosteric sigmoidal fit. Blue: Peak 1 at the day of purification (day 1). Black: Peak 2 at day 1. Red: Peak 1 after 7 days storage at 4°C. Grey: Hydrolytic inactive mutant (E663Q H664A) of HlyB* from Ap. Grey triangles: Peak 1 of this mutant. Grey upside down triangles: Peak 2 of this mutant. Kinetic parameters are listed in Supplementary Table 2. (**F**) Western Blot against HlyA of culture supernatants from secretion experiment with HlyB* from Xf, Ap, Mh, Cv and Kk. Only HlyB from *E. coli* (Ec) was able to secrete HlyA. (**G**) Western Blots against HlyA of culture supernatants from secretion experiment with chimeric HlyB* from Ap (ch-Ap) and Kk (ch-Kk), where their native CLD was exchanged to *E. coli* HlyB CLD. Upper panel: Secretion of full-length HlyA (110 kDa). Lower panel: Secretion of truncated HlyA1 (23 kDa). Numbers in (**F**) and (**G**) represent hours after induction of protein expression. Pink pixels represent saturation of the detector. M: molecular weight marker.

Secretion of HlyA

Multiple studies have shown that the HlyA T1SS is able to export fusion proteins when they are expressed with a C-terminal fragment of (pro-)HlyA at their C-terminus (Hess et al., 1990, Kenny et al., 1991, Bakkes et al., 2010, Schwarz et al., 2012, Lenders et al., 2015). Since HlyBD is able to transport fusion proteins, we were interested if homologous transporters can export pro-HlyA as well.

To test this, all five expressable HlyB* were cloned into a secretion competent vector via Gibson Assembly (Gibson et al., 2009). On this vector (pK184-HlyB(*)HlyD), HlyB* (from the respective homologous organism) and HlyD (from *E. coli*) are expressed under the same promotor and the plasmid is compatible with pSU2726, which encodes for (pro-)HlyA (Soloaga et al., 1996).

No pro-HlyA could be detected for any of the five HlyB* in the culture supernatant by this approach although all necessary components were expressed, which was confirmed by Western Blot (Figure 2 F and Supplementary Figure 2 A, B).

Based on domain specific alignments (see next section) and the expression level of HlyD in combination with the homolog, two chimeric HlyB* were cloned into the secretion competent vector (pK184-chHlyB*HlyD). HlyB* of Ap and Kk were chosen for this approach and their native CLD was exchanged to the CLD of *E. coli* HlyB. Still neither pro-HlyA, nor the truncated version (HlyA1), could be detected in the culture supernatant (Figure 2 G) suggesting a more specific role of the NBDs than simply hydrolyzing ATP or a more specific role for the TMDs than simply forming a channel through the inner membrane.

In silico analyses of T1SS components

For identification of the homologs used in this study, pBLAST was used which relies on alignments and sequence identity. This can be difficult for proteins containing multiple domains and information can be lost or misinterpreted. To minimize this problem we divided both components of the inner membrane into three domains each and aligned them separately. For HlyB and its homologs the domains are the CLD (residue 1 – 125), the TMD (transmembrane domain, residue

158 – 436) and the NBD (nucleotide binding domain, residue 468 – 703) and for HlyD we divided the protein into the CD (cytosolic domain, residue 1 - 60), TM-helix (residue 61-80) and PPD (periplasmic domain, residue 81-478). We relied on UniProt's domain predictions when separating each protein into these domains. Exceptions to this were the MFPs of Xf, Kk and Re (*Rhizobium etli*) as well as BtrA (group 1 T1SS ABC transporter (Michiels et al., 2001)) from Re and since no UniProt entry was present. These MFPs were subjected to a secondary structure prediction tool (Zimmermann et al., 2018) to identify the TM-helix, which separates the CD from the PPD. In order to divide BtrA in its domains we relied on the annotated regions in the sequence report published at NCBI (WP_040111896.1).

ABC transporters

The sequence identities of the domains of the ABC transporters are the lowest for the CLD in each case (37-60%) and quite similar for TMD and NBD (66-88%) in each organism (Table 2). This supports the hypothesis, that the CLD confers specificity by binding the respective substrate. However, exchanging the CLD of HlyB* was still not sufficient to allow transport of pro-HlyA or the truncated version HlyA1 pointing towards a more complex mechanism of recognition and/or transport.

The TMDs of the group 2 T1SS ABC transporters analyzed in this study show an overall identity between 66% and 88%, which is quite high (Table 2). When aligning the TMDs of group 1 T1SS ABC transporters (peptidase-active N-terminal extension), the identity is only around 25%, between 19% and 25% for group 3 T1SS ABC transporters (no N-terminal extension) and between 45% and 57% when analyzing other group 2 T1SS ABC transporters (see Table 2 for details on the different T1SS origins). Especially interesting is the comparison to CvaB, the ABC transporter for colicin V or MccV, since it is derived from *E. coli* as well (Gilson et al., 1987, Gilson et al., 1990). The identity of the TMD compared to HlyB is only 26% although the lipid composition of the membranes should be relatively similar considering CvaB and HlyB are both derived from *E. coli* strains. The high identity of the TMDs of group 2 T1SS ABC transporters across the organisms points to an

important role of the TMDs and probably a different transport mechanism compared to group 1 and 3.

The same trend is visible when aligning the NBDs (Table 2). While the homologs in this study show a sequence identity of 68% to 86%, the NBDs of group 3 T1SS ABC transporters show an identity between 32% and 33% and the NBDs of group 1 between 38% and 41% identity. All of them show the conserved Walker A and Walker B motifs as well as the Q-loop, C-loop, Pro-loop, D-loop, Gly-loop and switch II (also called H-loop) (Schmitt et al., 2003) (see Supplementary Figure 3 for full alignment). Nevertheless group 2 NBDs also show high conservation between the motifs, which is less pronounced for group 1 and 3.

Table 2: Identities of ABC transporter domains compared to HlyB from *E. coli*. Upper group 2 represents HlyB* researched in this study. Groups are according to (Kanonenberg et al., 2013). IDs marked with (*) can be found at NCBI instead of UniProt.

group	ABC protein	host	identity to HlyB [%]			
			CLD	TMD	NBD	UniProt-ID
	HlyB*	*X. fastidiosa*	37.1	66.3	68.3	A0A060H1A8
	HlyB*	*A. paragallinarum*	50.0	74.2	70.8	A0A0F5EW97
	HlyB* (LktB)	*M. haemolytica*	60.3	87.5	86.2	P0C087
2	HlyB*	*C. valvarum*	39.8	77.4	71.7	G9ZEE5
	HlyB* (RtxB)	*K. kingae*	50.4	79.9	72.4	A1YKX0
	HlyB* (LktB)	*B. trehalosi*	60.3	87.1	86.7	Q933E0
	HlyB* (MbxB)	*M. bovis*	42.9	77.4	73.8	Q7X2A4
1	CvaB	*E. coli*	18.3	26.0	41.1	A7DT68
	BtrA	*R. etli*	23.4	24.2	37.6	WP_040111896.1*
2	CyaB	*B. pertussis*	17.2	56.6	61.3	P0DKX6
	RtxB	*V. cholerae*	19.2	44.8	54.9	A0A2P1AHA9
	RsaD	*C. crescentus*	-	25.1	33.2	O85350
3	HasD	*S. marcescens*	-	18.6	31.7	Q53368
	AprD	*P. aeruginosa*	-	21.6	31.6	Q03024

MFPs

The CDs of the homologs show a sequence identity of 44 – 54% while Mh stands out with 71% (Table 3). Their similarity becomes more striking when secondary structure is predicted (Combet et al., 2000, Sapay et al., 2006). All of them show an amphipathic helix (AH) of 26 ± 1 residues at the N-terminus with 4-6 aromatic residues on the non-polar site of the AH (Figure 3 A). In a helical wheel projection all of them display an Arg at the non-polar site of the AH which is also visible as a conserved residue in the alignment (Arg26) (Figure 3). If this Arg is actually part of the helix, reducing the non-polarity of this side, or the helix stops earlier *in vivo*, cannot be concluded based on *in silico* analysis alone.
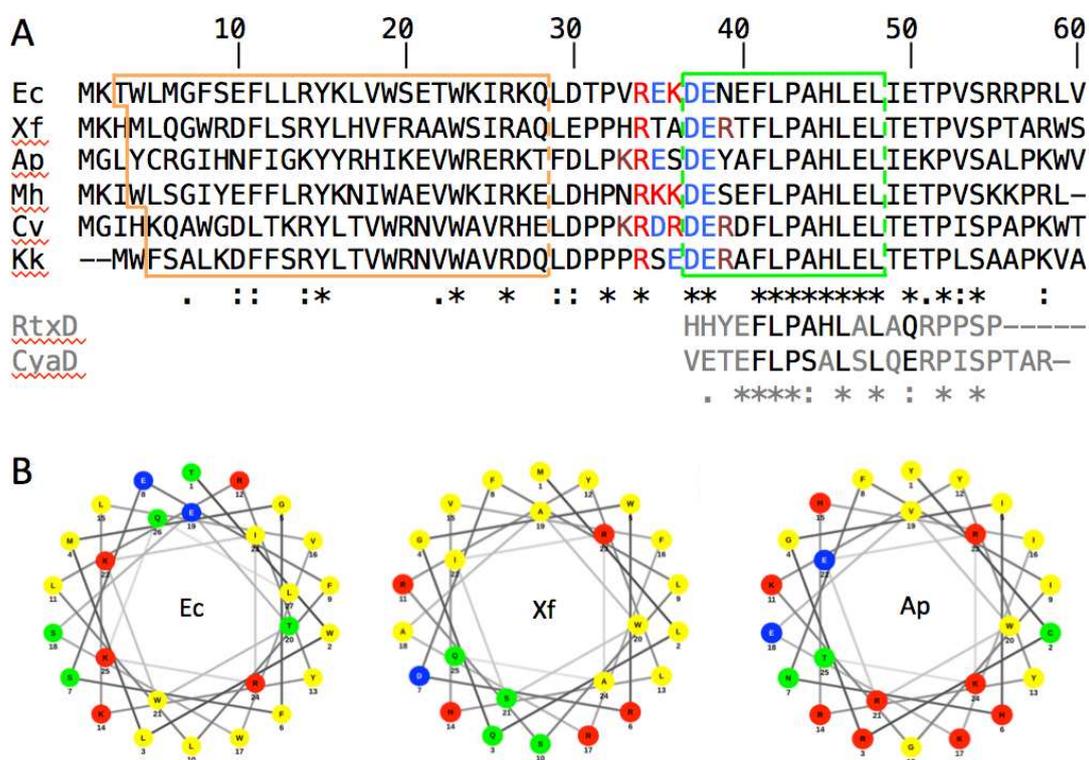

Figure 3: Analysis of cytosolic domain (CD) of group 2 T1SS MFPs. A) shows the alignment of group 2 MFPs. The orange box marks an amphipathic helix (Combet et al., 2000, Sapay et al., 2006) and the green box a second helical region (Jones, 1999, Wang et al., 2016, Heffernan et al., 2017, Zimmermann et al., 2018, Klausen et al., 2019). A charged cluster is marked in red (basic residues) and blue (acidic residues) (Balakrishnan et al., 2001). (*) marks the same residue, while (:) and (.) mark similar residues. The alignment was performed with Clustal Omega (Madeira et al., 2019). For details on sequence origin see Table 3. RtxD and CyaD are MFPs that were not part of this study and are shown in grey. **B)** Helical wheel projections of the amphipathic helix of *E. coli* (Ec), *X. fastidiosa* (Xf) and *A. paragallinarum* (Ap). Basic residues are marked red, acidic residues are marked blue, polar residues are marked green, non-polar residues are marked yellow (Mól et al., 2018).

Following the AH, a charged cluster of five residues can be found. The first (Arg34) and last two residues of this cluster (Asp37 and Glu38) are especially conserved (Figure 3 A). This cluster overlaps with a second helical region, which shows different helical predictions based on the used prediction tools (Jones, 1999, Dosztányi et al., 2005, Yan et al., 2013, Jones and Cozzetto, 2015, Hanson et al., 2016, Wang et al., 2016, Heffernan et al., 2017, Zimmermann et al., 2018, Klausen et al., 2019). However, this region shows a highly conserved motif of 8-10 residues: FLPAHLEL-(I/T)-E. This motif is not detected in MFPs of group 1 and 3 T1SSs but is also less conserved in comparison to other MFPs from group 2 T1SSs like CyaD and RtxD (FLP-(A/S)-xLxLx-(E/Q)) (Figure 3 A).

Table 3: Identities of MFP domains compared to HlyD from *E. coli*. Upper group 2 represents the MFPs that belong to the ABC transporters researched in this study. Groups are according to (Kanonenberg et al., 2013). IDs marked with (*) can be found at NCBI instead of UniProt.

			identity to HlyD [%]			
group	MFP	host	CD	TM-helix	PPD	UniProt-ID
	HlyD*	*X. fastidiosa*	54.2	12.5	37.0	ALR08405.2*
	HlyD*	*A. paragallinarum*	47.5	42.1	48.5	A0A0F5EXY0
2	HlyD* (LktD)	*M. haemolytica*	71.2	50.0	58.4	P16534
	HlyD*	*C. valvarum*	44.1	26.3	40.3	G9ZEE6
	HlyD*	*K. kingae*	49.1	36.8	38.9	WP_019390474.1*
1	CvaA	*E. coli*	25.0	35.7	25.5	P22519
	BtrB	*R. etli*	16.7	5.6	24.5	WP_040111953.1*
2	CyaD	*B. pertussis*	29.1	26.3	29.5	J7QCA7
	RtxD	*V. cholerae*	26.9	11.1	37.9	Q9X4W5
	RsaE	*C. crescentus*	5.0	16.7	21.4	Q84DJ5
3	HasE	*S. marcescens*	32.0	33.3	23.8	Q57387
	AprE	*P. aeruginosa*	28.6	27.8	24.3	Q03025

The lowest sequence identity in regards to group 2 MFPs is always found in the TM-helix (Table 3). Differences in membrane crossing domains (TMD and TM-helix) might be a consequence of the different lipid compositions of these organisms, however, the sequence identity for the TM-helix of the MFP is considerably lower in each case than for the TMD of the ABC transporter, which might be due to the shorter sequence.

The PPD of HlyD can be separated into multiple domains as well. In general the PPD of MFPs like AcrA, CusB, MacA and EmrA contain a membrane proximal (MP) domain, a β-barrel domain, a lipoyl domain and an α-helical domain with a helix-turn-helix motif (Mikolosko et al., 2006, Su et al., 2009, Yum et al., 2009, Hinchliffe et al., 2014, Kim et al., 2016). The crystal structure of a part of the PPD from HlyD shows an α-helical domain and a lipoyl domain, but is missing the last ~100 residues (Kim et al., 2016). In an earlier study, Lee *et al.* identified an RLT motif whose importance was shown by mutational studies and suggested to facilitate the contact to TolC (Lee et al., 2012). However, based on the crystal structure Kim *et al.* re-identified this motif as a DLA motif in the tip region and suggested that the RLT motif is important for oligomerization instead of contact to TolC.

The DLA motif can be found in all five homologs with Ala250 being the most promiscuous residue of this motif (Figure 4 A). Leu190 of the RLT motif is also conserved among these homologs. In an alignment of the RLT motif Lee *et al.* also highlighted Trp198 following Thr197 of the RLT motif. This aromatic residue was not part of their mutational studies but an aromatic residue (Trp or Phe) can be found at this position in all five homologs (Figure 4 A).

Figure 4: Alignment of periplasmic domain (PPD) of group 2 T1SS MFPs. Two letter abbreviations stand for genus and species of the respective organism (Table 3). **A)** The RLT motif (Lee et al., 2012) is shown in orange and the DLA motif (Kim et al., 2016) in red. Colored lines in (**A**) and (**B**) over the sequence show structure elements from HlyD crystal structure (Kim et al., 2016), with helix 2 in blue, helix 3 in purple and lipoyl domain in green. **B)** Residue V334 and V349 (both yellow) in the lipoyl domain were shown to influence HlyA folding (Pimenta et al., 2005). Yellow residues in (**C**) (K404 and D411) were part of the same study and are important for HlyA secretion. **C)** Last 120 residues of HlyD. Conserved motifs are colored in pink (specific to HlyD(-like) proteins) and brown (not specific to HlyD(-like) proteins). Bright green residues (L475, E477, R478) were shown to be important for HlyA secretion (Schülein et al., 1994). Colored lines over the sequence represent secondary structure predictions (Jones, 1999, Dosztányi et al., 2005, Yan et al., 2013, Wang et al., 2016, Heffernan et al., 2017, Zimmermann et al., 2018, Klausen et al., 2019). Lines are dashed in areas where different prediction tools were in disagreement. Solid lines represent predictions that are in agreement. Light blue shows predicted β-sheets, dark red shows predicted α-helices. Alignments were performed with Clustal Omega (Madeira et al., 2019). (*) marks the same residue, while (:) and (.) mark similar residues. Alignments of other group 2 T1SS MFPs (CyaD and RtxD) are shown in grey.

The lipoyl domain is not easy to identify by sequence alignments since it is made up of residues 96-131 and 326-361 and therefore interrupted by the α-helical domain (Kim et al., 2016). While the sequence identity of the α-helical domain is quite low, it is higher in the parts of the lipoyl domain. This domain contains two conserved Val residues (Val334 and Val349), which are important for correct folding of the substrate HlyA (Pimenta et al., 2005) and are present in all five homologs (Figure 4 B).

The extreme C-terminus of the MFPs displays the highest sequence identity with multiple conserved motifs (Figure 4 C). Especially the sequence FPYTRYGY (residues 393-400) seems to be specific for HlyD-like proteins, since pBLAST with this motif, in- or excluding *E. coli*, only showed HlyD-like proteins in the first 100 results. The region AEIKTG (residues 450-455) and YLLSPL (residues 462-467) showed more diverse pBLAST results including proteins and enzymes from eukaryotes and a variety of hypothetical and annotated proteins of different protein families. As for the conserved motif in the CD, these three motifs were not found in the MFPs of other T1SSs (CvaA, LipC, AprE, HasE, RsaE, PrtE), but are also less conserved in RtxD and CyaD (group 2 MFPs, Figure 4 C).

Finally the last 118 residues of HlyD, which include the conserved motifs, were subjected to secondary structure prediction since they are lacking in the published structure of HlyD (Kim et al., 2016). A high content of β-strands was identified by multiple tools (Jones, 1999, Yan et al., 2013, Wang et al., 2016, Heffernan et al., 2017, Zimmermann et al., 2018, Klausen et al., 2019) (Figure 4 C) and the prediction is very similar to the structure of, for example, MacA (Yum et al., 2009).

Discussion

T1SSs are widespread throughout Gram-negative bacteria (Linhartová et al., 2010). The recent database analysis performed here also suggested their presence in archaea. Based on the N-terminal extension of the ABC transporter component, T1SSs can be further subdivided into three groups depending on the presence and function of an N-terminal extension (Kanonenberg et al., 2013).

We were able to identify 25 group 2 T1SSs in Gram-negative bacteria (Supplementary Table 1) and successfully cloned the DNA sequence of seven ABC transporters in expression plasmids (Geertsma, 2013). Utilizing different variants of *E. coli* C41 (DE3) and C43 (DE3) (Miroux and Walker, 1996, Kanonenberg et al., 2019a), we were able to express six of these heterologously without codon optimization (Figure 1 A-D and Supplementary Figure 2 D). Five of them were subjected to detailed solubilization assays showing different behavior although they were expressed in the same *E. coli* strain, therefore holding the same lipid composition (Figure 1 E-I). These five HlyB* were purified with different levels of successes whereby the purification of HlyB* from Ap was the most successful in yield, stability and activity of the transporter (Figure 2 A-E). Two of these five were subjected to crystallization trials (HlyB* from Ap and Mh) but did not yield any crystals yet.

Interestingly none of the five sufficiently expressed HlyB* (from Xf, Ap, Mh, Cv and Kk) were able to secrete pro-HlyA (Figure 2 F). Furthermore two chimeric HlyB* (from Ap and Kk), who carried the CLD from *E. coli* HlyB instead of their native CLD, were unable to secrete pro-HlyA or the truncated version HlyA1 (Figure 2 G). This shows, that TMD and/or NBD of the ABC transporter play a crucial role in formation of the inner membrane complex (IMC) or during substrate recognition.

Substrate recognition

Surface plasmon resonance (SPR) studies with isolated HlyB NBD and fragments of HlyA showed that the NBD is able to interact with the secretion signal of HlyA (Benabdelhak et al., 2003). The NBDs of all five homologs used in the secretion assay show very high sequence identity to HlyB (68-86%), with HlyB* from Ap and Kk (used for chimeric approaches) showing a sequence identity of 71% and 72%, respectively (Table 2). Still, it is possible that the homologous NBDs do not interact with pro-HlyA.

The CLD has also been shown to interact with (pro-)HlyA in *in vitro* pull-down assays (Lecher et al., 2012). Interestingly the CLD and NBD seem to interact with different parts of the substrate, which makes simultaneous interaction possible.

The domain organization of CLD, TMD and NBD of HlyB towards each other is not known yet. It is possible that the chimeric HlyB* failed to interact in a simultaneous fashion with the substrate, since their domain organization could be disturbed.

Finally HlyA has also been shown to interact with HlyD in the absence of HlyB by cross-linking experiments (Thanabalu et al., 1998). However, in the secretion approaches HlyD and HlyA were wild type proteins from *E. coli* and should therefore be able to interact.

IMC formation

HlyB and HlyD form a stable IMC in the absence of the substrate (Thanabalu et al., 1998). By which interactions this IMC is formed or stabilized is not known.

To date there is no structural information available for an assembled T1SS, however, tripartite efflux pumps share common features with T1SSs and structures of these in complex with the OMP are available. Tripartite efflux pumps, similar to T1SSs, are made up of an OMP, a periplasmic adaptor protein (also called membrane fusion protein) and an inner membrane protein (Hinchliffe et al., 2013, Jo et al., 2019). This inner membrane protein can belong to the superfamily of resistance-nodulation-cell division (RND) proteins or to the family of ABC transporters. An example for the latter is the MacAB-TolC complex from *E. coli* whose structure was solved (Fitzpatrick et al., 2017). Examples for RND-type tripartite efflux pumps, whose structures were solved, are AcrAB-TolC from *E. coli* (Wang et al., 2017) and MexAB-OprM from *P. aeruginosa* (Tsutsumi et al., 2019). In each of these structures the lipoyl domain of the MFP does not interact with the inner membrane protein but the contact is made by the β-barrel and membrane proximal (MP) domain of the MFP (Fitzpatrick et al., 2017, Wang et al., 2017, Tsutsumi et al., 2019). Based on structural comparisons of the α-helical domain of HlyD to other MFPs, Kim *et al.* suggested that HlyD lacks the MP domain but still displays a β-barrel domain (Kim et al., 2016). Since such a domain was not identified in the fragment that Kim *et al.* were able to crystallize, it is likely present in the extreme C-terminus of HlyD, which is supported by secondary structure prediction, which showed high β-sheet content in

this region (Figure 4 C). Similar to tripartite efflux pumps this region could mediate the contact between HlyD and HlyB. The high sequence similarity in the last ~100 residues of the MFPs indicates that the contact between ABC transporter and MFP is mediated in the same way in these organisms and that HlyD could be able to form the IMC with HlyB* as well. This however has to be elucidated in detail.

At least five residues in this C-terminal region of HlyD were shown to be essential for HlyA secretion by two different mutational studies: Lys404 and Asp411 (Pimenta et al., 2005) as well as Leu475, Glu477 and Arg478 (Schülein et al., 1994). Asp411, Leu475, Glu477 and Arg478 are all conserved in the homologs, while Lys404 is not conserved in HlyD* from Ap but in the other homologs. Interestingly, the regions between Lys404 and Asp411 and following Asp411 show low sequence similarity (Figure 4 C). This might be the region where the MFP makes a specific contact to the respective ABC component, which would have resulted in a lack of assembly of the IMC in the secretion approach.

Without further investigation it is not possible to pinpoint why exactly the secretion of HlyA by homologous transporters failed. Secretion experiments with chimeras where the NBD or the TMD is exchanged could shed more light on this subject. Furthermore it would be interesting to investigate the assembly of the IMC by for example cross-linking experiments with these different chimeras.

Material and Methods

In silico tools

In order to identify the homologs pBLAST was employed (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

The sequences that were used were taken from UniProt (https://www.uniprot.org/): HlyB: Q1R2T6, HlyD: Q1R2T7, HlyA: Q1R2T5, HlyC: Q1R2T4. All alignments as well as the phylogenetic tree were created with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Madeira et al., 2019). The helical wheel projections were performed with NetWheel (http://lbqp.unb.br/NetWheels/) (Mól et al., 2018). Two different tools were employed for secondary structure

predictions. For the amphipathic helix of HlyD and its homologs AmphipaSeeK was used (https://npsa-prabi.ibcp.fr/) (Combet et al., 2000, Sapay et al., 2006). For all other secondary structure predictions Quick2D was utilized, which employs multiple prediction algorithms (https://toolkit.tuebingen.mpg.de/tools/quick2d) (Jones, 1999, Dosztányi et al., 2005, Yan et al., 2013, Jones and Cozzetto, 2015, Hanson et al., 2016, Wang et al., 2016, Heffernan et al., 2017, Zimmermann et al., 2018, Klausen et al., 2019). All tools are available online free of charge.

Cloning and transformation

The genomic DNA of the organisms listed in Table 1 was ordered from the DSMZ (German Collection of Microorganisms and Cell Cultures GmbH) and used for cloning with the FX cloning kit (Addgene) according to manufactures instructions (Geertsma, 2013). The DNA was first cloned into a cloning vector (pINITIAL) and then by digestion with SapI cloned into an expression vector derived from pBAD, which introduced a 10xhis-tag at the N-terminus of the homologs. Primers were designed with the designated online tool (https://www.fxcloning.org/) (Geertsma, 2013).

For secretion experiments the homologs or their domains were cloned into a secretion competent vector by Gibson Assembly (Gibson et al., 2009). The vector utilizes the backbone of pK184 and is often used for the expression of HlyB and HlyD under the control of the same promotor (Jenewein, 2008). The HlyB sequence was replaced with the sequence of the respective homolog. The primers were designed with the help of a primer design tool used for Gibson Assembly reactions (https://nebuilder.neb.com/).

Chemically competent *E. coli* cells were obtained by incubation in 0.1 M $CaCl_2$ and all plasmids were transformed into the respective strains by the heat-shock method (Mandel and Higa, 1970, Froger and Hall, 2007).

Test- and overexpression of HlyB*

In order to identify the best growth conditions small scale test expressions were performed. After the respective plasmid was transformed into either *E. coli* C43(DE3)Δ(*acrAB*), *E. coli* C41(DE3)Δ(*ompF-acrAB*) or *E. coli* C43(DE3)Δ(*ompF-acrAB*) pre-cultures of 5 mL LB-media with 100 μg/mL ampicillin were inoculated from transformation agar plates (Miroux and Walker, 1996, Kanonenberg et al., 2019a). They were incubated for 16 h at 37°C and 180 rpm before being used to inoculate either 50 mL LB-media or 50 mL 2xYT-media cultures supplemented with 100 μg/mL ampicillin to an OD_{600} of 0.1. These cultures were grown to OD_{600} 0.5-0.7 at 37°C and 180 rpm before being induced with 1 mM arabinose. They were subsequently incubated at either 18°C, 25°C or 37°C for 24 h. For analysis samples of 1 mL were taken at different time points after induction and centrifuged for 1 min at 11,000xg. Cell pellets were resuspended in water according to their OD_{600} and 80 μL cell suspension was mixed with 20 μL SDS-sample buffer (100 mM Tris pH 6.8, 3.3% (w/v) SDS, 0.02% (w/v) bromophenol blue, 40% (v/v) glycerol). Samples were analyzed on 10% SDS-PAGEs and transferred to polyvinylidene difluoride (PVDF) membranes with a Trans-Blot SD Semi-Dry Transfer Cell by Bio Rad according to manufactures instructions. The membranes were blocked with 10% (w/v) milk powder overnight at 4°C, incubated with the respective primary antibody for 1 h at room temperature (RT), then incubated with a secondary antibody coupled to horse radish peroxidase (HRP) for 1 h at RT, washed three times in between each step with TBS-T buffer (20 mM Tris pH 8, 250 mM NaCl, 0.1% (v/v) Tween-20) and analyzed using Amersham™ Imager 600 (GE Healthcare) or Chem Genius[2] bio imaging system (Syngene). For purification of HlyB* the volume of the pre-cultures and the main cultures was upscaled to 50 mL and 2 L respectively.

In general the best protein expression was reached in *E. coli* C41(DE3)Δ(*ompF-acrAB*) at 25°C after 3-5 h after induction with only HlyB* from Mh still showing strong bands on Western Blots after 24 h of expression.

Membrane preparation and solubilization

E. coli C41(DE3)Δ(_ompF-acrAB_) cells were transformed with pBAD derived expression plasmids and grown as described above. Cells were harvested by centrifugation at 8000xg for 20 min at 4°C, and cells from 1 L cell culture were resuspended in 45-50 mL resuspension buffer (50 mM Na_2HPO_4 pH 8, 300 mM NaCl). Cells were disrupted by passing through French press at 1.5 kBar four times. Cell debris was collected by centrifugation at 18,000xg for 30 min at 4°C and discarded. The membranes were collected by centrifugation at 120,000xg for 1 h and 30 min at 4°C. The supernatant was discarded and membrane pellets were homogenized with a hand held potter in resuspension buffer. Concentration was adjusted to ~10 mg/mL measured by NanoDrop and membranes were either used directly for solubilization or stored with 10 % (v/v) glycerol at -20°C. For small scale solubilizations 160 µL membrane suspension were mixed with 1% (w/v) of the respective detergent (2% (w/v) for n-octyl-β-D-glucopyranoside) and incubated for 1 h at 8°C under slow shaking. For purification the whole membrane suspension from 1 L cell culture was used and either incubated for 1 h or 16 h with 1% (w/v) of the respective detergent as indicated in the results section. After solubilization soluble and unsoluble fractions were separated by centrifugation at 120,000xg for 25 min (small scale) or 1 h and 30 min (purification scale) at 4°C. For analysis of the small scale solubilizations 80 µL of the soluble fractions were mixed with 20 µL SDS-sample buffer. The unsoluble membrane fractions were resuspended in 100 µL resuspension buffer and 80 µL of this solution were mixed with 20 µL SDS-sample buffer. Both samples were analyzed with 10% SDS-PAGEs and subsequent Western Blotting as described above.

Purification and ATPase activity

Membranes containing the protein of interest were prepared and solubilized as described. After removal of unsolubilized membrane fragments 2-10 mM imidazole were added to the membrane solution and it was loaded to a HiTrap Chelating IMAC column by GE Healthcare with an ÄKTAprime system (GE Healthcare) at 1 mL/min

and fractions of 10 mL were collected. The column had been loaded with Ni^{2+} and equilibrated with five column volumes resuspension buffer containing the respective detergent and imidazole concentration prior to that. If a change of detergent was anticipated, the membrane solution was diluted 1:4 with resuspension buffer prior to loading to the IMAC column to reduce detergent concentration. After loading, columns were washed at 2 mL/min with resuspension buffer containing the respective detergent until baseline was reached or at least 20 column volumes of buffer had passed the column. An imidazole gradient from 2 or 10 mM to 300 mM imidazole was applied to elute the protein over 1 h at 1 mL/min and 1 mL fraction size. In later purifications the columns were washed in a stepwise manner with imidazole concentrations that still allowed binding of the protein to the column (e.g. 50 mM and 100 mM imidazole for washing when purifying HlyB* from Ap). Subsequently, the protein was eluted with 300 mM imidazole (400 mM for HlyB* from Ap) at 1 mL/min and 1 mL fraction size. Fractions containing the protein of interest were collected and concentrated to ~500 µL or 2.5 mL depending on the downstream application by centrifugation in a protein concentrator with 50 kDa molecular weight cut-off at 4000xg at 4°C.

For SEC analysis 500 µL protein solution was applied to a Superose 6 10/300 GL column (GE Healthcare) at 0.5 or 0.25 mL/min and fractions of 1 mL were collected. To allow detection of free phosphate in the ATPase assay, the column was equilibrated with 50 mM HEPES pH 8 (HlyB* from Ap, Xf, Mh) or 50 mM TRIS pH 8 (HlyB* from Mh, Kk, Cv) and 250 mM NaCl. The SEC buffer also contained the respective detergent at 1.5x the cmc (critical micellar concentration) of this detergent. In some purifications SEC was replaced by a PD-10 column (GE Healthcare) according to manufacturer's instructions.

Fractions containing the protein of interest were collected and concentrated in the same protein concentrator that had been equilibrated with the respective SEC buffer. The yield was determined by protein concentration measurements, either via NanoDrop or BCA assay and activity was tested via an ATPase assay in which the free phosphate after ATP hydrolysis is stained with malachite green. The assay has been

adapted from Reimann *et al*. (Reimann et al., 2016). In brief, ATP (final concentrations 0-8 mM) and $MgCl_2$ (final concentration 10 mM) were mixed and the reaction was started by adding HlyB* (final concentration 0.1 mg/mL). Samples were incubated at 37°C and 550 rpm for 20 min and the reaction was stopped with cold 20% (v/v) H_2SO_4. Negative controls contained 1 mM EDTA instead of 10 mM $MgCl_2$. Free phosphate was stained by incubation with staining solution (0.096% malachite green, 1.48% (w/v) ammonium molybdate and 0.173% (v/v) Tween 20 in 2.36 M H_2SO_4) for 15 min. The absorbance at 595 nm was measured and compared to a phosphate standard stained in the same way. Data points were analyzed with GraphPad Prism 7 software and plotted with an allosteric sigmoidal fit. The assay was performed in the respective SEC buffer.

Secretion of HlyA

Secretion experiments were performed with *E. coli* BL21 (DE3) cells transformed with pSU2726, encoding for HlyA (Soloaga et al., 1996) or HlyA1, and pK184 encoding for HlyD and HlyB (Jenewein, 2008) or the respective HlyB*. Cultures containing 50 mL LB-media supplemented with 100 μg/mL ampicillin and 30 μg/mL kanamycin were inoculated from 5 mL pre-cultures (same media and antibiotic concentration) to an OD_{600} of 0.1. Cells were grown to an OD_{600} of 0.5-0.7 at 37°C and 180 rpm and were induced with 1 mM IPTG and 1 mM arabinose. Before induction (0 h) and several hours after induction (1 h, 2 h and 3 h) the OD_{600} was measured and 1 mL sample was taken from the culture. The sample was centrifuged for 1 min at 11,000xg. Supernatants were diluted to the lowest OD_{600} measured, mixed with SDS-sample buffer and analyzed for the presence of HlyA or HlyA1 by SDS-PAGE and Western Blot as described above. Cell pellets were resuspended in water according to their OD_{600}, mixed with SDS-sample buffer and analyzed in the same manner but for the presence of HlyB, HlyB* and HlyD.

Acknowledgments

References

Akatsuka, H., Kawai, E., Omori, K. and Shibatani, T. (1995) 'The three genes lipB, lipC, and lipD involved in the extracellular secretion of the *Serratia marcescens* lipase which lacks an N-terminal signal peptide', *Journal of Bacteriology,* 177(22), pp. 6381-6389.

Bakkes, P. J., Jenewein, S., Smits, S. H., Holland, I. B. and Schmitt, L. (2010) 'The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin type 1 secretion system', *J Biol Chem,* 285(52), pp. 40573-80.

Balakrishnan, L., Hughes, C. and Koronakis, V. (2001) 'Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*', *Journal of Molecular Biology,* 313(3), pp. 501-510.

Baumann, U., Wu, S., Flaherty, K. M. and McKay, D. B. (1993) 'Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif', *The EMBO Journal,* 12(9), pp. 3357-3364.

Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B. and Schmitt, L. (2003) 'A Specific Interaction Between the NBD of the ABC-transporter HlyB and a C-Terminal Fragment of its Transport Substrate Haemolysin A', *Journal of Molecular Biology,* 327(5), pp. 1169-1179.

Combet, C., Blanchet, C., Geourjon, C. and Deléage, G. (2000) 'NPS@: network protein sequence analysis', *Trends Biochem Sci,* 25(3), pp. 147-50.

Cowan, D. A. (1992) 'Biotechnology of the Archaea', *Trends in Biotechnology,* 10, pp. 315-323.

Danson, M. J. and Hough, D. W. (1998) 'Structure, function and stability of enzymes from the Archaea', *Trends in Microbiology,* 6(8), pp. 307-314.

Dosztányi, Z., Csizmók, V., Tompa, P. and Simon, I. (2005) 'The Pairwise Energy Content Estimated from Amino Acid Composition Discriminates between Folded and Intrinsically Unstructured Proteins', *Journal of Molecular Biology,* 347(4), pp. 827-839.

Duong, F., Lazdunski, A., Carni, B. and Murgier, M. (1992) 'Sequence of a cluster of genes controlling synthesis and secretion of alkaline protease in *Pseudomonas aeruginosa*: relationships to other secretory pathways', *Gene,* 121(1), pp. 47-54.

Ellinger, P., Kluth, M., Stindt, J., Smits, S. H. and Schmitt, L. (2013) 'Detergent screening and purification of the human liver ABC transporters BSEP (ABCB11) and MDR3 (ABCB4) expressed in the yeast *Pichia pastoris*', *PLoS One,* 8(4), pp. e60620.

Felmlee, T. and Welch, R. A. (1988) 'Alterations of amino acid repeats in the *Escherichia coli* hemolysin affect cytolytic activity and secretion', *Proceedings of the National Academy of Sciences,* 85(14), pp. 5269.

Fitzpatrick, A. W. P., Llabrés, S., Neuberger, A., Blaza, J. N., Bai, X. C., Okada, U., Murakami, S., van Veen, H. W., Zachariae, U., Scheres, S. H. W., Luisi, B. F. and Du, D. (2017) 'Structure of the MacAB-TolC ABC-type tripartite multidrug efflux pump', *Nat Microbiol,* 2, pp. 17070.

Frauenfeld, J., Löving, R., Armache, J.-P., Sonnen, A. F. P., Guettou, F., Moberg, P., Zhu, L., Jegerschöld, C., Flayhan, A., Briggs, J. A. G., Garoff, H., Löw, C., Cheng, Y. and Nordlund, P. (2016) 'A saposin-lipoprotein nanoparticle system for membrane proteins', *Nature Methods,* 13(4), pp. 345-351.

Froger, A. and Hall, J. E. (2007) 'Transformation of plasmid DNA into *E. coli* using the heat shock method', *J Vis Exp,* (6), pp. 253.

Geertsma, E. R. (2013) 'FX cloning: a versatile high-throughput cloning system for characterization of enzyme variants', *Methods Mol Biol,* 978, pp. 133-48.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A. and Smith, H. O. (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nature Methods,* 6(5), pp. 343-345.

Gilson, L., Mahanty, H. K. and Kolter, R. (1987) 'Four plasmid genes are required for colicin V synthesis, export, and immunity', *J Bacteriol,* 169(6), pp. 2466-70.

Gilson, L., Mahanty, H. K. and Kolter, R. (1990) 'Genetic analysis of an MDR-like export system: the secretion of colicin V', *Embo j,* 9(12), pp. 3875-84.

Glaser, P., Sakamoto, H., Bellalou, J., Ullmann, A. and Danchin, A. (1988) 'Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of *Bordetella pertussis*', *Embo j,* 7(12), pp. 3997-4004.

Goebel, W. and Hedgpeth, J. (1982) 'Cloning and functional characterization of the plasmid-encoded hemolysin determinant of *Escherichia coli*', *J Bacteriol,* 151(3), pp. 1290-8.

Gray, L., Baker, K., Kenny, B., Mackman, N., Haigh, R. and Holland, I. B. (1989) 'A novel C-terminal signal sequence targets *Escherichia coli* haemolysin directly to the medium', *Journal of Cell Science,* 1989(Supplement 11), pp. 45.

Gray, L., Mackman, N., Nicaud, J. M. and Holland, I. B. (1986) 'The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*', *Mol Gen Genet,* 205(1), pp. 127-33.

Gulati, S., Jamshad, M., Knowles, Timothy J., Morrison, Kerrie A., Downing, R., Cant, N., Collins, R., Koenderink, Jan B., Ford, Robert C., Overduin, M., Kerr, Ian D., Dafforn, Timothy R. and Rothnie, Alice J. (2014) 'Detergent-free purification of ABC (ATP-binding-cassette) transporters', *Biochemical Journal,* 461(2), pp. 269-278.

Hanson, J., Yang, Y., Paliwal, K. and Zhou, Y. (2016) 'Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks', *Bioinformatics,* 33(5), pp. 685-692.

Heffernan, R., Yang, Y., Paliwal, K. and Zhou, Y. (2017) 'Capturing non-local interactions by long short-term memory bidirectional recurrent neural networks for improving prediction of protein secondary structure, backbone angles, contact numbers and solvent accessibility', *Bioinformatics,* 33(18), pp. 2842-2849.

Helenius, A., McCaslin, D. R., Fries, E. and Tanford, C. (1979) 'Properties of detergents', *Methods Enzymol,* 56, pp. 734-49.

Helenius, A. and Simons, K. (1975) 'Solubilization of membranes by detergents', *Biochim Biophys Acta,* 415(1), pp. 29-79.

Hess, J., Gentschev, I., Goebel, W. and Jarchau, T. (1990) 'Analysis of the haemolysin secretion system by PhoA-HlyA fusion proteins', *Mol Gen Genet,* 224(2), pp. 201-8.

Hinchliffe, P., Greene, N. P., Paterson, N. G., Crow, A., Hughes, C. and Koronakis, V. (2014) 'Structure of the periplasmic adaptor protein from a major

facilitator superfamily (MFS) multidrug efflux pump', *FEBS Letters,* 588(17), pp. 3147-3153.

Hinchliffe, P., Symmons, M. F., Hughes, C. and Koronakis, V. (2013) 'Structure and Operation of Bacterial Tripartite Pumps', *Annual Review of Microbiology,* 67(1), pp. 221-242.

Jarchau, T., Chakraborty, T., Garcia, F. and Goebel, W. (1994) 'Selection for transport competence of C-terminal polypeptides derived from *Escherichia coli* hemolysin: the shortest peptide capable of autonomous HlyB/HlyD-dependent secretion comprises the C-terminal 62 amino acids of HlyA', *Molecular and General Genetics MGG,* 245(1), pp. 53-60.

Jenewein, S. (2008) *The Escherichia coli haemolysin transporter - A paradigm for Type I secretion.* PhD Doctoral Dissertation, Heinrich-Heine-University Duesseldorf.

Jo, I., Kim, J. S., Xu, Y., Hyun, J., Lee, K. and Ha, N. C. (2019) 'Recent paradigm shift in the assembly of bacterial tripartite efflux pumps and the type I secretion system', *J Microbiol,* 57(3), pp. 185-194.

Jones, D. T. (1999) 'Protein secondary structure prediction based on position-specific scoring matrices*Edited by G. von Heijne', *Journal of Molecular Biology,* 292(2), pp. 195-202.

Jones, D. T. and Cozzetto, D. (2015) 'DISOPRED3: precise disordered region predictions with annotated protein-binding activity', *Bioinformatics,* 31(6), pp. 857-863.

Kanonenberg, K., Royes, J., Kedrov, A., Poschmann, G., Angius, F., Solgadi, A., Spitz, O., Kleinschrodt, D., Stühler, K., Miroux, B. and Schmitt, L. (2019a) 'Shaping the lipid composition of bacterial membranes for membrane protein production', *Microb Cell Fact,* 18(1), pp. 131.

Kanonenberg, K., Schwarz, C. K. and Schmitt, L. (2013) 'Type I secretion systems - a story of appendices', *Res Microbiol,* 164(6), pp. 596-604.

Kanonenberg, K., Smits, S. H. J. and Schmitt, L. (2019b) 'Functional Reconstitution of HlyB, a Type I Secretion ABC Transporter, in Saposin-A Nanoparticles', *Scientific Reports,* 9(1), pp. 8436.

Kenny, B., Haigh, R. and Holland, I. B. (1991) 'Analysis of the haemolysin transport process through the secretion from *Escherichia coli* of PCM, CAT or beta-galactosidase fused to the Hly C-terminal signal domain', *Mol Microbiol,* 5(10), pp. 2557-68.

Kim, J. S., Song, S., Lee, M., Lee, S., Lee, K. and Ha, N. C. (2016) 'Crystal Structure of a Soluble Fragment of the Membrane Fusion Protein HlyD in a Type I Secretion System of Gram-Negative Bacteria', *Structure,* 24(3), pp. 477-85.

Klausen, M. S., Jespersen, M. C., Nielsen, H., Jensen, K. K., Jurtz, V. I., Sønderby, C. K., Sommer, M. O. A., Winther, O., Nielsen, M., Petersen, B. and Marcatili, P. (2019) 'NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning', *Proteins,* 87(6), pp. 520-527.

Koronakis, V., Eswaran, J. a. and Hughes, C. (2004) 'Structure and Function of TolC: The Bacterial Exit Duct for Proteins and Drugs', *Annual Review of Biochemistry,* 73(1), pp. 467-489.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B. and Hughes, C. (2000) 'Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export', *Nature,* 405(6789), pp. 914-9.

Lecher, J., Schwarz, Christian K. W., Stoldt, M., Smits, Sander H. J., Willbold, D. and Schmitt, L. (2012) 'An RTX Transporter Tethers Its Unfolded Substrate during Secretion via a Unique N-Terminal Domain', *Structure,* 20(10), pp. 1778-1787.

Lee, M., Jun, S. Y., Yoon, B. Y., Song, S., Lee, K. and Ha, N. C. (2012) 'Membrane fusion proteins of type I secretion system and tripartite efflux pumps share a binding motif for TolC in gram-negative bacteria', *PLoS One,* 7(7), pp. e40460.

Lenders, M. H. H., Weidtkamp-Peters, S., Kleinschrodt, D., Jaeger, K.-E., Smits, S. H. J. and Schmitt, L. (2015) 'Directionality of substrate translocation of the hemolysin A Type I secretion system', *Scientific Reports,* 5(1), pp. 12470.

Létoffé, S., Ghigo, J. M. and Wandersman, C. (1994) 'Secretion of the *Serratia marcescens* HasA protein by an ABC transporter', *J Bacteriol,* 176(17), pp. 5372-7.

Lin, D. Y.-w., Huang, S. and Chen, J. (2015) 'Crystal structures of a polypeptide processing and secretion transporter', *Nature,* 523(7561), pp. 425-430.

Lin, S.-H. and Guidotti, G. (2009) 'Chapter 35 Purification of Membrane Proteins', in Burgess, R.R. & Deutscher, M.P. (eds.) *Methods in Enzymology*: Academic Press, pp. 619-629.

Linhartová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., Kamanová, J., Procházková, K., Adkins, I., Hejnová-Holubová, J., Sadílková, L., Morová, J. and Šebo, P. (2010) 'RTX proteins: a highly diverse family secreted by a common mechanism', *FEMS Microbiology Reviews,* 34(6), pp. 1076-1112.

Mackman, N., Nicaud, J. M., Gray, L. and Holland, I. B. (1985) 'Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*", *Mol Gen Genet,* 201(3), pp. 529-36.

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D. and Lopez, R. (2019) 'The EMBL-EBI search and sequence analysis tools APIs in 2019', *Nucleic acids research,* 47(W1), pp. W636-W641.

Mandel, M. and Higa, A. (1970) 'Calcium-dependent bacteriophage DNA infection', *J Mol Biol,* 53(1), pp. 159-62.

Michiels, J., Dirix, G., Vanderleyden, J. and Xi, C. (2001) 'Processing and export of peptide pheromones and bacteriocins in Gram-negative bacteria', *Trends in Microbiology,* 9(4), pp. 164-168.

Mikolosko, J., Bobyk, K., Zgurskaya, H. I. and Ghosh, P. (2006) 'Conformational flexibility in the multidrug efflux system protein AcrA', *Structure,* 14(3), pp. 577-87.

Miroux, B. and Walker, J. E. (1996) 'Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels', *J Mol Biol,* 260(3), pp. 289-98.

Mól, A. R., Castro, M. S. and Fontes, W. (2018) 'NetWheels: A web application to create high quality peptide helical wheel and net projections', *bioRxiv,* pp. 416347.

Nicaud, J. M., Mackman, N., Gray, L. and Holland, I. B. (1985a) 'Characterisation of HlyC and mechanism of activation and secretion of haemolysin from *E. coli* 2001', *FEBS Lett,* 187(2), pp. 339-44.

Nicaud, J. M., Mackman, N., Gray, L. and Holland, I. B. (1985b) 'Regulation of haemolysin synthesis in *E. coli* determined by HLY genes of human origin', *Mol Gen Genet,* 199(1), pp. 111-6.

Oswald, C., Jenewein, S., Smits, S. H., Holland, I. B. and Schmitt, L. (2008) 'Water-mediated protein-fluorophore interactions modulate the affinity of an ABC-ATPase/TNP-ADP complex', *J Struct Biol,* 162(1), pp. 85-93.

Pimenta, A. L., Racher, K., Jamieson, L., Blight, M. A. and Holland, I. B. (2005) 'Mutations in HlyD, Part of the Type 1 Translocator for Hemolysin Secretion, Affect the Folding of the Secreted Toxin', *Journal of Bacteriology,* 187(21), pp. 7471-7480.

Pimenta, A. L., Young, J., Holland, I. B. and Blight, M. A. (1999) 'Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator', *Molecular and General Genetics MGG,* 261(1), pp. 122-132.

Reimann, S., Poschmann, G., Kanonenberg, K., Stühler, K., Smits, Sander H. J. and Schmitt, L. (2016) 'Interdomain regulation of the ATPase activity of the ABC transporter haemolysin B from Escherichia coli', *Biochemical Journal,* 473(16), pp. 2471-2483.

Sapay, N., Guermeur, Y. and Deléage, G. (2006) 'Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier', *BMC Bioinformatics,* 7(1), pp. 255.

Satchell, K. J. (2011) 'Structure and function of MARTX toxins and other large repetitive RTX proteins', *Annu Rev Microbiol,* 65, pp. 71-90.

Schmitt, L., Benabdelhak, H., Blight, M. A., Holland, I. B. and Stubbs, M. T. (2003) 'Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains', *J Mol Biol,* 330(2), pp. 333-42.

Schülein, R., Gentschev, I., Schlör, S., Gross, R. and Goebel, W. (1994) 'Identification and characterization of two functional domains of the hemolysin translocator protein HlyD', *Mol Gen Genet,* 245(2), pp. 203-11.

Schwarz, C. K. W., Landsberg, C. D., Lenders, M. H. H., Smits, S. H. J. and Schmitt, L. (2012) 'Using an *E. coli* Type 1 secretion system to secrete the mammalian, intracellular protein IFABP in its active form', *Journal of Biotechnology,* 159(3), pp. 155-161.

Soloaga, A., Ostolaza, H., Goñi, F. M. and de la Cruz, F. (1996) 'Purification of *Escherichia coli* pro-haemolysin, and a comparison with the properties of mature alpha-haemolysin', *Eur J Biochem,* 238(2), pp. 418-22.

Springer, W. and Goebel, W. (1980) 'Synthesis and secretion of hemolysin by *Escherichia coli*', *J Bacteriol,* 144(1), pp. 53-9.

Stanley, P., Packman, L. C., Koronakis, V. and Hughes, C. (1994) 'Fatty acylation of two internal lysine residues required for the toxic activity of *Escherichia coli* hemolysin', *Science,* 266(5193), pp. 1992-6.

Su, C.-C., Yang, F., Long, F., Reyon, D., Routh, M. D., Kuo, D. W., Mokhtari, A. K., Van Ornam, J. D., Rabe, K. L., Hoy, J. A., Lee, Y. J., Rajashankar, K. R. and Yu, E. W. (2009) 'Crystal Structure of the Membrane Fusion Protein CusB from *Escherichia coli*', *Journal of Molecular Biology,* 393(2), pp. 342-355.

Thanabalu, T., Koronakis, E., Hughes, C. and Koronakis, V. (1998) 'Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore', *Embo j,* 17(22), pp. 6487-96.

Thomas, S., Bakkes, P. J., Smits, S. H. and Schmitt, L. (2014) 'Equilibrium folding of pro-HlyA from *Escherichia coli* reveals a stable calcium ion dependent folding intermediate', *Biochim Biophys Acta,* 1844(9), pp. 1500-10.

Tsutsumi, K., Yonehara, R., Ishizaka-Ikeda, E., Miyazaki, N., Maeda, S., Iwasaki, K., Nakagawa, A. and Yamashita, E. (2019) 'Structures of the wild-type MexAB–OprM tripartite pump reveal its complex formation and drug efflux mechanism', *Nature Communications,* 10(1), pp. 1520.

Urh, M., Simpson, D. and Zhao, K. (2009) 'Chapter 26 Affinity Chromatography: General Methods', in Burgess, R.R. & Deutscher, M.P. (eds.) *Methods in Enzymology*: Academic Press, pp. 417-438.

Wandersman, C. and Delepelaire, P. (1990) 'TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion', *Proc Natl Acad Sci U S A,* 87(12), pp. 4776-80.

Wang, S., Peng, J., Ma, J. and Xu, J. (2016) 'Protein Secondary Structure Prediction Using Deep Convolutional Neural Fields', *Scientific Reports,* 6(1), pp. 18962.

Wang, Z., Fan, G., Hryc, C. F., Blaza, J. N., Serysheva, II, Schmid, M. F., Chiu, W., Luisi, B. F. and Du, D. (2017) 'An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump', *Elife,* 6.

Yan, R., Xu, D., Yang, J., Walker, S. and Zhang, Y. (2013) 'A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction', *Scientific Reports,* 3(1), pp. 2619.

Yum, S., Xu, Y., Piao, S., Sim, S. H., Kim, H. M., Jo, W. S., Kim, K. J., Kweon, H. S., Jeong, M. H., Jeon, H., Lee, K. and Ha, N. C. (2009) 'Crystal structure of the

periplasmic component of a tripartite macrolide-specific efflux pump', *J Mol Biol,* 387(5), pp. 1286-97.

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B. and Schmitt, L. (2005a) 'H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB', *Embo j,* 24(11), pp. 1901-10.

Zaitseva, J., Jenewein, S., Wiedenmann, A., Benabdelhak, H., Holland, I. B. and Schmitt, L. (2005b) 'Functional characterization and ATP-induced dimerization of the isolated ABC-domain of the haemolysin B transporter', *Biochemistry,* 44(28), pp. 9680-90.

Zaitseva, J., Oswald, C., Jumpertz, T., Jenewein, S., Wiedenmann, A., Holland, I. B. and Schmitt, L. (2006) 'A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer', *Embo j,* 25(14), pp. 3432-43.

Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A. N. and Alva, V. (2018) 'A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core', *Journal of Molecular Biology,* 430(15), pp. 2237-2243.

Supplement

Supplementary Table 1: Identity and characteristics of putative T1SS components of 25 different organisms. Proteins were identified by pBLAST search. The number (#) of GG repeats sometimes shows a range. This depends on how strict the motif (GGxGxDxUx, where U stands for a large hydrophobic residue and x stands for any amino acid) is applied (Linhartová et al., 2010). For some organisms multiple RTX proteins were identified. aa = amino acids.

organism	identity compared to [%]				putative RTX toxin	
	HlyB	HlyD	HlyA	HlyC	size [aa]	# of GG repeats
Enterobacter cloacae	98	98	97	98	1024	6-7
Vibrio parahaemolyticus	92	81	82	87	986	8-9
Proteus vulgaris	92	95	46	-	598	6
Morganella morganii	90	81	80	85	1024	5-6
Actinobacillus equuli subsp. haemolyticus	86	64	47	59	987	4
Aggregatibacter actinomycetemcomitans	84	68	50	71	1051	7
Pasteurella aerogenes	83	62	52	58	1049	6-9
Mannheimia haemolytica	82	61	43	55	953	5
Bibersteinia trehalosi	82	59	42	50	955	6
Kingella kingae	72	40	44	62	956	6-7
Snodgrassella alvi	71	45	50	31	895	24
Gallibacterium anatis	70	49	29	40	2038	5-6
Vitreoscilla sp. SN6	70	42	44	-	444	10
Neisseria sp. oral taxon 020	70	43	38, 46, 49	-	1605, 636, 188	14, 3-5, 4-5
Alysiella crassa	70	41	31	-	248	4
Cardiobacterium valvarum	70	41	34, 35, 41, 49	-	217, 569, 665, 558	4, 8-9, 2, 5-6
Avibacterium paragallinarum	69	48	32	-	2286	15
Moraxella bovis	69	41	43	56	927	5
Cronobacter malonaticus	69	46	36	59	866	6
Acinetobacter baumannii	69	44	52	28	3298	49-58
Serratia sp. Leaf51	69	42	33, 34	44	965, 2893	9, 12-14
Aeromonas diversa CDC 2478-85	65	38	43	-	351	9-10
Lysobacter antibioticus	64	37	42	-	574	9-11
Xanthomonas axonopodis	63	38	38	31	2512	36-38
Xylella fastidiosa	61	41	37	32	1814	16-18
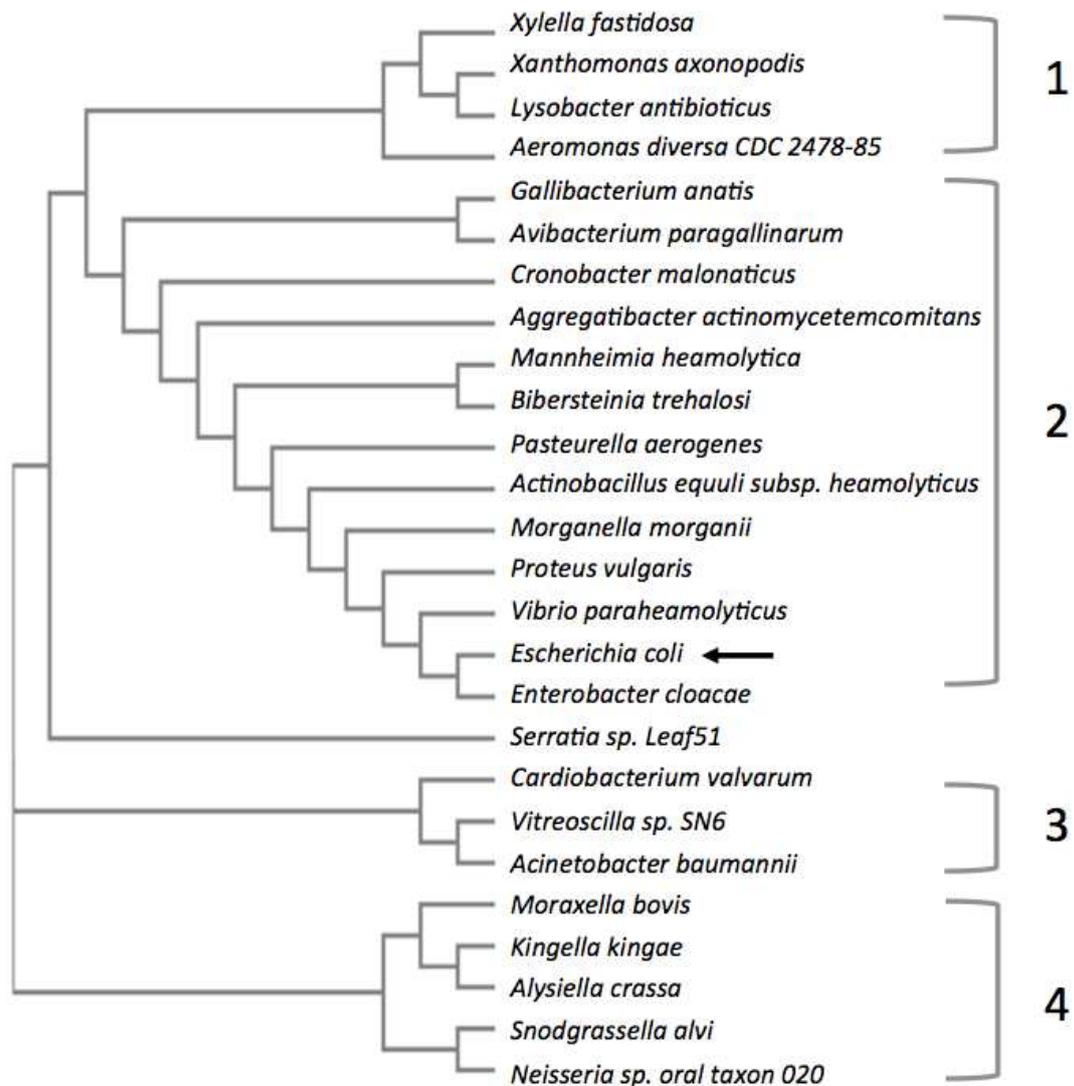
Supplementary Figure 1: Phylogenetic tree based on HlyB* from these organisms. The primary sequence of HlyB* from these organisms was used in an alignment done by Clustal Omega (Madeira et al., 2019). The organisms are clustered in four groups. *E. coli* is marked with a black arrow. The length of the lines does not correspond to the degree of relatedness.

Supplementary Table 2: Kinetic parameters of ATPase measurements of HlyB* from Ap compared to HlyB from *E. coli* (Ec) (Reimann et al., 2016).

	Ap				Ec
	peak 1, day 1	peak 2, day 1	peak 1, day 7	EQ HA, peak 2	HlyB
v_{max} [nmol P_i/(mg*min)]	27.1	19.2	22.1	2.7	8.1
h	1.6	1.9	1.4	0.9	1.5
K_{half} [mM]	1.0	6.5	0.6	0.9	0.3
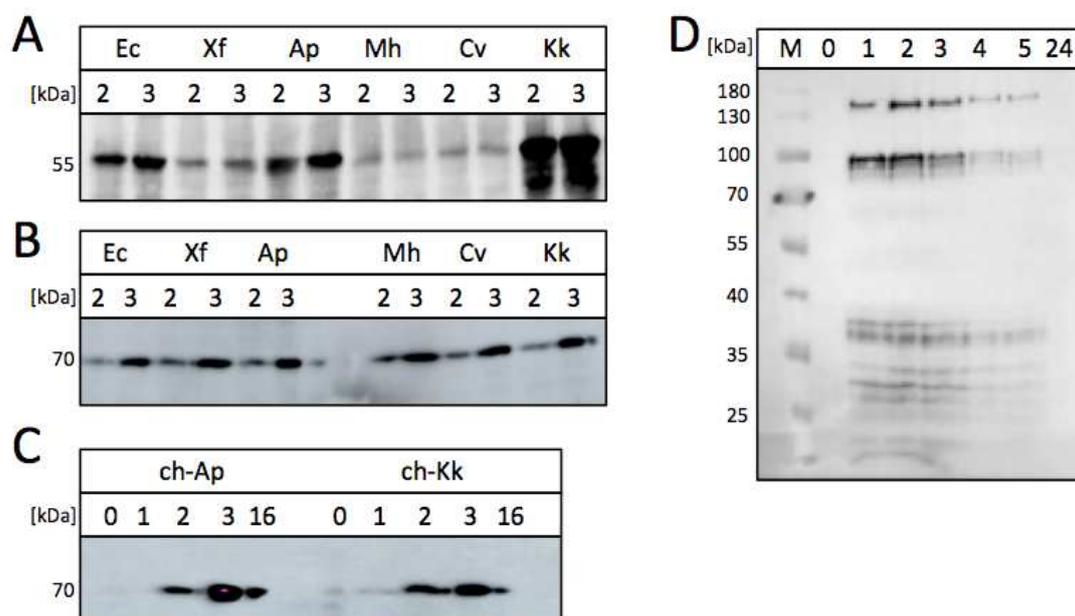
Supplementary Figure 2: A) Western Blot of whole cell samples from secretion experiments against HlyD. Samples were diluted to the same OD_{600}. **B)** Western Blot of the same samples against HlyB NBD. The antibody was able to bind the homologs. Two letter abbreviations stand for genus and species of the respective organism from Table 1 including *E. coli* as Ec. **C)** Western Blot of whole cell samples from secretion with chimeric HlyB* from Ap (ch-Ap) and Kk (ch-Kk) against HlyB NBD. **D)** Western Blot of whole cell samples from a test expression of HlyB* from Bt in *E. coli* C43(DE3)Δ(*ompF-acrAB*) at 37°C. Numbers indicate hours after induction of protein expression. M: molecular weight marker.

Walker A
```
Ec  ITFRNIRFRYKPDSPVILDNINLSIKQGEVIGIVGRSGSGKSTLTKLIQRFYIPENGQVL
Xf  ITFERLIFRYRPDTPDVLSGIDLDIQPGEVIGIVGRSGSGKSTLTKLVQRMYVPERGRVL
Ap  IVFENVKFRYKPEDRDIISDFSLKLSAGEVVGIVGRSGSGKSTIAKLIQRLYIPQSGRIY
Mh  ISFKNIRFRYKPDAPTILNNVNLEIRQGEVIGIVGRSGSGKSTLTKLLQRFYIPENGQVL
Cv  IVFDHVHFRYRPDAQPVLTDLSLSIRAGEVIGIVGRSGSGKSTLTKLVQRLYVPEQGRVL
Kk  ITFEHVDFRYKADGRLILQDLNLQIQAGEVLGIVGRSGSGKSTLTKLVQRLYTPENGRVL
    * * .: ***: :    :: ...*.:   ***:***********:**:**:* *: *::

                              Q-loop
Ec  IDGHDLALADPNWLRRQVGVVLQDNVLLNRSIIDNISLANPGMSVEKVIYAAKLAGAHDF
Xf  VDSHDLSLADPAWLRRQIGVVLQENFLFNRSVRENIAMADPGIPLERVIHAATLAGAHTF
Ap  IDGQDLSVVDPNWLRRQIGVVLQDNVLLNRSIRDNIAISEPGISMEKVIIAAKLAGAHDF
Mh  IDGHDLALADPNWLRRQIGVVLQDNVLLNRSIRENIALSDPGMPMERVIYAAKLAGAHDF
Cv  IDGNDLALADPAWLRRQVGVVLQENVLMNASVRDNIALSDPGMPLEHVMQAARLAGAHDF
Kk  VDGNDLALADPAWLRRQVGVVLQENVLLNRSIRDNIALTDTGMPLEQIIQAAKLAGAHDF
    :*.:**::.** *****:*****:*.*:* *: :**:::: *: :*::: ** ***** *

              C-loop                   P-loop Walker B  D-loop
Ec  ISELREGYNTIVGEQGAGLSGGQRQRIAIARALVNNPKILIFDEATSALDYESEHIIMRN
Xf  ISELPEGYDTKIGEHGTGLSGGQRQRIAIARALIGDPRILIFDEATSALDYESEHAVMQN
Ap  ITQLSEGYDTLVGEQGAGLSGGQRQRIAIARALISNPKILIFDEATSALDYESERIIMQN
Mh  ISELREGYNTIVGEQGAGLSGGQRQRIAIARALVNNPKILIFDEATSALDYESEHIIMQN
Cv  IMELREGYDTIVGEQGAGLSGGQRQRIAIARALVNNPRILILDEATSALDYESERAIMEN
Kk  IMELSEGYDTMVGEQGAGLSGGQRQRIAIARALITNPRILIFDEATSALDYESERAIMQN
    * :* ***:* :**:*:***************: :*:***:************: :*.*

            switch II               G-loop
Ec  MHKICKGRTVIIIAHRLSTVKNADRIIVMEKGKIVEQGKHKELLSEPESLYSYLYQLQSD
Xf  MRAICKGRTVLIIAHRLSTVRNADRIVVIDKGKIVESGSHETLLNRNNGHYAHLYRQQGT
Ap  MNSICRGRTVIIIAHRLSTIRMANKIVVMEQGNIVESGSHDMLLENENGLYSYLNQLQSG
Mh  MQKICQGRTVILIAHRLSTVKNADRIIVMEKGEIVEQGKHHELLQNSNGLYSYLHQQLN-
Cv  MDAICQNRTVLIIAHRLSTVRRAHRIIAMDKGRIIEEGSHDELLRRKDGYYRYLYQLQNG
Kk  MQAICQGRTVLIIAHRLSTVRHAHRIIAMDKGKIVEQGTHQALLQKADGYYRYLYDQNG-
    *  **:.***:********:: *.:*:.:::*.*:*.*.*. ** . :. * :*     .
```

Supplementary Figure 3: Alignment of HlyB* NBDs from group 2 T1SSs. (*) marks the same residue. (:) and (.) mark similar residues. Conserved motifs are labeled (Schmitt et al., 2003). Alignment was performed with Clustal Omega (Madeira et al., 2019). Two letter abbreviations stand for genus and species of the respective organism (see Table 1).

4. Discussion

The T1SS can be used as a biotechnological platform for the secretion of various proteins and peptides. Understanding the secretion process in detail can help to increase secretion efficiency and for this the structure of the whole complex during different stages of secretion is desirable. Obtaining such a structure is a complex task that is supported by structural approaches towards single components of the assembled system. In regards to the HlyA T1SS the last missing components are parts of HlyD and the transmembrane domains (TMDs) of HlyB. Since crystallization of HlyB was unsuccessful so far a homology approach was initiated during this work. Although no crystal was obtained until now, the heterologous expression of five HlyB homologs and purification protocols for three of these were successfully established (chapter 3.5). The second structural aim was directed towards the cytoplasmic domain (CD) of HlyD. Purification trials of this domain were most successful under denaturing conditions, which render structural approaches and interaction studies pointless. Dialysis into another buffer system led to aggregation of the domain under all conditions tested (88 buffer systems). Additionally, the yield from purifications under native conditions was too low to perform NMR studies or start crystallization trials.

Fortunately, structural information can also be gained by transferring information from one system to another, if they are similar enough. This principle is used by prediction and modeling tools, which become more accurate with increasing databases (chapter 3.3). The information gained from such tools should be validated by mutational studies.

The last aim of this work was gaining a more detailed understanding of the domain interactions that lead to the assembly of the HlyA T1SS. This was hindered by the unfavorable purification method for HlyD CD, which would have allowed *in vitro* interaction studies with HlyA, and the inability to transfer the proteinase K susceptibility assay to the HlyA T1SS (chapter 3.4). Nevertheless, during this work two motifs that are unique to HlyD-like membrane fusion proteins (MFPs) were identified, as well as two possible binding pockets (pbp's) for HlyA in the NBD

(nucleotide binding domain) of HlyB (chapter 3.3 and 3.5). Furthermore, two amphipathic helices (AHs) were confirmed: One in the secretion signal of HlyA and a second one in the CD of HlyD (chapter 3.3). The following sections will put the information gained during this work into the context of T1 mediated secretion and possible consequences are discussed.

4.1 The membrane fusion protein (MFP) – HlyD

The MFP HlyD represents the least investigated component of the HlyA T1SS. Its function is mostly described as simply connecting HlyB and TolC. However, Pimenta *et al.* found that mutations in the C-terminus of HlyD can affect the folding of secreted HlyA pointing towards a direct interaction between HlyD and HlyA and also towards an active role of HlyD in the secretion process (Pimenta et al., 2005). The identified residues (K404 and D411) are located in the periplasmic domain (PPD) close to a highly conserved motif (residue 393 - 400) that is present in all MFPs of group 2 and to some extent in MFPs of BTLCP-linked T1SSs. A pBLAST search with this motif only shows 'HlyD-like proteins' among the first 100 results (chapter 3.5). Furthermore, motif search tools such as 'prosite' and 'MOTIF' do not recognize this motif in HlyD nor can they assign a function when only the motif (FPYRYGY) is used as an input sequence.

A second conserved motif was identified in the CD of HlyD (residue 41 – 48) and this motif is exclusively found in MFPs of group 2 T1SSs. The CD of HlyD is essential to HlyA secretion as deletion of the first 40 amino acids as well as deletion of residue 26 to 45 completely abolishes HlyA secretion (Balakrishnan et al., 2001). MFPs of RND-type efflux pumps are mostly lipoproteins that have no CD, while MFPs of T1SSs hold CDs of different sizes. However, only MFPs of group 2 T1SSs hold a CD of 50-60 residues with an AH of ~25 residues followed by a charged cluster and the aforementioned conserved motif (residue 41 – 48) (chapter 3.5). So far it is unclear how the CD is involved in the secretion process. Cross-link experiments of Thanabalu *et al.* showed that HlyA and HlyD can interact in the absence of HlyB pointing towards an involvement of the CD in the substrate recognition process

(Thanabalu et al., 1998). Localization studies during this work showed a tendency of the CD to localize to the membrane fraction, which is likely caused by the AH (Figure 5 in chapter 3.3). A change in conformation (insertion in the membrane) points towards an involvement in signal transduction that might follow a direct interaction with the substrate HlyA. However, an interaction between HlyA and HlyD is not sufficient to recruit TolC; for this HlyB is also needed (Thanabalu et al., 1998). Recently, Alav *et al.* modeled an assembled HlyA T1SS and pointed out that the TM-helices of HlyD and HlyB could interact with each other (Alav et al., 2021). This interaction could stabilize the conformational change in HlyD that leads to TolC recruitment.

4.2   Interaction between HlyA and HlyB

HlyA C-terminus

There are two publications showing direct interactions between truncated versions of HlyA and isolated domains of HlyB. Measurements on a model of HlyB showed that these interactions can take place simultaneously (chapter 3.3) with the CLD (C39-like domain) of HlyB interacting with the RTX (repeats in toxin) domain of HlyA and the NBD of HlyB with the secretion signal of HlyA (Benabdelhak et al., 2003, Lecher et al., 2012). The secretion signal of HlyA is located in the C-terminal 60 residues and presents different features (Holland et al., 2016). During this work the main focus was placed on the N-terminal part of the secretion signal where the presence of an AH between residue P975 and A986 could be confirmed (chapter 3.3). Interestingly, heterologous substrates that can be secreted by HlyBD also show AHs in their C-terminal secretion signal when analyzed with the prediction tool AmphipaSeeK (described in chapter 3.3). Furthermore, these heterologous substrates display the same distance between the predicted AH and a GG-repeat of the RTX domain as HlyA (Table 1). A linker of a conserved length between the two interaction sites further strengthens the theory that the RTX domain and the secretion signal interact simultaneously with the ABC transporter, which might be important for the correct orientation of the substrate. Additionally, an aromatic residue (F990) close to

the AH was shown to be important for HlyA secretion and a Phe residue close to the predicted AHs can be found in the heterologous substrates as well (Table 1).

Table 1: Comparison of RTX proteins that can be secreted by HlyBD-TolC. The references list the publications that showed the heterologous secretion of the respective RTX toxin by HlyBD-TolC. The position of the first glycine of a GG repeat as well as the beginning of a predicted amphipathic helix (AH) are noted. The AHs were predicted with AmphipaSeeK (Sapay et al., 2006).

RTX protein	host	position of GG repeat	AH	distance GG - AH	aromatic residue	reference
HlyA	*E. coli*	844	974	130	F990	
FrpA	*N. meningitides*	1125	1257	132	F1274	(Thompson and Sparling, 1993)
FrpA	*K. kingae*	600	720	120	F733	(Erenburg, 2020)
MbxA	*M. bovis*	756	886	130	F901	(Erenburg, 2020)
HlyIA	*A. pleuropneumoniae*	840	969	129	F987	(Gygi et al., 1990)
LktA	*M. haemolytica*	779	909	130	F923	(Highlander et al., 1990)
PaxA	*P. aerogenes*	845	974	129	F991	(Kuhnert et al., 2000)

Mutations in the secretion signal of HlyA that cause the most severe defects in secretion cluster in the region of the AH (Holland et al., 2016). Thus, under the assumption that the AH of HlyA's secretion signal is the feature that interacts with the NBD of HlyB, two hydrophobic patches were identified that represent possible binding pockets (pbp-in and pbp-out) for the hydrophobic side of the AH (chapter 3.3). Both pbp's were mapped onto a model of dimeric HlyB leading to four possible interaction sites in the NBDs. This raises the question of how many HlyA molecules can or need to bind to HlyB to ensure efficient transport. The recently published structure of PCAT1 from *Clostridium thermocellum* in complex with its substrate CtA showed that two substrate molecules bind to PCAT1 (Kieuvongngam et al., 2020). The substrates bind with their N-terminal leader peptide to the C39 peptidase domain (termed PEP domain) of the transporter but only one of the substrates is positioned for cleavage. PCAT1 is derived from a Gram-positive organism so it cannot

be grouped within the introduced groups of T1SSs (see 1.2.2). However, considering that the ABC transporter displays an active C39 peptidase domain on its N-terminus, the substrate presents an N-terminal leader peptide that is cleaved and is relatively small (uncleaved: 10.2 kDa), PCAT1 shares most features with group 1 T1SSs and is thought to follow the classical 'alternating access mechanism' (introduced in 1.2.2.1) (Kieuvongngam et al., 2020). Most importantly, in the structure the substrates do not interact with the NBDs and therefore the transferability of the substrate-ABC transporter ratio from CtA-PCAT1 to HlyA-HlyB is very limited. However, there are some indirect indications that more than one HlyA molecule interacts with HlyB. Firstly, if only one HlyA molecule should bind to dimeric HlyB, either the amount of HlyA has to be strictly controlled or the second interaction site has to be disrupted by binding of the first HlyA. The latter could be envisioned for the NBDs that can dimerize but is harder to rationalize for the CLDs that show no tendency to interact with each other and are separated from each other in the modeled structure (chapter 3.3). Secondly, the ATPase activity of HlyB is inhibited at low HlyA concentrations but is stimulated once a certain concentration threshold of HlyA is surpassed (Reimann et al., 2016). Thirdly, the secretion efficiency of HlyA is increased by an untranslated region upstream of the *hlyA* gene. This region improves the interaction with ribosomal protein S1 leading to more HlyA mRNA and consequently to more HlyA protein (Khosa et al., 2018). Taken together this suggests that more than one HlyA molecule binds to HlyB at the same time, which is not considered when describing the secretion process. For simplicity the secretion process of HlyA is described for only one HlyA molecule. Most descriptions are in agreement that after secretion of one HlyA molecule ATP hydrolysis or ADP and P_i release reset HlyB and TolC disengages from the IM complex (Lenders et al., 2013, Kanonenberg et al., 2018, Smith et al., 2018b, Alav et al., 2021). When describing the disassembly of the system a paper is cited, which showed that HlyBD and TolC can disengage after secretion showing that the assembly is reversible. However, this disassembly was not observed after secretion of one HlyA but after the intracellular pool of HlyA had been depleted (Thanabalu et al., 1998); a detail that is frequently overlooked.

If more than one HlyA binds to HlyB but similar to PCAT1, only one of them is positioned for transport, it seems more efficient to keep the HlyBD-TolC channel assembled as long as the substrate is available opposed to a continuous assembly and disassembly after each round of secretion. This adds another layer of complexity to the secretion process as the initial interaction of the first HlyA molecule would induce channel assembly but following substrates would interact with the already assembled system. This intriguing question could have been investigated with the proteinase K susceptibility assay introduced in chapter 3.4. For this assay, the T1SS was stalled by fusing eGFP to the N-terminus of HlyA. Since directionality of secretion is C-terminal, the fast folding eGFP folds in the cytoplasm and clogs the transport channel, while the C-terminus of HlyA reaches the outside of the cell (Lenders et al., 2015). Permeabilization of the OM allowed proteinase K to enter the periplasm and digest TolC, which shows different digestion patterns depending on the assembly state (Masi and Wandersman, 2010). For the HlyA T1SS stalling of the system resulted in a decreased susceptibility of TolC to proteinase K compared to the non-assembled system (chapter 3.4). Initial experiments with wild type HlyA (without eGFP) showed the same trend implying that the system is constantly assembled even when not stalled. Unfortunately, many problems were encountered during the establishment of the assay and the decreased susceptibility of TolC to proteinase K was gradually lost over time for unknown reasons (chapter 3.4). Therefore, information gained from this assay should be treated with extreme care and validated by another approach.

HlyA N-terminus

The secretion signal of HlyA only represents ~6% of the complete molecule and emerges from the ribosome last. The second domain that was shown to interact with HlyB, the RTX domain, is also located in the C-terminus and therefore also emerges quite late from the ribosome (Lecher et al., 2012). So far it is unclear how the emerging protein chain is protected from degradation prior to secretion as no secreted degradation products of HlyA have been reported to date (Holland et al., 2016). This implies that either a mechanism of quality control is in place, where

partially degraded HlyA is not recognized by the transporter, or a mechanism of protection is in place that protects the emerging protein chain from cytoplasmic proteases. Such protection can be conferred by a chaperon and the CLD is suggested to take on this role for HlyA (Lecher et al., 2012, Kanonenberg et al., 2013). However, the CLD has been shown to interact with the RTX domain, which is located in the C-terminal part of HlyA. Therefore, additional interactions between the N-terminal part of HlyA and the transporter components seem plausible. This has been shown for HasA, which displays multiple linear regions throughout the whole protein that can interact with HasDE and induce TolC recruitment (Masi and Wandersman, 2010). However, HasA is quite different from HlyA as HasA interacts with the general chaperon SecB, holds no RTX domain and might even be secreted with its N-terminus first (Alav et al., 2021). Still, a stabilizing effect of HlyA's N-terminus on HlyB was observed during secretion experiments with wild type HlyA in comparison to the truncated version HlyA1 (Figure 4.1). Additionally, cells expressing HlyBD and eGFP-HlyA showed stronger HlyB signals on Western Blots than cells only expressing HlyBD implying a general stabilization effect of HlyA on HlyB (Figure 3 A in chapter 3.4).
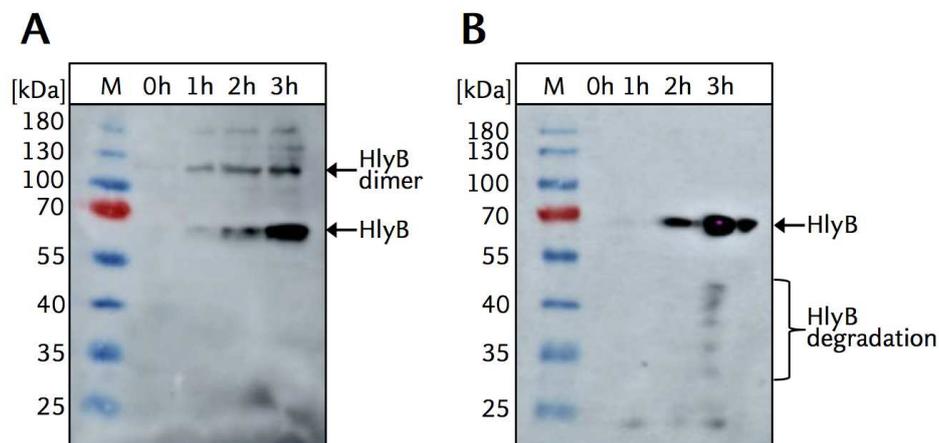
Figure 4.1 Comparison of HlyB stability during secretion of HlyA (A) and HlyA1 (B). Secretion experiments were performed in *E. coli* BL21(DE3). Before induction of protein expression (0 h) and at 1 h, 2 h and 3 h after induction, samples were taken, adjusted to the same OD_{600} and analyzed via SDS-PAGE and Western Blot as described in the material and methods section of chapter 3.5. The used antibody was directed against HlyB NBD. In (**A**) the full-length HlyA was used and in (**B**) the truncated version HlyA1 (see **Figure 1.14**). Degradation bands of HlyB are only visible in combination with HlyA1 (**B**), while dimer bands are only visible in combination with HlyA (**A**). Although the SDS-PAGE was conducted in the presence of 3.6% SDS, dimeric HlyB is visible in (**A**). The tendency to stay dimerized in the presence of SDS was also observed for some homologs of HlyB (Figure 1 in chapter 3.5). The Western Blots shown here represent one of three independent secretion experiments.

4.3   One mechanism for all?

T1SSs are widespread and the amount of information for each single system differs drastically. Grouping the systems based on their similarities helps to identify features and information that can be transferred from one system to another. In 2013 Kanonenberg *et al.* described three distinct groups of T1SSs based on the identity of the N-terminal extension of the ABC transporter (Kanonenberg et al., 2013) (see 1.2.2). Although this was the only criterion, they noted that also several substrate related features cluster to each of the groups: Substrates of group 1 are small bacteriocins without RTX domains but with an N-terminal leader peptide, which is cleaved prior to secretion and they are likely transported by the 'alternating access mechanism'. Substrates of group 2 and 3 have C-terminal secretion signals that are not cleaved and mostly display RTX domains, with one exception in group 3 (HasA). Substrates of group 2 tend to outweigh substrates of group 3 in size. Furthermore during this work it became apparent that the organization of the CD of MFPs of group 2 is unique to this group (chapter 3.5). The CDs of MFPs from group 1 and 3 are smaller (14-20 residues compared to 50-60 residues for group 2) and they do not display an AH, a charged cluster nor the conserved motif (FLP-(A/S)-xLxLx-(E/Q)).

Domain specific alignments revealed that the NBDs as well as the TMDs among each group display unexpectedly high sequence identities as well (chapter 3.5). NBDs hold several conserved motifs for binding and hydrolyzing ATP which make up ~20 % of the NBD. However, the lowest sequence identity among group 2 NBDs was still 55 %. For the TMDs no conserved motifs have been described so far but the lowest sequence identity among the group 2 TMDs was 45 %. Such high conservation explains why several heterologous substrates of group 2 can be transported by the HlyBD-TolC system (Table 1). Interestingly, HlyBD-TolC is able to transport LktA from *Mannheimia haemolytica* (Highlander et al., 1990) but LktB-HlyD-TolC (LktB was referred to as HlyB* from Mh) is not able to transport HlyA (chapter 3.5). Considering that HlyB can recognize LktA, which displays the same features as HlyA (Table 1), it would be expected that LktB can also recognize HlyA.

Additionally, LktB (HlyB* from Mh) shows the highest sequence identity among the analyzed homologs compared to HlyB in all domains (CLD: 60 %, TMD: 88 %, NBD: 86 %).

The secretion approach with chimeric ABC transporters further underlines the importance of the NBD and TMD for the secretion process (chapter 3.5). In these set ups, the substrate HlyA met the CD of HlyD, CLD of HlyB and NBD of a homolog in the cytoplasm, where the substrate recognition takes place. The NBDs of HlyB* from Ap and Kk, which were used for the chimeric approach, display sequence identities of 71 % and 72 % respectively. Furthermore, when their structures are modeled, they also display the identified pbp's, both pbp-in and pbp-out. All these similarities make it likely that recognition of HlyA in the cytoplasm was possible but secretion still failed. After substrate recognition the signal has to be transmitted in order to recruit TolC. For this a correctly assembled IM complex is necessary (Thanabalu et al., 1998). Unfortunately it is still unclear, which features are involved in assembling the IM complex of a T1SS. The information gained from assembled structures of tripartite efflux pumps are not helpful in this context, as the contact between an RND-type transporter and the respective MFP is facilitated by periplasmic domains (see 1.2.1), which are not present in an ABC transporter. Nevertheless, upon signal transmission, HlyD would recruit TolC and the actual secretion process would start during which HlyA would have to pass the TMDs of the chimeric ABC transporter before engaging with its native secretion channel (made up of HlyD and TolC). Taken together, the secretion of HlyA by chimeric ABC transporters likely failed for one of two reasons: Either the IM complex could not be assembled correctly, which would have let to the inability to recruit TolC, or HlyA got stuck in the TMDs of the chimeric ABC transporter. A reliable assay regarding the assembly of the HlyA T1SS could help to answer this question.

4.4   Summary

The T1SS is frequently described as one of the most simple secretion systems and was among the first identified systems as well. Still, after 40 years of research on this system by multiple research groups many questions remain unanswered and new ones have emerged.

This work had the ambitious aim to complement the already available structural information by performing NMR studies on the CD of HlyD and crystallization studies on homologs of HlyB. Unfortunately, the purification method for HlyD CD did not meet the demands of structural approaches and the homologs of HlyB that were submitted to crystallization trials did not yield any crystals yet (chapter 3.5). However, smaller structural features, such as an AH in the secretion signal of HlyA and the CD of HlyD, could be identified by combining the results of *in silico* methods with (published) experimental data (chapter 3.3). The now confirmed presence of an AH in the secretion signal of HlyA together with the known importance of this region (Holland et al., 2016) led to the identification of two pbp's in the NBDs of HlyB (chapter 3.3). This raised the question of how many HlyA molecules interact with the IM complex at the same time. Analyzing the published literature on this topic showed that more HlyA leads to a more efficient secretion (Khosa et al., 2018) and higher ATPase activity (Reimann et al., 2016) implying that more than one HlyA molecule binds to HlyB at the same time. A plausible consequence of this is that the HlyA T1SS channel stays assembled as long as substrate is available, which adds a layer of complexity to this 'simple' system. This complexity is further underlined by the secretion approaches with homologs of HlyB and chimeric ABC transporters (chapter 3.5), which showed that either the NBDs or TMDs of the ABC transporter are crucial in forming a functional IM complex. Continuous research on each group of T1SSs will help to answer the remaining question and the here identified motifs in HlyD (chapter 3.5) can help future researches to sort their specific T1SS into these groups.

5. References

Akeda, Y. and Galán, J. E. (2005) 'Chaperone release and unfolding of substrates in type III secretion', *Nature,* 437(7060), pp. 911-5.

Akita, M., Sasaki, S., Matsuyama, S. and Mizushima, S. (1990) 'SecA interacts with secretory proteins by recognizing the positive charge at the amino terminus of the signal peptide in *Escherichia coli*', *J Biol Chem,* 265(14), pp. 8164-9.

Alav, I., Kobylka, J., Kuth, M. S., Pos, K. M., Picard, M., Blair, J. M. A. and Bavro, V. N. (2021) 'Structure, Assembly, and Function of Tripartite Efflux and Type 1 Secretion Systems in Gram-Negative Bacteria', *Chemical Reviews.*

Alegria, M. C., Souza, D. P., Andrade, M. O., Docena, C., Khater, L., Ramos, C. H., da Silva, A. C. and Farah, C. S. (2005) 'Identification of new protein-protein interactions involving the products of the chromosome- and plasmid-encoded type IV secretion loci of the phytopathogen *Xanthomonas axonopodis pv. citri*', *J Bacteriol,* 187(7), pp. 2315-25.

Anes, J., McCusker, M. P., Fanning, S. and Martins, M. (2015) 'The ins and outs of RND efflux pumps in *Escherichia coli*', *Front Microbiol,* 6, pp. 587.

Bachmann, J., Bauer, B., Zwicker, K., Ludwig, B. and Anderka, O. (2006) 'The Rieske protein from *Paracoccus denitrificans* is inserted into the cytoplasmic membrane by the twin-arginine translocase', *Febs j,* 273(21), pp. 4817-30.

Bakkes, P. J., Jenewein, S., Smits, S. H., Holland, I. B. and Schmitt, L. (2010a) 'The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin type 1 secretion system', *J Biol Chem,* 285(52), pp. 40573-80.

Bakkes, P. J., Jenewein, S., Smits, S. H. J., Holland, I. B. and Schmitt, L. (2010b) 'The Rate of Folding Dictates Substrate Secretion by the *Escherichia coli* Hemolysin Type 1 Secretion System*', *Journal of Biological Chemistry,* 285(52), pp. 40573-40580.

Balakrishnan, L., Hughes, C. and Koronakis, V. (2001) 'Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*', *Journal of Molecular Biology,* 313(3), pp. 501-510.

Baron, C. (2010) 'Antivirulence drugs to target bacterial secretion systems', *Current Opinion in Microbiology,* 13(1), pp. 100-105.

Basler, M., Pilhofer, M., Henderson, G. P., Jensen, G. J. and Mekalanos, J. J. (2012) 'Type VI secretion requires a dynamic contractile phage tail-like structure', *Nature,* 483(7388), pp. 182-186.

References

Baumann, U., Wu, S., Flaherty, K. M. and McKay, D. B. (1993) 'Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif', *The EMBO Journal,* 12(9), pp. 3357-3364.

Bavro, V. N., Pietras, Z., Furnham, N., Pérez-Cano, L., Fernández-Recio, J., Pei, X. Y., Misra, R. and Luisi, B. (2008) 'Assembly and channel opening in a bacterial drug efflux machine', *Mol Cell,* 30(1), pp. 114-21.

Beer, T. (2020) *Isolation and cellular characterization of the hemolysin A type I secretion system from Escherichia coli.* PhD Doctoral Dissertation, Heinrich-Heine-University Duesseldorf.

Beis, K. and Rebuffat, S. (2019) 'Multifaceted ABC transporters associated to microcin and bacteriocin export', *Research in Microbiology,* 170(8), pp. 399-406.

Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B. and Schmitt, L. (2003) 'A Specific Interaction Between the NBD of the ABC-transporter HlyB and a C-Terminal Fragment of its Transport Substrate Haemolysin A', *Journal of Molecular Biology,* 327(5), pp. 1169-1179.

Bennion, D., Charlson, E. S., Coon, E. and Misra, R. (2010) 'Dissection of β-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of *Escherichia coli*', *Mol Microbiol,* 77(5), pp. 1153-71.

Berks, B. C. (1996) 'A common export pathway for proteins binding complex redox cofactors?', *Mol Microbiol,* 22(3), pp. 393-404.

Berks, B. C., Sargent, F. and Palmer, T. (2000) 'The Tat protein export pathway', *Molecular Microbiology,* 35(2), pp. 260-274.

Bhakdi, S., Mackman, N., Nicaud, J. M. and Holland, I. B. (1986) '*Escherichia coli* hemolysin may damage target cell membranes by generating transmembrane pores', *Infect Immun,* 52(1), pp. 63-9.

Boudaher, E. and Shaffer, C. L. (2019) 'Inhibiting bacterial secretion systems in the fight against antibiotic resistance', *MedChemComm,* 10(5), pp. 682-692.

Bountra, K., Hagelueken, G., Choudhury, H. G., Corradi, V., El Omari, K., Wagner, A., Mathavan, I., Zirah, S., Yuan Wahlgren, W., Tieleman, D. P., Schiemann, O., Rebuffat, S. and Beis, K. (2017) 'Structural basis for antibacterial peptide self-immunity by the bacterial ABC transporter McjD', *Embo j,* 36(20), pp. 3062-3079.

Boyd, C. D., Smith, T. J., El-Kirat-Chatel, S., Newell, P. D., Dufrêne, Y. F. and O'Toole, G. A. (2014) 'Structural Features of the *Pseudomonas fluorescens* Biofilm Adhesin LapA Required for LapG-Dependent Cleavage, Biofilm Formation, and Cell Surface Localization', *Journal of Bacteriology,* 196(15), pp. 2775-2788.

Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y. and Attree, I. (2009) 'Dissecting the bacterial type VI secretion system by a genome wide *in silico* analysis: what can be learned from available microbial genomic resources?', *BMC Genomics,* 10, pp. 104.

Brunet, Y. R., Zoued, A., Boyer, F., Douzi, B. and Cascales, E. (2015) 'The Type VI Secretion TssEFGK-VgrG Phage-Like Baseplate Is Recruited to the TssJLM Membrane Complex via Multiple Contacts and Serves As Assembly Platform for Tail Tube/Sheath Polymerization', *PLOS Genetics,* 11(10), pp. e1005545.

Bumba, L., Masin, J., Macek, P., Wald, T., Motlova, L., Bibova, I., Klimova, N., Bednarova, L., Veverka, V., Kachala, M., Svergun, Dmitri I., Barinka, C. and Sebo, P. (2016) 'Calcium-Driven Folding of RTX Domain β-Rolls Ratchets Translocation of RTX Proteins through Type I Secretion Ducts', *Molecular Cell,* 62(1), pp. 47-62.

Busch, A. and Waksman, G. (2012) 'Chaperone-usher pathways: diversity and pilus assembly mechanism', *Philos Trans R Soc Lond B Biol Sci,* 367(1592), pp. 1112-22.

Byrd, D. R. and Matson, S. W. (1997) 'Nicking by transesterification: the reaction catalysed by a relaxase', *Mol Microbiol,* 25(6), pp. 1011-22.

Cescau, S., Debarbieux, L. and Wandersman, C. (2007) 'Probing the in vivo dynamics of type I protein secretion complex association through sensitivity to detergents', *J Bacteriol,* 189(5), pp. 1496-504.

Chatzi, K. E., Sardis, M. F., Economou, A. and Karamanou, S. (2014) 'SecA-mediated targeting and translocation of secretory proteins', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1843(8), pp. 1466-1474.

Cherak, S. J. and Turner, R. J. (2017) 'Assembly pathway of a bacterial complex iron sulfur molybdoenzyme', *Biomolecular Concepts,* 8(3-4), pp. 155-167.

Chernyatina, A. A. and Low, H. H. (2019) 'Core architecture of a bacterial type II secretion system', *Nat Commun,* 10(1), pp. 5437.

Chetrit, D., Hu, B., Christie, P. J., Roy, C. R. and Liu, J. (2018) 'A unique cytoplasmic ATPase complex defines the *Legionella pneumophila* type IV secretion channel', *Nature Microbiology,* 3(6), pp. 678-686.

Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J. and Knight, S. D. (1999) 'X-ray Structure of the FimC-FimH Chaperone-Adhesin Complex from Uropathogenic *Escherichia coli*', *Science,* 285(5430), pp. 1061-1066.

Choudhury, H. G., Tong, Z., Mathavan, I., Li, Y., Iwata, S., Zirah, S., Rebuffat, S., van Veen, H. W. and Beis, K. (2014) 'Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state', *Proc Natl Acad Sci U S A,* 111(25), pp. 9145-50.

Clarke, M., Maddera, L., Harris, R. L. and Silverman, P. M. (2008) 'F-pili dynamics by live-cell imaging', *Proceedings of the National Academy of Sciences,* 105(46), pp. 17978-17981.

Cline, K. and Mori, H. (2001) 'Thylakoid DeltapH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport', *J Cell Biol,* 154(4), pp. 719-29.

Cohen, L. J., Han, S., Huang, Y. H. and Brady, S. F. (2018) 'Identification of the Colicin V Bacteriocin Gene Cluster by Functional Screening of a Human Microbiome Metagenomic Library', *ACS Infect Dis,* 4(1), pp. 27-32.

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M. and Waksman, G. (2015) 'Secretion systems in Gram-negative bacteria: structural and mechanistic insights', *Nature Reviews Microbiology,* 13(6), pp. 343-359.

Costa, T. R. D., Harb, L., Khara, P., Zeng, L., Hu, B. and Christie, P. J. (2020) 'Type IV secretion systems: Advances in structure, function, and activation', *Mol Microbiol.*

Crepin, V. F., Shaw, R., Abe, C. M., Knutton, S. and Frankel, G. (2005) 'Polarity of enteropathogenic *Escherichia coli* EspA filament assembly and protein secretion', *J Bacteriol,* 187(8), pp. 2881-9.

Dabney-Smith, C., Mori, H. and Cline, K. (2006) 'Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport', *J Biol Chem,* 281(9), pp. 5476-83.

Dalbey, R. E., Wang, P. and van Dijl, J. M. (2012) 'Membrane Proteases in the Bacterial Protein Secretion and Quality Control Pathway', *Microbiology and Molecular Biology Reviews,* 76(2), pp. 311-330.

De Buck, E., Lammertyn, E. and Anné, J. (2008) 'The importance of the twin-arginine translocation pathway for bacterial virulence', *Trends Microbiol,* 16(9), pp. 442-53.

De Buck, E., Vranckx, L., Meyen, E., Maes, L., Vandersmissen, L., Anné, J. and Lammertyn, E. (2007) 'The twin-arginine translocation pathway is necessary for correct membrane insertion of the Rieske Fe/S protein in *Legionella pneumophila*', *FEBS Lett,* 581(2), pp. 259-64.

Debarbieux, L. and Wandersman, C. (2001) 'Folded HasA inhibits its own secretion through its ABC exporter', *Embo j,* 20(17), pp. 4657-63.

Dehio, C. (2008) 'Infection-associated type IV secretion systems of *Bartonella* and their diverse roles in host cell interaction', *Cellular Microbiology,* 10(8), pp. 1591-1598.

Delepelaire, P. and Wandersman, C. (1990) 'Protein secretion in gram-negative bacteria. The extracellular metalloprotease B from *Erwinia chrysanthemi* contains a C-terminal secretion signal analogous to that of *Escherichia coli* alpha-hemolysin', *J Biol Chem,* 265(28), pp. 17118-25.

Delepelaire, P. and Wandersman, C. (1998) 'The SecB chaperone is involved in the secretion of the *Serratia marcescens* HasA protein through an ABC transporter', *Embo j,* 17(4), pp. 936-44.

Delgado, M. A., Solbiati, J. O., Chiuchiolo, M. J., Farías, R. N. and Salomón, R. A. (1999) '*Escherichia coli* Outer Membrane Protein TolC Is Involved in Production of the Peptide Antibiotic Microcin J25', *Journal of Bacteriology,* 181(6), pp. 1968-1970.

Deng, W., Marshall, N. C., Rowland, J. L., McCoy, J. M., Worrall, L. J., Santos, A. S., Strynadka, N. C. J. and Finlay, B. B. (2017) 'Assembly, structure, function and regulation of type III secretion systems', *Nature Reviews Microbiology,* 15(6), pp. 323-337.

Denks, K., Vogt, A., Sachelaru, I., Petriman, N.-A., Kudva, R. and Koch, H.-G. (2014) 'The Sec translocon mediated protein transport in prokaryotes and eukaryotes', *Molecular Membrane Biology,* 31(2-3), pp. 58-84.

Dodson, K. W., Jacob-Dubuisson, F., Striker, R. T. and Hultgren, S. J. (1993) 'Outer-membrane PapC molecular usher discriminately recognizes periplasmic

chaperone-pilus subunit complexes', *Proceedings of the National Academy of Sciences,* 90(8), pp. 3670-3674.

Douzi, B., Ball, G., Cambillau, C., Tegoni, M. and Voulhoux, R. (2011) 'Deciphering the Xcp *Pseudomonas aeruginosa* type II secretion machinery through multiple interactions with substrates', *J Biol Chem,* 286(47), pp. 40792-801.

Draper, O., César, C. E., Machón, C., de la Cruz, F. and Llosa, M. (2005) 'Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells', *Proc Natl Acad Sci U S A,* 102(45), pp. 16385-90.

Du, D., van Veen, H. W., Murakami, S., Pos, K. M. and Luisi, B. F. (2015) 'Structure, mechanism and cooperation of bacterial multidrug transporters', *Current Opinion in Structural Biology,* 33, pp. 76-91.

Du, D., Wang, Z., James, N. R., Voss, J. E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W. and Luisi, B. F. (2014) 'Structure of the AcrAB-TolC multidrug efflux pump', *Nature,* 509(7501), pp. 512-5.

Duquesne, S., Destoumieux-Garzón, D., Peduzzi, J. and Rebuffat, S. (2007) 'Microcins, gene-encoded antibacterial peptides from enterobacteria', *Nat Prod Rep,* 24(4), pp. 708-34.

Durand, E., Nguyen, V. S., Zoued, A., Logger, L., Péhau-Arnaudet, G., Aschtgen, M.-S., Spinelli, S., Desmyter, A., Bardiaux, B., Dujeancourt, A., Roussel, A., Cambillau, C., Cascales, E. and Fronzes, R. (2015) 'Biogenesis and structure of a type VI secretion membrane core complex', *Nature,* 523(7562), pp. 555-560.

Eisenbrandt, R., Kalkum, M., Lai, E. M., Lurz, R., Kado, C. I. and Lanka, E. (1999) 'Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits', *J Biol Chem,* 274(32), pp. 22548-55.

Erenburg, I. N. (2020) *Funtional and structural characterization of the RTX proteins MbxA from Moraxella bovis and FrpA from Kingella kingae.* PhD Dissertation, Heinrich-Heine-University Duesseldorf.

Erhardt, M., Mertens, M. E., Fabiani, F. D. and Hughes, K. T. (2014) 'ATPase-independent type-III protein secretion in *Salmonella enterica*', *PLoS Genet,* 10(11), pp. e1004800.

Eswaran, J., Koronakis, E., Higgins, M. K., Hughes, C. and Koronakis, V. (2004) 'Three's company: component structures bring a closer view of tripartite drug efflux pumps', *Current Opinion in Structural Biology,* 14(6), pp. 741-747.

Fath, M. J., Skvirsky, R., Gilson, L., Mahanty, H. K. and Kolter, R. 'The Secretion of Colicin V'. *Bacteriocins, Microcins and Lantibiotics.* Berlin, Heidelberg: Springer Berlin Heidelberg, 331-348.

Flaugnatti, N., Le, T. T., Canaan, S., Aschtgen, M. S., Nguyen, V. S., Blangy, S., Kellenberger, C., Roussel, A., Cambillau, C., Cascales, E. and Journet, L. (2016) 'A phospholipase A1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery', *Mol Microbiol,* 99(6), pp. 1099-118.

Fralick, J. A. (1996) 'Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of *Escherichia coli*', *Journal of Bacteriology,* 178(19), pp. 5803-5805.

Fridman, C. M., Keppel, K., Gerlic, M., Bosis, E. and Salomon, D. (2020) 'A comparative genomics methodology reveals a widespread family of membrane-disrupting T6SS effectors', *Nature Communications,* 11(1), pp. 1085.

Gangola, P. and Rosen, B. P. (1987) 'Maintenance of intracellular calcium in *Escherichia coli*', *J Biol Chem,* 262(26), pp. 12570-4.

Gentle, I. E., Burri, L. and Lithgow, T. (2005) 'Molecular architecture and function of the Omp85 family of proteins', *Molecular Microbiology,* 58(5), pp. 1216-1225.

Gentschev, I. and Goebel, W. (1992) 'Topological and functional studies on HlyB of *Escherichia coli*', *Molecular and General Genetics MGG,* 232(1), pp. 40-48.

Ghigo, J. M. and Wandersman, C. (1992) 'A fourth metalloprotease gene in *Erwinia chrysanthemi*', *Res Microbiol,* 143(9), pp. 857-67.

Ginalski, K., Kinch, L., Rychlewski, L. and Grishin, N. V. (2004) 'BTLCP proteins: a novel family of bacterial transglutaminase-like cysteine proteinases', *Trends in Biochemical Sciences,* 29(8), pp. 392-395.

Gohlke, U., Pullan, L., McDevitt, C. A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H. R. and Berks, B. C. (2005) 'The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter', *Proceedings of the National Academy of Sciences of the United States of America,* 102(30), pp. 10482-10486.

Gray, L., Mackman, N., Nicaud, J. M. and Holland, I. B. (1986) 'The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*', *Mol Gen Genet,* 205(1), pp. 127-33.

Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N. G., Stansfeld, P. J., Wang, Z., Zhang, Y., Wang, W. and Dong, C. (2016) 'Structural basis of outer membrane protein insertion by the BAM complex', *Nature,* 531(7592), pp. 64-69.

Guo, S., Stevens, C. A., Vance, T. D. R., Olijve, L. L. C., Graham, L. A., Campbell, R. L., Yazdi, S. R., Escobedo, C., Bar-Dolev, M., Yashunsky, V., Braslavsky, I., Langelaan, D. N., Smith, S. P., Allingham, J. S., Voets, I. K. and Davies, P. L. (2017) 'Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice', *Science Advances,* 3(8), pp. e1701440.

Gygi, D., Nicolet, J., Frey, J., Cross, M., Koronakis, V. and Hughes, C. (1990) 'Isolation of the *Actinobacillus pleuropneumoniae* haemolysin gene and the activation and secretion of the prohaemolysin by the HlyC, HlyB and HlyD proteins of Escherichia coli', *Mol Microbiol,* 4(1), pp. 123-8.

Hartmann, M. D., Ridderbusch, O., Zeth, K., Albrecht, R., Testa, O., Woolfson, D. N., Sauer, G., Dunin-Horkawicz, S., Lupas, A. N. and Alvarez, B. H. (2009) 'A coiled-coil motif that sequesters ions to the hydrophobic core', *Proc Natl Acad Sci U S A,* 106(40), pp. 16950-5.

Håvarstein, L. S., Holo, H. and Nes, I. F. (1994) 'The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria', *Microbiology (Reading),* 140 (Pt 9), pp. 2383-9.

Hay, I. D., Belousoff, M. J. and Lithgow, T. (2017) 'Structural Basis of Type 2 Secretion System Engagement between the Inner and Outer Bacterial Membranes', *mBio,* 8(5).

Hepp, C. and Maier, B. (2017) 'Bacterial Translocation Ratchets: Shared Physical Principles with Different Molecular Implementations: How bacterial secretion systems bias Brownian motion for efficient translocation of macromolecules', *Bioessays,* 39(10).

Hess, J., Gentschev, I., Goebel, W. and Jarchau, T. (1990) 'Analysis of the haemolysin secretion system by PhoA-HlyA fusion proteins', *Mol Gen Genet,* 224(2), pp. 201-8.

Highlander, S. K., Engler, M. J. and Weinstock, G. M. (1990) 'Secretion and expression of the *Pasteurella* haemolytica Leukotoxin', *J Bacteriol,* 172(5), pp. 2343-50.

Hinchliffe, P., Greene, N. P., Paterson, N. G., Crow, A., Hughes, C. and Koronakis, V. (2014) 'Structure of the periplasmic adaptor protein from a major

facilitator superfamily (MFS) multidrug efflux pump', *FEBS Letters,* 588(17), pp. 3147-3153.

Hinsa, S. M., Espinosa-Urgel, M., Ramos, J. L. and O'Toole, G. A. (2003) 'Transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein', *Mol Microbiol,* 49(4), pp. 905-18.

Hirst, T. R. and Holmgren, J. (1987) 'Conformation of protein secreted across bacterial outer membranes: a study of enterotoxin translocation from *Vibrio cholerae*', *Proc Natl Acad Sci U S A,* 84(21), pp. 7418-22.

Ho, B. T., Dong, T. G. and Mekalanos, J. J. (2014) 'A view to a kill: the bacterial type VI secretion system', *Cell Host Microbe,* 15(1), pp. 9-21.

Hobbs, E. C., Yin, X., Paul, B. J., Astarita, J. L. and Storz, G. (2012) 'Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance', *Proceedings of the National Academy of Sciences,* 109(41), pp. 16696-16701.

Holland, I. B., Peherstorfer, S., Kanonenberg, K., Lenders, M., Reimann, S. and Schmitt, L. (2016) 'Type I Protein Secretion-Deceptively Simple yet with a Wide Range of Mechanistic Variability across the Family', *EcoSal Plus,* 7(1).

Holtkamp, W., Lee, S., Bornemann, T., Senyushkina, T., Rodnina, M. V. and Wintermeyer, W. (2012) 'Dynamic switch of the signal recognition particle from scanning to targeting', *Nat Struct Mol Biol,* 19(12), pp. 1332-7.

Hu, B., Khara, P. and Christie, P. J. (2019a) 'Structural bases for F plasmid conjugation and F pilus biogenesis in *Escherichia coli*', *Proceedings of the National Academy of Sciences,* 116(28), pp. 14222-14227.

Hu, B., Khara, P., Song, L., Lin, A. S., Frick-Cheng, A. E., Harvey, M. L., Cover, T. L. and Christie, P. J. (2019b) '*In Situ* Molecular Architecture of the *Helicobacter pylori* Cag Type IV Secretion System', *mBio,* 10(3), pp. e00849-19.

Huang, L., Boyd, D., Amyot, W. M., Hempstead, A. D., Luo, Z. Q., O'Connor, T. J., Chen, C., Machner, M., Montminy, T. and Isberg, R. R. (2011) 'The E Block motif is associated with *Legionella pneumophila* translocated substrates', *Cell Microbiol,* 13(2), pp. 227-45.

Hubber, A., Vergunst, A. C., Sullivan, J. T., Hooykaas, P. J. J. and Ronson, C. W. (2004) 'Symbiotic phenotypes and translocated effector proteins of the *Mesorhizobium loti* strain R7A VirB/D4 type IV secretion system', *Molecular Microbiology,* 54(2), pp. 561-574.

Huber, D., Rajagopalan, N., Preissler, S., Rocco, M. A., Merz, F., Kramer, G. and Bukau, B. (2011) 'SecA Interacts with Ribosomes in Order to Facilitate Posttranslational Translocation in Bacteria', *Molecular Cell,* 41(3), pp. 343-353.

Huddleston, J. R. (2014) 'Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes', *Infect Drug Resist,* 7, pp. 167-76.

Hwang, J., Zhong, X. and Tai, P. C. (1997) 'Interactions of dedicated export membrane proteins of the colicin V secretion system: CvaA, a member of the membrane fusion protein family, interacts with CvaB and TolC', *Journal of Bacteriology,* 179(20), pp. 6264-6270.

Ieva, R. and Bernstein, H. D. (2009) 'Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane', *Proceedings of the National Academy of Sciences,* 106(45), pp. 19120-19125.

Ieva, R., Tian, P., Peterson, J. H. and Bernstein, H. D. (2011) 'Sequential and spatially restricted interactions of assembly factors with an autotransporter β domain', *Proceedings of the National Academy of Sciences,* 108(31), pp. E383-E391.

Ikonomidis, A., Tsakris, A., Kanellopoulou, M., Maniatis, A. N. and Pournaras, S. (2008) 'Effect of the proton motive force inhibitor carbonyl cyanide-m-chlorophenylhydrazone (CCCP) on *Pseudomonas aeruginosa* biofilm development', *Lett Appl Microbiol,* 47(4), pp. 298-302.

Isaac, D. T. and Isberg, R. (2014) 'Master manipulators: an update on *Legionella pneumophila* Icm/Dot translocated substrates and their host targets', *Future Microbiol,* 9(3), pp. 343-59.

Jarchau, T., Chakraborty, T., Garcia, F. and Goebel, W. (1994) 'Selection for transport competence of C-terminal polypeptides derived from *Escherichia coli* hemolysin: the shortest peptide capable of autonomous HIyB/HIyD-dependent secretion comprises the C-terminal 62 amino acids of HlyA', *Molecular and General Genetics MGG,* 245(1), pp. 53-60.

Jin, F. (2020) 'Structural insights into the mechanism of a novel protein targeting pathway in Gram-negative bacteria', *FEBS Open Bio,* 10(4), pp. 561-579.

Jo, I., Kim, J. S., Xu, Y., Hyun, J., Lee, K. and Ha, N. C. (2019) 'Recent paradigm shift in the assembly of bacterial tripartite efflux pumps and the type I secretion system', *J Microbiol,* 57(3), pp. 185-194.

Jones, H. E., Holland, I. B., Baker, H. L. and Campbell, A. K. (1999) 'Slow changes in cytosolic free Ca2+ in *Escherichia coli* highlight two putative influx mechanisms in response to changes in extracellular calcium', *Cell Calcium,* 25(3), pp. 265-274.

Jose, J., Jähnig, F. and Meyer, T. F. (1995) 'Common structural features of IgA1 protease-like outer membrane protein autotransporters', *Mol Microbiol,* 18(2), pp. 378-80.

Junker, M., Besingi, R. N. and Clark, P. L. (2009) 'Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion', *Mol Microbiol,* 71(5), pp. 1323-32.

Kanonenberg, K., Schwarz, C. K. and Schmitt, L. (2013) 'Type I secretion systems - a story of appendices', *Res Microbiol,* 164(6), pp. 596-604.

Kanonenberg, K., Smits, S. H. J. and Schmitt, L. (2019) 'Functional Reconstitution of HlyB, a Type I Secretion ABC Transporter, in Saposin-A Nanoparticles', *Scientific Reports,* 9(1), pp. 8436.

Kanonenberg, K., Spitz, O., Erenburg, I. N., Beer, T. and Schmitt, L. (2018) 'Type I secretion system-it takes three and a substrate', *FEMS Microbiol Lett,* 365(11).

Kapitein, N., Bönemann, G., Pietrosiuk, A., Seyffer, F., Hausser, I., Locker, J. K. and Mogk, A. (2013) 'ClpV recycles VipA/VipB tubules and prevents non-productive tubule formation to ensure efficient type VI protein secretion', *Molecular Microbiology,* 87(5), pp. 1013-1028.

Karst, J. C., Sotomayor Pérez, A. C., Guijarro, J. I., Raynal, B., Chenal, A. and Ladant, D. (2010) 'Calmodulin-Induced Conformational and Hydrodynamic Changes in the Catalytic Domain of *Bordetella pertussis* Adenylate Cyclase Toxin', *Biochemistry,* 49(2), pp. 318-328.

Kater, L., Frieg, B., Berninghausen, O., Gohlke, H., Beckmann, R. and Kedrov, A. (2019) 'Partially inserted nascent chain unzips the lateral gate of the Sec translocon', *EMBO reports,* 20(10), pp. e48191.

Kenny, B., Chervaux, C. and Holland, I. B. (1994) 'Evidence that residues −15 to −46 of the haemolysin secretion signal are involved in early steps in secretion, leading to recognition of the translocator', *Molecular Microbiology,* 11(1), pp. 99-109.

Kenny, B., Haigh, R. and Holland, I. B. (1991) 'Analysis of the haemolysin transport process through the secretion from *Escherichia coli* of PCM, CAT or beta-

galactosidase fused to the Hly C-terminal signal domain', *Mol Microbiol,* 5(10), pp. 2557-68.

Khosa, S., Scholz, R., Schwarz, C., Trilling, M., Hengel, H., Jaeger, K. E., Smits, S. H. J. and Schmitt, L. (2018) 'An A/U-Rich Enhancer Region Is Required for High-Level Protein Secretion through the HlyA Type I Secretion System', *Appl Environ Microbiol,* 84(1).

Kieuvongngam, V., Olinares, P. D. B., Palillo, A., Oldham, M. L., Chait, B. T. and Chen, J. (2020) 'Structural basis of substrate recognition by a polypeptide processing and secretion transporter', *Elife,* 9.

Kihara, A., Akiyama, Y. and Ito, K. (1995) 'FtsH is required for proteolytic elimination of uncomplexed forms of SecY, an essential protein translocase subunit', *Proceedings of the National Academy of Sciences,* 92(10), pp. 4532-4536.

Kim, J. S., Song, S., Lee, M., Lee, S., Lee, K. and Ha, N. C. (2016) 'Crystal Structure of a Soluble Fragment of the Membrane Fusion Protein HlyD in a Type I Secretion System of Gram-Negative Bacteria', *Structure,* 24(3), pp. 477-85.

Kim, K. H., Aulakh, S. and Paetzel, M. (2012) 'The bacterial outer membrane β-barrel assembly machinery', *Protein Science,* 21(6), pp. 751-768.

Kim, Y. R., Lee, S. E., Kook, H., Yeom, J. A., Na, H. S., Kim, S. Y., Chung, S. S., Choy, H. E. and Rhee, J. H. (2008) '*Vibrio vulnificus* RTX toxin kills host cells only after contact of the bacteria with host cells', *Cell Microbiol,* 10(4), pp. 848-62.

Kinch, L. N., Saier, J. M. H. and Grishin, N. V. (2002) 'Sec61β – a component of the archaeal protein secretory system', *Trends in Biochemical Sciences,* 27(4), pp. 170-171.

Kingsman, A. and Willetts, N. (1978) 'The requirements for conjugal DNA synthesis in the donor strain during Flac transfer', *Journal of Molecular Biology,* 122(3), pp. 287-300.

Koch, S., Seinen, A.-B., Kamel, M., Kuckla, D., Monzel, C., Kedrov, A. and Driessen, A. J. M. (2021) 'Single-molecule analysis of dynamics and interactions of the SecYEG translocon', *The FEBS Journal,* 288(7), pp. 2203-2221.

Koronakis, V., Cross, M. and Hughes, C. (1989) 'Transcription antitermination in an *Escherichia coli* haemolysin operon is directed progressively by *cis*-acting DNA

sequences upstream of the promoter region', *Molecular Microbiology,* 3(10), pp. 1397-1404.

Koronakis, V., Hughes, C. and Koronakis, E. (1991) 'Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes', *Embo j,* 10(11), pp. 3263-72.

Koronakis, V., Li, J., Koronakis, E. and Stauffer, K. (1997) 'Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals', *Mol Microbiol,* 23(3), pp. 617-26.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B. and Hughes, C. (2000) 'Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export', *Nature,* 405(6789), pp. 914-9.

Krampen, L., Malmsheimer, S., Grin, I., Trunk, T., Lührmann, A., de Gier, J.-W. and Wagner, S. (2018) 'Revealing the mechanisms of membrane protein export by virulence-associated bacterial secretion systems', *Nature Communications,* 9(1), pp. 3467.

Kubori, T., Sukhan, A., Aizawa, S. I. and Galán, J. E. (2000) 'Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system', *Proc Natl Acad Sci U S A,* 97(18), pp. 10225-30.

Kudryashova, E., Heisler, D., Zywiec, A. and Kudryashov, D. S. (2014) 'Thermodynamic properties of the effector domains of MARTX toxins suggest their unfolding for translocation across the host membrane', *Mol Microbiol,* 92(5), pp. 1056-71.

Kuhnert, P., Heyberger-Meyer, B., Nicolet, J. and Frey, J. (2000) 'Characterization of PaxA and its operon: a cohemolytic RTX toxin determinant from pathogenic *Pasteurella aerogenes*', *Infect Immun,* 68(1), pp. 6-12.

Lara-Tejero, M. and Galán, J. E. (2009) '*Salmonella enterica serovar typhimurium* pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells', *Infect Immun,* 77(7), pp. 2635-42.

Lecher, J., Schwarz, Christian K. W., Stoldt, M., Smits, Sander H. J., Willbold, D. and Schmitt, L. (2012) 'An RTX Transporter Tethers Its Unfolded Substrate during Secretion via a Unique N-Terminal Domain', *Structure,* 20(10), pp. 1778-1787.

Lee, M., Jun, S. Y., Yoon, B. Y., Song, S., Lee, K. and Ha, N. C. (2012) 'Membrane fusion proteins of type I secretion system and tripartite efflux pumps share a binding motif for TolC in gram-negative bacteria', *PLoS One,* 7(7), pp. e40460.

Lee, P.-C., Zmina, S. E., Stopford, C. M., Toska, J. and Rietsch, A. (2014) 'Control of type III secretion activity and substrate specificity by the cytoplasmic regulator PcrG', *Proceedings of the National Academy of Sciences,* 111(19), pp. E2027-E2036.

Leiman, P. G., Basler, M., Ramagopal, U. A., Bonanno, J. B., Sauder, J. M., Pukatzki, S., Burley, S. K., Almo, S. C. and Mekalanos, J. J. (2009) 'Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin', *Proceedings of the National Academy of Sciences,* 106(11), pp. 4154-4159.

Lenders, M. H., Reimann, S., Smits, S. H. and Schmitt, L. (2013) 'Molecular insights into type I secretion systems', *Biol Chem,* 394(11), pp. 1371-84.

Lenders, M. H. H., Beer, T., Smits, S. H. J. and Schmitt, L. (2016) '*In vivo* quantification of the secretion rates of the hemolysin A Type I secretion system', *Scientific Reports,* 6(1), pp. 33275.

Lenders, M. H. H., Weidtkamp-Peters, S., Kleinschrodt, D., Jaeger, K.-E., Smits, S. H. J. and Schmitt, L. (2015) 'Directionality of substrate translocation of the hemolysin A Type I secretion system', *Scientific Reports,* 5(1), pp. 12470.

Leo, J. C., Grin, I. and Linke, D. (2012) 'Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane', *Philos Trans R Soc Lond B Biol Sci,* 367(1592), pp. 1088-101.

Létoffé, S., Ghigo, J. M. and Wandersman, C. (1993) 'Identification of two components of the *Serratia marcescens* metalloprotease transporter: protease SM secretion in *Escherichia coli* is TolC dependent', *Journal of Bacteriology,* 175(22), pp. 7321-7328.

Létoffé, S., Ghigo, J. M. and Wandersman, C. (1994a) 'Iron acquisition from heme and hemoglobin by a *Serratia marcescens* extracellular protein', *Proc Natl Acad Sci U S A,* 91(21), pp. 9876-80.

Létoffé, S., Ghigo, J. M. and Wandersman, C. (1994b) 'Secretion of the *Serratia marcescens* HasA protein by an ABC transporter', *J Bacteriol,* 176(17), pp. 5372-7.

Li, X. Z. and Poole, K. (1999) 'Organic solvent-tolerant mutants of *Pseudomonas aeruginosa* display multiple antibiotic resistance', *Can J Microbiol,* 45(1), pp. 18-22.

Linhartová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., Kamanová, J., Procházková, K., Adkins, I., Hejnová-Holubová, J., Sadílková, L., Morová, J. and Šebo, P. (2010) 'RTX proteins: a highly diverse family secreted by a common mechanism', *FEMS Microbiology Reviews,* 34(6), pp. 1076-1112.

Llosa, M. and Alkorta, I. (2017) 'Coupling Proteins in Type IV Secretion', *Curr Top Microbiol Immunol,* 413, pp. 143-168.

Llosa, M., Zunzunegui, S. and de la Cruz, F. (2003) 'Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes', *Proceedings of the National Academy of Sciences,* 100(18), pp. 10465-10470.

Lobedanz, S., Bokma, E., Symmons, M. F., Koronakis, E., Hughes, C. and Koronakis, V. (2007) 'A periplasmic coiled-coil interface underlying TolC recruitment and the assembly of bacterial drug efflux pumps', *Proc Natl Acad Sci U S A,* 104(11), pp. 4612-7.

Locher, K. P. (2016) 'Mechanistic diversity in ATP-binding cassette (ABC) transporters', *Nature Structural & Molecular Biology,* 23(6), pp. 487-493.

Lorenz, C. and Büttner, D. (2009) 'Functional characterization of the type III secretion ATPase HrcN from the plant pathogen *Xanthomonas campestris pv. vesicatoria*', *J Bacteriol,* 191(5), pp. 1414-28.

Low, H. H., Gubellini, F., Rivera-Calzada, A., Braun, N., Connery, S., Dujeancourt, A., Lu, F., Redzej, A., Fronzes, R., Orlova, E. V. and Waksman, G. (2014) 'Structure of a type IV secretion system', *Nature,* 508(7497), pp. 550-553.

Luirink, J. and Sinning, I. (2004) 'SRP-mediated protein targeting: structure and function revisited', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1694(1), pp. 17-35.

Luirink, J., ten Hagen-Jongman, C. M., van der Weijden, C. C., Oudega, B., High, S., Dobberstein, B. and Kusters, R. (1994) 'An alternative protein targeting pathway in *Escherichia coli*: studies on the role of FtsY', *The EMBO Journal,* 13(10), pp. 2289-2296.

Lüke, I., Handford, J. I., Palmer, T. and Sargent, F. (2009) 'Proteolytic processing of *Escherichia coli* twin-arginine signal peptides by LepB', *Arch Microbiol,* 191(12), pp. 919-25.

Lycklama a Nijeholt, J. A. and Driessen, A. J. M. (2012) 'The bacterial Sec-translocase: structure and mechanism', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367(1592), pp. 1016-1028.

Mackman, N., Nicaud, J. M., Gray, L. and Holland, I. B. (1985) 'Identification of polypeptides required for the export of haemolysin 2001 from *E. coli*', *Mol Gen Genet,* 201(3), pp. 529-36.

Marshall, R. L. and Bavro, V. N. (2020) 'Mutations in the TolC Periplasmic Domain Affect Substrate Specificity of the AcrAB-TolC Pump', *Frontiers in Molecular Biosciences,* 7(166).

Masi, M., Duret, G., Delcour, A. H. and Misra, R. (2009) 'Folding and trimerization of signal sequence-less mature TolC in the cytoplasm of *Escherichia coli*', *Microbiology (Reading),* 155(Pt 6), pp. 1847-1857.

Masi, M. and Wandersman, C. (2010) 'Multiple signals direct the assembly and function of a type 1 secretion system', *J Bacteriol,* 192(15), pp. 3861-9.

Masuda, N., Sakagawa, E., Ohya, S., Gotoh, N., Tsujimoto, H. and Nishino, T. (2000) 'Substrate Specificities of MexAB-OprM, MexCD-OprJ, and MexXY-OprM Efflux Pumps in *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy,* 44(12), pp. 3322-3327.

Mehmood, S., Corradi, V., Choudhury, H. G., Hussain, R., Becker, P., Axford, D., Zirah, S., Rebuffat, S., Tieleman, D. P., Robinson, C. V. and Beis, K. (2016) 'Structural and Functional Basis for Lipid Synergy on the Activity of the Antibacterial Peptide ABC Transporter McjD', *J Biol Chem,* 291(41), pp. 21656-21668.

Meier, R., Drepper, T., Svensson, V., Jaeger, K.-E. and Baumann, U. (2007) 'A Calcium-gated Lid and a Large β-Roll Sandwich Are Revealed by the Crystal Structure of Extracellular Lipase from *Serratia marcescens**', *Journal of Biological Chemistry,* 282(43), pp. 31477-31483.

Menestrina, G., Pederzolli, C., Dalla Serra, M., Bregante, M. and Gambale, F. (1996) 'Permeability Increase Induced by *Escherichia coli* Hemolysin A in Human Macrophages is Due to the Formation of Ionic Pores: A Patch Clamp Characterization', *The Journal of Membrane Biology,* 149(2), pp. 113-121.

Meyer, T. H., Ménétret, J.-F., Breitling, R., Miller, K. R., Akey, C. W. and Rapoport, T. A. (1999) 'The bacterial SecY/E translocation complex forms channel-like structures similar to those of the eukaryotic sec61p complex*Edited by W. Baumeister', *Journal of Molecular Biology,* 285(4), pp. 1789-1800.

Mikolosko, J., Bobyk, K., Zgurskaya, H. I. and Ghosh, P. (2006) 'Conformational flexibility in the multidrug efflux system protein AcrA', *Structure,* 14(3), pp. 577-87.

Montagner, C., Arquint, C. and Cornelis, G. R. (2011) 'Translocators YopB and YopD from *Yersinia enterocolitica* Form a Multimeric Integral Membrane Complex in Eukaryotic Cell Membranes', *Journal of Bacteriology,* 193(24), pp. 6923-6928.

Morgan, J. L. W., Acheson, J. F. and Zimmer, J. (2017) 'Structure of a Type-1 Secretion System ABC Transporter', *Structure,* 25(3), pp. 522-529.

Mori, H. and Cline, K. (2002) 'A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid ΔpH/Tat translocase', *J Cell Biol,* 157(2), pp. 205-10.

Mori, H. and Ito, K. (2006) 'Different modes of SecY–SecA interactions revealed by site-directed *in vivo* photo-cross-linking', *Proceedings of the National Academy of Sciences,* 103(44), pp. 16159-16164.

Mougous, J. D., Cuff, M. E., Raunser, S., Shen, A., Zhou, M., Gifford, C. A., Goodman, A. L., Joachimiak, G., Ordoñez, C. L., Lory, S., Walz, T., Joachimiak, A. and Mekalanos, J. J. (2006) 'A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus', *Science,* 312(5779), pp. 1526-30.

Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T. and Yamaguchi, A. (2006) 'Crystal structures of a multidrug transporter reveal a functionally rotating mechanism', *Nature,* 443(7108), pp. 173-179.

Murakami, S., Nakashima, R., Yamashita, E. and Yamaguchi, A. (2002) 'Crystal structure of bacterial multidrug efflux transporter AcrB', *Nature,* 419(6907), pp. 587-93.

Nagai, H., Cambronne, E. D., Kagan, J. C., Amor, J. C., Kahn, R. A. and Roy, C. R. (2005) 'A C-terminal translocation signal required for Dot/Icm-dependent delivery of the *Legionella* RalF protein to host cells', *Proceedings of the National Academy of Sciences of the United States of America,* 102(3), pp. 826-831.

Nakashima, R., Sakurai, K., Yamasaki, S., Nishino, K. and Yamaguchi, A. (2011) 'Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket', *Nature,* 480(7378), pp. 565-569.

Nelson, W. C., Howard, M. T., Sherman, J. A. and Matson, S. W. (1995) 'The traY gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid oriT', *J Biol Chem,* 270(47), pp. 28374-80.

Newell, P. D., Boyd, C. D., Sondermann, H. and O'Toole, G. A. (2011) 'A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage', *PLoS Biol,* 9(2), pp. e1000587.

Newell, P. D., Monds, R. D. and O'Toole, G. A. (2009) 'LapD is a bis-(3′,5′)-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0–1', *Proceedings of the National Academy of Sciences,* 106(9), pp. 3461-3466.

Ng, T. W., Akman, L., Osisami, M. and Thanassi, D. G. (2006) 'The Usher N Terminus Is the Initial Targeting Site for Chaperone-Subunit Complexes and Participates in Subsequent Pilus Biogenesis Events', *Journal of Bacteriology,* 188(6), pp. 2295-2295.

Nguyen, V. S., Douzi, B., Durand, E., Roussel, A., Cascales, E. and Cambillau, C. (2018) 'Towards a complete structural deciphering of Type VI secretion system', *Curr Opin Struct Biol,* 49, pp. 77-84.

Nicaud, J. M., Mackman, N., Gray, L. and Holland, I. B. (1985a) 'Characterisation of HlyC and mechanism of activation and secretion of haemolysin from *E. coli* 2001', *FEBS Lett,* 187(2), pp. 339-44.

Nicaud, J. M., Mackman, N., Gray, L. and Holland, I. B. (1985b) 'Regulation of haemolysin synthesis in *E. coli* determined by HLY genes of human origin', *Mol Gen Genet,* 199(1), pp. 111-6.

Nikaido, H. (1998) 'Antibiotic Resistance Caused by Gram-Negative Multidrug Efflux Pumps', *Clinical Infectious Diseases,* 27(Supplement_1), pp. S32-S41.

Nikaido, H. and Pagès, J. M. (2012) 'Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria', *FEMS Microbiol Rev,* 36(2), pp. 340-63.

Nishiyama, K., Hanada, M. and Tokuda, H. (1994) 'Disruption of the gene encoding p12 (SecG) reveals the direct involvement and important function of SecG in the protein translocation of *Escherichia coli* at low temperature', *The EMBO Journal,* 13(14), pp. 3272-3277.

Nivaskumar, M. and Francetic, O. (2014) 'Type II secretion system: a magic beanstalk or a protein escalator', *Biochim Biophys Acta,* 1843(8), pp. 1568-77.

Noegel, A., Rdest, U., Springer, W. and Goebel, W. (1979) 'Plasmid cistrons controlling synthesis and excretion of the exotoxin alpha-haemolysin of *Escherichia coli*', *Mol Gen Genet,* 175(3), pp. 343-50.

Nuccio, S.-P. and Bäumler, A. J. (2007) 'Evolution of the Chaperone/Usher Assembly Pathway: Fimbrial Classification Goes Greek', *Microbiology and Molecular Biology Reviews,* 71(4), pp. 551-575.

Oliver, D. C., Huang, G., Nodel, E., Pleasance, S. and Fernandez, R. C. (2003) 'A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain', *Mol Microbiol,* 47(5), pp. 1367-83.

Paetzel, M., Karla, A., Strynadka, N. C. J. and Dalbey, R. E. (2002) 'Signal Peptidases', *Chemical Reviews,* 102(12), pp. 4549-4580.

Pagès, J.-M., Masi, M. and Barbe, J. (2005) 'Inhibitors of efflux pumps in Gram-negative bacteria', *Trends in Molecular Medicine,* 11(8), pp. 382-389.

Palmer, T. and Berks, B. C. (2012) 'The twin-arginine translocation (Tat) protein export pathway', *Nat Rev Microbiol,* 10(7), pp. 483-96.

Palomino, C., Marín, E. and Fernández, L. Á. (2011) 'The Fimbrial Usher FimD Follows the SurA-BamB Pathway for Its Assembly in the Outer Membrane of *Escherichia coli*', *Journal of Bacteriology,* 193(19), pp. 5222-5230.

Pattis, I., Weiss, E., Laugks, R., Haas, R. and Fischer, W. (2007) 'The *Helicobacter pylori* CagF protein is a type IV secretion chaperone-like molecule that binds close to the C-terminal secretion signal of the CagA effector protein', *Microbiology,* 153(9), pp. 2896-2909.

Paul, K., Erhardt, M., Hirano, T., Blair, D. F. and Hughes, K. T. (2008) 'Energy source of flagellar type III secretion', *Nature,* 451(7177), pp. 489-492.

Peterson, J. H., Tian, P., Ieva, R., Dautin, N. and Bernstein, H. D. (2010) 'Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment', *Proceedings of the National Academy of Sciences,* 107(41), pp. 17739-17744.

Phan, G., Remaut, H., Wang, T., Allen, W. J., Pirker, K. F., Lebedev, A., Henderson, N. S., Geibel, S., Volkan, E., Yan, J., Kunze, M. B., Pinkner, J. S., Ford, B., Kay, C. W., Li, H., Hultgren, S. J., Thanassi, D. G. and Waksman, G. (2011) 'Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate', *Nature,* 474(7349), pp. 49-53.

Pimenta, A. L., Racher, K., Jamieson, L., Blight, M. A. and Holland, I. B. (2005) 'Mutations in HlyD, Part of the Type 1 Translocator for Hemolysin Secretion, Affect the Folding of the Secreted Toxin', *Journal of Bacteriology,* 187(21), pp. 7471-7480.

Pimenta, A. L., Young, J., Holland, I. B. and Blight, M. A. (1999) 'Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator', *Molecular and General Genetics MGG,* 261(1), pp. 122-132.

Pineau, C., Guschinskaya, N., Robert, X., Gouet, P., Ballut, L. and Shevchik, V. E. (2014) 'Substrate recognition by the bacterial type II secretion system: more than a simple interaction', *Molecular Microbiology,* 94(1), pp. 126-140.

Portaliou, A. G., Tsolis, K. C., Loos, M. S., Balabanidou, V., Rayo, J., Tsirigotaki, A., Crepin, V. F., Frankel, G., Kalodimos, C. G., Karamanou, S. and Economou, A. (2017) 'Hierarchical protein targeting and secretion is controlled by an affinity switch in the type III secretion system of enteropathogenic *Escherichia coli*', *Embo j,* 36(23), pp. 3517-3531.

Pugsley, A. P. (1992) 'Translocation of a folded protein across the outer membrane in *Escherichia coli*', *Proc Natl Acad Sci U S A,* 89(24), pp. 12058-62.

Pugsley, A. P. (1993) 'The complete general secretory pathway in gram-negative bacteria', *Microbiol Rev,* 57(1), pp. 50-108.

Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D. and Mekalanos, J. J. (2007) 'Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin', *Proceedings of the National Academy of Sciences,* 104(39), pp. 15508-15513.

Pukatzki, S., Ma, A. T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W. C., Heidelberg, J. F. and Mekalanos, J. J. (2006) 'Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system', *Proceedings of the National Academy of Sciences,* 103(5), pp. 1528-1533.

Radics, J., Königsmaier, L. and Marlovits, T. C. (2014) 'Structure of a pathogenic type 3 secretion system in action', *Nature Structural & Molecular Biology,* 21(1), pp. 82-87.

Randall, L. L., Crane, J. M., Liu, G. and Hardy, S. J. (2004) 'Sites of interaction between SecA and the chaperone SecB, two proteins involved in export', *Protein Sci,* 13(4), pp. 1124-33.

Randall, L. L., Topping, T. B., Smith, V. F., Diamond, D. L. and Hardy, S. J. S. (1998) 'SecB: A chaperone from *Escherichia coli*', *Methods in Enzymology*: Academic Press, pp. 444-459.

Raymond, B., Young, J. C., Pallett, M., Endres, R. G., Clements, A. and Frankel, G. (2013) 'Subversion of trafficking, apoptosis, and innate immunity by type III secretion system effectors', *Trends Microbiol,* 21(8), pp. 430-41.

Rebuffat, S. (2011) 'Bacteriocins from Gram-Negative Bacteria: A Classification?', in Drider, D. & Rebuffat, S. (eds.) *Prokaryotic Antimicrobial Peptides: From Genes to Applications*. New York, NY: Springer New York, pp. 55-72.

Rehman, S., Li, Y. G., Schmitt, A., Lassinantti, L., Christie, P. J. and Berntsson, R. P.-A. (2019) 'Enterococcal PcfF Is a Ribbon-Helix-Helix Protein That Recruits the Relaxase PcfG Through Binding and Bending of the oriT Sequence', *Frontiers in Microbiology,* 10(958).

Reimann, S. (2017) *Regulation of the activity of the Escherichia coli ABC transporter haemolysin B.* PhD Doctoral Dissertation, Heinrich-Heine-University Duesseldorf.

Reimann, S., Poschmann, G., Kanonenberg, K., Stühler, K., Smits, S. H. and Schmitt, L. (2016) 'Interdomain regulation of the ATPase activity of the ABC transporter haemolysin B from *Escherichia coli*', *Biochem J,* 473(16), pp. 2471-83.

Ristow, L. C., Tran, V., Schwartz, K. J., Pankratz, L., Mehle, A., Sauer, J. D. and Welch, R. A. (2019) 'The Extracellular Domain of the β (2) Integrin β Subunit (CD18) Is Sufficient for *Escherichia coli* Hemolysin and *Aggregatibacter actinomycetemcomitans* Leukotoxin Cytotoxic Activity', *mBio,* 10(4).

Robinson, G. L. (1951) 'The haemolysin of Bacterium coli', *J Gen Microbiol,* 5(4), pp. 788-92.

Roman-Hernandez, G., Peterson, J. H. and Bernstein, H. D. (2014) 'Reconstitution of bacterial autotransporter assembly using purified components', *Elife,* 3, pp. e04234.

Rose, R. J., Welsh, T. S., Waksman, G., Ashcroft, A. E., Radford, S. E. and Paci, E. (2008) 'Donor-strand exchange in chaperone-assisted pilus assembly revealed in atomic detail by molecular dynamics', *J Mol Biol,* 375(4), pp. 908-19.

Ryu, J., Lee, U., Park, J., Yoo, D.-H. and Ahn, J. H. (2015) 'A Vector System for ABC Transporter-Mediated Secretion and Purification of Recombinant Proteins in

Pseudomonas Species', *Applied and Environmental Microbiology,* 81(5), pp. 1744-1753.

Samudrala, R., Heffron, F. and McDermott, J. E. (2009) 'Accurate prediction of secreted substrates and identification of a conserved putative secretion signal for type III secretion systems', *PLoS Pathog,* 5(4), pp. e1000375.

Sana, T. G., Lugo, K. A. and Monack, D. M. (2017) 'T6SS: The bacterial "fight club" in the host gut', *PLoS Pathog,* 13(6), pp. e1006325.

Sánchez-Magraner, L., Viguera, A. R., García-Pacios, M., Garcillán, M. P., Arrondo, J. L., de la Cruz, F., Goñi, F. M. and Ostolaza, H. (2007) 'The calcium-binding C-terminal domain of *Escherichia coli* alpha-hemolysin is a major determinant in the surface-active properties of the protein', *J Biol Chem,* 282(16), pp. 11827-35.

Sapay, N., Guermeur, Y. and Deléage, G. (2006) 'Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier', *BMC Bioinformatics,* 7(1), pp. 255.

Satchell, K. J. (2011) 'Structure and function of MARTX toxins and other large repetitive RTX proteins', *Annu Rev Microbiol,* 65, pp. 71-90.

Sauer, F. G., Fütterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J. and Waksman, G. (1999) 'Structural Basis of Chaperone Function and Pilus Biogenesis', *Science,* 285(5430), pp. 1058-1061.

Sauri, A., Soprova, Z., Wickström, D., de Gier, J. W., Van der Schors, R. C., Smit, A. B., Jong, W. S. P. and Luirink, J. (2009) 'The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease', *Microbiology (Reading),* 155(Pt 12), pp. 3982-3991.

Schiebel, E., Driessen, A. J. M., Hartl, F.-U. and Wickner, W. (1991) '$\Delta \mu$ H+ and ATP function at different steps of the catalytic cycle of preprotein translocase', *Cell,* 64(5), pp. 927-939.

Schmitt, L., Benabdelhak, H., Blight, M. A., Holland, I. B. and Stubbs, M. T. (2003) 'Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains', *J Mol Biol,* 330(2), pp. 333-42.

Schmitt, L. and Tampé, R. (2002) 'Structure and mechanism of ABC transporters', *Current Opinion in Structural Biology,* 12(6), pp. 754-760.

Schraidt, O. and Marlovits, T. C. (2011) 'Three-dimensional model of *Salmonella's* needle complex at subnanometer resolution', *Science,* 331(6021), pp. 1192-5.

Schülein, R., Gentschev, I., Schlör, S., Gross, R. and Goebel, W. (1994) 'Identification and characterization of two functional domains of the hemolysin translocator protein HlyD', *Mol Gen Genet,* 245(2), pp. 203-11.

Sebo, P. and Ladant, D. (1993) 'Repeat sequences in the *Bordetella pertussis* adenylate cyclase toxin can be recognized as alternative carboxy-proximal secretion signals by the *Escherichia coli* alpha-haemolysin translocator', *Mol Microbiol,* 9(5), pp. 999-1009.

Sennhauser, G., Bukowska, M. A., Briand, C. and Grütter, M. G. (2009) 'Crystal structure of the multidrug exporter MexB from *Pseudomonas aeruginosa*', *J Mol Biol,* 389(1), pp. 134-45.

Shi, X., Chen, M., Yu, Z., Bell, J. M., Wang, H., Forrester, I., Villarreal, H., Jakana, J., Du, D., Luisi, B. F., Ludtke, S. J. and Wang, Z. (2019) '*In situ* structure and assembly of the multidrug efflux pump AcrAB-TolC', *Nature Communications,* 10(1), pp. 2635.

Shneider, M. M., Buth, S. A., Ho, B. T., Basler, M., Mekalanos, J. J. and Leiman, P. G. (2013) 'PAAR-repeat proteins sharpen and diversify the type VI secretion system spike', *Nature,* 500(7462), pp. 350-353.

Silverman, J. M., Agnello, D. M., Zheng, H., Andrews, B. T., Li, M., Catalano, C. E., Gonen, T. and Mougous, J. D. (2013) 'Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates', *Mol Cell,* 51(5), pp. 584-93.

Simon, S. M., Peskin, C. S. and Oster, G. F. (1992) 'What drives the translocation of proteins?', *Proceedings of the National Academy of Sciences,* 89(9), pp. 3770-3774.

Skillman, K. M., Barnard, T. J., Peterson, J. H., Ghirlando, R. and Bernstein, H. D. (2005) 'Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter', *Mol Microbiol,* 58(4), pp. 945-58.

Smith, T. J., Font, M. E., Kelly, C. M., Sondermann, H. and O'Toole, G. A. (2018a) 'An N-Terminal Retention Module Anchors the Giant Adhesin LapA of *Pseudomonas fluorescens* at the Cell Surface: a Novel Subfamily of Type I Secretion Systems', *J Bacteriol,* 200(8).

Smith, T. J., Sondermann, H. and O'Toole, G. A. (2018b) 'Type 1 Does the Two-Step: Type 1 Secretion Substrates with a Functional Periplasmic Intermediate', *J Bacteriol,* 200(18).

Solbiati, J. O., Ciaccio, M., Farías, R. N., González-Pastor, J. E., Moreno, F. and Salomón, R. A. (1999) 'Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25', *J Bacteriol,* 181(8), pp. 2659-62.

Souza, R., Quispe Saji, G., Costa, M., Netto, D., Lima, N., Klein, C., Vasconcelos, A. and Nicolás, M. (2012) 'AtlasT4SS: A curated database for type IV secretion systems', *BMC microbiology,* 12, pp. 172.

Spitz, O., Erenburg, I. N., Beer, T., Kanonenberg, K., Holland, I. B. and Schmitt, L. (2019) 'Type I Secretion Systems-One Mechanism for All?', *Microbiol Spectr,* 7(2).

Stanley, P., Packman, L. C., Koronakis, V. and Hughes, C. (1994) 'Fatty acylation of two internal lysine residues required for the toxic activity of *Escherichia coli* hemolysin', *Science,* 266(5193), pp. 1992-6.

Staron, P., Forchhammer, K. and Maldener, I. (2014) 'Structure–function analysis of the ATP-driven glycolipid efflux pump DevBCA reveals complex organization with TolC/HgdD', *FEBS Letters,* 588(3), pp. 395-400.

Stebbins, C. E. and Galán, J. E. (2001) 'Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion', *Nature,* 414(6859), pp. 77-81.

Stietz, M. S., Liang, X., Wong, M., Hersch, S. and Dong, T. G. (2018) 'The double tubular contractile structure of the type VI secretion system displays striking flexibility and elasticity', *bioRxiv,* pp. 470229.

Sulavik, M. C., Houseweart, C., Cramer, C., Jiwani, N., Murgolo, N., Greene, J., DiDomenico, B., Shaw, K. J., Miller, G. H., Hare, R. and Shimer, G. (2001) 'Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes', *Antimicrob Agents Chemother,* 45(4), pp. 1126-36.

Symmons, M. F., Bokma, E., Koronakis, E., Hughes, C. and Koronakis, V. (2009) 'The assembled structure of a complete tripartite bacterial multidrug efflux pump', *Proceedings of the National Academy of Sciences,* 106(17), pp. 7173-7178.

Symmons, M. F., Marshall, R. L. and Bavro, V. N. (2015) 'Architecture and roles of periplasmic adaptor proteins in tripartite efflux assemblies', *Front Microbiol,* 6, pp. 513.

Tamura, N., Murakami, S., Oyama, Y., Ishiguro, M. and Yamaguchi, A. (2005) 'Direct interaction of multidrug efflux transporter AcrB and outer membrane channel TolC detected via site-directed disulfide cross-linking', *Biochemistry,* 44(33), pp. 11115-21.

Thanabalu, T., Koronakis, E., Hughes, C. and Koronakis, V. (1998) 'Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore', *Embo j,* 17(22), pp. 6487-96.

Thomas, S., Holland, I. B. and Schmitt, L. (2014) 'The Type 1 secretion pathway - the hemolysin system and beyond', *Biochim Biophys Acta,* 1843(8), pp. 1629-41.

Thompson, S. A. and Sparling, P. F. (1993) 'The RTX cytotoxin-related FrpA protein of *Neisseria meningitidis* is secreted extracellularly by *meningococci* and by HlyBD+ *Escherichia coli*', *Infect Immun,* 61(7), pp. 2906-11.

Tomasek, D., Rawson, S., Lee, J., Wzorek, J. S., Harrison, S. C., Li, Z. and Kahne, D. (2020) 'Structure of a nascent membrane protein as it folds on the BAM complex', *Nature,* 583(7816), pp. 473-478.

Tomkiewicz, D., Nouwen, N., van Leeuwen, R., Tans, S. and Driessen, A. J. M. (2006) 'SecA Supports a Constant Rate of Preprotein Translocation*', *Journal of Biological Chemistry,* 281(23), pp. 15709-15713.

Touzé, T., Eswaran, J., Bokma, E., Koronakis, E., Hughes, C. and Koronakis, V. (2004) 'Interactions underlying assembly of the *Escherichia coli* AcrAB-TolC multidrug efflux system', *Mol Microbiol,* 53(2), pp. 697-706.

Tsai, J. C., Yen, M. R., Castillo, R., Leyton, D. L., Henderson, I. R. and Saier, M. H., Jr. (2010) 'The bacterial intimins and invasins: a large and novel family of secreted proteins', *PLoS One,* 5(12), pp. e14403.

Tsukazaki, T. (2019) 'Structural Basis of the Sec Translocon and YidC Revealed Through X-ray Crystallography', *The Protein Journal,* 38(3), pp. 249-261.

Tsukazaki, T., Mori, H., Echizen, Y., Ishitani, R., Fukai, S., Tanaka, T., Perederina, A., Vassylyev, D. G., Kohno, T., Maturana, A. D., Ito, K. and Nureki, O. (2011) 'Structure and function of a membrane component SecDF that enhances protein export', *Nature,* 474(7350), pp. 235-238.

References

Tsutsumi, K., Yonehara, R., Ishizaka-Ikeda, E., Miyazaki, N., Maeda, S., Iwasaki, K., Nakagawa, A. and Yamashita, E. (2019) 'Structures of the wild-type MexAB–OprM tripartite pump reveal its complex formation and drug efflux mechanism', *Nature Communications,* 10(1), pp. 1520.

van den Berg, B. (2010) 'Crystal structure of a full-length autotransporter', *J Mol Biol,* 396(3), pp. 627-33.

van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. and Rapoport, T. A. (2004) 'X-ray structure of a protein-conducting channel', *Nature,* 427(6969), pp. 36-44.

Veenendaal, A. K., Hodgkinson, J. L., Schwarzer, L., Stabat, D., Zenk, S. F. and Blocker, A. J. (2007) 'The type III secretion system needle tip complex mediates host cell sensing and translocon insertion', *Mol Microbiol,* 63(6), pp. 1719-30.

Verger, D., Miller, E., Remaut, H., Waksman, G. and Hultgren, S. (2006) 'Molecular mechanism of P pilus termination in uropathogenic *Escherichia coli*', *EMBO reports,* 7(12), pp. 1228-1232.

Vergunst, A. C., van Lier, M. C. M., den Dulk-Ras, A., Grosse Stüve, T. A., Ouwehand, A. and Hooykaas, P. J. J. (2005) 'Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*', *Proceedings of the National Academy of Sciences of the United States of America,* 102(3), pp. 832-837.

Vignon, G., Köhler, R., Larquet, E., Giroux, S., Prévost, M. C., Roux, P. and Pugsley, A. P. (2003) 'Type IV-like pili formed by the type II secreton: specificity, composition, bundling, polar localization, and surface presentation of peptides', *J Bacteriol,* 185(11), pp. 3416-28.

von Heijne, G. (1990) 'The signal peptide', *The Journal of Membrane Biology,* 115(3), pp. 195-201.

Wagner, A., Tittes, C. and Dehio, C. (2019) 'Versatility of the BID Domain: Conserved Function as Type-IV-Secretion-Signal and Secondarily Evolved Effector Functions Within *Bartonella*-Infected Host Cells', *Front Microbiol,* 10, pp. 921.

Wagner, S., Grin, I., Malmsheimer, S., Singh, N., Torres-Vargas, C. E. and Westerhausen, S. (2018) 'Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells', *FEMS Microbiol Lett,* 365(19).

Wandersman, C. and Delepelaire, P. (1990) 'TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion', *Proc Natl Acad Sci U S A,* 87(12), pp. 4776-80.

Wang, J., Brackmann, M., Castaño-Díez, D., Kudryashev, M., Goldie, K. N., Maier, T., Stahlberg, H. and Basler, M. (2017a) 'Cryo-EM structure of the extended type VI secretion system sheath–tube complex', *Nature Microbiology,* 2(11), pp. 1507-1512.

Wang, Y., Wang, R., Jin, F., Liu, Y., Yu, J., Fu, X. and Chang, Z. (2016) 'A Supercomplex Spanning the Inner and Outer Membranes Mediates the Biogenesis of β-Barrel Outer Membrane Proteins in Bacteria*', *Journal of Biological Chemistry,* 291(32), pp. 16720-16729.

Wang, Z., Fan, G., Hryc, C. F., Blaza, J. N., Serysheva, II, Schmid, M. F., Chiu, W., Luisi, B. F. and Du, D. (2017b) 'An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump', *Elife,* 6.

Wattiau, P., Woestyn, S. and Cornelis, G. R. (1996) 'Customized secretion chaperones in pathogenic bacteria', *Molecular Microbiology,* 20(2), pp. 255-262.

Welch, R. A. (1991) 'Pore-forming cytolysins of gram-negative bacteria', *Mol Microbiol,* 5(3), pp. 521-8.

Werner, J., Augustus, A. M. and Misra, R. (2003) 'Assembly of TolC, a structurally unique and multifunctional outer membrane protein of *Escherichia coli* K-12', *J Bacteriol,* 185(22), pp. 6540-7.

Werner, J. and Misra, R. (2005) 'YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*', *Mol Microbiol,* 57(5), pp. 1450-9.

Weston, N., Sharma, P., Ricci, V. and Piddock, L. J. V. (2018) 'Regulation of the AcrAB-TolC efflux pump in *Enterobacteriaceae*', *Res Microbiol,* 169(7-8), pp. 425-431.

Whitaker, N., Bageshwar, U. K. and Musser, S. M. (2012) 'Kinetics of precursor interactions with the bacterial Tat translocase detected by real-time FRET', *J Biol Chem,* 287(14), pp. 11252-60.

Wiles, T. J., Dhakal, B. K., Eto, D. S. and Mulvey, M. A. (2008) 'Inactivation of host Akt/protein kinase B signaling by bacterial pore-forming toxins', *Mol Biol Cell,* 19(4), pp. 1427-38.

Woida, P. J. and Satchell, K. J. F. (2018) 'Coordinated delivery and function of bacterial MARTX toxin effectors', *Mol Microbiol,* 107(2), pp. 133-141.

Wolff, N., Sapriel, G., Bodenreider, C., Chaffotte, A. and Delepelaire, P. (2003) 'Antifolding activity of the SecB chaperone is essential for secretion of HasA, a quickly folding ABC pathway substrate', *J Biol Chem,* 278(40), pp. 38247-53.

Wu, K.-H., Hsieh, Y.-H. and Tai, P. C. (2012) 'Mutational Analysis of Cvab, an ABC Transporter Involved in the Secretion of Active Colicin V', *PLOS ONE,* 7(4), pp. e35382.

Wu, K. H. and Tai, P. C. (2004) 'Cys32 and His105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter', *J Biol Chem,* 279(2), pp. 901-9.

Yamamoto, K., Tamai, R., Yamazaki, M., Inaba, T., Sowa, Y. and Kawagishi, I. (2016) 'Substrate-dependent dynamics of the multidrug efflux transporter AcrB of *Escherichia coli*', *Scientific Reports,* 6(1), pp. 21909.

Yamashita, E., Zhalnina, M. V., Zakharov, S. D., Sharma, O. and Cramer, W. A. (2008) 'Crystal structures of the OmpF porin: function in a colicin translocon', *Embo j,* 27(15), pp. 2171-80.

Yan, Z., Yin, M., Xu, D., Zhu, Y. and Li, X. (2017) 'Structural insights into the secretin translocation channel in the type II secretion system', *Nat Struct Mol Biol,* 24(2), pp. 177-183.

Yoneyama, H., Maseda, H., Kamiguchi, H. and Nakae, T. (2000) 'Function of the membrane fusion protein, MexA, of the MexA, B-OprM efflux pump in *Pseudomonas aeruginosa* without an anchoring membrane', *J Biol Chem,* 275(7), pp. 4628-34.

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B. and Schmitt, L. (2005) 'H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB', *Embo j,* 24(11), pp. 1901-10.

Zaitseva, J., Oswald, C., Jumpertz, T., Jenewein, S., Wiedenmann, A., Holland, I. B. and Schmitt, L. (2006) 'A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer', *Embo j,* 25(14), pp. 3432-43.

Zavialov, A. V., Tischenko, V. M., Fooks, L. J., Brandsdal, B. O., Aqvist, J., Zav'yalov, V. P., Macintyre, S. and Knight, S. D. (2005) 'Resolving the

energy paradox of chaperone/usher-mediated fibre assembly', *Biochem J*, 389(Pt 3), pp. 685-94.

Zechner, E. L., Lang, S. and Schildbach, J. F. (2012) 'Assembly and mechanisms of bacterial type IV secretion machines', *Philos Trans R Soc Lond B Biol Sci*, 367(1592), pp. 1073-87.

Zgurskaya, H. I., Krishnamoorthy, G., Ntreh, A. and Lu, S. (2011) 'Mechanism and Function of the Outer Membrane Channel TolC in Multidrug Resistance and Physiology of *Enterobacteria*', *Front Microbiol*, 2, pp. 189.

Zgurskaya, H. I. and Nikaido, H. (1999a) 'AcrA is a highly asymmetric protein capable of spanning the periplasm', *J Mol Biol*, 285(1), pp. 409-20.

Zgurskaya, H. I. and Nikaido, H. (1999b) 'Bypassing the periplasm: reconstitution of the AcrAB multidrug efflux pump of *Escherichia coli*', *Proc Natl Acad Sci U S A*, 96(13), pp. 7190-5.

Zhang, F., Greig, D. I. and Ling, V. (1993) 'Functional replacement of the hemolysin A transport signal by a different primary sequence', *Proceedings of the National Academy of Sciences*, 90(9), pp. 4211-4215.

Zhong, X., Kolter, R. and Tai, P. C. (1996) 'Processing of colicin V-1, a secretable marker protein of a bacterial ATP binding cassette export system, requires membrane integrity, energy, and cytosolic factors', *J Biol Chem*, 271(45), pp. 28057-63.

Zoued, A., Durand, E., Bebeacua, C., Brunet, Y. R., Douzi, B., Cambillau, C., Cascales, E. and Journet, L. (2013) 'TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system', *J Biol Chem*, 288(38), pp. 27031-27041.

6.   List of abbreviations

Adenosine diphosphate (ADP)

Adenosine triphosphate (ATP)

Ångström (Å)

Aquifex aeolicus (Aa)

ATP binding cassette (ABC)

Avibacterium paragallinarum (Ap)

Bacterial transglutaminase-like cysteine proteinase (BTLCP)

Basic local alignment search tool (BLAST)

β-barrel assembly machinery (BAM)

Bibersteinia trehalosi (Bt)

Bordetella pertussis (*B. pertussis*)

C39-like domain (CLD)

Carbonyl cyanide m-chlorophenyl hydrazone (CCCP)

Cardiobacterium valvarum (Cv)

Caulobacter crescentus (*C. crescentus*)

Cryogenic electron microscopy (cryo-EM)

Cytoplasm (CP)

Cytoplasmic domain (CD)

Donor strand complementation (DSC)

Donor strand exchange (DSE)

Enhanced green fluorescent protein (eGFP)

Escherichia coli (*E. coli* or Ec)

Green fluorescent protein (GFP)

Helix-turn-helix (HTH)

Hemolysin (Hly)

Inner membrane (IM)

Inner membrane complex (IMC)

Kingella kingae (Kk)

Klebsiella pneumonia (*K. pneumonia*)

Legionella pneumophilia (*L. pneumophilia*)

Lipopolysaccharide (LPS)

Loose (L)

Major facilitator superfamily (MFS)

Mannheimia haemolytica (Mh)

Membrane proximal (MP)

Moraxella bovis (Mb)

Multidrug and toxic compound extrusion (MATE)

Multidrug resistance (MDR)

Multifunctional auto-processing repeat in toxin (MARTX)

Negative stained electron microscopy (ns-EM)

Nuclear magnetic resonance (NMR)

Nucleotide binding domain (NBD)

Open (O)

Origin of transfer (oriT)

Outer membrane (OM)

Outer membrane complex (OMC)

Outer membrane protein (OMP)

Periplasm (PP)

Periplasmic domain (PPD)

Polypeptide transport-associated (POTRA)

Possible binding pocket (pbp)

Proteobacterial antimicrobial compound efflux (PACE)

Proton motive force (pmf)

Pseudomonas aeruginosa (*P. aeruginosa*)

Repeat in toxin (RTX)

Resistance nodulation cell division (RND)

Retention module (RM)

Rhizobium etli (Re)

Root mean square deviation (RMSD)

Serratia marcescens (*S. marcescens*)

Signal recognition particle (SRP)

Single stranded deoxyribonucleic acid (ssDNA)

Small multidrug resistance (SMR)

Sodium dodecylsulfate (SDS)

Surface plasmon resonance (SPR)

Tight (T)

Translocation signal (TS)

Transmembrane (TM)

Transmembrane complex (TMC)

Transmembrane domain (TMD)

Trigger factor (TF)

Twin-arginine-dependent translocation (TAT)

Type 1 secretion system (T1SS)

Type 2 secretion system (T2SS)

Type 3 secretion system (T3SS)

Type 4 secretion system (T4SS)

Type 5 secretion system (T5SS)

Type 6 secretion system (T6SS)

Vibrio cholerae (*V. cholerae*)

Xanthomonas VirD4-interacting protein conserved domain (XVIPCD)

Xylella fastidiosa (Xf)

7. List of figures

8.   List of tables

8.   List of tables

9. Acknowledgement

Zuallererst möchte ich mich bei meiner Familie bedanken, die mich von Anfang an auf die akademische Laufbahn vorbereitet und mich bis zum Ende und darüber hinaus unterstützt hat. Ein besonderer Dank geht dabei an meine ältere Schwester Charlotte, die den Weg zum Dr. rer. nat. vor mir gegangen ist. Dadurch blieben mir viele Überraschungen und langwieriges recherchieren größtenteils erspart.

Als nächstes möchte ich mich bei allen meinen Freunden bedanken, die ich nicht durch die Universität kennen gelernt habe. Obwohl unsere Leben durch die verschiedensten Berufswahlen sehr unterschiedlich verlaufen, finden wir glücklicherweise immer noch Zeit füreinander. Ein besonderer Dank geht dabei an meinen Partner Sinan, der am meisten unter meinen Stress während der Promotion zu leiden hatte. Trotzdem hatte er immer Verständnis und ein offenes Ohr für alle arbeitsbetreffenden Probleme und hat mich gleichzeitig an die nötige work-life-balance erinnert.

Diese Promotion wäre ohne das Institut für Biochemie nicht möglich gewesen. Daher geht ein großer Dank an meinen Doktorvater Lutz, der dieses ganze Projekt ermöglicht hat. Und an meinen zweiten Betreuer Johannes Hegemann, der durch einen anderen Blickwinkel auf die Projekte und kritische Nachfragen interessante und wissenschaftlich wertvolle Ideen beigetragen hat. Glücklicherweise stand Lutz' Bürotür immer offen und lag schräg gegenüber meines Labors, sodass ich immer die Möglichkeit hatte, schnelle Rückmeldungen und auch neue Ideen für meine Experimente zu erfragen. Die regelmäßigen subgroup-meetings, in denen natürlich auch Sander sehr wichtig war (und immer noch ist), waren durch die kritischen wissenschaftlichen Diskussionen sehr aufschluss- und hilfreich. Daher geht mein Dank natürlich auch an alle ehemaligen und momentanen Mitglieder der T1-gruppe! Trotzdem möchte ich auch Sander nochmal besonders danken. Zum einen für das nötige Antreiben und zum anderen für das Verteilen von „Zwangsurlaub", wenn nötig. Und wenn wir schon bei den Chefs sind, darf Uli nicht fehlen. Nicht nur weil er den Studiengang, der mich zur Promotion geführt hat, zum großen Teil gestaltet

hat, sondern auch, weil er immer Rücksicht und Fairness bei der Verteilung der Lehrverantwortungen gezeigt hat.

Besonders hervorheben möchte ich außerdem Martina! Danke für alles. Und damit meine ich alles: Du weißt wo alles steht, wie alles geht und wie man das neuen (noch dummen) Doktoranden vermittelt! Ohne dich wurden die Spaziergänge zur Mensa außerdem ein wenig langweiliger und ich kann gar nicht mehr zählen wie oft Du mich schon aus einer Kriese gerettet hast.

Obwohl ich natürlich allen ehemaligen und momentanen Mitgliedern des Instituts für Biochemie dankbar für die schöne Zeit und tolle Arbeitsatmosphäre bin, verdienen ein paar Mitglieder besondere Hervorhebung. Dazu gehören natürlich Julia, Isa, Rebecca und Katja, mit denen ich auch außerhalb der Arbeit schon viele schöne Erinnerungen gesammelt habe. Ich freue mich schon sehr darauf euch wieder persönlich und nicht nur über Skype zu sehen! Und danke nochmal für all die tollen Kaffee-Runden mit mimimi und ohne.

An diesen oftmals unterhaltsamen Kaffee-Runden haben sich natürlich auch andere beteiligt. Und so geht mein Dank auch an meinen Bench-Nachbarn und MALS-Mitverantwortlichen Martin und an Tobi, den ersten Biochemiker, den ich an der HHU kennen gelernt habe. Genauso wie an Tim, der ebenfalls seit dem ersten Semester dabei war und sich das Labor mit Martin, Manuel W. und mir geteilt hat. Aber es gibt noch mehr Namen mit M. Wie zum Beispiel Marcel, der mich schon in der Masterarbeit betreut hat und zwei Michaels, die zwar sehr unterschiedlich aber beide wunderbare Arbeitskollegen sind, ebenso wie Maryna. Und natürlich mein ehemaliger Masterstudent, später Kollege, Manuel A., der hervorragende Arbeit mit der Pipette und dem Pinsel leisten kann. Besonders danken möchte ich Dir natürlich für die Mithilfe am Homologie-Projekt und dafür, dass Du die lästigen ATPasen gemacht hast! Methodisch möchte ich mich auch bei Sven bedanken, der mir gezeigt hat wie man HlyB und damit Membranproteine reinigt. Und bei Sakshi, die mir gezeigt hat, wie man lösliche Proteine reinigt. Und bei Diana, die mir gezeigt hat wie man kloniert, was für diese Promotion sehr nützlich war. Zu nennen sind außerdem noch meine Bürokollegen Katharina, Kalpana, Kerstin, Zohreh, Steffi und Vivien.

10. Curriculum vitae

Personal Data

Name	Olivia Spitz
Year of birth	1991
Place of birth	Germany

Education

since 04/2016	Doctorate Heinrich-Heine-Universität Düsseldorf Supervisor: Prof. Dr. Lutz Schmitt Institute of Biochemistry Title: Approaches to replace the HlyA T1SS by homologs
10/2013 – 04/2016	Master of Science (overall grade 1.3) Heinrich-Heine-Universität Düsseldorf Supervisor of Master thesis: Prof. Dr. Lutz Schmitt Institute of Biochemistry Title: *In vivo* and *in vitro* characterization of the ABC transporter NisT
10/2010 – 09/2013	Bachelor of Science (overall grade 1.5) Heinrich-Heine-Universität Düsseldorf Supervisor of Bachelor thesis: Prof. Dr. Eckhard Lammert Institute of Metabolic Physiology Title: Expression analysis of Sema3A and Nrp1 in the hypothalamus, the pituitary gland and the adrenal gland
06/2001 – 06/2010	General qualification for university entrance (Abitur, overall grade 1.2) Gymnasium Adolfinum in Moers, Germany
06/1997 – 06/2001	Elementary school Eichendorff Grundschule in Moers, Germany

11. Declaration

Ich versichere an Eides statt, dass die vorliegende Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist. Diese Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen oder erfolgreichen Promotionsversuche unternommen.

---------------------------------------   -----------------------------------

(Ort, Datum)   Olivia Spitz