Studies of structure-function relationship of components of multidrug efflux pumps and type I secretion systems

Dissertation

zur Erlangung des naturwissenschaftlichen Doktorgrades

der Bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Johann Georg Polleichtner

Aschaffenburg

Würzburg 2006

Eingereicht am: _____

Mitglieder der Promotionskommission: Vorsitzender: _____

1. Gutachter: Prof. Dr. R. Benz

2. Gutachter: Prof. Dr. J. Reidl

Tag des Promotionskolloquiums: _____

Doktorurkunde ausgehändigt am: _____

Diese Dissertation wurde von mir selbstständig und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt.

Die von mir vorgelegte Dissertation hat noch in keinem früheren Prüfungsverfahren in ähnlicher oder gleicher Form vorgelegen.

Ich habe zu keinem früheren Zeitpunkt versucht, einen akademischen Grad zu erwerben.

Publications

- Maier,E., G.Polleichtner, B.Boeck, R.Schinzel and R.Benz. 2001. Identification of the outer membrane porin of *Thermus thermophilus* HB8: the channel-forming complex has an unusually high molecular mass and an extremely large single-channel conductance. *J Bacteriol.* **183:**800-3.

- Knapp,O., E.Maier, G.Polleichtner, J.Masin, P.Sebo and R.Benz. 2003. Channel formation in model membranes by the adenylate cyclase toxin of *Bordetella pertussis*: effect of calcium. *Biochemistry.* **42:**8077-84.

- Polleichtner,G. and C.Andersen. 2006. The channel-tunnel HI1462 of *Haemophilus influenzae* reveals differences to *Escherichia coli* TolC. *Microbiology.* **152:** 1639-1647.

Contents

Chapter 1

Introduction

1.1. The cell envelope of Gram-negative bacteria serves as barrier to noxious substances

The *Eubacteria* can be divided into two large groups, the Gram-positive and the Gram-negative bacteria. This classification is based on the phenomenological observation that the first can be stained dark blue in a reaction called Gram-staining, while this is not possible with the latter (Beveridge, 2001). This phenotypical difference is correlated to the different molecular structure of the cell envelope of the two groups. The cell-wall of Gram-positive bacteria is formed by peptidoglycan, a meshwork of strands of peptides and glycans that is covalently cross-linked (Walsh, 2000; Beveridge, 2001). This peptidoglycan or murein sacculus confers mechanical strength to osmotic lysis but is not very selective and allows diffusion of relatively large substances. Gram-negative bacteria, however, are characterized by a low content of peptidoglycan and the presence of a second, the outer membrane at the cell surface (see Figure 1.1). The space between the cytoplasmic or inner membrane and the outer membrane is the periplasm. Membranes, in general, form permeability barriers for hydrophilic substances. The outer membrane of Gram-negative bacteria, however, must at least allow influx of nutrients and efflux of waste products, requiring the presence of selective transport proteins (Nikaido and Vaara, 1985). The outer membrane is asymmetrically composed of phospholipids in the inner leaflet and lipopolysaccharides (LPS) in the outer leaflet. Additionally, it contains proteins, which constitute nearly half of its mass. Because of its special composition and organization, the outer membrane shows very low permeability

toward lipophilic solutes and high permeability toward hydrophilic solutes. Therefore, it acts as a very effective permeation barrier especially to many antibiotics, whereas at the same time it allows efficient diffusion of nutrients (Nikaido and Vaara, 1985). To the outer membrane the Gram-negative bacteria owe their higher tolerance to noxious agents compared to the Gram-positives, which is obviously related to the prevalence of Gram-negative infections in the modern hospital environment (Nikaido and Vaara, 1985).

The mentioned transport of substances across the inner membrane is energized by cellular ATP or by ion gradients across the membrane. Therefore special inner membrane transporter proteins can actively accumulate substances on one side of the membrane (Saier, 2000). Across the outer membrane, in contrast, transport occurs by diffusion through water-filled channel-proteins like porins, following a concentration gradient between the periplasmic space and the external medium. Active transport across the outer membrane is only possible by interaction of outer membrane transport proteins with energized proteins of the inner membrane (Andersen, 2003).
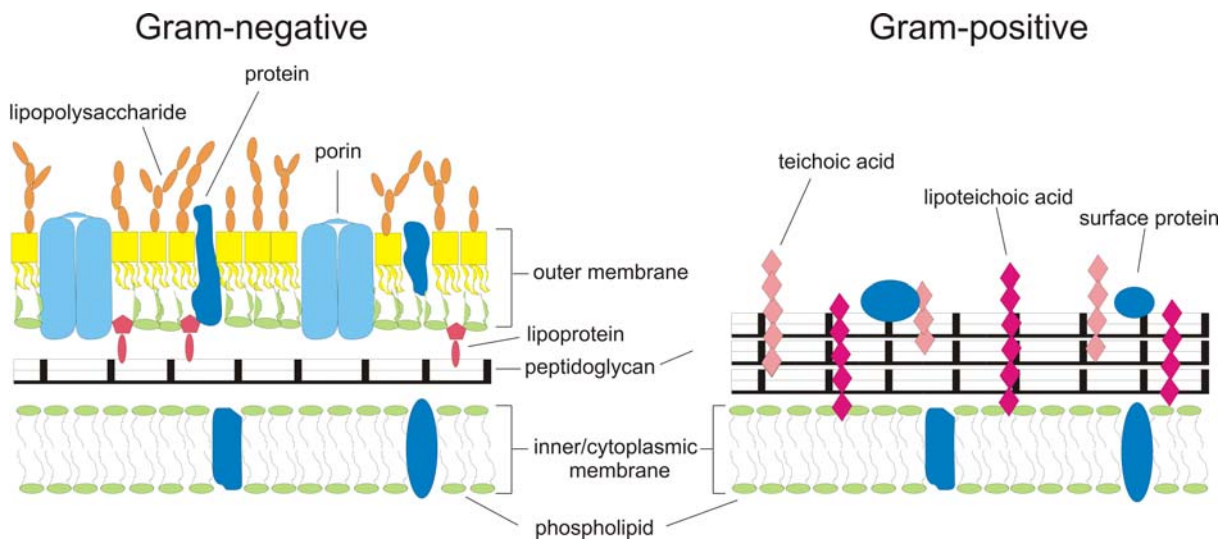
Figure 1.1: Schematical presentation of Gram-negative and Gram-positive cell envelope.

1.2. Targets and effects of antibiotics

The discovery of penicillin in 1928 by Alexander Fleming (1881 – 1955) was one of the most important steps in modern medicine to fight bacterial pathogens. The development and use of antibiotics in the 20th century apparently made it possible to control infectious diseases

(Putman *et al.*, 2000). However, once the effectiveness of an antibiotic is proven and it enters widespread therapeutic use, its days are numbered (Walsh, 2000). Appearance of clinically significant resistance is observed in periods of months to years (Davies, 1996). The subsequent spread of resistance in pathogenic organisms have made many currently available antibiotics ineffective (Moellering, 1998; Neu, 1992). The reemergence of tuberculosis and other infectious diseases is the consequence and causes a serious public health problem (Cohen, 1992; Culliton, 1992). For understanding how bacteria can evolve resistance it is necessary to know, how antibiotics work. There are three proven targets for the main classes of antibacterial drugs: (1) bacterial cell-wall biosynthesis, (2) bacterial protein synthesis and (3) bacterial DNA replication and repair (Walsh, 2000).

As mentioned above, the bacterial cell-wall consists of peptidoglycan, a meshwork of strands of peptides and glycans, which are connected and extended by action of transpeptidases and transglycosylases. Bifunctional enzymes containing both transpeptidase and transglycosylase domains are the targets for the β-lactam-containing penicillins and cephalosporins. They act as pseudosubstrates and occupy the enzyme active sites, preventing normal cross-linking of peptide chains (Spratt and Cromie, 1988). Vancomycin also targets the cell-wall biosynthesis by complexation of the D-Ala-D-Ala termini of peptidoglycans and thereby preventing it from reacting with either the transpeptidases or the transglycosylases (Williams, 1996). The effect is the same for both the β-lactams and vancomycin: failure to make cross-links leaves the peptidoglycan layer mechanically weak and susceptible to lysis on changes in osmotic pressure (Walsh, 2000).

The bacterial RNA and protein machinery is sufficiently distinct from the analogous eukaryotic machinery so that there are many inhibitors of protein synthesis with selective antibacterial action (Walsh, 2000). These include antibiotics as the macrolides, tetracyclines and aminoglycosides (e.g. streptomycin and kanamycin) (Chopra, 1998).

Finally, the fluoroquinolones (e.g. ciprofloxacin) are a group of synthetic antibiotics that kill bacteria by targeting DNA replication and repair. The molecular target of quinolone antibiotics is the bacterial DNA gyrase, the enzyme responsible for uncoiling the intertwined circles of double-stranded bacterial DNA and thus reducing the number of superhelical twists (Maxwell, 1997; Walsh, 2000). Bacterial DNA gyrases have a topoisomerase II function introducing transient double-strand breaks in the DNA prior to uncoiling. Fluoroquinolones act by forming a complex with the enzyme and the cleaved DNA and thereby inhibiting the religation of DNA. As a consequence, double-strand breaks accumulate and ultimately set off the SOS repair system leading to cell death (Walsh, 2000).

1.3.   Bacterial mechanisms to resist antibiotic action

Bacteria have developed various ways to nullify the toxic effects of antibiotics and other drugs (Hayes and Wolf, 1990; Neu, 1992) (see Figure 1.2). The first possibility for bacteria to gain resistance is to destroy the chemical warhead in the antibiotic by hydrolysis or the formation of inactive derivatives (Davies, 1994). Well-known examples are β-lactamases, which deactivate penicillins and cephalosporins by hydrolysis of the β-lactam ring, and enzymes that phosphorylate, adenylate or acetylate aminoglycoside antibiotics (Bush *et al.*, 1995; Rasmussen and Bush, 1997; Shaw *et al.*, 1993; Putman *et al.*, 2000).

A second resistance mechanism is the alteration of the target by mutation or enzymatic modification in such a way that the affinity of the antibiotic for the target is reduced (Spratt, 1994; Putman *et al.*, 2000). Examples are the mutations in the DNA gyrase gene to quinolone resistance (Yoshida *et al.*, 1990), the methylation of a specific adenine residue in the 23S rRNA leading to resistance against all members of the erythromycin class of drugs (Weisblum, 1995) or the reprogramming of the peptidoglycan termini from D-Ala-D-Ala to D-Ala-D-Lac lowering the binding affinity of vancomycin 1,000-fold (Walsh *et al.*, 1996; Bugg *et al.*, 1991).

A third, more general mechanism of resistance is the inhibition of drug entry into the cell (Putman *et al.*, 2000). As mentioned above, diffusion of drugs across the cell envelope is reduced by the low permeability of the outer membrane of Gram-negative bacteria (Nikaido, 1989). However, once the drugs have entered the cell and accumulated at certain concentrations, these barriers cannot prevent them from exerting their toxic action. Therefore active efflux of drugs is essential to ensure significant levels of drug resistance (Nikaido, 1994; Putman *et al.*, 2000). This fourth mechanism is described in more detail in the next section.
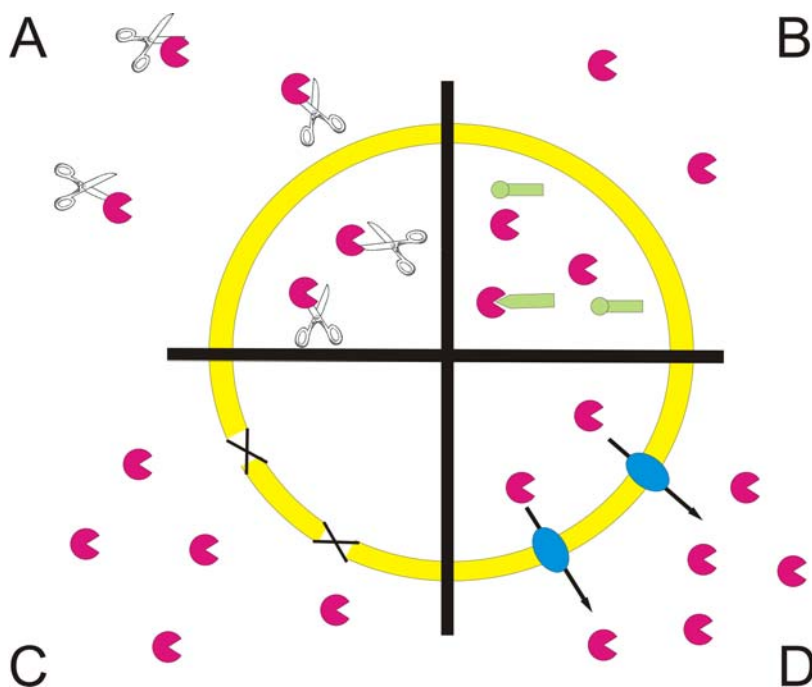
Figure 1.2: Resistance mechanisms in bacteria: (A) drug inactivation, (B) target alteration, (C) prevention of drug influx, and (D) active extrusion of drug from the cell.

1.4. Drug efflux as resistance mechanism to antibiotics

Extrusion of a wide variety of structurally unrelated compounds out of the cell is mediated by the so-called multidrug transporters, which actively (by consumption of energy) reduce the intracellular drug concentration to subtoxic levels (Putman *et al.*, 2000).

1.4.1. *Multidrug transporters – the engines of drug efflux*

In general, multidrug transporters fall into five superfamilies: the ABC (ATP-binding cassette), the MF (major facilitator), the SMR (small multidrug resistance), the RND (resistance nodulation cell-division) and the MATE (multidrug and toxic compound extrusion) superfamily (Andersen, 2003, Putman *et al.*, 2000). ABC-type multidrug transporters use the free energy of ATP hydrolysis to pump drugs out of the cell, while transporters of the other superfamilies utilize the transmembrane electrochemical gradient of protons or sodium ions to extrude substances out of the cell. The MF superfamily comprises sym-, anti- or uniporters, while members of the SMR and RND superfamilies are proton antiporters (Saier *et al.*, 1994, 2000; Griffith *et al.*, 1992; Marger and Saier, 1993). Multidrug transporters of any superfamily are found in bacteria and archaea as well as in higher

eukaryotes with the exception of drug transporters of the SMR superfamily which are exclusively found in prokaryotes (Saier *et al.*, 1998; Higgins, 1992; Tseng *et al.*, 1999).

Hydropathy analysis and alignment of conserved motifs of drug transporters of the MF superfamily revealed that these proteins can be divided into two subfamilies with either 12 or 14 membrane-spanning helices (Paulsen and Skurray, 1993; Putman *et al.*, 2000). Well known representatives are the class B tetracycline transporter TetA(B) of *E. coli* (12-helix transporter) and the class K tetracycline transporter TetA(K) of *Staphylococcus aureus* (14-helix transporter). The MF transporters consist of two halves with usually related sequences containing six and seven transmembrane helices, respectively, connected by a cytoplasmic loop. Therefore, it is assumed that the MF transporters evolved by gene duplication (Andersen, 2003).

Multidrug transporters of the SMR superfamily, the smallest drug efflux proteins known, are about 100 amino acid residues in length (Yerushalmi and Schuldiner, 2000). The proteins comprise four tightly packed α-helical transmembrane segments and arrange as dimers or trimers (Tate *et al.*, 2001; Muth and Schuldiner, 2000; Yerushalmi and Schuldiner, 2000).

The MATE family was first described by Brown *et al.* (1999). It contains more than 30 proteins, including NorM, a drug transporter of *Vibrio parahaemolyticus*, and its *E. coli* homologue YdhE, which were previously suggested to be members of the MF superfamily (Morita *et al.*, 1998).

The structures of MsbA and BtuCD have been solved as representatives of the ABC superfamily. They crystallized as dimers revealing that the membrane-spanning part of the functional unit consists of twelve transmembrane helices (six per monomer). Linked to the cytoplasmic side of the membrane domain, two ATP-binding domains (one per monomer) provide energy for the translocation process (Chang and Roth, 2001; Locher *et al.*, 2002).

The RND transporter AcrB of *E. coli* was crystallized by Murakami and coworkers in 2002 (see Figure 1.3). It is assembled to a homotrimer comprising two distinct domains, a 50 Å thick transmembrane region composed of twelve helices per monomer and a large headpiece domain formed by loops between the helices H1 and H2, and H7 and H8 protruding 70 Å into the periplasmic space. Located at the base of the headpiece a central cavity is connected by a central pore with a funnel-like opening at the top. It opens laterally to the periplasm via three vestibules near the membrane plane. These openings are assumed to serve as entrance for transported molecules from the periplasm or the outer leaflet of the cytoplasma membrane (Murakami *et al.*, 2002).
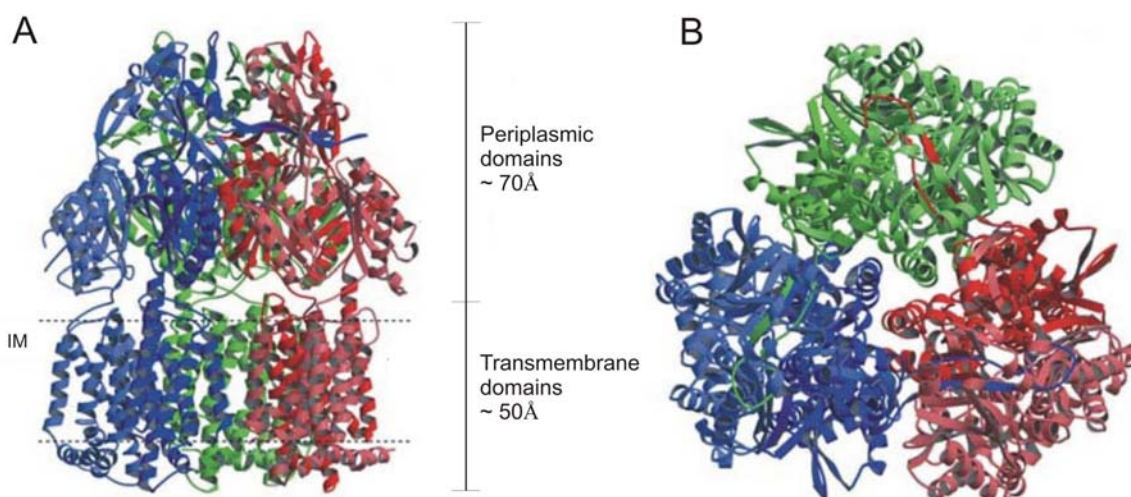
Figure 1.3: Ribbon representation of AcrB. Three protomers are individually coloured (blue, green and red). The N-terminal and C-terminal halves of the protomers are depicted as dark and pale colours, respectively. (A) side view, (B) top view (Adapted by permission from Macmillan Publishers Ltd: Nature, Murakami *et al.*, 2002).

The continual progress in genome sequencing makes it possible today to study the origin of multidrug transporters. The fact that they are found in the genome of pathogenic as well as in nonpathogenic bacteria in comparable numbers indicates, that these export systems did not evolve recently as a result of extensive exposure to medically relevant drugs (Saier *et al.*, 1998). Instead, they are assumed to be important for the extrusion of naturally occurring noxious substances.

1.4.2. *The efflux systems of Gram-negative bacteria: multidrug efflux pumps*

As mentioned above, Gram-negative bacteria show a higher resistance to a large number of antibiotics and chemotherapeutic agents than Gram-positive bacteria, because of the outer membrane acting as an additional permeability barrier. Because of this outer membrane, however, Gram-negative bacteria require additional systems to translocate drugs to the cell exterior. Therefore, Gram-negative bacteria evolved special export systems termed multidrug efflux pumps. These multidrug efflux pumps consist of the inner membrane drug transporter of the ABC, MF or RND superfamilies, a periplasmic adaptor protein (previously termed membrane fusion protein) and an outer membrane protein of the TolC family (Andersen, 2003). The best-characterised example is the AcrAB/TolC multidrug efflux pump consisting of the AcrB inner membrane RND transporter cooperating with AcrA as adaptor protein and TolC as outer membrane component. This tripartite export apparatus is able to extrude substances directly into the external medium. Because drugs are pumped out faster than they

can reenter the cells, the intracellular concentrations are kept low and significant resistance levels can be achieved in Gram-negative bacteria.

The importance of multidrug efflux pumps for Gram-negative bacteria becomes visible in several investigations, where the main efflux systems of different species were inactivated. Thus inactivation of AcrAB/TolC of *E. coli* (Fralick, 1996; Sulavik *et al.*, 2001), MtrCDE of *Neisseria gonorrhoe* (Lucas *et al.*, 1995), AmrAB/OprA of *Burkholderia pseudomallei* (Moore *et al.*, 1999), SmeDEF of *Stenotrophomonas maltophilia* (Zhang *et al.*, 2001), HI0894/HI0895 of *Haemophilus influenzae* (Sánchez *et al.*, 1997), or MexAB/OprM of *Pseudomonas aeruginosa* (Li *et al.*, 1995) resulted in an increased susceptibility of the strains to various drugs. This demonstrates how highly efficient multidrug efflux pumps work in cooperation with the outer membrane barrier (Nikaido, 1996; Thanassi *et al.*, 1995).

All these given examples of bacterial multidrug efflux pumps are expressed in wild-type strains and contribute to the intrinsic antibiotic resistance level. Multidrug transporters, however, are associated with both intrinsic and acquired resistance to antibiotics and particularly the latter causes serious problems in the pharmacological treatment of patients with infectious diseases, since the substrate spectra of many multidrug efflux systems include clinically relevant antibiotics (Putman *et al.*, 2000). Acquired resistance based on drug efflux can arise via three mechanisms: (I) amplification and mutations of genes encoding multidrug transporters, which change the expression level (Ng *et al.*, 1994) or activity (Klyachko and Neyfakh, 1998), (II) mutations in specific or global regulatory genes, which lead to the increased expression of multidrug transporters (Poole *et al.*, 1996; Dean *et al.*, 2005), and (III) intercellular transfer of transporter genes on transposons or plasmids (Kazama *et al.*, 1998).

It should be mentioned here that Gram-negative bacteria evolved homologue systems for the export of proteins (for example the HlyBD/TolC Typ I secretion apparatus of *E. coli* exporting hemolysin A) and cations such as Cd, Zn, Ni, or Co. These are also tripartite systems arranged of an inner membrane transporter, a periplasmic adaptor protein and an outer membrane TolC family member (Wandersman and Delepelaire, 1990; Nies *et al.*, 1989; Hassan *et al.*, 1999). The Typ I secretion system will be described later in more detail.

1.5. The outer membrane exit: channel-tunnels of the TolC family

1.5.1. A brief history of TolC

The prototypical and best-studied representative for the outer membrane component of multidrug efflux pumps is TolC of *E. coli*. It was first named according to the observation that TolC-deficient mutants are tolerant of certain colicins (Nagel de Zwaig and Luria, 1967). Colicins are toxins, or more precisely bacteriocins, that target *E. coli*. They are lethal weapons of bacteria against other bacterial competitors in the battle for resources. Colicins comprise three diverse protein domains, which catalyze three sequential actions in the implementation of toxicity: binding, translocation and killing (Cao and Klebba, 2002). The central receptor domain is responsible for the binding at a receptor on the surface of the target cell, and the N-terminal translocation domain somehow manages the transport of the C-terminal killing domain through the outer membrane. The killing domain acts either by forming pores in the cytoplasmic membrane, or by digesting DNA or RNA or inhibiting protein synthesis in the cytoplasm, or by degrading peptidoglycan (Cao and Klebba, 2002). The mechanism of translocation of the killing domain across the outer membrane is not understood. For colicin E1 and colicin 10, it is known that specific interaction of the translocation domain with TolC is prerequisite for translocation (de Zwaig and Luria, 1967; Pilsl and Braun, 1995). The fact, however, that other colicins interact with other outer membrane proteins such as porins indicates, that the uptake process is not related to the special shape or function of TolC. Besides for colicins, TolC was shown to act as a cell-surface receptor also for the TLS bacteriophage (German and Misra, 2001), and it could be demonstrated that extracellular domains are responsible for this interaction (Etz *et al.*, 2001).

In the 1980s, several studies showed an influence of the *tolC* gene on the expression of other outer membrane proteins (Morona and Reeves, 1982; Misra and Reeves, 1987; Dorman *et al.*, 1989). A direct involvement of TolC in gene regulation, however, can be excluded because of the local separation from the genes (Andersen, 2003). Dorman and coworkers suggest an indirect influence of TolC via DNA supercoiling. If TolC influences the supercoiling state by a possible direct contact site to the origin of replication or again indirectly via an altered membrane integrity in TolC-deficient strains, which somehow leads to altered responses to environmental conditions, remains unclear (Andersen, 2003).

A further phenotype of TolC-deficiency is the occurrence of anucleate cells (Hiraga *et al.*, 1989). The finding that TolC expression is downregulated if SeqA, a protein shown to be involved in sequestration during chromosome segregation, is absent, is also supporting a

possible role of TolC in cell division (Bahloul *et al.*, 1996). The exact role of TolC in the cell division, however, is not known (Andersen, 2003).

In 1990, Wandersman and Delepelaire for the first time demonstrated the involvement of TolC in the export of hemolysin (HlyA) and evidence that TolC forms a multidrug efflux pump with AcrA and AcrB was offered by Fralick in 1996, but the exact mechanism of interaction with the secretion apparatus was unknown.

The ability of TolC to form channels of 80 pS in 1M KCl in black lipid bilayer experiments and the fact, that the channels could be blocked with small peptides suggested that TolC is a peptide-specific porin (Benz *et al.*, 1993). The assumption of Benz and his coworkers that TolC is functioning as a trimer was confirmed by results of 2D crystallization of TolC in 1997 (Koronakis *et al.*, 1997). The results, however, pointed towards a single pore formed by three monomers rather than three individual pores as observed for porins.

Concerning the elucidation of export mechanisms associated with TolC and its cooperating partners it was a breakthrough when Koronakis and coworkers solved the structure of TolC by x-ray crystallography with a resolution of 2.1 Å in 2000 (Koronakis *et al.*, 2000).

1.5.2. *TolC – structure and biophysical characteristics*

Until the TolC structure was solved, it seemed mysterious, how inner membrane complexes like AcrAB or HlyBD connect with TolC to bypass the periplasm, believed to measure at least 130 Å across, if TolC had a porin-like structure and only spanned the outer membrane. TolC, however, revealed a fundamentally different structure as it is a homotrimer forming a hollow tapered cylinder 140 Å in length (Koronakis *et al.*, 2000) (see Figure 1.4). It comprises a 40 Å long outer membrane β-barrel, the so-called channel domain, and a contiguous 100 Å long α-helical barrel, the tunnel domain. These two domains form a continuous tube projecting across the periplasmic space and gave the name for the structure: channel-tunnel. A third, mixed α/β-structure forms a "strap" around the midsection of the helical barrel and is called the equatorial domain. TolC thus provides a water-filled exit duct with an internal diameter of 35 Å, a volume of roughly 43,000 $Å^3$ and a cross section area of 960 $Å^2$ (which is 15-fold larger than that of the general diffusion pore OmpF). However, the tunnel diameter, which is uniform for a length of 100 Å, decreases to a virtual close at its periplasmic entrance (Koronakis *et al.*, 2001).
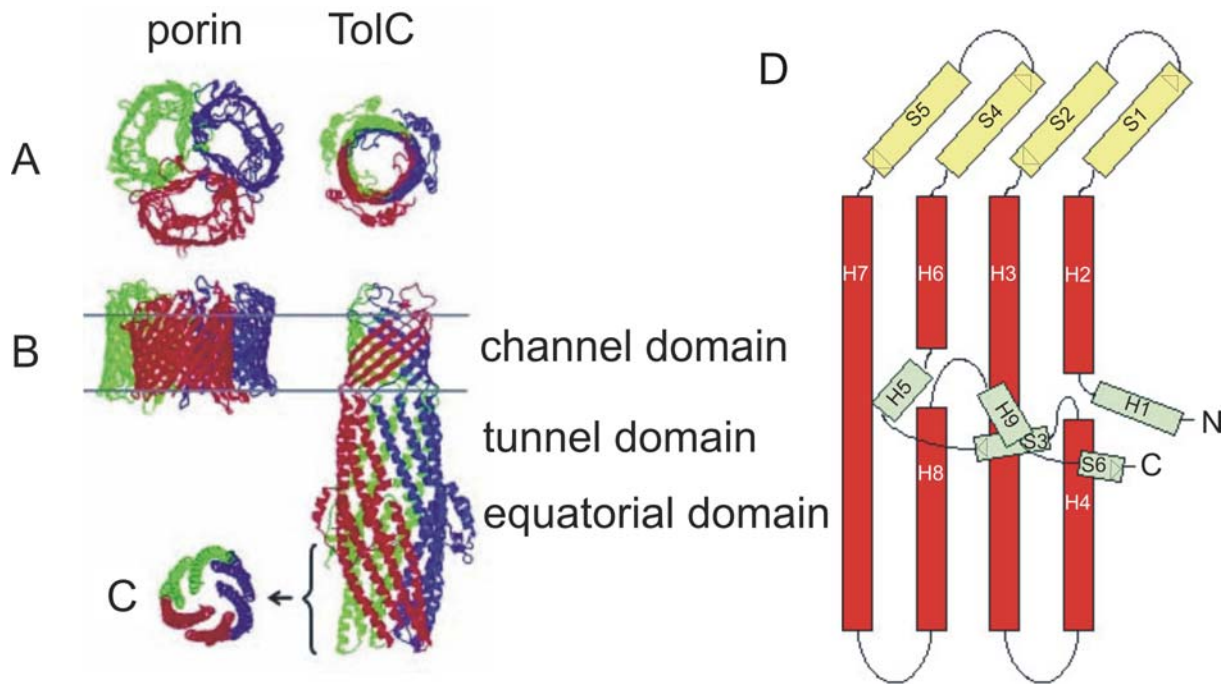
Figure 1.4: The structures of porin and TolC. (A) Top view of the proteins down the OM channel. In the case of TolC this extends to ~2/3 of the height, i.e. the top 100 Å. (B) Side view, at right angles to the plane of the outer membrane. (C) Cross-section of TolC near the tunnel entrance, i.e. the bottom 40 Å. Red, green and blue indicate individual monomers (Adapted by permission from Macmillan Publishers Ltd: EMBO reports, Andersen *et al.*, 2000). (D) Schematical presentation of TolC. Red, yellow and green indicate the α-helical tunnel, the channel and the equatorial domain, respectively.

TolC forms a 12-stranded barrel, where each of the three monomers adds four antiparallel β-strands (S1, S2, S4, and S5) to the channel domain, and four α-helical strands [formed by two long helices (H3 and H7) and four shorter helices (H2, H4, H6 and H8)] to the tunnel domain. The equatorial domain is formed by small N- and C-terminal α-helical and β-strand structures (H1, H5, H9, S3, and S6) (Koronakis *et al.*, 2000).

TolC is anchored in the outer membrane with the channel domain consisting of 12 β-strands forming a right twisted barrel, seen from the top. Alternating polar residues pointing to the inner side of the channel and apolar residues at the outside of the barrel facing the lipids lead to typical amphipathic β-strands. To form the barrel, the β-strands must both curve and twist. The dense packing of the inner residues necessary to accommodate the curvature is enabled by small or unbranched side chains facing the interior. The region of the 8- and 21-residues-long extracellular loops between the strands S1 and S2, and S4 and S5, respectively, show high crystallographic thermal disorder factors and a weaker electron density corresponding with a high flexibility of the loops (Andersen, 2003). A restriction of the channel entrance

from the extracellular side by these loops as observed in other channel-forming proteins (Koebnik *et al.*, 2000) is unlikely (Andersen, 2003).

The tunnel domain is a unique structure and consists entirely of α-helices. The upper part of the tunnel forms an α-barrel, assembled by each helix packing laterally with two neighbouring helices. The resulting "knobs-into-holes" interactions (Crick, 1953) at the two interfaces reflect overlapping seven-residues motifs, so-called heptadic repeats, in the primary sequence (Koronakis *et al.*, 2000; Calladine *et al.*, 2001). Unlike helices in conventional two-stranded coiled coils, the helices in this part of the TolC tunnel domain have to untwist to assemble into the barrel. As a consequence, the helices can lie on a cylindrical surface and this may be further facilitated by the bulkier sidechains tending to lie on the outside of the barrel (Koronakis *et al.*, 2001). The α-barrel structure, together with the small helices H2 and H6, ends with the equatorial domain. Below the equatorial domain, the long helices H3 and H7 pair with helices H4 and H8 and form conventional coiled coils. The long helix H3 is straight and helix H4 coils around it, forming the outer coiled coil. The other coiled coil formed by helix pair H7/H8 bends inwards and is responsible for the tapering of the periplasmic end of TolC. Seen from the top, the arrangement of the coiled coils resembles an iris, which is almost sealing the periplasmic entrance (Andersen, 2003).

The remarkable structure of TolC correlates with its properties in planar lipid bilayer experiments. The asymmetric architecture of TolC leads to a directed insertion with the channel domain first into artificial membranes, which allows the interpretation of the channel characteristics in respect to its membrane orientation (Andersen, 2003). The narrow periplasmic entrance with an inner diameter of almost 4 Å is responsible for the low single channel conductance of around 80 pS although the inner diameter of the upper part of the tunnel with a diameter of 35 Å shows a bigger cross sectional area than OmpF (Koronakis *et al.*, 2001). The TolC channel-tunnel is cation-selective particularly due to negatively charged aspartate residues lining the tunnel entrance. The substitution of alanines for these residues leads to anion-selective TolC channel-tunnels (Andersen *et al.*, 2002c). Additionally the aspartate residues are shown to form an ion-binding site, and to be responsible for pH-dependent closure of the TolC channel-tunnel. The ion-binding site was also proven to bind reversibly di- and trivalent cations like Ca^{2+}, Zn^{2+} or Tb^{3+} resulting in blockage of the potassium ion flux through TolC. The binding site was accessible for the di- and trivalent cations only from the channel mouth and not through the narrow tunnel entrance (Andersen *et al.*, 2002c).

If the tunnel entrance is too small even for big cations, then it undoubtedly has to be opened for the passage of exported substances in transport complexes. The closed state of TolC was observed to be very stable in electrophysiological measurements. A circular network of interactions between three residues was identified to maintain this configuration: aspartate 153 at the outer coiled coil forms an intramolecular hydrogen bond with tyrosine 362 located at the inner coiled coil of the same monomer. Additionally, the aspartate residue also forms an intermolecular salt bridge with arginine 367 of the adjacent monomer. The exchange of tyrosine 362 for phenylalanine and of arginine 367 for serine leads to disruption of these connections and to the opening of the tunnel entrance. Measurements of the mutant TolC revealed a tenfold higher single channel conductance of the channel-tunnels compared to the wild-type TolC, but the recordings showed a low stability of the open configuration (Andersen *et al.*, 2002b).

1.5.3. The TolC family

The TolC family is widespread among Gram-negative bacteria (Andersen *et al.*, 2000), which underlines the importance of this outer membrane protein for the cells. On the basis of a sequence alignment of 36 TolC homologues, Andersen *et al.* (2000) proposed a phylogenetic tree of the TolC family that reflects the amino acid sequence relatedness of the homologues. Interestingly, the sequence similarity among the TolC protein family correlates with the substrate specificity of the exporter system the TolC homologue is belonging to. As one can see in Figure 1.5, the proteins can be grouped into three subfamilies corresponding to their roles in protein export, drug efflux and cation efflux (Andersen *et al.*, 2000).

TolC of *E. coli* is an exception, as it belongs to the protein secretion subfamily, but as mentioned above, TolC of *E. coli* is functional as outer membrane component of both protein secretion and multidrug efflux systems. It was shown to cooperate with the hemolysin A inner membrane secretion complex HlyBD (Wandersman and Delepelaire, 1990) and the colicin V secretion complex CvaAB (Hwang *et al.*, 1997). Additionally, TolC is employed by the multidrug efflux systems AcrAB (AcrB: RND transporter, AcrA: adaptor protein; Fralick, 1996), EmrAB (EmrB: MF transporter, EmrA: adaptor protein; Lomovskaya and Lewis, 1992), most likely EmrJK (ErmK: MF transporter, ErmJ: adaptor protein; Nishino and Yamaguchi, 2001), and MacAB (MacA: ABC transporter, MacB adaptor protein; Kobayashi *et al.*, 2001). This exceptional functionality profile coincides with an exceptional genomic organization. While TolC homologues and their inner membrane partners are usually encoded

within the same operons, TolC of *E. coli* is a member of the stress-induced *mar-sox* regulon (Aono *et al.*, 1998) and not related to any efflux/exporter operon.
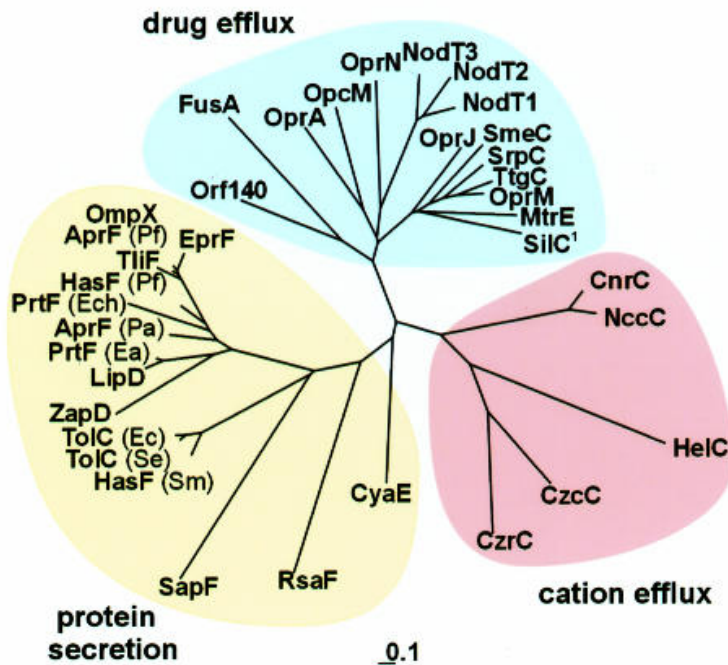
Figure 1.5: Phylogenetic tree of the TolC family. Abbreviations: Ea, *Erwinia amylovora*; Ec, *Escherichia coli*; Ech, *Erwinia chrysanthemi*; Se, *Salmonella enteritidis*; Sm, *Serratia marcescens*; Pa, *Pseudomonas aeruginosa*; Pf, *Pseudomonas fluorescens*. Scale '0.1' indicates 0.1 nucleotide substitutions (Reprinted by permission from Macmillan Publishers Ltd: EMBO reports, Andersen *et al.*, 2000).

In the early evolution of the TolC family a gene duplication is supposed, because for all homologues the N- and the C-terminal halves of the proteins can be sequentially aligned and also structurally superimposed. This was first recognized for CyaE, the TolC homologue of *Bordetella pertussis* (Gross, 1995), and further investigated for the whole TolC family (Johnson and Church, 1999). It is remarkable that of all known homologues CyaE shows the highest primary sequence identity between the two halves of the protein (29,1%). It is also nearest to the root of the family tree, suggesting that it is closest to the family progenitor (Andersen *et al.*, 2000).

Highly conserved characteristics of TolC homologues are the length of the tunnel-forming helices and several structurally important residues at the transitions between the different parts of the structure. These are glycine residues located at the tip of the coiled-coils, necessary to enable the tight turn of the helices, proline and glycine residues at the transition

between the α-helices of the tunnel domain and the β-strands of the channel domain, and aromatic residues, forming an aromatic ring around the bottom of the channel domain. Other conserved residues common to all homologues are small residues, such as alanine and serine, at the interface of the coiled coils that allow a very dense packing that determines the tapering and closure of the tunnel entrance. Conserved residues are also aspartates, which line the tunnel entrance and have strong influence on the electrophysiological behavior of the channel-tunnels (Andersen, 2003). A special characteristic of the drug efflux subfamily members is a conserved cysteine residue at the N-terminus of the mature protein, that constitutes an acylation site and serves as a membrane anchor (Andersen *et al.*, 2000). Studies with mutants of OprM of *P. aeruginosa*, however, showed that this acylation site is not needed for functionality (Li and Poole, 2001; Yoneyama *et al.*, 2000). Variability of the overall length of the TolC homologues from 414 to 541 amino acids is mostly due to variable extensions at the N- and C-terminus. Significant sequence gaps or insertions only occur in the equatorial domain or the extracellular loops (Andersen *et al.*, 2000).

1.6. Fusing, bridging or linking? The role of the adaptor proteins

The adaptor proteins were previously named membrane fusion proteins, based on weak similarities to a protein of paramyxovirus involved in virus penetration, hemolysis and cell fusion (Dinh *et al.*, 1994). As mentioned, it seemed quite mysterious how the TolC-dependent multidrug efflux pump or protein secretion systems bridge the periplasmic space for substrate transport. Therefore, without knowledge of the structures of the involved proteins, the adaptor proteins were supposed to form the bridge between the inner and outer membrane or to pull the two membranes together to enable the contact of the transporter and TolC (Dinh *et al.*, 1994; Johnson and Church, 1999). Solving of the structures of representatives of all three involved protein families revealed that the periplasmic domain of the inner membrane transporter and the tunnel domain of TolC could bridge the periplasmic space without any fusion or constriction of the two membranes. The role of the membrane fusion proteins is now seen as linking the inner membrane transporter to the outer membrane channel-tunnel and to stabilize this assembly. Furthermore, they have to induce the opening of the periplasmic entrance of the cannel-tunnel to enable substrate transport (Andersen, 2003). Regarding the membrane fusion proteins as dynamic adapters between channel-tunnels and inner membrane transporters, they are called adaptor proteins throughout this work.

The first crystallized member of the adaptor protein family was MexA of *P. aeruginosa* (Akama *et al.*, 2004a; Higgins *et al.*, 2004) (see Figure 1.6). In parts the structure of the MexA monomers verified the former structural model of adaptor proteins (Johnson and Church, 1999) showing a 47 Å long α-helical hairpin domain connected to a flattened β-sandwich domain folded like the known lipoyl domain from biotinyl/lipoyl carrier proteins. Beside these already predicted domains, a third domain, the α/β domain could be solved showing a six-stranded β-barrel with a short α-helix. It is expected that there exists at least a fourth domain comprising the N- and C-terminus of the protein because the structure of the 28 N-terminal and 101 C-terminal residues could not be solved (Higgins *et al.*, 2004). The second representative adaptor protein that could be crystallized recently is AcrA of *E. coli* (Mikolosko *et al.*, 2006) (see Figure 6). The three-dimensional structure of AcrA strongly resembles MexA as expected considering their high sequence identity (Mikolosko *et al.*, 2006). The oligomerisation state in the assembled export complexes, however, is not clear. MexA crystallized as tri-decamer, AcrA packs as an apparent dimer of dimers. It is unlikely that one of these assemblies represents the native form of the proteins and thus the stoichiometry of the adaptor protein in the efflux/export complex will be a point of discussion in this work.
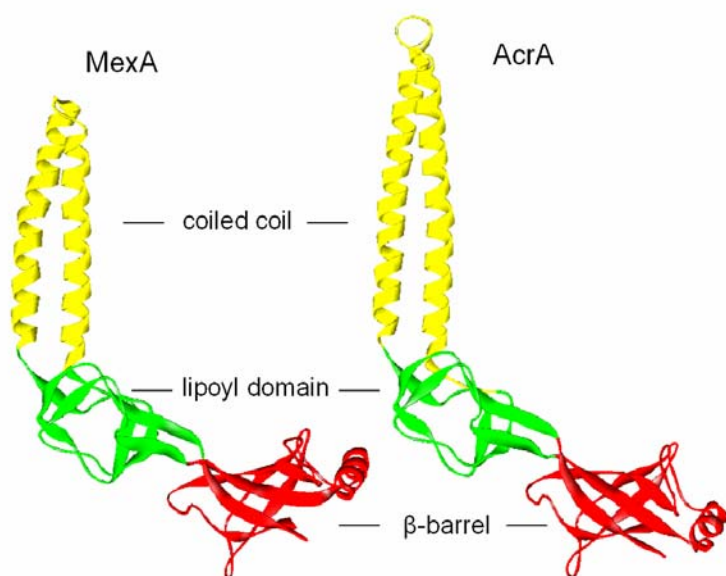
Figure 1.6: Ribbon representation of monomers of MexA and AcrA (by Andersen).

1.7.   A second mission: the homologue type I protein secretion system

Besides multidrug efflux pumps, Gram-negative bacteria posses a homologuous system for the secretion of proteins, called type I secretion system. Protein secretion itself is necessary and useful for bacteria in several fields such as nutrient acquisition or expression of virulence factors. Six groups of protein secretion pathways are known in Gram-negative bacteria (Thanassi and Hultgren, 2000). The type II and type IV secretion system, the chaperone/usher and the autotransporter secretion pathway are sec-dependent transport pathways. They use the general secretory pathway (GSP) for transport across the inner membrane, where proteins with a cleavable N-terminal secretion signal are directed to the sec-system, which transports them into the periplasmic space (Fekkes and Driessen, 1999; Manting and Driessen, 2000). The signal sequence is cleaved after or during the translocation. The subsequent transport of the proteins across the outer membrane into the cell exterior differs between the four groups, but the existence of a periplasmic intermediate is a common characteristic of these secretion pathways. In contrast, the translocation of substrates by type I and type III secretion systems are sec-independent and no periplasmic intermediates are found, indicating a direct transport of the substrates across both inner and outer membrane. The complex type III secretion apparatus consists of approximately 20 - 25 proteins spanning both membranes and ending in a needle-shaped structure projecting out of the cell surface and providing direct access from the bacterial to the host cell cytoplasm (Cornelis, 2002; Johnson *et al.*, 2005). It translocates antihost factors into the cytosol of target eucaryotic cells and is essential for the pathogenicity of several bacteria such as *Yersinia* or *Salmonella* (Hueck, 1998; Cheng and Schneewind, 2000). The composition of the type I system is less complicated as it is – like the multidrug efflux pumps – composed of only three proteins: an inner membrane ABC transporter, which forms a complex with an adaptor protein in the periplasmic space, and an outer membrane protein of the TolC family. The type I secretion system can be divided into two subtypes dependent of the location of the secretion signal in the transported protein (Andersen, 2003). The classic type I secretion system translocates proteins with a secretion signal located within the last 60 C-terminal residues, which is not cleaved after secretion (Stanley *et al.*, 1991). Substrates of these subtype systems are for example RTX (repeats in toxin) toxins such as hemolysin of *E. coli* or adenylate cyclase of *B. pertussis*, extracellular enzymes such as proteases or lipases, or S-layer proteins or certain glycanases, which remain attached to the cell surface (Wandersman and Delepelaire, 1990; Glaser *et al.*, 1988; Duong *et al.*, 1994; Thompson *et al.*, 1998; Awram and Smit, 1998; Finnie *et al.*, 1997). The size of the

transported proteins varies between a few hundred and more than 8,000 residues (Delepelaire, 2004). The second subtype is described to be homologous to the classic subtype and responsible for the export of colicins and microcins (Hwang *et al.*, 1997; Garrido *et al.*, 1988; Azpiroz *et al.*, 2001). The substrates are too short to carry a C-terminal secretion signal of 45 residues and it is known that they are processed and a N-terminal signal sequence is cleaved (Havarstein *et al.*, 1994; Lagos *et al.*, 1999).

The hemolysin export apparatus HlyBD/TolC of *E. coli*, however, is the prototype of the type I secretion system. The genes for the 110 kDa hemolysin HlyA, the ABC transporter HlyB, the adaptor protein HlyD and the acyl transferase HlyC, which activates the HlyA protoxin by fatty acylation at two lysine residues (Stanley *et al.*, 1994), are organized in the *hly*-operon (Andersen, 2003). For the HlyB transporter, only the structure of the cytoplasmic domain, containing the nucleotide-binding site, could be solved (Kranitz *et al.*, 2002). The second domain is the membrane domain consisting presumably of six transmembrane helices (Andersen, 2003). The structure of this membrane domain eventually resembles that of the crystallized distantly related homologue MsbA (Chang and Roth, 2001). The adaptor protein HlyD has three domains: a 59-residue-long N-terminal cytoplasmic domain, which is linked by a 21-residue-long transmembrane domain to a large periplasmic domain that comprises residues 81-478 (Andersen, 2003). HlyB and HlyD form a stable inner membrane complex and both proteins are shown in *in vivo* cross-linking experiments to bind independently the substrate HlyA. Furthermore, the binding of HlyA induces the bridging of the HlyBD complex to TolC via HlyD (Thanabalu *et al.*, 1998; Balakrishnan *et al.*, 2001). This bridging seems to be dynamic as TolC and the inner membrane HlyBD complex disengage after hemolysin passage. This is a clear difference to the AcrAB/TolC multidrug efflux pump, which seems to be assembled permanently, independent of the presence or absence of substrates (Touze *et al.*, 2004; Tamura *et al.*, 2005).

This work deals with channel-tunnel dependent multidrug efflux pump and type I secretion systems, more concrete with the improved classification of the adaptor protein family, the characterization of the TolC-homologue protein HI1462 of *H. influenzae*, and the molecular characterization of the interaction between TolC and AcrA of *E. coli*.

Chapter 2

Classification of adaptor protein family reveals two types of channel-tunnel dependent machineries in Gram-negative bacteria

2.1. Abstract

Adaptor proteins (APs), also known as membrane fusion proteins mediate the contact between outer membrane channel-tunnels and diverse transporters in the inner membrane to form functional export machineries in Gram-negative bacteria. The classification of APs in the database is unordered and in parts incorrect. By sequence analysis, we have grouped the AP superfamily in seven families and 32 subfamilies, which can be identified by specific sequence signatures. An alignment of the consensus sequences of all subfamilies shows that the seven families divide into two groups distinguishable by the C-terminus and the length of the hairpin domain. Interestingly, APs with an extra C-terminal domain and short hairpin domains form complexes with transporters that are characterized by periplasmic extensions, whereas APs with short C-terminus and long hairpin domains pair with transporters lacking periplasmic domains. Based on the consensus sequence alignment and the solved structure of MexA of *Pseudomonas aeruginosa* we built structural models of APs, which explain how APs connect different transporter types with the outer membrane channel tunnel. The oligomeric state of APs and transporters is discussed and models of two distinct export systems are presented.

2.2.   Introduction

In Gram-negative bacteria, exported substances have to cross two membranes. The channel-tunnel dependent export machineries provide a pathway from the inner membrane across the outer membrane bypassing the periplasmic space. Examples of channel-tunnel dependent export machineries are the type I secretion systems for the export of proteins and multidrug efflux pumps, which expel noxious substances providing resistance (Delepelaire, 2004; Zgurskaya and Nikaido, 2000b; Andersen *et al.*, 2001). Channel-tunnels are the outer membrane component of the tripartite export machineries. The 140 Å long canon shaped structure protrudes into the periplasmic space providing an exit duct through the periplasm and the outer membrane for exported substances (Koronakis *et al.*, 2000). Channel-tunnels interact with inner membrane complexes, which are composed of proteins belonging to two different protein superfamilies. One is a transporter, which energizes the export process. Different types of transporters can be involved in channel-tunnel dependent export machineries. In the case of the protein secretion systems, the transporter belongs to the ABC (ATP binding cassette) transporter superfamily. Efflux pumps are driven by transporters belonging either to the RND (resistance nodulation cell division) superfamily, to the MF (major facilitator) superfamily, to the putative extrusion transporters (PET), or to the superfamily of ABC transporters (Andersen 2003).

The second protein, which forms a complex with the transporter, is a so-called adapter protein (AP), also known as membrane fusion proteins (MFP). The AP superfamily is very divergent. In some cases, there is almost no evidence for homology between different APs. In the databases their classification is not uniform and in some cases misleading. Here we have analysed the AP superfamily. We have determined sequence motifs, which helps to identify different subfamilies of APs. In respect of the growing number of sequenced genomes, this study contributes to an improved classification of APs concomitant with a better prediction for the secretion process they are involved in. Additionally, this detailed classification gives insights into structural differences of APs and provides the basis for structure prediction for whole export apparatus assemblies.

2.3. Materials and Methods

2.3.1. Database search

Sequences were collected in the PIR-NREF Database (Protein Information Resource Non-Redundant Reference Sequence Database; Wu *et al.*, 2003). Signatures were designed using the PROSITE syntax and the PIR signature search tool PIR Pattern/Peptide Match (http://pir.georgetown.edu/pirwww/search/patmatch.html) was the tool for collecting sequences matching the designed signature.

2.3.2. Multialignments

Sequences were aligned using the MULTIALIN-platform of PBIL (Pole Bio-Informatique Lyonnaise) (Corpet 1988) and the Homology Module of the InsightII software (Accelrys). The matrix used is BLOSUM62 with variable parameters for gap length and gap penalty. At few positions, multialignments are manipulated subsequently by eye. Consensus sequences were taken from the resulting multialignments and are manipulated for further investigations (see text).

2.3.3. Protein modelling

Proteins were modelled using the homology module of the Insight II software. As template served the MexA structure (PDB entry: 1T5E, Higgins *et al.*, 2004). The models of the secretion systems were built including the structural data of AcrB (PDB entry 1IWG; Murakami *et al.*, 2002), of HlyB nucleotide binding domain (PDB entry 1MTO; Schmitt *et al.*, 2003), and of the channel-tunnel TolC (PDB entry 1EK9; Koronakis *et al.*, 2000).

2.4. Results and Discussion

2.4.1. The AP superfamily groups into seven families

The basis of the bioinformatical analysis is the collection of over 1000 sequences found after extensive search for several criteria of different classification systems in the PIR-NREF database. Sequence alignments helped to group the protein sequences into 7 families and 32 subfamilies. The division into seven families is in accordance with different transporter types, which interact with the APs. Therefore, we choose to name the AP families according to the name of the transporter family the APs are interacting with as annotated in the transporter classification database (http://tcdb.ucsd.edu/): HAE1 (hydrophobe/amphiphile efflux-1), HME (heavy metal efflux), ME (macrolide exporter), ProtE (protein exporter), DHA2 (drug:H+ Antiporter-1 (14 spanner)), PET (putative efflux transporter), and AMTS (ABC-type multidrug transport system). The subfamilies are named if possible after the best-characterised member. Most APs belong to the HAE1 AP family (37%), followed by the ProtE family (21%). The two AP families DHA2 and ME comprise each about 10% of the sequences followed by the ME, AMTS, and HME AP family (7%, 7%, and 6%, respectively). The majority of the AP sequences (84%) are from proteobacteria. However, APs are also found in evolutionary older bacterial species as the cyanobacteria or bacteria of the CFB (Cytophaga-Flavobacterium-Bacteroides) group (Table 2.1). It is astonishing that APs are also present in bacteria lacking an outer membrane.

Signatures were defined, which are characteristic for different groups. Thus, almost 90 percent of the sequences could be detected (see supplementary material). The remaining sequences are assigned by similarity to certain families and subfamilies. For this work, the signatures helped to detect new sequences not identified as APs before. In future, they will help to annotate genomes more precisely. Furthermore, the classification of the AP superfamily gives an insight into the evolution of the AP family and makes it easier to predict putative functions of AP dependent export apparatus. Here, we focus on the AP classification as the basis for structure predictions, structure modelling, and structure/function relationships.

Table 2.1: Characteristics of the different AP subfamilies

AP Family Subfamily prototype	Species	N-term.	hairpin length	heptads	Reference
ProtE					
LssD	*L. pneumophila*	α,β,γ,δ	cd	139-162	3+5(6) ; 8(9)	Jacobi & Heuner, 2003
LapC	*P. fluorescens*	α,β,γ	cd	189-209	3+8 ; 11	Hinsa et al., 2003
AprE	*P. aeruginosa*	α,γ,ε	cd	192-196	3+8 ; 11	Duong et al., 1992
HlyD	*E. coli*	α,β,γ	cd	162-201	4+7(5) ; 11(9)	Thanabalu et al., 1998
Alr4240	*Nostoc sp.*	c	cd	124-293	var.	Kaneko et al., 2001
CvaA	*E. coli*	γ	cd/ss	158	2 x 11	Hwang et al., 1997
RaxA	*X. oryzae*	α,β,γ	cd	158-159	2 x 11	da Silva et al., 2004
ComB	*S. pneumobiae*	f	cd	149-224	2 x 9 + insert	Hui et al., 1995
AMTS					
VPA0490	*V. parahaem.*	α,β,γ	cys	126	2 x 9	Makino et al., 2003
YbhG	*E. coli*	α,β,δ,γ,c, cfb	ss	126-129	2 x 9	Blattner et al., 1997
YhiI	*E. coli*	α,β,γ,δ,cfb	ss	130 (102)	2 x 9 (7)	Blattner et al., 1997
PA3402	*P. aeruginosa*	γ,ε,cfb	ss	126	2 x 9	Stover et al., 2000
DHA2					
EmrA	*E. coli*	α,β,γ	cd	120 (134)	2 x 8 (9)	Lomovskaya & Lewis, 1992
VceA	*V. cholera*	α,β,γ	cd	146-148	2 x 10	Colmer et al., 1998
RmrA	*R. etli*	α,β,γ,cfb	cd	130 (114-144)	2 x 9 (8 or 10)	Gonzalez & Martinez, 2000
PET					
AaeA	*E. coli*	α,β,γ	ss	70-76	2 x 5	Van Dyk et al., 2004
YjcR	*E. coli*	γ	ss	126	2 x 9	Blattner et al., 1997
PA3360	*P. aeruginosa*	γ	cd/ss	119-121	2 x 8	Stover et al., 2000
YiaV	*E. coli*	γ	ss	116-124	2 x 8	Blattner et al., 1997
HAE1					
AcrA	*E. coli*	α,β,γ,δ,ε,cfb	cys	74 (60)	2 x 5 (4)	Ma et al., 1995
YegM	*E. coli*	α,β,γ,δ,c	ss	72-74	2 x 5	Baranova & Nikaido, 2002
MexH	*P. aeruginosa*	α,β,γ,δ,c,cfb	ss	60 (74)	2 x 4 (5)	Aendekerk et al., 2002
All3144	*Nostoc sp.*	c	cys	74-186	2 x 5-8	Kaneko et al., 2001
MexJ	*P. aeruginosa*	α,β,γ	cys	74-75	2 x 5	Chuanchuen et al., 2002
VC1674	*V. cholerae*	γ	cys	60-61	2 x 4	Heidelberg et al., 2000
HP0606	*H. pylori*	ε		58-72	2 x 4-5	Tomb et al., 1997
HME					
CusB	*E. coli*	α,β,γ,δ,cfb	ss	51 (34-64)	2 x 3 (var.)	Franke et al., 2003
CzrB	*P. aeruginosa*	α,β,γ,δ,c,cfb,spi	ss	75 (5-155)	2 x 5 (var.)	Hassan et al., 1999
ME					
MacA	*E. coli*	α,β,γ,δ,ε,cfb	ss	88	2 x 6	Kobayashi et al., 2001
VP1999	*V. parahaem.*	γ,cfb	ss	81-83	2 x 6	Makino et al., 2003
YvrP	*B. subtilis*	f	ss	79-112	var.	Wipat et al., 1998
Alr1505	*Nostoc sp.*	c	ss	60-202	var.	Kaneko et al., 2001
Cg3322	*C. glutamicum*	f,act	ss	148-283	var.	Kalinowski et al., 2003
DevB	*Anabaena sp.*	c	cd/ss	75-233	var.	Fiedler et al., 1998

The subfamilies are named according to the best-characterized member of each subfamily. The species column lists the bacterial groups, which are represented in each AP subfamily (α-,β-,γ-,δ-,ε- proteobacteria, Cyanobacteria (c), Firmicutes (f), Actinomycetes (act), Cytophaga-Flavobacterium-Bacteroides (cfb), and Spirochaetes (spi). The N-term. column shows the nature of the N-terminus, either a cytoplasmic domain (cd), a signal sequence (ss), or a signal sequence ending with a cysteine residue (cys). The column heptads lists the number of predicted heptadic repeats in the hairpin domain. In all families except for the ProtE AP family two helices with identical number of heptadic repeats form the hairpin. In the ProtE AP family the first helix of the hairpin can be interrupted by a proline rich domain. For several subfamilies it was not possible to predict the number of heptadic repeats because of the variability (var.) in hairpin length.

2.4.2. The consensus sequence alignment

Multiple sequence alignments give information about the grade of conservation of individual parts of the protein and detect positions where deletions or insertions occur. The consensus sequence of each AP subfamily alignment is shown schematically as a bar in Figure 2.1. The thickness of the bar corresponds to the relative number of sequences present at the position and the colour indicates the grade of conservation from yellow (low conservation) to dark red (high conservation). In a next step, we aligned all consensus sequences of the different subfamilies to get information about the relationship between the individual subfamilies (see supplementary material). Therefore, we manipulated the consensus sequences in a way that regions of insertions, which are present in only a few members of the subfamilies, were omitted. The advantage of this approach is that these insertions do not interfere when comparing the subfamilies. The information about insertions, which is lost by this procedure, is integrated again by taking the consensus sequence alignment as basis to arrange the individual subfamily consensus sequences leading to gaps within the alignments (Figure 2.1).
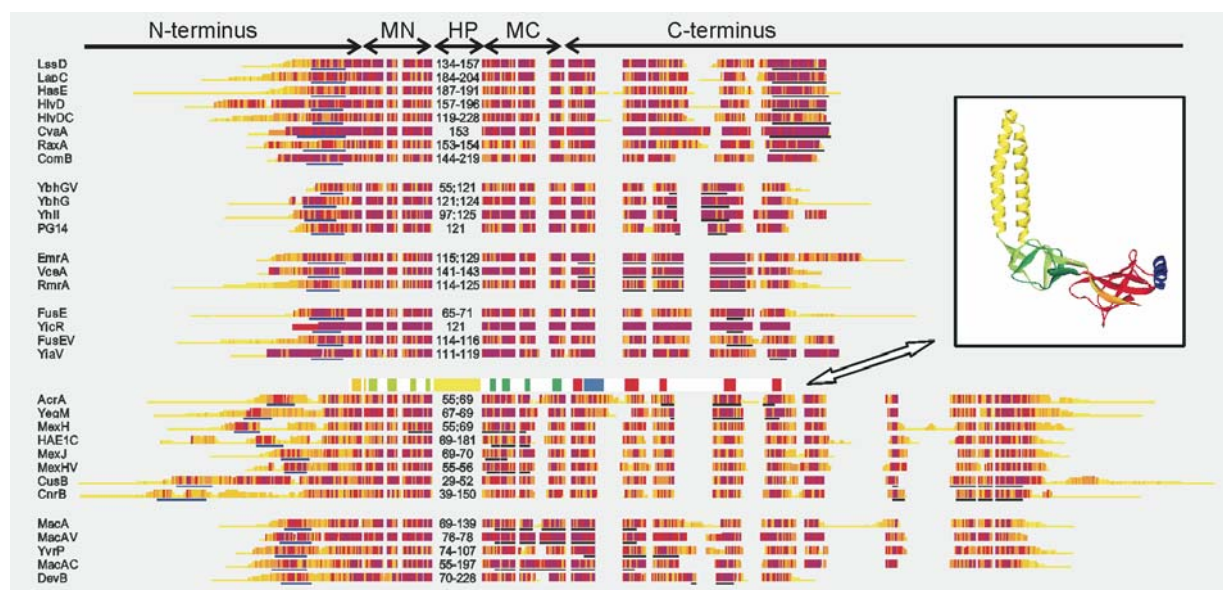
Figure 2.1: Alignment of the consensus sequences of AP subfamilies. The consensus sequences are shown schematically as a bar except the hairpin domain, which size is given as numbers of residues. The thickness of the bar corresponds to the relative number of sequences present at the position and the colour indicates the grade of conservation from yellow (low conservation) to dark red (high conservation). The blue bar below the consensus sequences close to the N-terminus marks the position of the hydrophobic region, the black bar marks the area used to define the signature of each AP subfamily. The insert shows the structure of MexA of *P. aeruginosa* (Higgins et al., 2004), which is divided in three domains, the lipoyl domain (green), the α-helical hairpin (yellow), and the β-barrel domain (red) comprising a short helix (blue). The coloured bar above the AcrA subfamily consensus sequence shows the overlay of the secondary structure elements from MexA with the multialignment of the consensus sequences.

What all APs have in common are two highly conserved motifs MN and MC with a length of about 30 residues, each homologue to a half of lipoyl domain found in 2-oxo acid dehydrogenases or biotinyl carboxyl carrier proteins, respectively (Neuwald *et al.*, 1997; Johnson and Church, 1999). In the following the region upstream of MN is referred as the N-terminus, downstream of MC as the C-terminus, and the domain between the two motifs is referred to as the hairpin domain. The length of the hairpin domain is highly variable and varies between 34 and 293 residues in the AP superfamily. For this reason, the sequence of the hairpin domain was left out when comparing the consensus sequences and it is not presented as bar in Figure 2.1. The importance of this domain and the reason for its variable size is analysed below.

The seven families divide in two groups according to the length of the C-terminus. The C-termini of the three AP families HAE1, HME, and ME are in average about 60 residues longer than the C-termini of the other four AP families. Another feature, which is also different within the two groups, is the size of the peptide chain between the hydrophobic domain in the N-terminus and the MN motif. Within the four AP families ProtE, DHA2, AMTS, and PET the number of residues between the hydrophobic domain and the start of the lipoyl domain lies between 40 and 50 residues whereas in the other three families the peptide chain between these two positions is at least 20 residues longer.

Generally, the N-terminus is the most variable area within the whole AP family. One has to consider that the alignment includes also sequences with extended N-terminus due to incorrect initiation codon assignment, which might enhance this observation. However, it is not clear for all APs how and if at all, they are anchored by their N-terminus in the inner membrane. One possibility present e.g. in AcrA of *E. coli,* is a cysteine residue acylated by a fatty acid, which serves as a membrane anchor (Zgurskaya and Nikaido, 1999). Conserved cysteine residues at the end of the N-terminal hydrophobic signal sequence are characteristic i.a. for several HAE1 AP subfamilies. Alternatively, the N-terminus forms a cytoplasmic domain, which is connected with the rest of the protein by a transmembrane helix as e.g. observed for EmrA (DHA2 AP family) and HlyD (ProtE AP family) of *E. coli* (Lomovskaya and Lewis, 1992; Schulein *et al.*, 1992). For HlyD it could be shown that the cytoplasmic domain comprises a binding site for the exported substrate (Balakrishnan *et al.*, 2001). Topology predictions reveal that almost all members of the ProtE and DHA2 AP families possess a cytoplasmic domain with a length between 10 and over 80 residues (see Table 2.1). All other APs have an N-terminal signal sequence. However, it is not known, if the signal sequence is cleaved after transport or if it remains connected and serves as membrane anchor.

The grouping into families and subfamilies brings light into the till now disordered AP superfamily. Furthermore, the alignment of consensus sequences shows the relations of all individual AP subfamilies and provides the basis for transferring structural and functional information from one subfamily to others.

2.4.3. *Conserved structure of adaptor proteins*

The alignment of the AP consensus sequences is the basis for structural investigations of the AP superfamily. In the last years, there was good progress in solving structures of proteins involved in channel-tunnel dependent export systems, which lead to a deeper understanding of the transport mechanisms. Until now the structure of the channel tunnels TolC of *E. coli*, OprM of *P. aeruginosa*, and VceC of *Vibrio cholera* are solved, as well as the RND transporter AcrB of *E. coli* (Koronakis, *et al.* 2000; Akama *et al.*, 2004b; Federici *et al.*, 2005). As the missing link between these inner and outer membrane components, the structures of two AP, namely MexA of *P. aeruginosa* (Higgins *et al.*, 2004; Akama *et al.*, 2004a) and AcrA of *E. coli* (Mikolosko et al., 2006) were solved recently. However, it was just possible to solve approximately two-thirds of mature protein. About 30 residues of the N-terminus and about 100 residues of the C-terminus could not be solved. The structures of MexA and AcrA are very similar and the proteins can be divided in three domains (insert Figure 1). As predicted, the two highly conserved motifs MN and MC adopt a lipoyl domain like structure (green) and the domain in between forms an α-helical hairpin (yellow). Unpredicted was the third domain, a β-barrel formed by six β-sheets, one located upstream of the MN-motif (orange), the other five downstream of the MC motif (red).

We have overlaid the structural information taken from the MexA structure with the multialignment of the consensus sequences of the AP subfamilies. It is evident that regions comprising secondary structures are conserved in all subfamilies except the helix in the β-barrel domain (blue), which is found exclusively in the AcrA and YegM subfamily. Gaps in the multialignment are located at positions of transitions between secondary structure elements in the MexA structure. This means that these regions are not so restricted for mutations than the regions forming the β-barrel. The overlay of the MexA structure with the multialignment also shows that the ending of the known MexA structure corresponds with the ending of the APs of the ProtE, AMTS, DHA2, and PET AP family. Exceptional is the DevB subfamily, which has also a short C-terminus, although the rest of the sequence implies a clear membership to the ME AP family. One can suppose that the structures of these APs end with

the C-terminal β-barrel. The C-terminal extension found in the remaining APs forms most likely an independent extra domain.

It is striking that this additional C-terminal domain is found almost exclusively in APs, which interact with transporters having large periplasmic extensions. The structure of the RND transporter as well as the topology prediction of ME transporters show extensive periplasmic domains in contrast to any other transporter interacting with APs. Models of multidrug efflux pumps assume that the hairpin domain interacts with the tunnel domain of the channel tunnel and that the lipoyl and β-barrel domain interact with the upper part of the periplasmic extension of the transporter (Eswaran *et al.*, 2004; Fernandez-Recio *et al.*, 2004; Akama *et al.*, 2004a). The proposed assembly demands a longer peptide chain between the transmembrane anchor and the MN motif of the lipoyl domain to bridge the distance between the membrane plane and the lipoyl domain. In fact, the APs with a C-terminal extension are characterized by such longer N-terminal region. It is tempting to speculate that the extra C-terminal domain assembles with the lower part of the periplasmic extension of the transporter. With the intention to get information about the possible structure of this extra domain, we performed Blast searches in the PDB using consensus sequences of the C-terminal extra domain. All hits revealed areas of proteins with compact domains formed by β-sheets. Therefore, we propose that the extra C-terminal domain is also composed of β-sheets. It might be that the folding also integrates parts of the N-terminus as seen in the β-barrel domain. According to the consensus sequence alignment, which shows five conserved regions in the additional C-terminal part, we propose that the folding might be similar to β-barrel domain including also parts of the N-terminus.

In the case of proteins of the ProtE, AMTS, DHA2, and PET AP family interacting with transporters without periplasmic extensions, one can assume that the distance from the lipoyl domain and the membrane plane is much shorter, which explains the shorter peptide chain in between. Concerning the assembly of the AP with the transporter, it is very likely that the β-barrel domain is interacting with the transporter close to the membrane plane and that an extra C-terminal domain would be a sterical hindrance.

However, the MexA structure can be used as template to built reliable structural models of all APs, at least from the lipoyl and the β-barrel domain based on the multialignment of all AP consensus sequences. The hairpin domain is somehow special because this region varies a lot within AP families. Bioinformatical analysis of the loop domains enlighten the characteristics of different AP families and allow structure predictions of this part of the protein, which could explain its putative function.

2.4.4. *Analysis of the hairpin domain*

The solved structure of MexA has confirmed the previous structure prediction for the hairpin domain (Johnson and Church, 1999). The 60 residues form a 47 Å long α-helical hairpin comprising a straight C-terminal helix and an N-terminal helix with a left-handed superhelical twist. Each helix has eight coils corresponding with two times four heptadic repeats in the amino acid sequence.

The 60 residues long hairpin domain of MexA is one of the smallest within the AP superfamily. APs with hairpins of the same size are found in the HAE AP subfamilies AcrA, MexH, and VC1674 as well as in the Alr1505 subfamily. The majority of APs pairing with RND transporter have hairpin domains, which are 14 residues longer (AcrA, YegM, MexJ, CzrB subfamilies). These hairpins possess one additional heptadic repeat per helix corresponding with a hairpin extension of about 10 Å. Even smaller than the MexA hairpins is the hairpin domain of the HME AP subfamily CusB. Interestingly, CusB belongs to an efflux pump, which is not a tri- but a tetrapartite resistance system involving a small periplasmic copper binding protein CusF, which interacts with the AP CusB and with the channel tunnel CusC (Franke *et al.*, 2003). In most operons comprising AP of the CusB subfamily are genes coding for such small periplasmic proteins. In contrast, no such genes were found near AP genes of the closely related CzrB subfamily. This let us assume that the exceptional small hairpin domain of CusB APs is related with the cooperation of these cation efflux pumps with the fourth component. The typical length of the hairpin domain of ME APs is about 88 residues, which corresponds to six heptadic repeats and a predicted hairpin length of about 67 Å. Exceptionally long are hairpins of APs found in Firmicutes, Actinomycetes or Cyanobacteria (YvrP, Cg3322, Alr1505, DevB, All3144 subfamilies), which is due to multiple insertion and duplication events. However, it is evident that the APs of the three AP families HME, HAE1, and ME, which interact with transporters with periplasmic extension, have generally shorter hairpin domains compared to the remaining APs. The function of the hairpins is most likely to provide contact with the tunnel domain of the outer membrane component. It is obvious that the contact site of the tunnel domain ranges only below the equatorial domain. Assuming a parallel arrangement of the helices of the tunnel and the hairpins the maximal length of the hairpin needed to pair with this part of the channel tunnel is about 50 Å. This means that hairpins of the ME family with a length of about 67 Å are not fully assembled with the lower tunnel domain. In other words, there is an extra part, which might be necessary to bridge the gap between the tunnel entrance and the exit of the transporter. However, the smaller size of the periplasmic loops of ME transporter compared to

RND transporter might be a hint that ME transporter extend not as far as RND transporter into the periplasmic space leading to a wider gap between transporter and channel tunnel.

Consequently, it is not surprising that APs interacting with transporter without periplasmic extension have hairpin domains, which are longer than those of proteins of the ME, HAE, and HME AP family. The PET AP subfamilies YiaV and PA3360 have a hairpin domain length of around 120 residues corresponding with eight heptadic repeats per helix. A similar length characterises proteins of the AMTS AP subfamilies. Eight, nine, and ten heptadic repeats per helix are found in the three DHA2 AP subfamilies EmrA, RmrA, and VceA. This means that APs belonging to the PET, AMTS, and DHA2 AP families have an approximate hairpin length between 90 and 110 Å. Assuming that the upper 40-50 Å of the hairpin are necessary to assemble with the tunnel domain of the channel tunnel, the remaining part of the hairpin has a length, which is similar to the height of the periplasmic domain of RND transporter. This means that it would be enough to bridge the distance between the membrane plane and the tunnel entrance.

The AaeA subfamily is exceptional because it does not match this theory. The hairpin domain is of similar size than the AcrA subfamily namely 70-76 residues corresponding to five heptadic repeats per helix and a hairpin length of around 57 Å, which would be too small to bridge the gap. However, the predicted topology of the PET transporters is ambiguous. Most of the transporters have two transmembrane domains each with six transmembrane helices connected by a long hydrophilic loop. However, there are also examples of transporter with each transmembrane domain predicted to consist of seven transmembrane helices, which would mean that the hydrophilic loop in between is not located in the cytosol but in the periplasm. It might form a similar periplasmic extension as known from RND or ME transporter, which could explain the short hairpins of the AaeA AP subfamily. A detailed analysis is necessary to clarify the topology of the PET transporter family.

In the hairpin domains of all AP families discussed above, the tip of the hairpin is always located in the middle and deletions or insertions are symmetric meaning that N-terminal and C-terminal half possess always equal numbers of amino acids. The pattern A-X(5)-R-X(3)-L-X(11-12)-[ED] matches to the region in the middle of the hairpin domain of the majority of APs characterizing the tip of these hairpins.

The hairpin domains of the ProtE AP family are different. Only in the case of the RaxA subfamily, the pattern characterizing the tip of the hairpin is found in the middle of the hairpin domain suggesting a hairpin like structure. According to the length of the peptide chain (158 residues), the hairpin consists of helices comprising approximately eleven heptadic repeats. It

can be assumed based on the high homology between the RaxA and CvaA subfamily in regions outside the hairpin domain that the CvaA hairpin folds similar. The hairpin domains of the remaining families are different in a way that they contain in the first half proline rich regions and insertion with no mirroring insertions in the second half. This indicates that this region of the hairpin adopt not a helical fold but forms more likely a random coil with variable length. However, the comparison of the hairpin domains of the LssD with the LapC and AprE subfamily shows that the shorter hairpin domains of the LssD subfamily is not because of larger deletions in the proline rich region but because of the deletion of two times 14, 21, or 25 residues in the region downstream of the proline rich region. This reminds on the symmetric deletions/insertion found in the hairpins of the HAE, HME, ME, or PET AP families. Therefore, we conclude that in the middle of this region, which comprises a conserved glycine residue, is the tip of a hairpin. This is further supported by two LapC homologues of *Bordetella* species, which comprises two cysteine residues, one 17 residues in front and another 18 residues behind the conserved glycine. It seems likely that the cysteine residues form a disulfide bond in the folded protein. Downstream of the potential tip are 8 to 9 (LssD) and 11 (AprE and LapC) heptadic repeats, which most likely correspond to the overall extension of the hairpin domain. The calculated length of about 90 to 120 Å would be in the same range as predicted for the PET, DHA2, and AMTS AP family. Upstream of the potential tip the 12 to 25 residues long proline rich region separates a region comprising three heptadic repeats from a region with 5 (LssD) to 8 (AprE and LapC) heptadic repeats. Thus, the number of heptadic repeats upstream of the potential tip matches the number of heptadic repeats downstream within the different ProtE AP subfamilies. The HlyD subfamily differ from the other subfamilies in a way that the proline rich region is shifted one heptadic repeat closer towards the potential tip. Another difference is that the HlyD subfamily has no conserved glycine residue at the tip. However, the sequence homologies with the other subfamilies make it possible to locate the tip. The peptide chain downstream of the potential tip comprises in the HlyD subfamily nine to eleven heptadic repeats. Thus, in contrast to the hairpin domains of all other APs, which are composed of two helices of equal length, one helix of the hairpins of the ProtE AP subfamilies HlyD, AprE, LssD, and LapC is interrupted by an insertion of a proline rich region with variable length. Hairpin domains of the two AP subfamilies found in Firmicutes and cyanobacteria (ComB and Alr4240) are highly variable in length ranging from 149 to 224 and 124 to 293 residues, respectively. A more detailed analysis is difficult because of the small number of proteins. However, for the ComB subfamily at least eight heptadic repeats can be assigned at the beginning and the end of the hairpin domain. The lack of a

proline rich region with variable length suggests that the ComB APs have hairpins similar to that of ProtE subfamilies CvaA or RaxA. The variable insertions at the tip of this hairpin might form extra domains. In this context, it should be considered that ComB APs are found in Firmicutes, which have no outer membrane.

Taken together, our analysis could explain the differences of these highly variable hairpin domains within the APs. The analysis reveals a clear relation between the existence of a periplasmic domain of the transporter and the length of the hairpin structure. Hairpins with a length of 47 to 67 Å (corresponding to 4 to 6 heptadic repeats per helix) characterize APs interacting with transporters with periplasmic domains whereas hairpins with 8 to 11 heptadic repeats are found in APs interacting with transporters without periplasmic extension. Thus, our analysis makes it possible to model hairpin structures of almost all APs. These models make it possible to narrow down the regions, which mediate interaction with the channel-tunnel, and help to identify contact sites of the two proteins.

2.4.5. *The oligomeric state of adaptor proteins and transporters*

It is obvious that neither the spiral assembly of MexA protomers into hexamers and heptamers in the MexA crystal (Akama *et al.*, 2004a) nor the packing into a dimer of dimers in the AcrA crystal (Mikolosko et al., 2006) represents the assembly *in vivo*. Thus, the oligomeric state of APs *in vivo* is still an open question and several models are proposed. One model suggests that three AP monomers are involved in functional drug efflux machineries (Fernandez-Recio *et al.*, 2004). A second model integrates three dimers in the efflux machinery (Akama *et al.*, 2004a). A third model proposes an assembly of nine APs (Higgins *et al.*, 2004). Supporting the first model cross-linking of *in vivo* complexes has identified trimers of the MexA homologue AcrA and HlyD, an AP belonging to the ProtE AP family (Zgurskaya and Nikaido, 2000a; Thanabalu *et al.*, 1998). In contrast, biochemical data determining the subunit stoichiometry of the MexAB/OprM efflux pump reveal that six molecules of MexA are present per trimer of MexB and OprM (Narita *et al.*, 2003), which would support the second model.

However, all models of the MexAB/OprM efflux pump have in common that the helices of the hairpins contact the tunnel domain of the outer membrane component. This means that each helix of the hairpin is in contact with at least two other helices. Indeed, the sequences of the hairpin domains have a specific type of heptadic repeat. Similar to helices found in the tunnel domain of channel-tunnels (Calladine *et al.*, 2001), the hairpin domains show two sequence signatures that are phased to match two interfaces (Figure 2.2). In the case of APs

interacting with transporters with periplasmic extensions, the helices of the hairpins interact with almost their full length with helices of the tunnel domain. Here, it is not astonishing that the sequences have this specific heptadic repeats. However, the heptadic repeats of APs with long hairpins also have this special pattern close to the lipoyl domain, a site, which is very likely not in contact with the channel-tunnel. We conclude that the helices have also two contact sites in this region and that the hairpins interact with each other to form a hollow, conduit like assembly. If APs would be trimeric, the conduit formed by six helices would have an inner diameter of approximately 10 Å, wide enough to transport small molecules but too narrow for the secretion of partly folded peptide chains. For interaction with the tunnel domain the three hairpins need to disassemble, which would lead to larger gaps in the conduit wall. Moreover, it is questionable if it is generally possible that six antiparallel helices form such an α-barrel with a hydrophilic interior. As far as we know no such structure is reported by now, nor is there a theoretical prediction that such an assembly is stable. All these reasons lead to the assumption that a trimeric AP assembly is improbable.


```
                          I       II     III      IV
                       defgabcdefgabcdefgabcdefgabcdefg
HlyD    125   LKLTALGAEADTLKTQSSLLQARLEQIRYQIL   156
YbhG     78   KPYEIALMQAKAGVSVAQAQYDLMLAGYRDEE   109
VceA     96   ARLALLQAEADLALAKRRVRSYLANDEGLSAM   127
YjcR     84   YEANLAKAEASLAALDKQIMLTQRSVDAQQFG   115
AcrA    101   YQATYDSAKGDLAKAQAAANIAQLTVNRYQKL   132
CzrB    219   PQLSDQRSEFAAAQRRLSLAQSTYKREQQLWK   250
MacA     97   AENQIKEVEATLMELRAQRQQAEAELKLARVT   128
```

Figure 2.2: The hairpin domain consists of a specific type of heptadic repeat. (A) Schematic model of two antiparallel helices in an α-barrel. The hairpin domains show two sequence signatures that are phased to match two interfaces. Residues of the heptadic repeat have been labelled in sequence as a, b … g. (B) Section of sequence taken from the beginning of the hairpin domain of a representative member of each AP family. Residues at position a and d, and at positions b and f, respectively form the interface between the helices. Residues at these positions are typically hydrophobic (marked in bold). Arginine and lysine residues are also present at these positions. It is known from the α-barrels of channel-tunnels that these amino acids with its long acyl chains can also be found at the interfaces between the helices.

Models suggesting hexameric or nonameric assemblies are more likely. Here the diameter of the conduit would be large enough for the passage of partly folded proteins. The model of nonameric APs is proposed based on the propensity of MexA to pack side-by-side as observed in the MexA crystal, in which seven monomers form a twisted spiral-arc. A ring consisting of nine monomers was modelled, which has an inner diameter wide enough to encompass the open state of the channel-tunnel (Higgins *et al.*, 2004). What speaks against a nonameric and prefers a hexameric assembly is the symmetry within the whole export machinery. It is predicted that both, RND transporters as well as channel-tunnels are evolved by gene duplication events, which can be seen in the primary sequence as well as the 3D structure (Andersen *et al.*, 2000). This means that there is hexameric symmetry within the outer and inner membrane component. In contrast to a hexameric assembly, nonameric APs would break this symmetry. By now, the only known structure of transporters, which are involved in channel-tunnel dependent export machineries, are RND transporters. However, it is speculated that ME transporter might act as hexamers (Kobayashi *et al.*, 2003). This would mean that the periplasmic extension would consist of six periplasmic loops analogous to the six periplasmic loops of assembled RND transporters. Following the hypothesis, that there is a common symmetry in all three components of the export machineries, we assume that all other transporter families, which form a complex with APs, have also a six-fold symmetry. This seems to conflict with the general opinion, that ABC transporters of type I secretion systems are dimers (Holland and Blight, 1999), which is supported by the solved structure of the nucleotide binding domain (NBD) of HlyB of *E. coli* (Schmitt *et al.*, 2003). However, we propose that ABC transporter involved in type I secretion systems exist as hexamers, composed of three dimers. This does not contradict the finding of dimeric assemblies. The assumption is strongly supported by 2D crystals of ArsA of *E. coli* (Wang *et al.*, 2000). ArsA protein is the soluble subunit of the ArsAB pump, which extrudes arsenite or antimonite from the cytoplasm. ArsA consists of two homologous halves each comprising a NBD (Chen *et al.*, 1986). The 2D crystals show ring shaped assemblies with a diameter of 8.5 nm with a clear threefold symmetry. ArsA interacts with the membrane-bound subunit ArsB comprising 12 transmembrane helices to transport the substrates across the membrane. It is very likely that ArsB also forms trimers. Therefore, the transmembrane part of the fully assembled ArsAB transporter would consist of 36 transmembrane helices, which corresponds well with the proposed hexameric assembly of ABC transporter, which interact with APs.

The bioinformatical analysis reveals good indications that APs form conduit like assemblies that bridge the distance between the exit of the transporter and the tunnel entrance. For

symmetrical reasons it is most likely that the oligomeric state of APs is hexameric. Such an assembly of 12 antiparallel helices forming an α-barrel is already known from channel-tunnels. We propose that all transporter interacting with APs are assemblies with a six-fold symmetry. Therefore, future models of ABC, MFS, AMTS, or PET transporter interacting with APs should assume that these proteins form hexamers or trimers, when the proteins contain internal duplications. By the ongoing progress in membrane protein crystallization, in the near future the solved structure of a transporter of any of the families might proof this hypothesis.

2.4.6. *Models of two types of export systems*

According to the different topology of inner membrane transporters, we modelled two types of export systems showing the different features of the APs in these assemblies (Figure 2.3). The RND driven efflux pump represents an export system involving a transporter with periplasmic extension, whereas the type I secretion system represents an example for export systems comprising transporter without periplasmic extension. In the case of the efflux pump, structural information for the RND transporter is available. The nature of ABC transporter involved in protein secretion is not known. However, based on the assumptions discussed above we have modelled an assembly consisting of six cytoplasmic NBDs and six transmembrane domains each comprising six membrane spanning helices. It is just a rough model but one gets an idea about the dimension of such a hexameric transporter and it becomes visible that it fits well with a hexameric AP assembly.
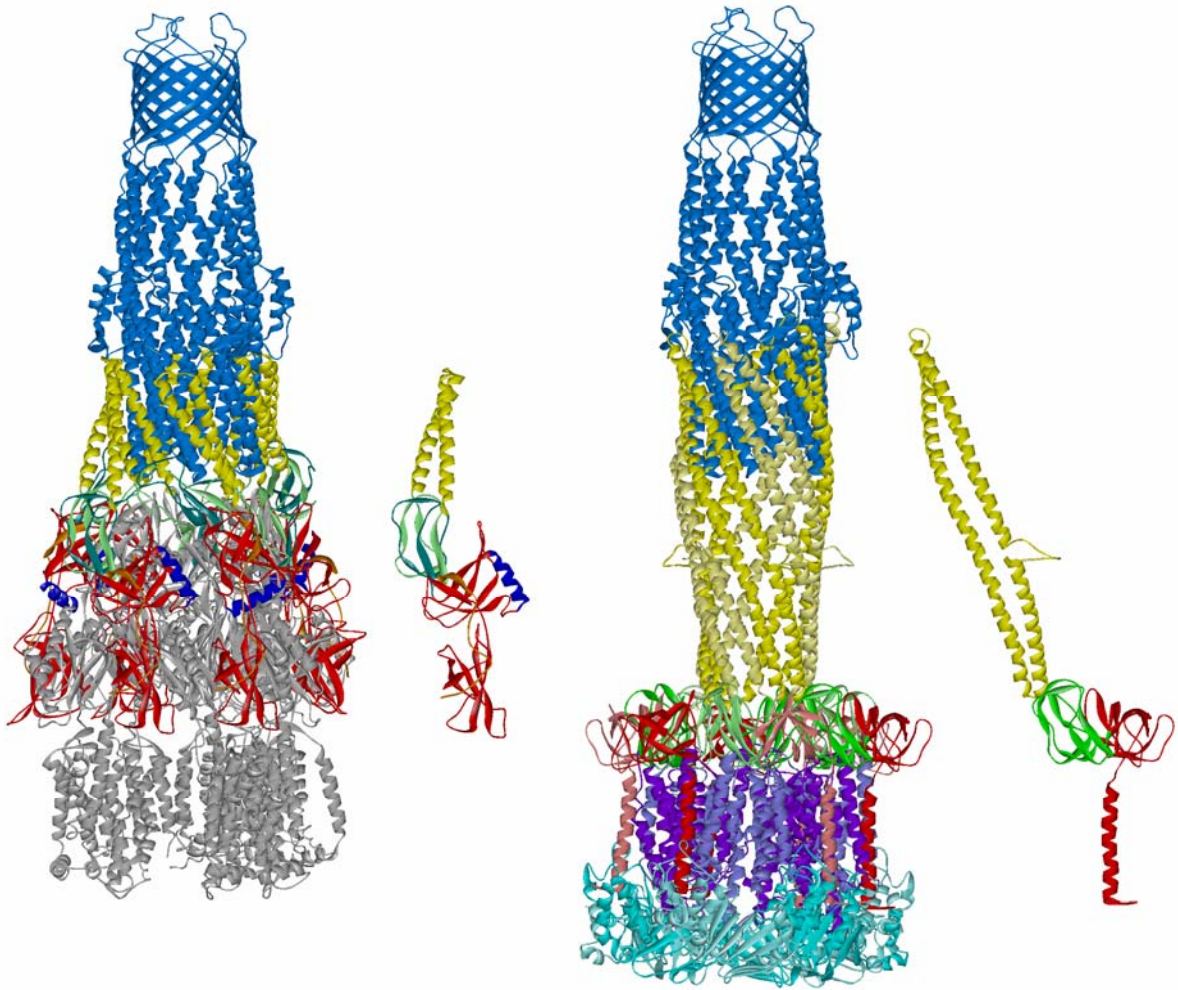
Figure 2.3: Models of two different types of export systems involving APs. The efflux pump (left) is driven by a trimeric RND transporter (grey), the type I protein secretion system (right) is driven by a hexameric ABC transporter composed of transmembrane domains (dark blue) and cytoplasmic NBDs (cyan). Both systems involve a trimeric outer membrane channel-tunnel (blue), which is connected to the inner membrane transporter by six APs. For clarity, a single AP of each system is shown. The AP colour code is the same as in Figure 1. Note that the APs of the efflux pump contain two β-barrel domains. For further information, see text.

As discussed above, the lipoyl- and β-barrel domain is close to the inner membrane. In our model AP monomers have intermolecular contact by these two domains and form a ring like basis with the conduit arising in the centre. In this context, it is remarkably that mutations at the very C-terminus of HlyD of *E. coli* lead to unstable proteins (Pimenta *et al.*, 1999). This supports our idea that the β-barrel domain is involved in protein/protein interactions and is relevant for the proper assembly of the APs. The close contact of the APs is also supported by the possibility to cross-link HlyD APs (Thanabalu *et al.*, 1998). The oligomeric complex has a molecular weight of around 190 kDa interpreted as HlyD trimer. However, the molecular

weight of HlyD is 54 kDa. Considering that compact protein complexes run in SDS-PAGE not according to their molecular weight but faster, it is rather possible that the high molecular band represents HlyD hexamers. In contrast to APs interacting with transporter lacking a periplasmic extension, no one has ever succeeded to isolate pure cross-linked AP complexes of systems with transporters with periplasmic domains. Our model of the efflux pump could explain this. Here, the AP monomers are not in tight contact. In fact, the attachment to the periplasmic domain of the transporter determines the position of the individual domains. It was necessary to twist the domains slightly in order to fit the MexA structure to the RND transporter. It seems valid to manipulate the structure in such a way because the individual domains are connected by a flexible linker and have beside that no further intramolecular interactions. One should also consider that the assembly found in the MexA crystal does not represent the native folding. Figure 3 also shows that the AP monomers assemble differently with the RND protein. The reason is that the two halves of the RND transporter protein are not fully homologues resulting in a different folding and variable attachment sites for the APs. This results also in slightly different positions of the hairpin domains. This agrees well with the nature of the channel tunnel, where the conformation of the two protein halves is also not identical (Koronakis *et al.*, 2000). As proposed above, the additional C-terminal domains, which might form a second β-barrel domain, interact with the lower part of the periplasmic extension of the transporter. The C-terminal β-domains fill the grooves, which are on the side of the transporter, and connect the membrane anchor with the remaining part of the proteins.

Both models clearly show the role of AP as connector between the inner and outer membrane component. However, in one case the two membrane components are already in close contact and the function of the AP is to fix this arrangement. In the other case, the APs form an independent assembly necessary to bridge the gap between the transporter and the channel-tunnel.

2.5. Conclusion and outview

The bioinformatical analysis has ordered the very diverse AP superfamily. The division in seven families and 32 subfamilies, which can be identified by specific sequence signatures, will help in future to annotate genomes correctly. The increasing information from transcriptomic and proteomic investigations makes a proper annotation of genes necessary for the correct interpretation of the data. Especially export apparatus as the multidrug efflux pumps have an enormous impact on the pathogenicity and are often found to be up-regulated in bacteria with increased resistance (Grkovic *et al.*, 2002). However, the systematic search for AP sequences revealed also APs not described as part of export apparatus before. Here, it needs future investigations, what kind of substances are exported and what physiological function the export apparatus provide. In this work, we focus on the AP classification as the basis for structure predictions, structure modelling, and structure/function relationships. Central for this attempt is the alignment of the consensus sequences of the different subfamilies. It gives insight into the relations of all individual AP subfamilies and provides the basis for structure modelling. It is evident that there exists a clear correlation between the existence of an extra C-terminal domain, the size of the hairpin domain and the topology of the inner membrane transporter. Bringing the information together the AP models presented here explain the function of the hairpin domain. One function is to provide contact with the channel tunnel. For this purpose, the 50 to 60 Å long hairpins of APs interacting with transporters with periplasmic extension are sufficient. In this case, one can suppose that the periplasmic exit of transporter and the tunnel entrance are already in close proximity. However, in export apparatus involving transporters without periplasmic extensions there is a major gap between the inner membrane transporter and the tunnel entrance. Here, the hairpins need to have a second function by forming a conduit like assembly bridging the gap. The APs interacting with transporters without periplasmic domains have predicted hairpin length of at least 90 Å, which agrees well with this model.

Of course, the presented models need to be proven by crystallization. However, until now, AP could not be crystallized in their natural assembly and it is questionable if APs will ever assemble in their natural form when uncoupled from their interacting partners. One solution for this problem is the stabilization of the natural conformation by introducing cysteine residues at strategic positions, which fix the natural assembly by forming disulfide bridges. Finding such strategic positions is difficult but structural models of APs provide a good help. Fluorescence resonance energy transfer (FRET) experiments could be a supporting tool to

demonstrate the neighbourhood of chosen positions. Another aim in future is of course to crystallize assemblies of complete export machineries. This ambitious task becomes complicated because of the weak interaction between the inner membrane complex and the outer membrane channel tunnel. Covalent linkage of the APs and the channel tunnels by disulfide bridges could be the clue for getting stable assemblies. Again, here are structural models of APs very helpful to identify promising positions for introducing cysteine residues. Furthermore, contact sites between the three proteins forming export machineries are potential target sites for new drugs. The identification of the interaction sites will help to design drugs, which inhibit the interaction and thus disarm the bacteria from these pathogenic factors.

2.6. Supplementary Material

2.6.1. Collecting adaptor protein sequences from databases

With the accelerated accumulation of genomic sequence data, a number of different classification systems have been developed to organize proteins. Several established classification systems are incomplete and partly misleading. Proteins encoded in complete genomes published in the database of the National Center for Biotechnology Information (NCBI) are phylogenetically classified in cluster of orthologous groups (COG; http://www.ncbi.nlm.nih.gov/COG (Tatusov et al., 1997; Tatusov et al., 2001)). The European Bioinformatics Institute (EBI) has installed the database InterPro (http://www.ebi.ac.uk/interpro (Apweiler et al., 2001)), which integrates the major signature databases PROSITE, PFAM, PRINTS, ProDom, SMART and TIGRFAMs (Falquet et al., 2002; Bateman et al., 2003; Attwood et al., 2003; Corpet et al., 2000; Letunic et al., 2002; Haft et al., 2001). In addition, the Protein Information Resource (PIR) provides an integrated public resource of protein informatics including the PIR-NREF, a Non-redundant REFerence protein database and the *i*ProClass database, which present comprehensive family relationships and structural/functional features of proteins. This database integrates information of all major classification systems and can be seen as basis of a single worldwide database of protein sequence and function UniProt (Apweiler et al.,2004), which is now installed by PIR, EBI and SIB (Swiss Institute of Bioinformatics) unifying the PIR, TrEMBL & Swiss-Prot database activities.

All these databases use to some extent different classification systems. Regarding the AP family, the classification is not uniform and in parts disordered leading to wrong annotations of potential genes in newly sequenced genomes (Table 2.I). Therefore, we have analysed the AP family to put light on this group of proteins. The first step of our bioinformatical analysis is to collect as much AP sequences as possible. We chose to use the search tool of the Protein Information Resource (PIR) to search in the PIR-NREF database. The criterion comprising most proteins is the PFAM domain PF00529 (HlyD family secretion protein). Some of the collected sequences are classified into PIR superfamilies, belong to classifications defined by a ProClass-Motif (PCM00543) or a PRINT signature (PR01490) or are sorted according to the COG-system either as COG0845 (AcrA, membrane-fusion protein) or COG1566 (EmrA, multidrug resistance efflux pump). Finally, a text search for the string 'membrane fusion' excluding eukaryotic and viral proteins detected proteins, which are

not classified by one of the system mentioned before. From the total collection, multiple sequences were removed resulting in over 1000 proteins.

For the sake of completeness, it should be mentioned that APs are categorized by even more classification systems. The institute for genomic research (TIGR) classifies APs into four families. This classification system is not so strict because there are many overlaps between the families (many proteins belong to multiple families). In the InterPro system by EBI MFPs are combined in one family (IPR006143) with three subgroups, one of them containing another subgroup (Mulder et al., 2003). APs are listed also in the transport commission (TC) database in the category 8.A. 'Auxiliary transport proteins' separated in seven subgroups (Ren et al., 2004).

2.6.2. *Sequence alignments and development of characteristic signatures*

In order to sort the AP collection the sequences were aligned using the MULTIALIN-program (Corpet, 1988). Previous analyses have identified motifs, which are present in all APs (Johnson and Church, 1999). There are two conserved motifs (MN and MC) with a length of about 30 residues, each homologue to a half of lipoyl domain found in 2-oxo acid dehydrogenases or biotinyl carboxyl carrier proteins, respectively (Johnson and Church, 1999; Neuwald et al., 1997). Compared to the lipoyl domains, where the two halves are connected by a three residues short turn, the connection between the two motifs in APs comprise with a few exceptions between 29 and over 200 residues. In the following the domain between the two motifs MN and MC is referred to as the loop domain, the region upstream of MN is referred as the N-terminus, downstream of MC as the C-terminus. Beside the MN and MC motifs, the domain comprising the most conserved regions is the C-terminus. Therefore, we have used alignments of the C-terminus to group the protein sequences into families and to define signatures, which are characteristic for the different families. Signatures were developed by try and error using the PROSITE syntax. The aim was to define signatures, which match specifically as much as possible members of a certain family. However, it was not possible for all families to define signatures recognizing specifically almost all family members without matching members of other families or non-related proteins. Several sequences deriving from species belonging to beta-, delta-, epsilon-Proteobacteria, Spirochaetes, Fusobacteria, Deionococci, or CFB (Cytophaga-Flavobacterium-Bacteroides) group were also not detected by the signatures and were assigned to certain subfamilies. The number of known genomes from these groups is still low, that it makes no sense to define own subfamilies with only few sequences. Subsequent searches of PIR-NREF using the

newly developed signatures revealed additional sequences, which were not detected before. Using this strategy, we have collected over 1000 sequences, which could be categorized into different AP families.

It should be mentioned that there exist already two signatures for APs. They are described as ProClass Motifs PCM00543 and PCM00013. PCM00543 is named HLYD_FAMILY hlyD, CvaA, EmrA, which detects about 116 sequences out of the newly defined HlyD and EmrA subfamilies. PCM00013 defines the signature for a prokaryotic membrane lipoprotein lipid attachment site (signature definition {DERK}(6)-[LIVMFWSTAG](2)-[LIVMFYSTAGCQ]-[AGS]-C). Over 550 proteins are marked and a search for the signature by the PIR signature search tool in the PIR-NREF revealed over 140,000 proteins, which makes this signature not suited to find specifically APs.

Table 2.I: Overview about several systems for classification of APs.

Classification sequences system	identification	description	number of
PFAM	PF00529	HlyD family secretion protein	689[a]
COG	COG0845	AcrA, Membrane-fusion protein	87[a]
	COG1566	EmrA, Multidrug resistance efflux pump	32[a]
PROSITE	PCM00543	HLYD_FAMILY (hlyD, CvaA, EmrA)	22[a] (116[f])
PRINTS	PR01490	Gram-negative bacterial RTX secretion protein D signature	56[a]
iProClass	SF001891	hemolysin secretion protein D	68[a]
	SF006211	multidrug resistance protein A	62[a]
	SF006319	Escherichia coli hypothetical protein b1644	36[a]
	SF005532	Lactococcus lactis mesY protein	12[a]
	SF034252	-	6[a]
	SF016738	-	4[a]
	SF018581	nickel-cobalt resistance determinant structural protein CnrB	4[a]
	SF018527	cation efflux system membrane protein	3[a]
	SF018797	-	2[a]
	SF019096	-	2[a]
	SF024638	Aquifex aeolicus hypothetical protein aq_986	1[a]
TIGR	TIGR00998	efflux pump membrane protein (IPR005694; TC 8 A.1.1)	976[b]
	TIGR00999	Membrane Fusion Protein cluster 2 protein (TC 8 A.1.2)	809[b]
	TIGR01730	RND_mfp; efflux transporter, RND family, MFP subunit	816[b]
	TIGR01843	subfamily; type I secretion membrane fusion protein, HlyD family	153[b]
Interpro Release	IPR006143	Secretion protein HlyD	916
	IPR003997[c]	Gram-negative bacterial RTX secretion protein D Function: type I protein secretor activity	253
	IPR010129[d]	Type I secretion membrane fusion protein, HlyD	153
	IPR005694[c]	Efflux pump membrane protein Emr; Signature: TIGR00998 Function: drug transporter activity	23
	IPR005695[c]	Membrane Fusion Protein cluster 2; Signature: TIGR00999 Function: metal ion transporter activity	8
	IPR006144[e]	HlyD secretion protein, C-terminal Function: protein transporter activity	83
TC-Blast	8.A.1.1.1	MFP cluster 1 (function with MFS + other porters) Substrate: Multiple drugs Example: EmrA of *E. coli*	
	8.A.1.2.1	MFP cluster 2 (function with RND porters) Substrate: Multiple drugs; heavy metals; oligosaccharides Example: CzcB of *A. eutrophus*	
	8.A.1.3.1	MFP cluster 3 (function with ABC porters) Substrate: Proteins, peptides Example: HlyD of *E. coli*	
	8.A.1.4.1	Mesenterecin Y105 secretion accessory protein, Full length Gram pos Substrate: Bacteriocins Example: MesE of *L. mesententeroides*	
	8.A.1.4.2	Competence factor transport accessory protein, Truncated Gram pos Substrate: Competence peptide Example: ComB of *Strept.pneumoniae*	
	8.A.1.5.1	Open reading frame Substrate: Bacteriocins Example: Orf2 of *Lactob. gasseri*	
	8.A.1.6.1	Acridine efflux pump constituent Substrate: Multiple drugs Example: AcrA of *E. coli*	

Table 2.I lists classification systems by the systems PFAM (Protein families database of alignments and HMMs, (Peterson et al., 2001)), COG (Cluster of ortholog groups (Tatusov et al., 1997)), PROSITE (Falquet et al., 2002), PRINTS (Attwood et al., 2003), iProClass (Apweiler et al., 2001), TIGR (Haft et al., 2001), Interpro Release (Mulder et al., 2003) and TC Blast Transport Commission (TC) system (Ren et al., 2004).

[a] number of sequences in the PIR-NREF database, which are marked by the identification parameter

[b] matches in the comprehensive microbial resources (CMR) database[16]

[c] subfamily of IPR006143

[d] subfamily of IPR003997

[e] domain (no family!)

[f] number of sequences found by the PIR signature search tool in the PIR-NREF, using the PCM00543 signature [LIVM]-x(2)-G-[LM]-x(3)-[STGAV]-x-[LIVMT]-x-[LIVMT]-[GE]- x-[KR]-x-[LIVMFYW](2)-x-[LIVMFYW](3)

Table 2.II: Signatures characterizing AP families and subfamilies

Fam.	Subfam.	Signature	Mat.	Ass.	Sum
ProtE		**G-[LM]-X-{CRP}-X(3)-[IVLF]-X(4)-[RKH]-X(4)-[YFW]-X(3)-P-{PRKGQ}-X(5,14)>**	**175**	**12**	217 (21.0 %)
	LssD	P-G-M-X-A-X(2)-[DE]-[IV]-[KRL]-[TS]-G-X(5)-[DQN]-Y-X(2)-[KR]-P-X(6,15)>	16	-	
	LapC	[IL]-[FI]-P-G-M-X(3)-[AV]-[DH]-[IV]-X-[TS]-G-[DEQK]-[KR]-[TS]-[VIL]-[LMF]-X-Y-X(2)-[KN]-P-X-{NI}-X(9,10)>	13	2	
	AprE	G-[MT]-[PANQ]-X(3)-[FHVILQT]-[IVLF]-X(4)-[RG]-X(4)-Y-X(3)-P-X(11,12)>	64	13	
	HlyD	G-M-X(5)-[IV]-X(4)-R-X(4)-[YF]-X(2)-[SGE]-P-X(6)-[EDK]-X(5)>	47	10	
	Alr4240*	[LIF]-X-[PSV]-G-[MQ]-X-[LIAV]-X-[GAV]-{F}-[FIV]-X(2)-[RS]-X(7)-[LIF]-X (10,16)>	11	1	
	CvaA	N-G-[LM]-X(10)-R-X(4)-W-X(3)-P-X(14,15)>	11	-	
	RaxA	{N}-G-[LM]-X(5)-[IVLF]-X(4)-[RKH]-X(4)-W-X(3)-P-{PRKGQ}-X(5,20)>	11	-	
	ComB*	[GS]-[VIL]-X(6)-F-X(3)-D-X-F-G-Y-X(9)-Q	18	-	
AMTS		**[DEQ]-X(4)-[AS]-[PTW]-X(2)-[GA]-X-[VIL]-X(7)-G-{Q}-X(7)-[VILA]-X(9)-[YWS]-X-{E}-X(13)-[GDQ]-X(7)-[PD]**	**60**	**11**	71 (6.9 %)
	YbhG	A-[ES]-F-T-P-[KRE]-X-[IV]-[EQY]-[FT]-X(3)-R	25	-	
	VPA0490	[VI]-[AS]-X(2)-[AP]-[SE]-[YF]-T-P-[YP]-X(7)-R-X(2)-L-X-Y	11	-	
	PA3402	G-X-[FY]-[AS]-[TA]-X-[RK]-[AS]-T-X(5)-[FY]-D-X-[RK]-[TS]-[FY]-E	16	-	
	YhiI	A-Q-F-T-P-K-X-V-E-X(3)-E-R-X-K-L	19	-	
DHA2		**[ASG]-[NA]-X(2)-E-X(8)-G-X(7)-[DM]-X(17,19)-G-X(12)-G-N-[FW]-X-[KR]-X(2)-[QE]-R-X(2)-[VIL]-X-[VIL]**	**102**	**1**	103 (10.0 %)
	EmrA	A-N-F-K-E-X(8)-G-X(7)-D-X(16)-{G}-G-X-G-X(12)-G-N-W-X-K-X(2)-[QE]-R-X(2)-V-X-[IV]	35	-	
	VceA	A-N-X(2)-E-X(8)-G-X(7)-D-X(16)-G-G-X-G-X(12)-G-N-W-X-K-X(2)-Q-R-X(2)-[VI]-X-I	6	-	
	RmrA	[ASG]-[NA]-X(2)-E-X(8)-G-X(7)-[DM]-X(17,18)-G-X(12)-G-N-[FW]-X-[KR]-X(2)-[QE]-R-X(2)-[VIL]-X-[VIL]	60	2	
PET					71 (6.9 %)
	AaeA*	W-[VI]-[RK]-L-A-Q-[RL]-X-[PR]-[VL]	43	-	
	YjcR*	W-[VI]-[RK]-V-A-Q-[RL]-X-[PR]-[VL]	5	-	
	PA3360*	[DKS]-A-Q-R-X-[PR]-[VL]-[NHI]-[LVI]-X(3)-[DER]	11	1	
	YiaV*	P-X-G-X(2)-[AVG]-[QE]-[VA]-A-V-Y	11	-	
HAE1		**[GA]-X-[LIMVFW]-X(10)-[STN]-[GDANS]-X-[VIFLM]-X(2)-[RKE]-X(2)-[VILFM]-X-[ND]-X(4)-[LIF]-{R}-[PDS]-[GN]**	**193**	**9**	385 (37.3 %)
	AcrA	[GA]-X-[LIMVFW]-X(10)-[STN]-[GDANS]-X-[VIFLM]-X(2)-[RKE]-X(2)-[VILFM]-X-[ND]-X(4)-[LIF]-{RWF}-[PDS]-[GN]	151	11	
	YegM	[DN]-N-X(5)-[TS]-[GA]-T-X(3)-[KR]-[AG]-X-[FI]-X-N-X(4)-L-[FWMK]-P-[GN]-[QL]-X-V	39	1	
	MexH*	[GN]-X(2)-[VI]-X(5)-[LI]-X(69,70)-[AS]-P-[FSI]-X-G-X(2)-[GA]-X(6)-G-{AG}	52	18	
	All3144*	[VI]-X-A-P-X(2)-G-X-V-G-X-[LI]-P-X(3)-G-D-X-V	13	2	
	MexJ*	[YF]-X(2)-L-X(5)-[GA]-X-[VI]-X(5)-[EDA]-X-G-[QLE]-[VT]-[VIA]-X-[AST]-[GAT]-X(2)-[VIA]-X(3)-[AS]-X(5)-[ED]	42	4	
	VC1674*	L-X-[YF]-T-X-[LI]-X-[AS]-P-X(2)-G-X-[VI]-[GS]-X(6)-{G}-X(2)-[VI]	19	-	
	HP0606*	F-G-[ED]-G-X-I-X-T-K	5	-	
	diverse*	no pattern	-	28	
HME		**[VL]-X(7,8)-[FY]-X(4)-[VIL]-X-[VILT]-[GFEL]-X(8)-[VIL]-X(2)-G-[LIV]-X(2)-G-[DEQAST]-X-[VIY]-X(7)-[VIL]-[DKNRY]**	**58**	**1**	62 (6.0 %)
	CusB	[VL]-X(7,8)-[FY]-X(4)-[VIL]-X-[VILT]-[GFEL]-X(8)-[VIL]-X(2)-G-[LIV]-X(2)-G-[DEQAST]-X-[VIY]-X(7)-[VIL]-D	19	-	
	CzrB	[VL]-X(7,8)-[FY]-X(4)-[VIL]-X-[VILT]-[GFEL]-X(8)-[VIL]-X(2)-G-[LIV]-X(2)-G-[DEQAST]-X-[VIY]-X(7)-[VIL]-[KNRY]	42	1	
ME		**[IVL]-X-[AS]-[PT]-X(2)-G-X-[VI]-X(7)-G-X(2)-[VIL]-X(23,29)-[VILM]-X-E-X(8)-[GK]-X(2)-[VLIAM]-X-[VLIAF]-X(5)-{KIA}**	**72**	**5**	104 (10.1 %)
	MacA	[IVL]-X-[AS]-[PT]-X(2)-G-X-[VI]-X(7)-G-{N}-X-[VIL]-X(6)-P-X(16,18)-[VILM]-X-E-X(8)-[GK]-X(2)-[VLIAM]-X-[VLIAF]	50	11	
	VP1999	[IVL]-X-[AS]-[PT]-X(2)-G-X-[VI]-X(7)-G-X(2)-[VIL]-X(24)-G-X-[VILM]-X-E-X(8)-[GK]-X(2)-[VLIAM]-X-[VLIAF]	6	-	
	YvrP*	E-X(11)-V-X(11)-W-X-G-[KTV]-[VIL]-X(2)-[VI]-[GS]-X(2)-[PK]	20	1	
	Alr1505	[IVL]-X-A-P-X(2)-G-X-[VI]-X(7)-G-X(2)-V-X-P-X(4)-S	11	-	
	Cg3322*	E-X(8)-G-X(9)-T-G-X(5)-G-X-V-X(3)-S	5	-	
	DevB*	P-x(4)-D-X-R-[VI]-[VI]-E-V-X-[VI]	18	-	
Total			**925**	**106**	

Table 2.II lists signatures characterizing AP families (Fam.) and subfamilies (Subfam.). Signatures were developed by try and error using the PROSITE syntax. The number of sequences, which matches (Mat.) a specific signature and the number of assigned sequences (Ass.) are given. Sum refers to the total number of sequences of each family. The common family signature printed in bold does not detect subfamilies marked by a star.

Figure 2.I

```
                                                         50
              |         |         |         |         |         |         |         |
LssD     ----------------------------------------------------------------------DAALTESAHPGASLILW
LapC     ---------------------------------------------------RNRKLTP----EELDFVDDKTAALLLNTPRSARLLLW
AprE     -----------------------------------------------------------------DIVTRSIRRHLRLGWL
HlyD     -------------------------MWLGLSEFLLRYKTVWREVWKIRKQLDTPVREKDENEFLPAHLELIETPVSPRPRLVAY
Alr4240  ---------------------------------------------------------SQNNKDLSFEQPVVLQQSRRWSRGLLW
CvaA     -------------------------------------------------MFRQDALENRKMKWQGRAILLPGIP-LWLIML
RaxA     ----------------------------------------------MDLFRQEVLDAQRQSRLGGISLIQPIS-LTLVLA
ComB     -----------------------------------------------MDPKFLESSEFYQRRYRNFSTLIII
VPA0490  --------------------------------------------------------------------------M
YbhG     ---------------------------------------------------------MMKKPVVIGLAVVAL
YhiI     --------------------------------------------------------MDKSKRHLAWWVVGL
PA3402   -----------------------------------------------------SMSDKNKALLTGAL
EmrA     -----------------------------------------------ETQTPQPPPKKKSKRKRLLLLLTL
VceA     ---------------------------------------------MSENAQQNNSNDEFDGARRKQSRKKLFLGLAA
RmrA     ----------------------------------------------------RPKRSKRRRKLLFLLLA
AaeA     -------------------------------------------------MKTLSRKLGRVALTLLVVV
YjcR     -------------------------------------------------MESTPKKAPRSKFPALLVV
PA3360   ---------------------------------------------------------MTPDQQFRSWTR
YiaV     -----------------------------------------MDLLLILTYVALCWAIFKIFKIPLNKWTVPTA
AcrA     -----------------------------------------RRLLAALLAL------ALALLAGCGDKQAA
YegM     -----------------------------KGSRRVIVIVVVIAALAAWWFWPARTDSQS------AAPGATGQAGQSGR
MexH     --------------------------------------KLIMLIVAALLFG------AVIGFNKKIAEAFA
All3144  ---------------------------------MVQPEPAFAQGLSLLALCLL------ALGGGGAAQPEAGA
MexJ     ----------------------------------------RARLALLALLAALALLLLASCGKEAA-
VC1674   -------------------------------------------MKKLLVALAAALLLTGCGKPPQP
CusB     -MKKIALIIGSLLAGGIISATGFQYFNSAEPPAETTSTAERKILFWYDPMKPNTKFDKPGKSPFMDMDLVPKYADEDSSSGGVRIDP
CzrB     ------------------------------------------MRRGYSVELGLAWLALGGKALANGGE-----AEVKSEEGKIELTA
MacA     ---------------------------------------------KRRKKLYIAALLIVLAGGLTLWRS---LNAPV
VP1999   -----------------------------------------------MKKYFMTKRWLVSAVSVAL-LGGGAYFYLQ---SSAEP
YvrP     ------------------------------------------------MKRKKIKKKLKLLIASVVALVLIAGGYFLWAQQQNTESAK
Alr1505  --------------------------------------------PLVGKKKKPLRWLSGLAGLGLGVVGGYFIYLR---ASPPP
DevB     ------------------------------------------------LDLEQRKIIWGLAVALPVAVGIGLYSLSQFQQLQP
```

```
           100                                   150
              |         |         |         |         |         |         |         |
LssD     ASLLLFLIFLIWAWFAPLDEVTTGQGKVIPS-SREQVIQSLEGGI---LQQLL-VKEGDIVEKGQILAQLDXXXQLTRTTIRSPVRG
LapC     VIVLFFLAAILWASFAEIDEVTVGEGKVIPS-SQIQVIQSLEGGI---VKEIL-VREGQQVQKGQPLLRIDXXXRVNRTVVRSPVRG
AprE     LILLLFGGFGGWAALAPLDSAVIASGTVVVS-GNRKTVQHLEGGI---VSEIL-VREGDRVKAGQVLIRLDXXXQLKRTDIRAPVDG
HlyD     FIMLFLVLALLWSVLGKVEIVATATGKLVPS-GRSKEIQPIENAI---VKEIH-VKDGQFVKKGDVLLELTXXXRQQLSVIRAPVSG
Alr4240  TLIGVTVFAIAWASIAKIEEVVPAQGKLEPQ-GTVKEVQAPVNGV---VKTIL-VKEGQRVKKGQVLLQLDXXXQLKYQELRAPVDG
CvaA     GSIVFITAFLMFIIVGTYSRRVNVSGEVTTW-PRAVNIYSGVQGF---VVRQF-VHEGQLIKKGDPVYLIDXXXVEGEIIIRALTDG
RaxA     SSLI-VLLIVVFLYVGSYTRRSTVTGQLVPS-KGLIKVYAPQAGV---VAQVH-VSEGQQVLAGQPLATLTXXXARGTLVVRAPESG
ComB     PIFLLLVFLVIFSFFAKREITITGQGELEPT-KILPIIQSTSNNP---IKQNY-LKEGKFVKKGDTLLTYKXXXDSQEYTVKAPESG
VPA0490  SRLCFIPLLALLLSACGTESPPQALGTVERD---RVTLTPTENEI---IRELP-VKEGDRVKQGDVLVTLDXXXKLAERTIVAPRDG
YbhG     AAVVAGGYWWY---QSRQDNGLTLYGNVDIR---TVNLSFRVGGR---VASLA-VDEGDAIKAGQVLGELDXXXNLQDTTLIAPSDG
YhiI     LAVAAYVAWWLLRPAGLPEGFASGNGRIEAT---EVDIATKIAGR---IDEIL-VKEGDFVRAGEVLAKMDXXXDIDDSELKAPRDG
PA3402   AVLLWLAYLFYLAYRPKPE---VLQGQIEAR---EYSVSSKVPGR---IEKVF-VRKGDQVKKGDLVFTLSXXXYLKDTKATAPIDG
EmrA     LFIIAAVAYGIYWFLVLRHHEETDDAYVAGN---QVQITPQVSGS---VTKVN-ADDTDFVKQGDVLVTLDXXXALQRTRIVSPVTG
VceA     VIVVAAGSYALYWHFIASRYVSTDNAYVAAE---IAQVTPAVGGT---IAEVR-VVDTQQVKKGDVLVQLDXXXDLERTVIRAPVSG
RmrA     LLGLAALWYGYYYWPFGRFYESTDDAYVQAD---ITTVAPQVSGY---VAKVL-VKDNQHVKAGDLLFRIDXXXDLSYTRIRAPVDG
AaeA     VAVLAGWWLWVYY---MQS-PWTRDGRVRAD---VVQIAPDVSGL---ITEVN-VKDNQLVKKGDVLFTIDXXXNLERTEVRAPVDG
YjcR     ALALVLALVFVIWR---VDSAPSTNDAYASAD---TIDVVPEVSGR---IVELA-VTDNQAVKQGDLLFRIDXXXKHLEMATVRAPFDG
PA3360   YIIVLFIALFLYTLWADRVTPLTPQARVQGY---VVQVAPEVSGK---ITQVL-VSNNQHVQAGDVLFRIDXXXDLANTQVLAPSDG
YiaV     VLGGVFLVGGLILLM-NYNHPYTFLAQKAVI---TIPIVPQVTGR---VIEVT-VKPNQPIKKGDVLFRLDXXXNLEQTVVRAPSDG
AcrA     QAQPPPPVGVVTVKPQPVTLTTTELPGRTEAY--RVAEVRPRVSGI---ILKRL-FTEGSDVKAGQLLYQIDXXXNLGYTKVTAPISG
YegM     GRGPPVPVQVATATEQDVPRYLTGLGTVTAA--NTVTVRSRVDGQ---LVALH-FQEGQQVKAGDLLAEIDXXXQLDYTRITAPIDG
MexH     NRPPPVPVTTAEVEPRPWPPVLTAIGTLEAN--QGVTLSAEVSGR---VTQIA-FESGQQVKAGDLVLQLDXXXALARKTIRAPFSG
All3144  P-PPAVPVKLATVQPATLEDTTEYTGTLQSR--RSVTLQPQVEGR---VTQIN-VDEGDRVQSGQPLAQIDXXXQLNQTRIRAPFSG
MexJ     ADPRIRPVVVVVVVPAPSAASRSFTGTVAAR--VQSDLGFRVSGK---IVERL-VDVGDRVKAGQPLARLDXXXQLSYAELRADADG
VC1674   AEPLSPPVKLVSVSDGEVSDSRTFPAVVNAA--ERAHLSFRVAGE---IQQLL-VKEGDRVKKGQVLATLDXXXQLSYTQLRAPFDG
CusB     TQVQNLGVKTAKVTRGPLNYSQTTPGNVSYNEYQFAIVQARADGF---VEKVYPLTVGDHVKKGTPLLDLTXXXIQTRFTIKAPIDG
CzrB     EQLKAAGIKVETAGPAPLATGVSLPGEIRFNPDRTAHVVPRVAGV---VESVQ-VNLGDAVKAGQVLAVISXXXAGNRYEVRAPIAG
MacA     PTYQTATVRRGDIEQ-----SVLATGTLQAL--RQVDVGAQVSGQ---LKSLS-VEVGDKVKKGQLLAEIDXXXNLGYTRIVAPMDG
VP1999   DAYPTLVVTTGTIEK-----QAVAVGKIVPA--HSVSIKSQIDGI---VGEIY-AKVGEKVTQGQPLIKVRXXXARLTSTIYAPIDG
YvrP     EAYSTVSVTEGSIAS-----STLLSGTVKAN-NEEYVYFDANKGN---DAEVT-VKVGDQVKKGQQLVQYNXXXALNDTVVTSDVSG
Alr1505  RQQADLSNTTAAVERGNLPIRISASGTVQPE--RSVNISPKQAGL---LKQLY-VDQGDEVKKGQILARMDXXXQLNDTVIRAPFSG
DevB     ESAPAVRAPVAPVIE-----AVTALGRLEPQGEVIKLSAPTSFGSGSRVAQLL-VKEGDRVKKGQVLAILDXXXDLELTYIRAPQDG
```

```
                    200                                                   250
                     |         |         |         |         |         |         |
LssD       IVKDIRVTTIGGVVQPG---------GELMEIVP--LD-DQLLIEARISPRDIAFIH----------PGQPALVKITAYDYSIYGGL
LapC       TVKTLLINTVGGVIQPG---------MDIIEIVP--TD-DTLLVEAKIAPQDIAFLH----------PGLPAVVKFTAYDFTIYGGL
AprE       TVVNLAVHTEGGVIRPG---------EPLMEIVP--DD-DPLEVEARVAPTDIDKVH----------VGLPVDLRFSAFNQRTTPRL
HlyD       TVQQLKVHTEGGVVTTA---------QTLMVIVP--ED-DVLEVEALVLNKDIGFVE----------PGQEVVIKVEAFPYTRYGYL
Alr4240    VVFDLKATNPGFVVQSG---------EPLLKIVP--QN-N-LVAKVSIPNRDIGFVR----------TGMRVDVRIDSFPYSEFGVL
CvaA       KVDSLSV-TVGQMVNTG---------DSLLQVIP--ENIENYYLILWVPNDAVPYIS----------AGDKVNIRYEAFPAEKFGQF
RaxA       IVAAQPV-ESGQTVDSG---------QPLLSLLP--ENSE-LEAELYAPSRAAGFIK----------PGDEVRLRYQAYPYQKFGQQ
ComB       VLHLNDEYSGIKYIPTG---------TTLAQIYPVLKKQTKLKITAYIPSTDISSLK----------VGQKVRFKVTR-NVPKPLIL
VPA0490    TLDDLPY-NLGERVPVS---------GPVASIQP----DGNPKLRVYVPETYRSKFA----------PGSTVTVHCDGCAD----PL
YbhG       TVLTRAV-EPGTMLNAG---------GTVFTLSL----TRPVWVRAYVDERNLGQAQ----------PGRKVLLYTDGRPD---KPY
YhiI       RVQYRVA-EPGEVLAAG---------GRVLNMVD----LSDVYMTFFLPTEQAGRLK----------LGGEARLVLDAAPD---YVI
PA3402     EVSNVLL-SGGELAPQG---------FPVVTLID----LKDAWAVFNVREDLLKEFK----------KGKEFEAYIPAL-D---KSI
EmrA       YVARRSV-QVGQRVSPG---------TPLMAVVP----ADQMWVDANFKETQLANMR----------IGQPVTIISDLYGD--DVVY
VceA       VVAKRQV-QVGQRVQPG---------TPLMTIVP----TDHLYVDANFKEVQLRDVR----------VGQPVTLTADLYGG--DVTY
RmrA       VVGNRSV-RVGQYVTAG---------TQLLSVVP----LDDLYVVANFKETQLTHMR----------PGQPVTITVDAFPG---QKL
AaeA       WVTNLNV-RAGDYATAG---------KPVVALVD----SNSFYVDGYFEETKLEGIR----------EGDRAEITLMGDN----KTL
YjcR       RVISLKT-SVGQFASAM---------RPIFTLID----TRHWYVIANFRETDLKNIR----------SGTPATIRLMSDSG--KTF
PA3360     VITNLQL-EVGSYANAG---------QPLLTFIP----NGSLWITADFRENSLAHVR----------PGDRVLIVFDAYPG---EVF
YiaV       YVTQLLL-RPGTYAVPLP-------LRPVMVFIP----DQKRQIVAQFRQNSLLRLQ----------PGDDAEVVFNALPG---KVF
AcrA       RIGRSSV-TEGALVT-G---------QATALATIQQ---LDPIYVDFTQSSADLLRLRRALASGKLQQGDAPVSLTEDGST---YPH
YegM       RVGLRQV-DVGNLISSG-------DTTGIVVITQ----THPIDVVFTLPESDLATVVQAQKAGKLV-----VEAWDRTNSK---KLS
MexH       VVGIRQV-SLGQYVTAG---------TPIVTLED----LSTMRVDFTVPEQYLSQLK----------LGQPVTVTVDAYPG---QTF
All3144    VVGNRPV-KVGDFVQPG---------TELLTLTE----NRPLEVNIQVPEERLPRLR----------PGQPVEVTDDAGQ----RVG
MexJ       VVTATLA-EPGQVVSAG---------QPVVRLAR----DGEREAVVDLPETLRPALGG---------VAVQVSLWGDPS-----ITA
VC1674     IISDVYV-ENFENVQAG---------QPIVNLHT----NDLLEVLIQLPDTLLASQSN---------AVKLEAVVRFDSGP---KTF
CusB       VITAFDL-RDGMNISKD---------QVVARIQG----LSPVWVTAAVPESIASLLK----------DGSQFEITVPAYPD---KTF
CzrB       VVVERHI-TLGEAVAPA---------NLFT-VAD----LSSVWVEAQVYEKDLGRVR----------VGQEVTVS-SASGG---VPL
MacA       TVVAITT-EEGQ------TVNAAQQAPTILTLAD----LSTMTVKAQVSEADVIKVK----------PGQPAYFTVLGDPD---KRY
VP1999     TILNQKV-EVGE---PIISTQSSQAATEMMSLAD----MNNLIFKGSVSEHDAAQLS---------PGMPVLLTVAPYPD---VAI
YvrP       TVVEVNN-DIDP---------SSKTSQTLVHIAT----EGKLQVKGTLSEYDLANVK----------VGQEVKIKSKVYPG---KEW
Alr1505    IITQKYA-NPGAFVTPTTSASATSSATSSSIVAL----ASGLEVVAKVPESDIGRIR----------PGQPVTIRADAYPG---ETF
DevB       QILKIHT-RPGE----------RISSDGIAELGQ----TQQMYAVAEVYESDIGRVR----------IGQKVTITSENGAF--SGEL


                    300
                     |         |         |         |         |         |         |
LssD       DGVVEQISPDTIQDEVDPEEF------------YYRVYIRTEQNYLGNKGK----RLPIIPGMVATVDIRTGQKTVLDYLLKPLNKA
LapC       KGKVEHISADTTQDEEGNS------------FYIVRVRTEKTTLGDQEK-----LPIIPGMTASVDIITGKRTVLSYLLKPILRA
AprE       NGEVTLISADRLTDERTGQP------------YYVVRVEVDPEELKLKG------LELKPGMPAEVFIRTGERTLLSYLFKPLTDR
HlyD       TGKVKNISLDAIEDEQLGL--------------VFNAIISLDKNTLSIDGK----EIPLSSGMAVTAEIKTGKRRVISYLLSPLEES
Alr4240    KGEVTSIGSDALPPDEQYQFY-----------RFPATIKLDQQYLLSGGR----QIPLQSGMSATANIKLRERTVLSFFTDPFTDK
CvaA       SATVKTISRTPASTQEMLTYKGAPQNTPGASVPWYKVIAMPEKQIIRYDEK----YLPLENGMKAESTLFLEKRRIYQWMLSPFYDM
RaxA       HGQVVRISRSALTPGELAALGGGTQE------PVYRVRVKLDRQSITAYGE----AEPLKPGMLLEADILLEKRRLYEWVLEPLYSL
ComB       TGTIKQISSAPTKTKEGN--------------FYKVTATTNISNED--------AELIHYGLQGKVTTITGKKTYFNYYKDKLLNR
VPA0490    QATVRWVAPEPEFTPPVIYSEEERSKL------VYLAEADLP-----------EPALPLPSGQPVQVDLPEE---------------
YbhG       HGQIGFVSPTAEFTPKTVETPDLRTDL------VYRLRIVVT----------DADDALRQGMPVTV--RFGDEAGHE---------
YhiI       PAKVSFVASVAQFTPKTVETSDERLKL------MFRVKARIPPELLQ------QHLEYVKTGLPGMAYVRLDPEAPWPDNLVVRLPQ
PA3402     EFKVTYISVMGDFATWRATDASGGFDL-----RTFEVEARPL-----------EPIEGLRVGMSVLVTL------------------
EmrA       HGKVVGLDMGT---------GSAFSLLPAQNATGNWIKVVQRVPVRIELDPKQLEEHPLRIGLSMTVTVDTKNQDG------PVLAS
VceA       HGKVAGFSGGT---------GSAFAMIPAQNATGNWIKVVQRLPVRIELDPKELQAHPLQVGLSMVATIDTSGTT------------
RmrA       KGHVESISPAT---------GSEFSLLPPDNATGNFTKIVQVRIVLDAGQPLAGRLRPGMSVVVVTVDTRSAP------------
AaeA       QGHVESIARGI----EDRSRSSDSSLLPNVNPNFNWVRLAQRVPVRIALD---EVPQDVTLVAGTTATVAVGQ---------------
YjcR       EGKVDSIGYGV----LPDDGGLVLGGLPKVSRSINWVRVAQRFPVKIMVD--KPDPEMFRIGASAVANLEPQ---------------
PA3360     EGSISSSIDYGV-----SAGQSDPNGALTGVETSNRWVRDAQRFRVNLEFD----DLPALRVGGRATVGLYPDE----SYFEWLA---
YiaV       KGKVISILPAV-----GGGQYQARGALQGLNALPGSGGVFATIELTDDLD-----IYALPDGIYAQVAVYSDH---------FSHVS
AcrA       KGTLDFSDVTVD---------------PTTGTVTLRAVFPN-----------PDGTLLPGMFVRVRLEEG--VQPDAILVPQQAV
YegM       EGTLLSLDNQID---------------ATTGTIKLKARFNN-----------QDDALFPNQFVNVRLLVD--TLQNAVVVPTAAV
MexH       KGKISAIDPRVD---------------AQTRNIQVRATFPN-----------PDGKLRPGMFARVRVALP--EEENALTVPETAI
All3144    EGRVTFIAPNAD----------------NNTQSILVKVLYDN----------SDGRLRAGQLARVRIIWD---QQPGLLVPTTAV
MexJ       TGKLRELSPAAD----------------PQTRTFRVRVVLDGA----------PAAAPL-GATVTVTLP-----AAPAIELPLSAL
VC1674     PLRFKEISTEPD----------------PKTGTYTVTLTMPRP----------EDEGILPGMAVSVHVDGQNVGTSSGITIPTTAL
CusB       QGEKWYILPSVD----------------PATRTLQVRLELSN----------KDEALKPGMNANLKLNTP---SQEMLLIPSQAV
CzrB       EGRVSYVGPLLD----------------EQTRTATVRVTLDN----------PNGALRPGMFVTVRVTTA---QEAALAVPESAV
MacA       EGTLRQILP--------------TPEKINNAVYYALFEV----------PNPDGLRPGMTAQVSIVLG--DVKNVLTIPLAAL
VP1999     SGVLTKVAIQSENLNSP---GGNASAKSFDNGFEVEVGQLNI-----------PSEVLLRSGFSSTAKITLK--KSENVLTLPERAL
YvrP       TGKVSYIGNYPTESN------GENTTPAGSDGQGYPYTIDLD-----------SSLNQLKQGFHVSVEVVNG----EKTALVPLSSV
Alr1505    KGRVRLVAPE------------AVVEQNVTSFEVKVS-LTD-----------PQEQ-LRSGMNVDVNFQAG--RLNNALLVPTVAI
DevB       RGTVTRIGPQIGKKDV-----LNTDPAADIDARVVEVKIRLD----------PEDSQKVAGLTNLKVIVAIEP------------
```

```
             350                                400
              |     |     |     |     |     |     |     |     |     |     |
LssD      KE-ALRER-----------------------------------------------------------------
LapC      KENALRER-----------------------------------------------------------------
AprE      AHRAFREE-----------------------------------------------------------------
HlyD      VTESLRER-----------------------------------------------------------------
Alr4240   TESLKRV------------------------------------------------------------------
CvaA      KHSATGPLND---------------------------------------------------------------
RaxA      RGKL---------------------------------------------------------------------
ComB      D------------------------------------------------------------------------
VPA0490   -------------------------------------------------------------------------
YbhG      -------------------------------------------------------------------------
YhiI      -------------------------------------------------------------------------
PA3402    -------------------------------------------------------------------------
EmrA      TVRSTPAY----TSTALEIDLS-PVDKLIDDIIQANAG-----------------------------------
VceA      -------------------------------------------------------------------------
RmrA      -------------------------------------------------------------------------
AaeA      -------------------------------------------------------------------------
YjcR      -------------------------------------------------------------------------
PA3360    -----DLQIWLWSLLHYVY------------------------------------------------------
YiaV      VMRKVLLRMTSWQN--YLYLDH---------------------------------------------------
AcrA      TRDQKGKATVLVVGADNKVELR-PVTTGRTVGDKWLVTSGLKAGDRVIVEGLQKVRPGAPVKPVEVTADAAAAA--------
YegM      QMGNEGHF-VYVVNSDNKVSKR-LVTVGIQDGDKVVIRAGLSAGDRVVTDGIDRLTEGAKVEVVEAQSAPTAAEKATSREAK
MexH      TYSLYGDY-VYVVK---VAKQV-FVKTGERVGGRVEILSGLKAGDRVVTSGQVRLRDGAAVKVVGSDAP-------------
All3144   TRLAGQQF-VFVVQGDQVARQR-PVQLGGIQGGQYQILSGLKPGERIVVSGILPLRDGVPIRPAS----------------
MexJ      LDAGGGPGVWVVDPQTSTVSLR-PVKVLRYEDDSVVVSGGLKAGERVVAAGVHLLRPGQKVRVAGGAA-------------
VC1674    FNQDGDY-VWVVNGDNQVVKKR-EVIL----NASGEVSSGLADGDQVVIAGVSRLREGIKVRVWVKERGL-----------
CusB      IDTGKEQR-VIVVDGEGRFVPK-RVHVGRESQGQVEIRSGLNEGEKVVTSGLFLIDSEANISGALERMRSEQPTHAH-----
CzrB      -QTVDGKS-VVFVRVQDGFEPR-PVKLGRRDGGYVEILSGLKAGDQVATAGSFLLKAELGKGSAEHGH-------------
MacA      --GGDGRYKVKVLENGEVRERE--VTTGIRNDVDVEILSGLKEGDQVVIGEATPGAA-----------------------
VP1999    -QFEGDAPHVLIPDSSEQGFHRQPVQLGLSDGINVEVLDGVTLDEEIIDNSMMGGAHG----------------------
YvrP      -IKKDGKNYVWVVDDGKAKKVE--VKLGNADADSQEITSGLTKGDIVISNPDKNLKDGMEVESISI--------------
Alr1505   -VTEDGKTGVLVPDAQNKPQFR-PVEIGTSSGDKTQILSGLRPGERVFIDLPPGARPQQT--------------------
DevB      -------------------------------------------------------------------------
```

Figure 2.I: Alignment of consensus sequences of all AP subfamilies. The consensus sequences of each subfamily derived from multialignments of all members of the subfamily. Regions not present in at least 50 percent of the subfamily members are omitted in the consensus sequences. The hairpin domain is marked by triple X.

The channel-tunnel HI1462 of *Haemophilus influenzae* reveals differences to *Escherichia coli* TolC

3.1. Abstract

Efflux pumps play a major role in multidrug resistance of pathogenic bacteria. The TolC-homologue HI1462 was identified as the single channel-tunnel in *Haemophilus influenzae* required to form a functional multidrug efflux pump. The outer membrane protein was expressed in *E. coli*, purified and reconstituted in black lipid membranes. It exhibits a comparatively small single-channel conductance of 43 pS in 1M KCl and is the first known TolC-homologue which is anion-selective. The HI1462 structure was modelled and an arginine residue lining the tunnel entrance was identified. The channel-tunnel of a mutant with the arginine substituted by an alanine residue was cation selective and had a sevenfold higher single-channel conductance compared to wild-type. These results confirm that the arginine is responsible for anion selectivity and forms a salt bridge with a glutamate residue of the adjacent monomer, establishing a circular network, which keeps the tunnel entrance in a tightly closed conformation. In *in vivo* experiments, both the wild-type HI1462 and the mutant were able to substitute for *E. coli* TolC in the hemolysin secretion system, but not in the AcrAB/TolC multidrug efflux pump. The structure-function relationship of HI1462 is discussed in the context of the well-studied TolC channel-tunnel of *E. coli*.

3.2. Introduction

Channel-tunnel-dependent export systems contribute to the pathogenicity of various Gram-negative bacteria (Nikaido, 1994; Delepelaire, 2004). Channel-tunnels are pore-forming proteins located in the outer membrane, which interact with diverse inner membrane complexes to form distinct export machinery (Andersen *et al.*, 2002a; Koronakis *et al.*, 2004). The type I secretion system is a channel-tunnel-dependent export system for proteins, such as toxins, S-layer proteins and diverse extracellular-acting enzymes. The paradigm for this export apparatus is the hemolysin secretion system of *Escherichia coli* (Wandersman & Delepelaire, 1990). In addition to protein secretion, channel-tunnels are also involved in export of noxious compounds, such as antibiotics, dyes, detergents, bile salts, or heavy metals (Poole, 2005; Borges-Walmsley *et al.*, 2003). These channel-tunnel-dependent efflux pumps play a major role in establishing the resistance of most Gram-negative bacteria to these compounds, and are becoming an increasing problem for the clinical use of antibiotics.

Currently, the structures of three channel-tunnels are known (Koronakis *et al.*, 2000; Akama *et al.*, 2004; Federici *et al.*, 2005). The trimeric assembly forms a 140 Å (14 nm) long, cannon-shaped structure. It is anchored in the outer membrane by a β-barrel domain and protrudes into the periplasm via a 100 Å (10 nm) long, α-helical tunnel domain, which assembles with inner membrane complexes to form a continuous export pathway across the cell envelope. In contrast to the wide-open extracellular entrance, the periplasmic entrance is almost closed, which explains the very low single-channel conductance of the best-characterized, representative TolC of *E. coli* (Andersen *et al.*, 2002a). For TolC, the role of individual residues in electrophysiological characteristics and opening of the periplasmic entrance has been determined (Andersen *et al.*, 2002b; Andersen *et al.*, 2002c). It can be shown that an iris-like outwards movement of the inward-folded tunnel helices is necessary to open the periplasmic entrance, allowing export of substrates (Eswaran *et al.*, 2003; Andersen *et al.*, 2002b).

In *Haemophilus influenzae*, a human pathogen responsible for significant morbidity and mortality in young children (Turk, 1984; Funkhouser *et al.*, 1991), a multidrug efflux pump has been identified, whose components are homologues of the *E. coli* AcrAB/TolC efflux pump (Sanchez *et al.*, 1997; Trepod & Mott, 2004). The AcrB homologous HI0895 is a transporter of the resistance nodulation cell division (RND) family, which forms a complex with the AcrA-homologous adaptor or membrane fusion protein HI0894 (Sanchez *et al.*, 1997). This inner membrane complex interacts with the TolC homologous, outer membrane

channel-tunnel HI1462 to assemble a functional multidrug efflux pump (Trepod & Mott, 2004). Disruption of any of the coding genes causes hypersusceptibility to antibiotic agents (Dean *et al.*, 2005; Sanchez *et al.*, 1997; Trepod & Mott, 2004). It can also be shown that a loss of pump repression by mutation or disruption of the AcrR repressor gene is the origin of increased resistance of several clinical *H. influenzae* isolates, illustrating the importance of this efflux pump (Dean *et al.*, 2005).

In this study, we have isolated the TolC-homologous HI1462 and characterized the biophysical properties of this pore-forming protein in lipid bilayer experiments. We present a computer model of the HI1462 channel-tunnel which explains its biophysical characteristics. Furthermore, we tested the ability of HI1462 to substitute the *E. coli* TolC in the AcrAB/TolC efflux pump and the HlyBD/TolC type I secretion system.

3.3. Methods

3.3.1 Construction of HI1462 and TolC expression vectors

At first, the *HI1462* gene was cloned into the pAraJS.2 vector using *H. influenzae* Rd chromosomal DNA as template and two oligonucleotides, 5´-GGTGCATTGGTGGGAGCTCCCAATATTGGCGATTCTTATC-3´ and 5´-CGGTTTCAGAAAAAGCAGCTCAGCCAGTAGATGCAGTGAAAG-3´, as primers (with restriction sites for SacI and Bpu1102I respectively). This vector is a pET12-a-based plasmid, with insertion of a DNA cassette donating a His_{10}-tag sequence directly after the OmpT leader sequence and a more variable multicloning site (SmaI/BamHI/NcoI/SacI/XhoI/BsrGI/Bpu1102I). The cassette was constructed wizh oligonucleotides JS2_up (5´-TCGACGCATCATCACCATCACCATCACCATCACCACGGCGCCGAAGGCCGCCCCGGGATCCATGGAGCTCGAGTGTACAGC-3´) and JS2_down (5´-AGTCGACATGTGAGCTCGAGGTACCTAGGGCCCCGCCGGAAGCCGCGGCACCACTACCACCACCACTACCACTACTACGC-3´), leading to SalI and Bpu1102I overhangs, and ligated into a SalI/Bpu1102I-digested pET12-a vector using T4 DNA Ligase, resulting in p12JS.2. Due to the tighter control of the araBAD promotor compared to the plac promotor of p12JS.2, the initially mentioned promotor, together with *araC*, was amplified by PCR using the oligonucleotides 5´-GCGCGAGGCAGCAGATCTATTCG-3´ and 5´-CGAGCTCGGATCCATATGTAATTCCTCCTGTTAGCCC-3´ as primers and pBAD/myc-

His A (Invitrogen) as template. The plac promotor of p12JS.2 was now replaced by araBADC using BglII/NdeI restriction sites resulting in pAraJS.2.

In the same way, the pAra21a vector was made by substituting the IPTG-inducible T7 promotor region of pET21-a (Novagen) with araBADC. Then, the *HI1462* gene, conjoined with the N-Terminal His-tag, was cloned from pAraJS.2HI1462 into pAra21a using the primers 5´-GAATTACATATGCGGGCGAAAC-3´ and 5´-GATTAATCCAATTTGATGAATTCAAATAATATCCCGCCATTG-3´ (with restriction sites for NdeI and EcoRI, respectively). The resulting vector contained the *HI1462* gene with N-terminal His$_{10}$-tag and a C-terminal His$_6$-tag under the control of an arabinose-inducible promotor.

The mutation of the arginine residue 396 to alanine was performed using the QuikChange Site-Directed Mutagenesis kit (Stratagene) on pAra21HI1462biHis with the mutation primers 5´-GCGTATCCGAATTGGCCGAATGGTTAGTTGC-3´ and 5´-GCAACTAACCATTCGGCCAATTCGGATACGC-3´, following the standard protocol.

TolC was cloned into pBAD/myc-His C using *E. coli* chromosomal DNA as template and the oligonucleotides 5´-CCACAAGGAATGCTCATGAAGAAATTGC-3´ and 5´-GTCGTCATCAGTTACGGAATTCGTTATGACCG-3´ as primers (with restriction sites for PagI and EcoRI, respectively). All constructs were confirmed by DNA sequencing (Seqlab, Germany).

3.3.2. *Bacterial strains and growth conditions*

The *E. coli* strains and plasmids used in this study are listed in Table 3.1. The *tolC* knockout in AG100 and DC14 was performed according to the method of Datsenko & Wanner (2000). A TolC-KO_up primer (5´-CGCGCTAAATACTGCTTCACCACAAGGAATGCAAATGAAG AA<u>GTGTAGGCTGGAGCTGCTTC</u>-3´), a TolC-KO_down primer (5´-CCGTTACTGGTGTAGTGCGTGCGGATGTTTGCTGAACGACTG<u>CATATGAATATCC TCCTTA</u>-3´) (underlined region binds to pKD3) and pKD3 as template were used for PCR to produce a *tolC* knockout fragment. The loss of *tolC* was verified by colony PCR using primers that bind up- and downstream of *tolC*. The resulting strains were denoted AG100TC and DC14TC, respectively.

E. coli strains were grown at 37°C with aeration in Luria-Bertani (LB) medium (Difco). When necessary, selective antibiotics were added.

Table 3.1: Strains and vectors

Strains	Genotype	Reference
AG100		Jellen-Ritter & Kern, 2001
AG100TC	Δ*tolC*	*
DC14	Δ*acrAB::kan*	Jellen-Ritter & Kern, 2001
DC14TC	Δ*tolC*, Δ*acrAB::kan*	*
BL21(DE3)Omp8	Δ*lamB ompF::Tn5* Δ*ompA* Δ*ompC*	Prilipov *et al.*, 1998

Vectors		
pET12-a		Novagen
p12JS.2	pET12-a + JS.2-cass.	*
pET21-a		Novagen
pAra21a	pET21-a + *araCaraBAD*	*
pAraJS.2	pET12-a + JS.2-cass. + *araCaraBAD*	*
pAraJS.2HI1462	*HI1462*	This work
pAra21HI1462biHis	*HI1462* (N-term. His$_{10}$- and C-term. His$_6$-tags)	This work
pAra21HI1462R396A	*HI1462RA* (N-term. His$_{10}$- and C-term. His$_6$-tags)	This work
pBAD/myc-His C		Invitrogen
pBADTolC	*TolC* (C-term. His$_6$-tag)	This work
pRSC2	Wt hly-cluster: *hlyR,C,A,B,D*; Cm^R	Schulein *et al.*, 1992

* Stegmeier & Andersen. Unpublished data

Hemolysis was tested on blood-agar plates consisting of 25ml LB agar supplemented with 0.75 ml defibrinated horse blood, spotted with 0.2 μl cell suspension (1 x 10^9 cells ml^{-1}) and incubated for 18 h at 37°C. MIC tests were performed in 96-well plates by serial dilution of the tested antimicrobial agent in LB medium with 2 x 10^4 cells per well as the inoculum. The plates were incubated overnight at 37°C and analysed photometrically after 24 h with an ELISA reader (Molecular Devices) at 600 nm.

3.3.3. Protein expression

E. coli BL21(DE3)Omp8 pAra21HI1462biHis was routinely grown at 37°C with shaking (150 rpm) in LB Medium with ampicillin. Expression of HI1462 was induced by adding 0.02% (w/v) arabinose at OD_{650} 0.5 - 0.6. The culture was grown at 30°C to prevent formation of inclusion bodies. After 3 h, cells were harvested by centrifugation (5,000 g for 15 min) and washed twice in 10 mM Tris/HCl (pH 8.0). Cells were passed three times through a French pressure cell at 900 p.s.i. (6210 kPa). Unbroken cells were removed by centrifugation at 5,000 g for 15 min. The cell envelope was obtained by centrifugation of the supernatant at 170,000 g for 60 min. The pellet was successively resuspended in 2 ml 1% (v/v) Triton (twice), 1% (v/v) LDAO and 2% (v/v) LDAO followed by centrifugation at 170,000 g for 60 min, respectively. The 2% (v/v) LDAO supernatant contained HI1462. For further purification, the supernatant was separated by FPLC using a MonoQ column (Pharmacia) and 0.5% (v/v) LDAO 10 mM Tris pH8.0 as buffer. For elution of bound proteins, the KCl concentration in the buffer was raised continuously from 0 to 1M. HI1462 eluted at 200 mM KCl. Finally, HI1462-containing fractions were applied to His-Select Ni-NTA spin-columns (Sigma).

3.3.4. SDS-PAGE and Western blotting

SDS-PAGE was performed according to the Laemmli gel system (Laemmli, 1970). The gels were stained with Coomassie brilliant blue or with silver stain (Blum *et al.*, 1987). Western blots were processed with the Bio-Rad Tank-Blot system with Protran nitrocellulose transfer membranes (Schleicher & Schuell BioScience) by applying a constant current of 350 mA for 30 min. The ECL Western Blotting Detection kit (Amersham Pharmacia Biotech) was used to detect binding of the Anti-His antibody according to the manufacturers instructions.

3.3.5. Lipid bilayer experiments

The methods used for the lipid bilayer experiments have been described in detail by Benz *et al.* (1978). Black lipid bilayer membranes were obtained from a 1% (w/v) diphytanoyl phosphatidylcholine (Avanti Polar Lipids) in *n*-decane. The temperature was maintained at 20°C during all experiments. Zero-current membrane potentials were measured by establishing a salt gradient across membranes containing 100 - 200 channels, as described by Benz *et al.* (1985).

3.3.6. Protein modelling

The HI1462 structure was modelled using the Homology module of the InsightII software package (Accelrys) based on the structure of the closest homologous channel-tunnel, OprM of *Pseudomonas aeruginosa* (Akama *et al.*, 2004). A refinement of the structure was performed using the CharmM module of the InsightII software package.

3.4. Results

3.4.1. Purification of HI1462

In a first attempt, we tried to purify the N-terminal His-tagged version of HI1462 expressed by pAraJS.2HI1462 in *E. coli* BL21(DE3)Omp8. The protein was extracted from the cell membrane by 2% (v/v) LDAO. However, purification by affinity chromatography revealed that binding of HI1462 to the Ni-NTA column was weak and the majority of the protein was already eluted when washing the column with 50 mM imidazole. HI1462 eluted by 250 mM imidazole showed a single band in SDS-PAGE but reconstitution in the very sensitive planar lipid bilayer setup revealed that the sample contained other pore-forming proteins, which were also observed when adding detergent extracts of BL21(DE3)Omp8 cells not expressing HI1462. In order to obtain purer protein samples, we constructed a double His-tagged version of HI1462. The protein bound much tighter to the Ni-NTA column, which made it possible to wash the column extensively, eliminating non-specifically bound protein. Elution of HI1462 by 500 mM imidazole (pH 8.0) resulted in highly pure protein (Fig. 3.1). When incubated at 100°C prior to SDS-PAGE, the protein ran at 48 kDa, concordantly with the calculated value for the HI1462 monomer. Without a denaturing step, the band shifted to ~ 110 kDa, corresponding to HI1462 trimer. By Western blot analysis, we could clearly confirm that both bands represented HI1462.
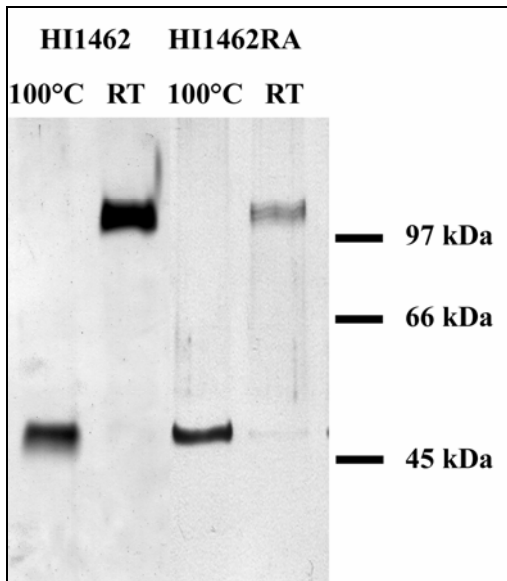
Figure 3.1: SDS-PAGE (10% polyacrylamide) of His-tagged HI1462 and HI1462RA samples, performed according to the method of Laemmli and silver-stained (Laemmli, 1970; Blum *et al.*, 1987). Purified protein samples (10 μl) were treated for 10 min at 100°C or at room temperature (RT) before loading.

3.4.2. Reconstitution of HI1462 in black lipid membranes

For biophysical characterization of purified HI1462, protein was reconstituted in black lipid membranes. After addition of 2 ng protein to the aqueous phase, a stepwise increase of conductance was observed; each step corresponding to membrane-insertion of a single HI1462 trimer (Fig. 3.2). We assumed that the double His-tag on both ends of the protein had no influence on the single-channel conductance. It is very unlikely that the N-terminal His-tag, which was located at the 37-residue-long tail outside the conduit structure, had any influence on the channel-forming characteristics. The double-tagged version led to the same conductance steps as the single-tagged version, therefore, we concluded that the C-terminal His-tag, which was closer to the tunnel structure but still not part of it, also had no influence.
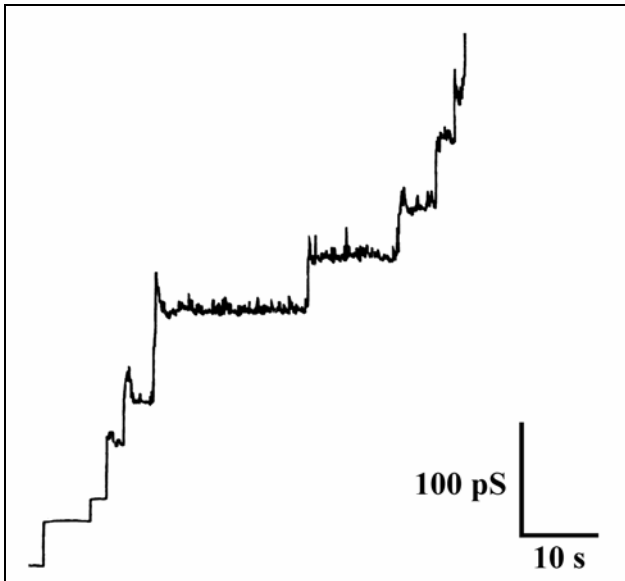
Figure 3.2: Single-channel recording of the HI1462 wild-type pores. Channel-forming activity was measured with lipid bilayer membranes made from a 1% (wt/vol) solution of diphytanoyl phosphatidylcholine (Avanti Polar Lipids) in *n*-decane (Benz et al., 1978). The aqueous phase contained 1 M KCl and 2 ng protein ml^{-1}. The applied membrane potential was 50 mV and the temperature was 20°C. The single-channel recordings were performed using Ag/AgCl electrodes (with salt bridges) connected in series to a voltage source and a Keithley 427 current amplifier. The amplified signal was monitored on a strip chart recorder.

The single-channel conductance in 1 M KCl was 43 pS. Measured in potassium acetate and LiCl, the single-channel conductance is 21 and 39 pS, respectively. This means that exchange of the cation had only a minor effect, while exchange of the anion led to a major decrease. This is the first hint that HI1462 is an anion-specific channel. Ion-specificity was quantified by zero-current potential measurements in the presence of salt gradients. A fivefold KCl gradient across a membrane with reconstituted HI1462 led to a membrane potential of 11.6 mV, negative on the diluted side. Analysis using the Goldman-Hodgkin-Katz equation resulted in a permeability ratio (P_K/P_{Cl}) of 0.5. Thus, the HI1462 channel is anion-selective, as suggested by the single-channel recordings.

Measuring the single-channel conductance in lower KCl concentrations revealed a linear dependence between 1 and 0.3 M KCl. A decrease to 0.1 M KCl resulted in almost no further reduction of the single-channel conductance. The single-channel conductance recorded in 3 M KCl showed major flickering, and conductance steps varied between 60 and 100 pS with a mean value of ~ 81 pS.

Further biophysical characterization of HI1462 channels revealed that the pores were not voltage-dependent in a range between +100 and -100 mV, meaning that the conductance remained unchanged, and no opening or closing was observed. Unlike TolC of *E. coli*, HI1462 channels could not be blocked by the divalent cation Zn^{2+}.

3.4.3. *Model of the HI1462 structure explains its biophysical characteristics*

The biophysical analysis revealed that HI1462 is anion-specific. This is a major difference to TolC of *E. coli*, which is highly selective for cations (Andersen *et al.*, 2002a). To obtain a structural insight at the molecular level, we modelled the structure of HI1462. At present, the structures of three channel-tunnels are known: TolC of *E. coli*, OprM of *P. aeruginosa* and VceC of *Vibrio cholerae* (Koronakis *et al.*, 2000; Akama *et al.*, 2004; Federici *et al.*, 2005). A sequence alignment of the HI1462 sequence showed the best homology with the sequences of OprM (data not shown). Therefore, we used OprM as a template to model the HI1462 structure (Fig. 3.3A). Of particular interest are residues at the periplasmic end of the tunnel domain. It is known from *E. coli* TolC that residues lining the periplasmic entrance have a major effect on the electrophysiological behavior of the channel-tunnel. Looking at the HI1462 model, there are valine residues (Val400) located on the second helical turn of the most inwardly directed helix, facing the lumen of the periplasmic entrance (Fig. 3.3B). These hydrophobic residues might be responsible for the low single-channel conductance observed for HI1462. Below this hydrophobic region, there are three charged residues per monomer, Glu394, Arg396, and Glu397, whereas the arginine residues are nearest to the lumen and are responsible for the electropositive lining of the periplasmic entrance. We suppose that these arginine residues are the origin for the anion-specificity of HI1462.
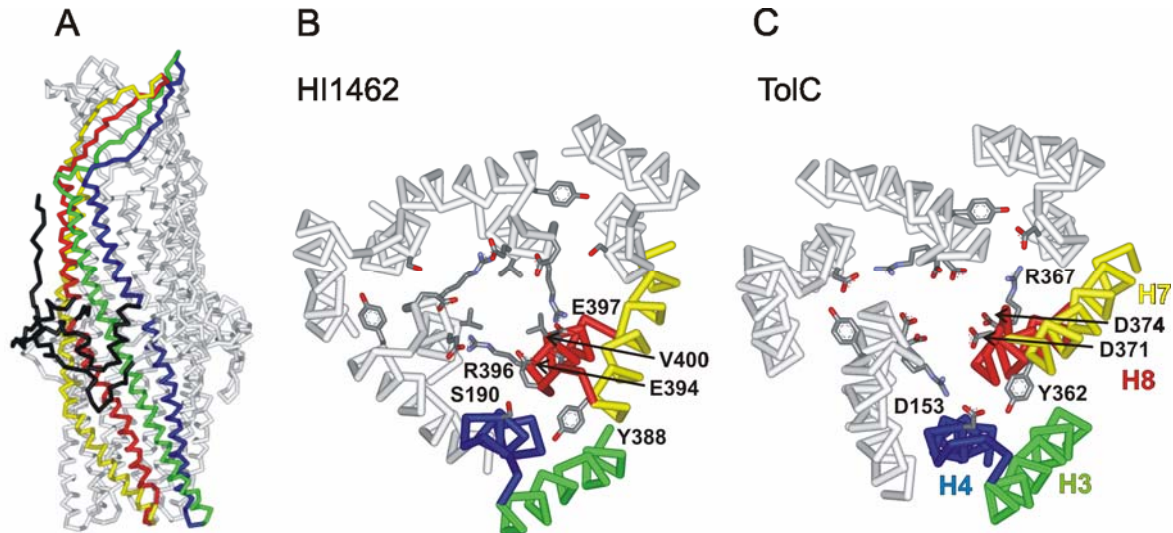
Figure 3.3: (A) Structural model of an HI1462 trimer shown as a backbone representation. One monomer is coloured. The backbone of the peptide chain outside of the tunnel domain is shown in black. The α-helices and β-strands forming the channel-tunnel are highlighted in the same colour scheme as in (B). Tunnel entrance of HI1462 (B) and of TolC (C) seen from the periplasmic side. Helices of one of the three monomers are coloured green (H3, outer coiled coil), blue (H4, outer), yellow (H7, inner coiled coil) and red (H8, inner). Important residues are shown in detail. Depictions were generated using WebLab Viewer (Accelrys).

3.4.4. A single-point mutation at the periplasmic entrance changes the biophysical properties of HI1462

To prove the dominant role of Arg396 in the electrophysiological properties of HI1462, we substituted the amino acid with alanine. The resulting mutant HI1462RA was purified and electrophysiologically characterized in the same way as the HI1462 wild-type. The single-channel conductance in 1 M KCl was 282 pS, which was almost sevenfold higher than that of HI1462 wild-type, showing that the amino acid substitution had, in fact, a massive effect on the single-channel conductance. In contrast to HI1462 wild-type, there was an almost linear dependency of the single-channel conductance with a KCl concentration of 0.03 - 1 M KCl. At a higher salt concentration, it was not possible to measure single insertion events. Here, the protein-induced conductance increase was very noisy and no defined step could be identified. The single-channel conductance measured in LiCl and potassium acetate gave the first hint that the ion selectivity was also changed in HI1462RA. In contrast to HI1462 wild-type, the change of cation from K^+ to Li^+ led to a massive decrease in single-channel conductance, whereas the single-channel conductance measured in potassium acetate was similar to that measured in KCl (Table 3.2). A quantitative determination of ion selectivity by zero-current measurements supported the assumption that exchange of Arg396 by alanine reversed ion

selectivity. The permeability ratio (P_K/P_{Cl}) was 8.6, showing that HI1462RA was cation-selective.

However, beside pores with a single-channel conductance of ~ 280 pS, we also observed a second type of pore for HI1462RA with a single-channel conductance between 1.9 and 2.1 nS in 1 M KCl. The pores were rarely observed and their lifetime was very short. They closed immediately and adopted a conformation with a single-channel conductance of ~ 280 pS corresponding to the small pores (Fig. 3.4). This observation let us conclude that both pore types can be attributed to the HI1462RA protein adopting two different conformations and not to another co-purified, pore-forming protein. In lower salt concentrations, the higher conductance state was not observed.

Table 3.2: Single-channel conductance of HI1462 and HI1462RA in different electrolytes and at different electrolyte concentrations

		HI1462	HI1462RA
Electrolyte	Concentration (M)	Gsc (pS)	Gsc (pS)
KCl	3	81	n.m.
	1	43	282
	0.3	13	89
	0.15		45
	0.1	12	34
	0.03	n.m.	11
LiCl	1	39	122
KAc	1	21	269

Each single-channel conductance (Gsc) value is based on at least 80 insertion events. The applied voltage was 20 mV and the temperature was 20 °C. n.m. means not measurable.
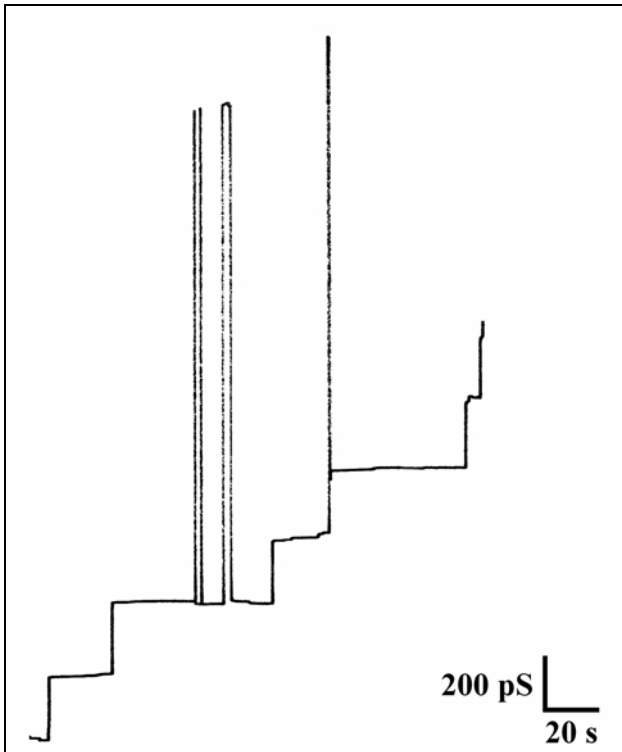
Figure 3.4: Single-channel recording of HI1462RA mutant pores. The aqueous phase contained 1 M KCl and 2 ng protein ml^{-1}. The applied membrane potential was 20 mV and the temperature was 20°C.

3.4.5. HI1462 is able to substitute E. coli *TolC in the hemolysin secretion apparatus but not in a multidrug efflux pump*

Channel-tunnels are involved in diverse export processes. We tested if HI1462 can substitute *E. coli* TolC as an outer membrane component of the type I hemolysin secretion system HlyBD/TolC, and of the multidrug efflux pump AcrAB/TolC. Hemolysin secretion was tested on blood agar plates with cells harbouring pRSC2 encoding the hemolysin operon (Fig. 3.5). Channel-tunnels TolC and HI1462 were expressed from an arabinose-inducible plasmid in the *tolC*-deficient strain AG100TC. An arabinose concentration of 0.1% (w/v) was chosen for induction of channel-tunnels because it resulted in the same zone of clearance as AG100 cells expressing TolC from the chromosome. Cells lacking TolC or HI1462 had a small zone of clearance due to hemolysin release by cell lysis. Interestingly, cells expressing HI1462 wild-type, as well as the HI1462 mutant, were able to secrete hemolysin. However, the halo was much smaller compared to cells expressing TolC. When HI1462 or TolC were induced by higher arabinose concentration, the size of the halo increased, but cells expressing HI1462 always had a smaller zone of clearance.
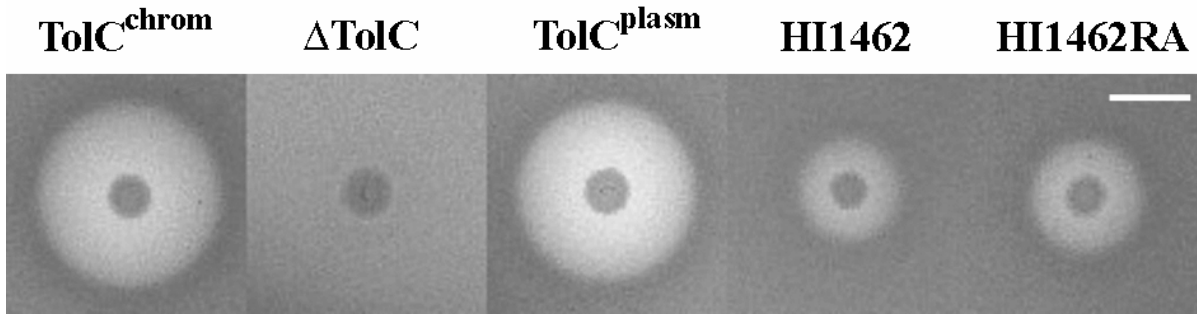
Figure 3.5: Test of hemolysin secretion on blood agar plates. All cells contained pRSC2 encoding the hemolysin operon. *E. coli* AG100 cells were used as a positive control with TolC expressed from the chromosome (TolC^{chrom}), and the TolC-deficient strain AG100TC harbouring the pBAD vector as a negative control (ΔTolC). Channel-tunnels TolC, HI1462, or HI1462RA were expressed from an arabinose-inducible plasmid in AG100TC cells. Plates were inoculated with 2x10^5 cells and grown for 18 h at 37 °C. The bar corresponds 1 cm.

To test if HI1462 is also able to substitute TolC in the multidrug efflux pump AcrAB/TolC, we determined the MIC of *tolC*-deficient AG100TC cells expressing HI1462 variants. Using TolC expressed from a plasmid-encoded, arabinose-inducible gene, we have determined the arabinose concentration necessary to reveal MIC values, which correspond to that of wild-type strain AG100. AG100TC cells expressing HI1462 or HI1462RA induced by the same arabinose concentration revealed MIC values which were indistinguishable from those of AG100TC cells expressing no channel-tunnel (Table 3.3). Therefore, we conclude that HI1462 is not able to form a functional, or at least an effective, efflux pump with the *E. coli* AcrAB complex.

Table 3.3: Susceptibility of strains depending on the channel-tunnel

	MIC (µg/ml)		
Strains and plasmids	VANC	ERM	R6G
AG100 pBad	>12	>50	>1000
AG100TC pBadTolC	>12	>50	1000
AG100TC pAra	>12	3.13	3.9
AG100TC pAra21HI1462biHis	>12	3.13	3.9
AG100TC pAra21HI1462biHisRA	>12	3.13	3.9

Abbreviations: VANC, vancomycin; ERM, erythromycin; R6G, rhodamine 6G.
All strains were induced with 0.01% arabinose

3.5. Discussion

According to a phylogenetic analysis of membrane transport systems, *H. influenzae* possesses six predicted multidrug efflux pumps (Paulsen *et al.*, 2000). Out of these, AcrAB/HI1462 has been characterized as the primary efflux pump, exporting a broad variety of toxic compounds (Sanchez *et al.*, 1997; Trepod & Mott, 2004). A decreased susceptibility of clinical isolates due to overexpression of efflux pump genes has confirmed the role of this efflux pump as a major determinant for drug resistance of *H. influenzae* (Dean *et al.*, 2005). Analysis of the chromosome has revealed that HI1462 is the only functional channel-tunnel in *H. influenzae*. A phylogenetic analysis separates the channel-tunnel family into three groups, corresponding to the export processes in which the proteins are involved: protein secretion, drug efflux and cation efflux (Andersen *et al.*, 2000). HI1462 - like OprM of *P. aeruginosa* - belongs to the group of channel-tunnels involved in drug efflux, whereas TolC of *E. coli* is a representative of the channel-tunnels involved in protein secretion. HI1462 reconstituted in planar lipid membranes revealed a single-channel conductance of 43 pS in 1 M KCl, almost half that of TolC, but 20-fold smaller than that of OprM (Andersen *et al.*, 2002a; Wong *et al.*, 2001). The low single-channel conductance of HI1462 accords well with our modelled structure (Fig. 3). The periplasmic entrance of HI1462 is lined by three valine residues (Val400) forming a hydrophobic constriction, which is most likely responsible for the lower single-channel conductance of HI1462 compared to that of TolC. In TolC, there is an aspartate at the corresponding position (Asp374). Together with a second aspartate (Asp371), which is located one helical turn closer to the periplasmic entrance, it is responsible for the cation selectivity of TolC (Andersen *et al.*, 2002c). In HI1462, the second central position is also covered by a negatively charged residue (Glu397). However, our measurements revealed that HI1462 is the first known TolC homologue which is anion-selective. This also explains why divalent cations do not block the channel as observed for TolC (Andersen *et al.*, 2002c). We showed that the anion selectivity originates from the Arg396, which also lines the tunnel entrance. The dominant role of Arg396 was investigated in the mutant HI1462RA. Substitution of arginine by alanine reversed the ion selectivity of the HI1462 channel, implying that the tunnel entrance became electronegative due to the glutamate residues at position 397 and 394 (Fig. 3). Additionally, the HI1462RA mutant had an increased single-channel conductance. However, the almost sevenfold higher value (282 pS), compared to wild-type, cannot be explained solely by the absence of the bulky arginine residues. Compared with the single-channel conductance of TolC (80 pS), which has no arginine

residue at the corresponding position, the HI1462RA entrance needs to be opened further. When looking at the HI1462 model, it becomes evident that Arg396 is involved in stabilization of the closed conformation of the HI1462 tunnel entrance. Arg396 is in close proximity to Glu397 of the adjacent monomer, allowing the formation of a circular network of salt bridges, which keeps the inward-folded helices in position. The absence of these connections in HI1462RA might allow the inner helices to move outwards, explaining the higher single-channel conductance. Interestingly, these were not the only conductance steps seen for HI1462RA. We also observed insertion events with a six- to sevenfold higher single-channel conductance, which could be undoubtedly assigned to the HI1462RA mutant protein. This high conductance level was unstable and observed only rarely at high salt concentrations. We assume that the lack of the circular network allows the tunnel entrance to temporarily adopt a fully open state. At lower salt concentrations, the high conductance state was absent. This could be explained by additional ionic interactions which keep the tunnel entrance in a narrow conformation. At high ionic strength, these interactions are destabilized, allowing transition into the fully open state, whereas at lower salt concentrations, they are stable enough to keep the entrance in the narrow conformation. For TolC, it is known that the open state can be achieved by disturbance of a circular network of hydrogen bonds and salt bridges close to the periplasmic entrance (Andersen *et al.*, 2002b). Sequence alignment shows that the aspartate (Asp153 in TolC) located at the outer coiled coil, which is central to the formation of the circular network, is not conserved in HI1462, which possesses a serine (Ser190) at the corresponding position. Therefore, it seems reasonable that there must be other connections, which keep the channel-tunnel in a closed conformation, to prevent unwanted in- or efflux through a wide-open channel-tunnel in the outer membrane. The electrophysiological characterization of the HI1462RA mutant shows that the salt bridge, which can be formed between Arg396 and Glu397 according to the modelled HI1462 structure, is most likely responsible for keeping the tunnel entrance in a tightly closed conformation.

It is surprising that OprM of *P. aeruginosa*, which shows a higher homology to HI1462 than TolC, has a single-channel conductance of 850 pS in 1 M KCl (Wong *et al.*, 2001), which is almost 20-fold higher than that of HI1462. However, the high single-channel conductance seems to conflict with the solved crystal structure of OprM, which shows a nearly closed periplasmic entrance (Akama *et al.*, 2004). It should be mentioned that the reconstituted channels are not stable and appear to form substates (Wong *et al.*, 2001), a characteristic also observed for the open state of TolC (Andersen *et al.*, 2002b). Additionally, the pore-forming activity of OprM in planar lipid bilayers is very low compared to other reconstituted pore-

forming proteins (R. Benz, personal communication), which lets us assume that only a minority of the OprM molecules reconstitutes as pores in a transient open state, whereas the majority of the molecules inserts into the membrane as closed pores, which allow no passage of ions and can therefore not be detected by electrophysiological measurements.

The genomic organization of the multidrug efflux pump in *H. influenzae* resembles that of *E. coli* to the extent that the gene encoding the channel-tunnel is not linked to those encoding the proteins of the inner membrane complex. Among the six predicted multidrug efflux pumps in *H. influenzae,* there is another export mechanism coded by a gene for the adaptor protein EmrA (HI0898) and a transporter of the major facilitator superfamily EmrB (HI0897), which also needs a channel-tunnel protein to assemble a functional efflux pump according to the known homologues in *E. coli* (Lomovskaya & Lewis, 1992). Therefore, the sole channel-tunnel in *H. influenzae* HI1462 needs to be compatible with two different inner membrane complexes. This is in accordance with TolC of *E. coli*, which is known to interact with at least eight different inner membrane complexes (Andersen, 2003). Among these is the HlyBD complex, which with TolC forms a secretion apparatus for hemolysin (Wandersman & Delepelaire, 1990). Genes encoding a type I secretion system are absent in *H. influenzae.* Therefore, it is surprising that HI1462 can substitute for TolC in the *E. coli* protein secretion apparatus. The secretion was not as efficient as that of the wild-type system, which suggests that either the amount of HI1462 is lower compared to TolC, or that the two proteins HlyD and HI1462 are not fully compatible. By higher induction of HI1462, it was possible to increase the amount of secreted hemolysin. However, we never reached the maximal secretion levels observed with cells expressing TolC. Therefore, reduced hemolysin secretion is probably caused by compatibility problems between HlyD and HI1462 in the hybrid system. A slightly bigger halo, observed for the hybrid system including the HI1462RA mutant, might lead to the supposition that facilitated opening of the tunnel entrance by interaction with HlyD plays a role in enhanced secretion. Other hybrid type I secretion systems have exchanged the outer membrane component with highly homologous channel-tunnels, e.g. TolC of *E. coli*, HasF of *Serratia marcescens*, and PrtF of *Erwinia chrysanthemii*. They can be interchanged between the different secretion systems without loss of functional secretion (Binet & Wandersman, 1996; Letoffe *et al.*, 1994; Akatsuka *et al.*, 1997). Recently, it has been shown that RaxC, a channel-tunnel of the plant pathogen *Xanthomonas oryzae* pv. *Oryzae*, complemented an *E. coli tolC* mutant in a type I secretion system (da Silva *et al.*, 2004). However, RaxC is also a channel-tunnel belonging to the protein secretion family. Thus,

HI1462 is the first example of a channel-tunnel belonging to the drug efflux family which is able to act as an outer membrane component for a type I secretion system.

Substitution of TolC by HI1462 as an outer membrane component of the *E. coli* AcrAB efflux pump did not restore resistance. Similar results have been observed when trying to complement TolC by OprM of *P. aeruginosa* (Tikhonova *et al.*, 2002). It should be mentioned that MIC values are generally much lower for *H. influenzae* than for *E. coli* (Trepod & Mott, 2004). If HI1462 is responsible for the less efficient efflux pump of *H. influenzae*, the possibility cannot be excluded that the hybrid efflux pump AcrAB/HI1462 in *E. coli* is functional but does not give rise to the same level of resistance as the wild-type efflux pump. However, there are no reports in the literature which point towards the channel-tunnel as a rate-limiting factor for the efflux. Therefore, we conclude that HI1462 is not compatible with the inner membrane complex of *E. coli*. This means that interaction between the inner membrane complexes of drug efflux pumps and the corresponding channel-tunnels is more specific than the interaction of the inner and outer components of type I secretion systems. Further research is needed to understand these interactions at the molecular level. This is essential for developing drugs that inhibit these export mechanisms and disarm harmful bacteria.

Chapter 4

Molecular characterization of interaction between AcrA and TolC

4.1   Abstract

The tripartite AcrAB/TolC multidrug efflux pump of *Escherichia coli* plays a major role in establishing resistance to a broad range of antibiotic substances. While there are several analogous systems identified and characterized in various Gram-negative bacteria and there is progress in the crystallization of representatives of its components, the exact interaction mechanisms are still not solved. Based on the known structures we designed a model of the interaction between TolC and AcrA and identified two residues K383 and R390 in TolC as conceivable contact points for residues E118 and D125 in AcrA. The determination of considerably lower MIC-values with a charge-exchange mutant of TolC revealed the importance of these residues for the correct function of the AcrAB/TolC efflux pump. It could be demonstrated that the susceptible phenotype of the mutant could be restored to wild-type level by exchanging the corresponding residues of AcrA. Furthermore, the model was confirmed by site-directed disulfide cross-linking between TolC and AcrA. The residues S339 of TolC and A93 of AcrA are in sufficient proximity in the assembled pump complex so that the formation of a disulfide bond occurred when these residues were both changed to cysteine. Integrating the results of this work and recent findings of other groups, a model of the assembling process of the AcrAB/TolC efflux pump is discussed.

4.2. Introduction

Tolerance to noxious substances in the environment is very important for bacteria, particularly for pathogens, and becomes more and more problematic in the medical treatment of bacterial diseases (Walsh, 2000). Gram-negative bacteria evolved a special pump-system to directly extrude invading antimicrobial agents out of the cells and thus reducing their intracellular concentration (Zgurskaya 2000b). Because of the broad spectrum of transported substances, these systems are called multidrug efflux (MDE) pumps. The two most extensively characterized MDE systems are AcrAB/TolC of *E. coli* and MexAB/OprM of *Pseudomonas aeruginosa* (Ma *et al.*, 1995; Hirakata *et al.*, 2002). These MDE pumps are tripartite complexes consisting of an energized inner membrane pump, an adaptor or membrane fusion protein and an outer membrane channel-tunnel (Andersen, 2003). The three subunits are assembled to a transporter unit bridging both the inner and the outer membrane. Representatives for all three subunits, each of them belonging to large families of proteins, could be crystallized in the last years. The inner membrane protein AcrB of *E. coli* is the first crystallized proton antiporter belonging to the resistance nodulation cell division (RND) family (Murakami, 2002). It comprises a homotrimer of around 110 kDa per monomer. Each subunit contains 12 transmembrane helices forming the 40 Å thick transmembrane region and two large periplasmic loops (between transmembrane helices 1 and 2 and between 7 and 8) forming a huge periplasmic domain protruding 70 Å into the periplasm. The three periplasmic domains form the headpiece of the trimer with a central pore composed of three α-helices. This pore connects a funnel-like opening at the top of the headpiece with a large central cavity at its base. This cavity has a diameter of 35 Å and opens into the periplasm via three vestibules located between the subunit interfaces. These openings probably allow direct capturing of drugs from the periplasm as well as the outer leaflet of the cytoplasmic membrane (Murakami, 2002).

Like AcrB the outer membrane component is also a homotrimer forming a cylindrical conduit. Nowadays the structures of the homologue proteins TolC of *E. coli*, OprM of *P. aeruginosa* and VceC of *Vibrio cholerae* have been determined (Koronakis *et al.*, 2000; Akama *et al.*, 2004b; Federici *et al.*, 2005). The trimers are anchored in the outer membrane by a 12 stranded β-barrel, the channel domain, and project across the periplasmic space with a 100 Å long α-helical barrel, the tunnel domain. Recently, a direct interaction between the rim of the AcrB funnel and the periplasmic end of the channel-tunnel could be detected by site

directed disulfide cross-linking (Tamura *et al.*, 2005). This means that a direct passage of exported compounds from the transporter into the channel-tunnel is possible.

For the assembly of a functional efflux pump, however, a third component, the adaptor protein, is also essential (Fralick, 1996). Whereas the inner and outer membrane proteins could be crystallized in their native oligomerisation state, the first member of the adaptor protein family MexA of *P. aeruginosa* crystallized as tri-decamer (Akama *et al.*, 2004a; Higgins *et al.* 2004). The second representative adaptor protein AcrA of *E. coli* however packs as an apparent dimer of dimers (Mikolosko et al., 2006). It seems implausible that one of these assemblies represents the native form of the proteins and, as discussed in detail in chapter 2, we prefer the model of a hexameric state of the adaptor protein in the pump complex. In parts, however, the structure of the monomers verified the former structural model of adaptor proteins (Johnson and Church, 1999) showing a 47 Å long α-helical hairpin domain connected to a flattened β-sandwich domain folded like the already known lipoyl domain from biotinyl/lipoyl carrier proteins for MexA. Beside these already predicted domains, a third domain, the α/β domain could be solved showing a six-stranded β-barrel with a short α-helix. It is expected that at least a fourth domain comprising the N- and C-terminus of the protein exists, because the structure of the 28 N-terminal and 101 C-terminal residues could not be solved (Higgins *et al.*, 2004). The three-dimensional structure of AcrA strongly resembles MexA as expected considering their high sequence identity (Mikolosko et al., 2006)

In all models of multidrug efflux pumps it is assumed that the adaptor protein mediates the contact between the inner membrane transporter and the outer membrane channel-tunnel (Akama *et al.*, 2004a; Fernandez-Recio *et al.*, 2004; Higgins *et al.*, 2004). Recent experimental evidence, showing that AcrA can be chemically cross-linked with both TolC and AcrB, independent of the presence of a pump substrate (Husain *et al.*, 2004; Tikhonova and Zgurskaya, 2004; Touze *et al.*, 2004), supported this hypothesis. This also means that the efflux pump assembly is permanent in contrast to the analogous type I secretion system, for which it is shown that interaction between the inner membrane complex and the outer membrane component is transient and depends on the presence of the substrate (Thanabalu *et al.*, 1998). Concerning the assembly of drug efflux pumps, it is still an open question, which domains of the adaptor protein are necessary for functional interaction with the channel-tunnel. The core fragment composed of residues 45-315 has been found to be insufficient to bind TolC suggesting a requirement for either or both the N- and C-terminus of AcrA (Touze et al., 2004; Mikolosko et al., 2006). Additionally, the findings that point mutations in the α/β

domain of AcrA could reverse the hypersensitive phenotype of a TolC mutant implicate that also this region might have contact with the channel-tunnel (Gerken and Misra, 2004). However, a direct influence of the termini and the α/β domain on the TolC association is not clear. Mikolosko and coworkers assume their influence as indirectly by mediating the AcrA oligomerization. And finally there is the open question if the two interacting proteins act as mutual chaperons within the assembling process of the efflux pump.

In the present study the possible interaction site between the TolC tunnel domain and the hairpin domain of AcrA was modeled. Based on this model it was attempted to verify the assumptions by analyzing the functionality of point-mutated TolC and AcrA and by screening for closely adjoining residues in disulfide cross-linking experiments.

4.3. Material and methods

Table 4.1: Strains and plasmids

Strains	Genotype	Reference
AG100		Jellen-Ritter & Kern, 2001
AG100TC	$\Delta tolC$	*
DC14	$\Delta acrAB::kan$	Jellen-Ritter & Kern, 2001
DC14TC	$\Delta tolC$, $\Delta acrAB::kan$	*

Vectors		
pACYC184	Cm^R	Chang & Cohen, 1978
pAX629	pACYC184 derivative + $tolC$	Hiraga et al., 1989
pAXGP1	pAX629 + $tolC\ K383E\ R390E$	This work
pBAD/myc-His C	Amp^R	Invitrogen
pBADacrAB	pBAD/myc-HisC + $acrAacrB$	*
pBADacrAB-CH1	pBAD/myc-HisC + $acrA\ E118K\ D125R\ acrB$	This work
pBADTolC	$TolC$ (C-term. 6xHis-Tag)	This work
pBADGP1		
pRSC2	Wt hly-cluster: $hlyR,C,A,B,D$; Cm^R	Schulein et al., 1992
pASK-IBA12	Amp^R	IBA
pASK-TolC	pASK-IBA12 + $tolC$	This work
pASK-TolC$_{S124C}$	pASK-IBA12 + $tolC\ S124C$	This work
pASK-TolC$_{A128C}$	pASK-IBA12 + $tolC\ A128C$	This work
pASK-TolC$_{A132C}$	pASK-IBA12 + $tolC\ A132C$	This work
pASK-TolC$_{S339C}$	pASK-IBA12 + $tolC\ S339C$	This work
pASK-TolC$_{A343C}$	pASK-IBA12 + $tolC\ A343C$	This work
pUC18acrA	pUC18 + $acrA$	*
pBADacrAB-cat	pBADacrABΔbla + cat, Cm^R	This work
pBADacrA$_{A93C}$B-cat	pBADacrAB-cat + $acrA\ A93C$	This work
pBADacrA$_{A94C}$B-cat	pBADacrAB-cat + $acrA\ A94C$	This work

* Stegmeier et al., unpublished data

Table 4.2: Primers

Name	Sequence in 5´-3´ direction
TolCK383ER390E_up	CGTTGTACAACGCCGAGCAAGAGCTCGCGAATGCGGAGTATAACTACCTG
TolCK383ER390E_down	CAGGTAGTTATACTCCGCATTCGCGAGCTCTTGCTCGGCGTTGTACAACG
pASKTolCup	ATGGTAGGTCTCACTCCGAGAACCTGATGCAAGTTTATCAG
pASKTolCdown	ATGGTAGGTCTCATATCATTCGTTATGACCGTTACTGGTGG
TolC_S124C_QCup	GCTATTGACGTTCTTTGCTATACACAG
TolC_S124C_QCdown	CTGTGTATAGCAAAGAACGTCAATAGC
TolC_A128C_QCup	CTTTCCTATACACAGTGCCAAAAAGAAGCG
TolC_A128C_QCdown	CGCTTCTTTTTGGCACTGTGTATAGGAAAG
TolC_A132C_QCup	CAGGCACAAAAAGAATGCATCTACCGTCAATTAG
TolC_A132C_QCdown	CTAATTGACGGTAGATGCATTCTTTTTGTGCCTG
TolC_S339C_QCup	GCATCTATCTGTAGCATTAACGCCTA
TolC_S339C_QCdown	TAGGCGTTAATGCTACAGATAGATGC
TolC_A343C_QCup	CAGTAGCATTAACTGCTACAAACAAGCCG
TolC_A343C_QCdown	CGGCTTGTTTGTAGCAGTTAATGCTACTG
AcrAB_up	CTCGAGGTTTACTCATGAACAAAAACAGAGG
AcrAB_down	GGCATGTCTTAACGGCTCGAGTTTAAGTTAAGACTTGG
E118KD125R_QCup	CAGTAAGCAAAAGTACGATCAGGCTCTGGCTCGTGCGCAACAGGC
E118KD125R_QCdown	GCCTGTTGCGCACGAGCCAGAGCCTGATCGTACTTTTGCTTACTG
Cat_BspHI_up	AATCATGAGCACCTCAAAAACACCATCATACAC
Cat_BspHI_down	AATCATGACAGGCGTTTAAGGGCACCAATAACT
AcrA_A93C_QCup	GAAAGCCCAGTGTGCAGCCAATATCGCGCAATTG
AcrA_A93C_QCdown	CAATTGCGCGATATTGGCTGCACACTGGGCTTTC
AcrA_A94C_QCup	GAAAGCCCAGGCTTGTGCCAATATCGCGCAATTG
AcrA_A94C_QCdown	CAATTGCGCGATATTGGCACAAGCCTGGGCTTTC

4.3.1. Construction of TolC knock-out strains and plasmids

The *tolC* knock-out in AG100 and DC14 was performed following the method of Datsenko and Wanner (2000) using a knock-out fragment produced by PCR with the primer pair TolC-KO_up and TolC-KO_down and pKD3 as template (see Chapter 3). The loss of *tolC* was verified by PCR. The resulting strains were denoted AG100TC and DC14TC, respectively.

Plasmids were constructed using standard recombinant DNA techniques (Sambrook *et al.*, 1988). PCRs were performed using Pfu Turbo DNA polymerase (Stratagene) according to the manufacturers' manual and DNA was purified using the Nucleospin Extract Kit (Macherey & Nagel). Plasmids were isolated using the Nucleospin plasmid Kit (Macherey & Nagel). All DNA modifying enzymes were supplied by MBI Fermentas or New England Biolabs. All constructed plasmids were verified by DNA sequencing (SEQLAB).

The plasmid pAXGP1 was constructed using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) on the pAX629 plasmid with the primers TolCK383ER390E_up and TolCK383ER390E_down following the standard protocol. The resulting vector carries the mutant *tolC K383E R390E* gene.

The construction of pBADTolC is described Chapter 3.

For cloning *tolC* into the pASK-IBA12 vector the gene from pBADTolC was amplified with the primers pASKTolCup and pASKTolCdown. Both the PCR product and the pASK-IBA12 vector were BsaI digested and ligated, resulting in pASK-TolC, carrying the wild-type *tolC* with N-terminal OmpA leader sequence and Strep-tag II. The cysteine mutations in TolC were introduced by QuikChange PCR on pASK-TolC with the primer pairs TolC_S124C_QCup/down, TolC_A128C_QCup/down, TolC_A132C_QCup/down, TolC_S339C_QCup/down and TolC_A343C_QCup/down, respectively.

Primers AcrAB_up and AcrAB_down were used for amplifying *acrAB* and the PagI/XhoI digested PCR product was inserted into the NcoI/XhoI digested pBAD/myc-His C vector resulting in pBADacrAB. For AcrA-CH1 mutation QuikChange PCR was performed using pBADacrAB as template and primers E118KD125R_QCup and E118KD125R_QCdown. To avoid unwanted mutations in the large *acrAB* genes the BamHI/AsiSI *acrA-CH1* fragment was religated into the BamHI/AsiSI digested pBADacrAB vector giving the pBADacrAB-CH1 plasmid.

For disulfide cross-linking experiments it was necessary to replace the *bla* resistance gene of pBADacrAB with the *cat* gene of pACYC184 transferring chloramphenicol instead of ampicillin resistance. Therefore, the *cat* gene was amplified from pACYC184 with the primers Cat_BspHI_up and Cat_BspHI_down. The PCR product and the vector pBADacrAB

were digested with PagI and ligated resulting in pBADacrAB-cat. Cysteine mutations were introduced in AcrA by QuikChange PCR using primer pairs AcrA_A93_QCup/down and AcrA_A94C_QCup/down, respectively on pUC18acrA. The mutated region of *acrA* was then excised by Bpu1102I digestion and inserted into comparably treated pBADacrAB.

4.3.2. *Growing conditions*

All bacterial strains were grown at 30°C or 37°C in LB medium (DIFCO) or on LB agar plates with appropriate antibiotics (Sigma). 100μg/ml ampicillin, 50μg/ml kanamycin and 40μg/ml chloramphenicol were used for selection of plasmids in the *E. coli* strains Top10F` (Invitrogen), AG100, DC14 (Jellen-Ritter and Kern, 2001). The *E. coli tolC*-knock out strains AG100TC, DC14TC were grown with reduced antibiotic concentrations (50μg/ml ampicillin, 25μg/ml kanamycin and 10μg/ml chloramphenicol).

4.3.3. *Expression of pump components*

E. coli DC14TC with pBADTolC or pBADGP1 was grown at 37°C with shaking (150 rpm) in LB-Medium with ampicillin. Expression of TolC was induced by adding 0.002% (w/v) arabinose at OD_{650} = 0.5 to 0.6. Subsequently the cultures were grown at 30°C to prevent formation of inclusion bodies. After 3 hours of induction the cells were harvested by centrifugation (5,000xg for 15 min) and washed twice in 10 mM Tris-HCl (pH 8.0).

DC14TC cells harbouring pAX629 or pAXGP1 were grown at 37°C with shaking (150 rpm) in LB-Medium with chloramphenicol until OD_{650} = 1. DC14TC cells with pAX629 or pAXGP1 and additionally pBADacrAB or pBADacrAB-CH1 in any combination were grown at 37°C with shaking (150 rpm) in LB-Medium supplemented with chloramphenicol, ampicillin and 0.01% (w/v) arabinose again until OD_{650} = 1. Harvesting of the cells was performed as described above.

4.3.4. *Isolation of the outer membrane by sucrose-step gradient centrifugation*

The harvested cells were passed three times through a French pressure cell at 900 p.s.i. (= 6210 kPa). Unbroken cells were removed by centrifugation at 5,000xg for 15 min. The cell envelope was obtained by centrifugation of the supernatant at 170,000xg for 60 min. The membrane pellets of 50 ml initial culture were carefully homogenized in 1 ml 10 mM Tris pH 8.0. Of each sample 0.5 ml were loaded on a sucrose-step gradient composed of 3 ml 30%, 4 ml 50% and 3 ml 70% sucrose. The gradient was centrifuged at 114,000xg at 4°C for 16 hours using a Beckman SW40Ti swinging bucket rotor. After centrifugation the inner and

outer membrane containing regions were visible as separated bands and were collected in different fractions.

4.3.5. In vivo *spontaneous disulfide cross-linking assay*

DC14TC cells were transformed with pBADacrA$_{A93C}$B-cat or pBADacrA$_{A94C}$B-cat and pASK-TolC or one of the five cysteine-mutant pASK-TolC plasmids, respectively in any combination. The strains were grown in 25 ml LB medium with ampicillin, chloramphenicol 25ng/ml anhydrotetracycline (AHT) and 0.01% arabinose at 30°C until an OD$_{600}$ of 1.

After harvesting the cells by centrifugation for 15 min at 5,000xg the pellets were resuspended in 3 ml 10 mM Tris pH 8.0, 40 mM *N*-ethylmaleimide. The cells were disrupted by repeated brief sonication on ice and cell-membranes were collected by centrifugation at 170,000xg for 1 hour at 4°C. The resulting pellets were resuspended in 8 M urea before SDS-PAGE and Western-blot analysis.

4.3.6. Gel-Electrophoresis and Western-Blot

Prior to electrophoresis equal numbers of cells or aliquots of purified proteins were incubated in sample loading buffer for 10min at 100°C if not described differently. SDS-PAGE was performed according to the Laemmli gel system (Laemmli, 1970). For Western blots, a tank blot system (Amersham Biosciences) was used as described previously (Towbin *et al.*, 1979). The anti-His and the HRP-linked anti-mouse, respectively HRP-linked anti-rabbit antibodies were purchased from Amersham Biosciences, the TolC and the AcrA specific antibodies from rabbit were produced by Pineda Antibody Services, Berlin. ECL Western Blotting Detection Reagents (Amersham Biosciences) were used for detection.

4.3.7. Determination of minimal inhibitory concentration (MIC)

The bacterial strains were grown overnight in LB media supplemented with 50µg/ml ampicillin and 10µg/ml chloramphenicol. Afterwards, the cultures were diluted 1:100 into 5ml fresh LB media supplemented with 50µg/ml ampicillin, 10µg/ml chloramphenicol and 0.01% arabinose and grown for 5 hours at 37°C. The cultures were diluted and added to a 96-well-plate (Cellstar) using an initial concentration of 100 cells per well. In every well 25µg/ml ampicillin, 5µg/ml chloramphenicol, 0.01% arabinose and antimicrobial agents in following concentrations were given: novobiocin 0, 5, 10, 20, 40, 60, 80, 100, 120, 150, 180, 210 µg/ml, rhodamine 6G 0, 7.8, 15.6, 31.3, 62.5, 93.8, 125, 187.5, 250, 375, 500, 750 µg/ml, erythromycin 0, 1.56, 3.13, 6.25, 12.5, 25, 37.5, 50, 75, 100, 150, 200 µg/ml, benzalkonium

chloride 0, 0.6, 1.3, 2.5, 4.1, 5, 8.1, 10, 16.3, 20, 32.5, 40 µg/ml and sodium dodecyl sulfate (SDS) 0, 1.25, 2.5, 5, 10, 12.5, 20, 25, 40, 50, 80, 100 mg/ml. The 96-well-plates were incubated for 24h at 37°C without shaking and growth was verified using an ELISA Reader (Molecular Devices). The MIC was determined by at least four independent MIC tests per strain.

4.3.8. *Hemolysis assay*

Hemolysis was tested on blood-agar plates consisting of 25ml LB-agar supplemented with 0.75 ml defibrinated horse-blood, 50µg/ml ampicillin, 10µg/ml chloramphenicol and 0.001% arabinose. To test hemolysis, 0.2 µl cell-suspension ($1*10^9$ cells ml^{-1}) was spottet on the blood-agar plates and incubated 18 h at 37°C.

4.3.9. *Modelling of the AcrA/TolC interaction site*

The structure of AcrA was taken from Mikolosko *et al.* (2006). The data of TolC came from Koronakis *et al.* (2000). The binding sites of the TolC channel tunnel with AcrA adaptor protein were first identified by bioinformatical approach and then supported by molecular modeling using the CharmM module of the InsightII software package (Accelrys) (as described by Stegmeier *et al.*, submitted).

4.4. Results and discussion

4.4.1. *Identification of potential interaction sites between AcrA and TolC*

The considerations about the possible interaction sites between the single components based on the three dimensional model of an assembled efflux pump (see Fig. 2.3, chapter 2). In nonassembled state the TolC trimer is nearly closed at the tunnel mouth by the inwardly twisted coiled coils formed by helices H7 and H8 (Koronakis *et al.*, 2000). For functional transport of substances or even large proteins like hemolysin A the tunnel has to be opened by an iris-like outward movement of the coiled coils (Andersen *et al.*, 2002b). This opening is most likely caused by interactions with the inner membrane transporter complex of AcrA and AcrB. Based on the modeled pump assembly we assume that the hairpin domain of AcrA enables opening of TolC. The importance of the hairpin was shown by Stegmeier *et al.* (unpublished data) by an AcrA mutant lacking the 74 residues (97-170), forming the hairpin. When this AcrAΔHP mutant was expressed together with AcrB and TolC, the amount of

AcrAΔHP detected by immunoblot was over 100-fold lower compared to AcrA wild-type and the construct failed to mediate resistance. This suggests that deletion of the hairpin domain destabilizes the protein explaining why the efflux pump is not functional. Stegmeier, however, could exclude that contact with the outer membrane component is necessary for correct folding of AcrA, because the protein level is not reduced in a TolC deficient background.

In the detailed interaction model the focus was set on the domain of the α-helical coiled-coils underneath the equatorial domain of TolC, which in our opinion is the most likely interaction site, because here the inwardly directed helices H7 and H8 have to carry out the widest movement for the opening of the channel-mouth. We searched for charges on the surface of the trimeric TolC in this region which are accessible for countercharges on the AcrA α-helical hairpin domain. This way the residues K383 and R390 on the helix H8 of TolC and E118 and D125 on the AcrA hairpin were identified as possible partners for electrostatic interactions, which probably help to open the channel-mouth (see Figure 4.1).
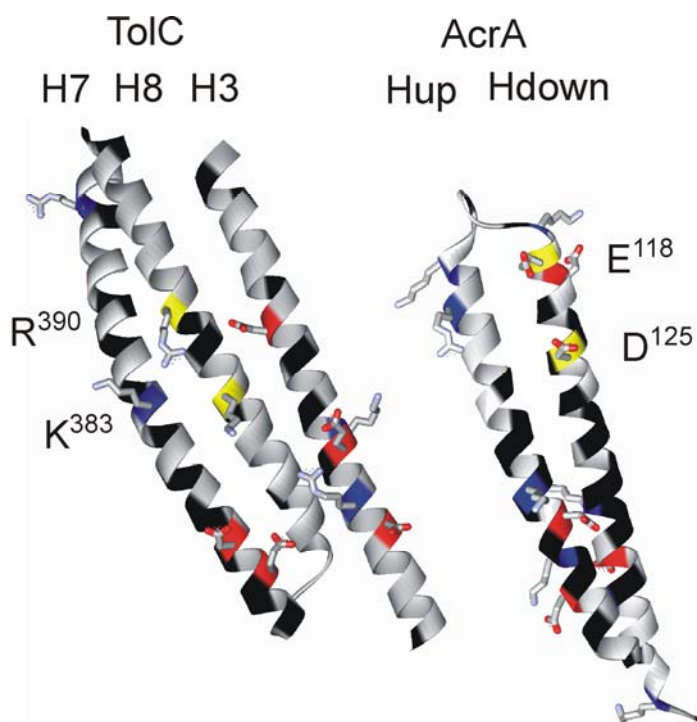
Figure 4.1: Charged residues on the surfaces of the TolC helices H7, H8 and H3 and the AcrA hairpin domain. The yellow marked, annotated sites represent residues mutated in this study. The numbering of the residues corresponds to the mature proteins.

4.4.1.1. Effect of GP1-mutation on MIC

To determine the effects of our mutations in TolC and AcrA we had to establish a TolC and AcrAB deficient strain, and complement the genes by an appropriate plasmid system to

restore the wild-type phenotype. The DC14TC strain lacking TolC, AcrA and AcrB showed hypersusceptibility in MIC-tests, which is in accordance to the findings of Sulavik and coworkers (Sulavik *et al.*, 2001). To reinsert the *tolC* gene the pAX629 plasmid was used which contains the TolC wild-type gene under the control of the native promotor so that its expression is controlled by cellular regulation. The genes *acrA* and *acrB* were cloned into the pBAD vector under the control of an arabinose inducible araBAD promotor, hence uncoupled from cellular regulation mechanisms. In MIC-tests against novobiocin the pAX629 pBADacrAB plasmid system was tested with variable arabinose concentrations to find conditions in which MIC test revealed similar resistance levels for DC14TC pAX629 pBADacrAB as obtained with wild-type strain AG100 pACYC184 pBAD. In this way an arabinose concentration of 0.01% was determined as adequate to complement the knock out phenotype.

Glutamate was substituted for both lysine 383 and arginine 390 in TolC by site-directed mutagenesis to investigate the role of these charges in AcrA/TolC interaction. In antibiotic susceptibility tests the resulting TolC K383E R390E mutant GP1 showed a considerable decrease of the MIC-value against novobiocin from 150 µg/ml for the wild-type to 60 µg/ml for the GP1 mutant, from 375 µg/ml to 31.3 µg/ml against rhodamine 6G, from over 200 µg/ml to 25 µg/ml against erythromycin and from 80,000 µg/ml to 1,250 µg/ml against SDS. This shows that the exchange of the charged residues has a massive influence on the functionality of the efflux pump. The mutant TolC, however, has to be functional because the MIC-values (except for SDS) of the GP1 mutation are clearly higher than of the TolC knock-out strain. The effect of the GP1 mutation on the SDS resistance seems to be dramatic. This observation will be discussed later in more detail.

4.4.1.2. Effect of GP1 on hemolysin secretion

In a next attempt, we tested the ability of GP1 to interact with the hemolysin secretion apparatus. Therefore the TolC deficient AG100TC strain was transformed with pBadTolC with TolC under the control of the arabinose inducible araBAD promotor and the pRSC2 plasmid which contains the complete hlyR,C,A,B,D cluster. As a negative control the AG100TC strain was transformed with the blank pBAD/myc-His C plasmid and pRSC2. Spotted and grown on blood agar plates supplemented with 0.001% arabinose, both the TolC wild-type containing strain and the GP1 containing strain showed clearly visible zones of clearance surrounding the cells indicating lysis of the erythrocytes by actively secreted hemolysin A (see Figure 4.2). Compared to the wild-type TolC the halo of GP1 was slightly

reduced, whereas the negative control strain without TolC showed just very weak hemolysis in the direct proximity of the cells most likely by release of hemolysin A by lysed cells.
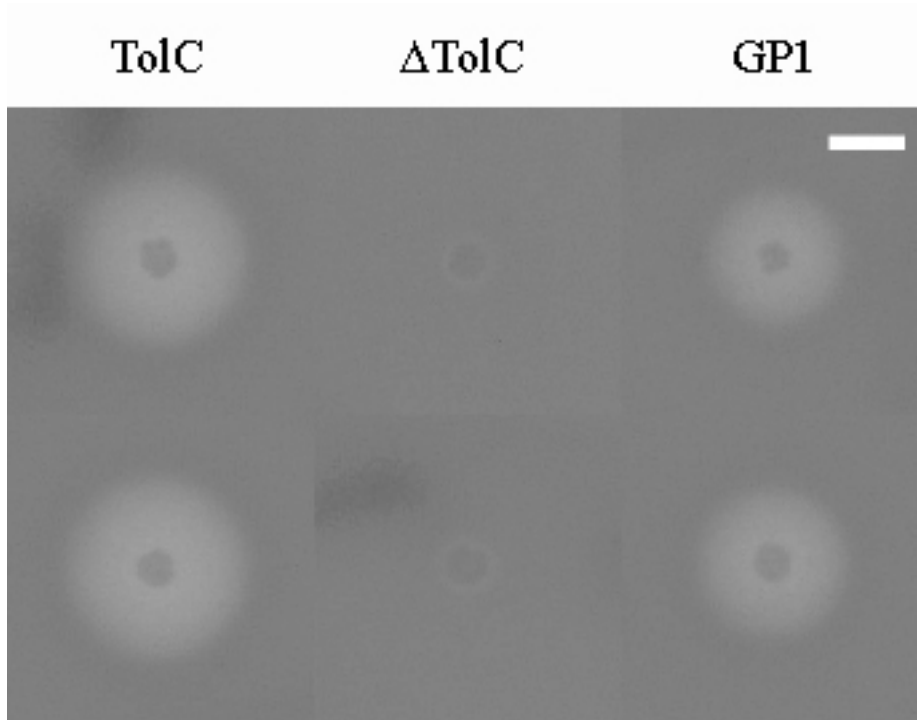
Figure 4.2: Test of hemolysin secretion on blood agar plates. All AG100TC strains contained the plasmid pRSC2 coding for the hemolysin operon. AG100TC cells harbouring the pBAD vector were used as negative control (ΔTolC). TolC and GP1 are expressed from the arabinose inducible plasmids pBADTolC and pBADGP1 respectively. Plates were inoculated with 2×10^5 cells and grown for 18 h at 37 °C. The white bar corresponds 1 cm.

4.4.1.3. Effect of the GP1 mutation on trimer stability in overexpression experiments

For further biophysical characterization, the mutated GP1 protein was overexpressed in *E. coli* using the arabinose inducible pBAD vector system. The expression itself was very successful, but attempts to purify GP1 for black lipid bilayer measurements revealed that in SDS-PAGE a high molecular weight band in unboiled samples corresponding to the trimeric TolC could not be detected for GP1 as for wild-type TolC. This also explains, why we never got activity for GP1 in bilayer experiments. To get to the bottom of this phenomenon we isolated the outer membrane after growing and induction of the cells by sucrose-step gradient centrifugation as described above. Analyzing these outer membrane fractions with or without preincubation at 100°C for 10 minutes prior to SDS-PAGE and Western-blotting only the wild-type TolC showed a shift of the band of the monomeric protein from 52 kDa in the preincubated probe to a 130 kDa trimer band in the unboiled probe. The absence of monomeric TolC in the unboiled sample means that all TolC molecules are assembled to

trimers. For GP1, however, the boiled as well as the unboiled probe just showed a band at 52 kDa corresponding to monomeric protein (see Figure 4.3).
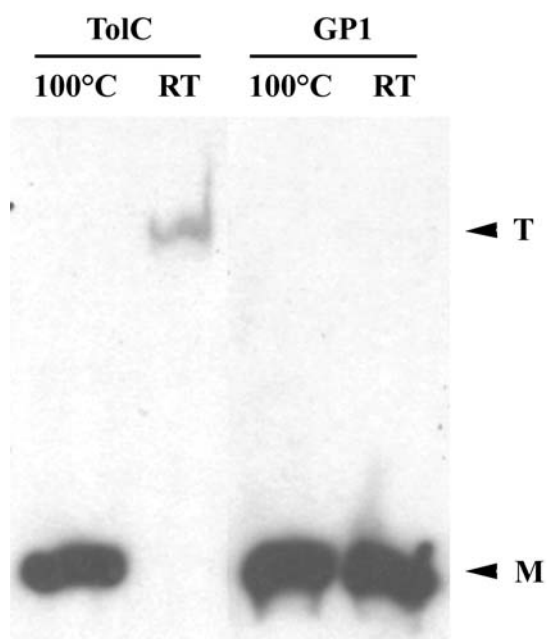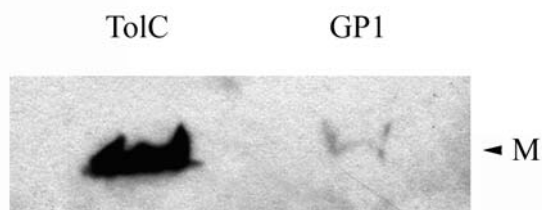

Figure 4.3: Western-Blot of outer membrane fractions of AG100TC cells with pBADTolC or pBADGP1 respectively. Probes were incubated at 100°C or room-temperature (RT) for 10 minutes prior to 10% SDS-PAGE. TolC and GP1 were detected with α-HIS antibodies (Amersham Bioscience). It should be mentioned here that the intensity of the trimer band was always reduced in comparison with the monomer band, most likely because of the more compact structure of the trimers. M indicates the molecular weight of 52 kDa corresponding to TolC monomers, T indicates the molecular weight of 130 kDa corresponding to TolC trimers.

The fact that in the case of GP1 no trimer band is detectable leads to the assumption that the exchange of the two residues lysine 383 and arginine 390 to glutamate respectively has serious impact on the stability of the protein.

The finding that GP1 is still functioning in the hemolysin secretion system, however, proves that there has to be functional trimeric protein in the outer membrane. Furthermore, the observation that GP1 appears in the outer membrane of the cells clearly indicates the presence of trimeric protein because TolC is inserted into the outer membrane only as trimer. We therefore conclude that the GP1 mutation located below the equatorial domain possibly does not influence trimerization and membrane insertion of the β-barrel.

4.4.1.4. GP1 shows lower protein level when controlled under native promotor

The experiment described above was performed with TolC or GP1 under the control of the araBAD promotor, which is not regulated very tightly, resulting in high expression rates even at low inductor concentrations. It is possible that overexpression leads to problems regarding the correct folding of GP1 and consequently influences the stability of the protein. Therefore the experiment was repeated with the *tolC* gene under its native promotor. DC14TC cells were transformed with the plasmids pAX629 and pAXGP1, respectively. After growing and harvesting of the cells, the outer membrane fraction was obtained by sucrose-step gradient centrifugation as described above. Again, the probes were analyzed by SDS-PAGE and Western-blotting showing a highly reduced amount of GP1 protein in comparison with wild-type TolC (see Figure 4.4).

Figure 4.4: Western-Blot of outer membrane fractions of DC14TC cells with pAX629 or pAXGP1, respectively. Probes were incubated at 100°C for 10 minutes prior to 10% SDS-PAGE. TolC and GP1 were detected with α-TolC antibodies (Pineda Antibody Services, Berlin). M indicates the molecular weight of 52 kDa corresponding to TolC monomers.

The same reduction of GP1 was observed when the DC14TC cells were transformed with pAX629 or pAXGP1 and additionally pBADacrAB, respectively.

All these results of the expression experiments indicate a considerable disturbance of the stability of the GP1 trimer independently of a possible interaction with the AcrAB complex. However, how can it be explained that the amount of GP1 is so much reduced in comparison to wild-type TolC, when expressed under native promotor in contrast to the overexpression experiment? An explanation might be that the instability of the GP1 trimers leads to an enhanced degradation of the proteins. At high expression rate, however, the degrading processes cannot keep up with the protein production. This could be the reason why there is so much protein in the outer membrane assembled to instable trimers.

At this point it seems reasonable to suspect that the observed phenotype could be explained by the lower level of GP1, but on the other hand it cannot be excluded that a disturbed

interaction is not at least partially responsible for the lower MIC-values of the GP1 mutation. A second question is, if AcrA is eventually acting as a chaperon helping to stabilize the TolC trimer by interaction and if this assistance is inhibited by the GP1 mutation. Therefore, we investigated if a mutation in the interacting protein AcrA could restore the GP1 phenotype.

4.4.2. *Mutation of residues in AcrA*

Following the model of interaction the negatively charged residues glutamate 118 and aspartate 125 in AcrA might function as counterparts to the positively charged arginine 390 and lysine 383 in TolC, respectively as described above. Thus, we mutated E118 to lysine and D125 to arginine with the intention to compensate the possible interaction deficiency of the GP1 mutant. The resulting AcrA E118K D125R mutant was named CH1.

4.4.2.1. The CH1 mutant does not stabilize GP1

After transformation of pAX629 or PAXGP1 and pBADacrAB or pBADacrAB-CH1 in any combination into DC14TC cells, the strains were grown and induced followed by separation of the outer membrane fraction by sucrose-step gradient centrifugation. Afterwards the outer membrane probes were examined in Western-blot analysis. Again the amount of GP1 is drastically decreased compared to wild-type TolC independent of the presence of AcrA as wild-type or CH1 mutant (see Figure 4.5). Also, in comparison with the combination GP1-AcrA, the amount of GP1 protein is not increased when GP1 is expressed together with CH1, meaning that the CH1 mutant does apparently not assist or stabilize the assembly of GP1.
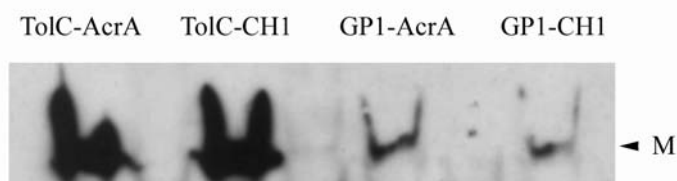
Figure 4.5: Western-Blot of outer membrane fractions of DC14TC cells with pAX629 or pAXGP1 and pBADacrAB or pBADacrAB-CH1, respectively. Samples were incubated at 100°C for 10 minutes prior to 10% SDS-PAGE. TolC and GP1 were detected with α-TolC antibodies (Pineda Antibody Services, Berlin). M indicates the molecular weight of 52 kDa corresponding to TolC monomers.

4.4.2.2. The CH1 mutation complements GP1 phenotype

In further experiments we investigated if the CH1 mutation in AcrA had any influence on the function of the AcrAB/TolC efflux pump. The CH1 mutant in combination with wild-type TolC has no effects in MIC tests, and this fact we interpret as indication that (I) the exchange of the two residues has no influence on the functionality of AcrA and (II) there have to be assumed alternative interaction sites in TolC allowing the contact even to mutated AcrA. In our opinion this seems conceivable because TolC is functioning as outer membrane component in several transport complexes and a flexible interaction site here would be reasonable (Koronakis *et al.*, 2004).

However, interestingly CH1 is able to complement the GP1 phenotype. The DC14TC strain transformed with pAXGP1 and pBADacrAB-CH1 showed wild-type MIC-values again for all tested agents (see Table 4.3).

Table 4.3: Minimal inhibitory concentrations (MIC) of five antimicrobial agents for *E. coli* DC14TC with the plasmids pAX629 or pAXGP1 and pBADacrAB or pBADacrAB-CH1 expressing efflux pump components in various combinations.

TolC	AcrA		MIC (µg/ml)				
			NOV	ERM	R6G	BENZ	SDS
–	–	*	2.5	7.8	4.7	1.3	1,250
TolC	AcrA		150	>200	375	20	80,000
TolC	CH1		150	>200	375	16.3	80,000
GP1	AcrA		60	25	31.3	16.3	1,250
GP1	CH1		150	>200	250	16.3	100,000

* determined by Stegmeier (subm.)
• Abbreviations: NOV, novobiocin; ERM, erythromycin; R6G, rhodamine 6G; BENZ, benzalkonium chloride; SDS, sodium dodecyl sulfate
in all strains AcrA was induced with 0.01% arabinose

The observation of a drastic effect of the GP1 mutant on SDS-resistance could be coherent with the instability of the protein. Eventually the detergent destabilizes the GP1 trimers even more resulting in a highly susceptible phenotype comparable to the knock-out. Then the restoring of the wild-type resistance level by complementing GP1 with CH1 might reflect a stabilization of GP1 trimers by CH1.

Together with the finding, however, that the amount of GP1 does not increase when expressed in combination with CH1, the higher resistance levels of GP1 in combination with CH1 in fact indicate an improved efficiency of the interaction of the two mutants. Obviously the smaller amount of GP1 trimers present in the outer membrane is already sufficient to ensure an effective efflux of the tested substances. Perhaps this amount of GP1 represents the number of assembled AcrAB/TolC efflux pumps in the outer membrane where GP1 is protected against degradation. It is likely that in the case of GP1 the disturbed interaction between TolC and the inner membrane transporter complex AcrAB and not the reduced number of channel-tunnels constitutes the limiting factor for the maintenance of the multidrug resistance.

4.4.3. *Investigation of interaction by site-directed disulfide cross-linking*

The general interaction between TolC, AcrA and AcrB could already be proven repeatedly as mentioned above by chemical cross-linking (Husain et al. 2004, Touze et al. 2004, Tikhonova and Zgurskaya 2004), but using this method it is not possible to make any statements about the domains or even the residues which are close enough to each other to be linked. The successful complementation of GP1 by CH1 was a first hint that our interaction model is correct also on molecular level. To further test the model in a more direct way, residues on the TolC and AcrA surfaces were identified which are in close proximity. These residues were changed to cysteines, which allows to prove the close contact of two sites by disulfide cross-linking. The cross-linked AcrA-TolC can then be detected by purification of one of the interaction partners by specific affinity chromatography and thereby copurification of the other linked protein. Because of its high specificity in affinity purification we used a Strep-tag system for TolC. The gene was cloned into the pASK-IBA12 vector as described above resulting in an anhydrotetracycline (AHT)-inducible plasmid expressing N-terminal Strep-tagged TolC. The tet-promotor is a very tight controlled promotor that can be tuned very fine. In MIC-tests against novobiocin we investigated the ability of pASK-TolC to compensate the TolC knock-out of AG100TC and determined a concentration of AHT of 25 ng/ml necessary to restore the wild-type resistance level of AG100 (data not shown).

The five residues serine 124, alanine 128, alanine 132, serine 339 and alanine 343 in TolC and the residues alanine 93 and alanine 94 in AcrA were selected as positions for single amino acid exchange mutations to cysteine, respectively (see Figure 4.6).
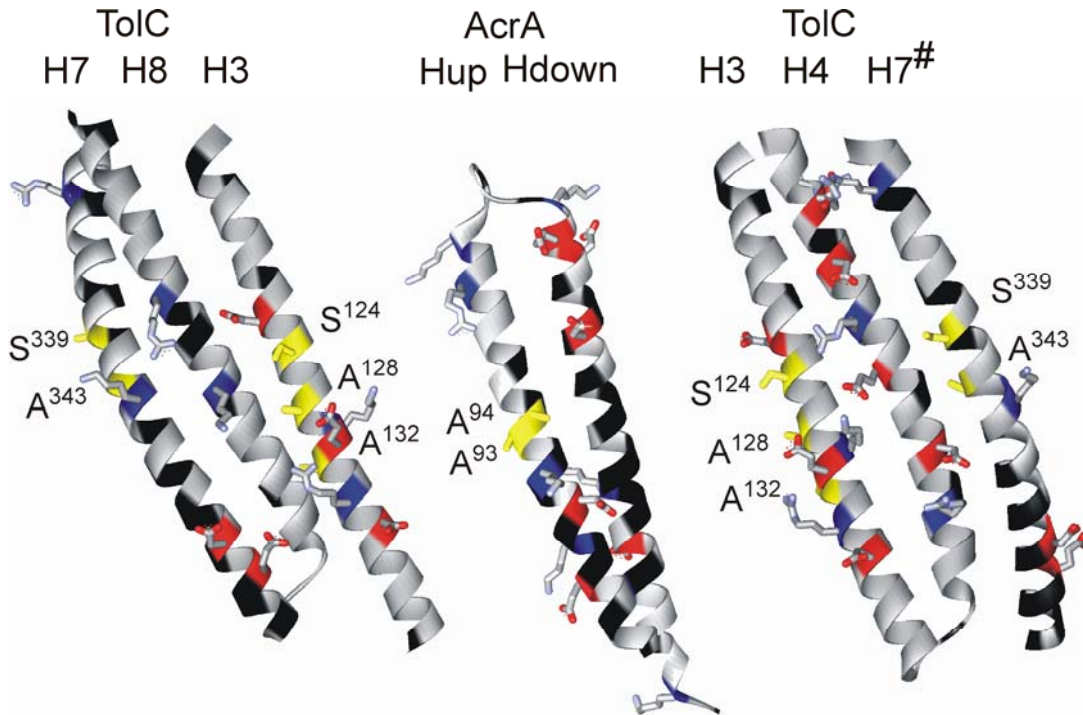

Figure 4.6: Structures of the interacting domains of TolC and AcrA. Sites of residues mutated to cysteine are named and marked in yellow. The numbering of the residues corresponds to the mature proteins.

For cross-link experiments DC14TC cells were transformed with pBADacrA$_{A93C}$B-cat or pBADacrA$_{A94C}$B-cat and pASK-TolC or one of the five cysteine-mutant pASK-TolC plasmids, respectively in any combination. The strains were grown and induced as described above. After ultracentrifugation of the disrupted cells, the pelleted membranes were resuspended in 8M urea and samples of every strain were analysed under non-reducing conditions in a 6% SDS-PAGE and following Western-blot (see Figure 4.7).

With the exception of TolC$_{S124C}$ + AcrA$_{A94C}$ all combinations of the TolC cysteine mutants form a complex band of 140 kDa, which is not present for wild-type TolC. Although this mass is apparently too high, we consider this band for TolC dimers formed by intermolecular disulfide bridging between the introduced cysteines. The running behaviour could probably be explained by the elongated form of the cross-linked product and a resulting low electrophoretic mobility. The same observations were made for TolC cysteine mutants by Tamura *et al.* (2005). In parallel the AcrA cysteine mutants show a complex band at around 115 kDa in every combination which we consider for AcrA dimers for the same reasons as

described for TolC. Just the combination $TolC_{S339C}$ + $AcrA_{A93C}$ shows additional bands both for TolC and AcrA at 130 kDa respectively and this intermediate mass would make sense for cross-linked TolC-AcrA.
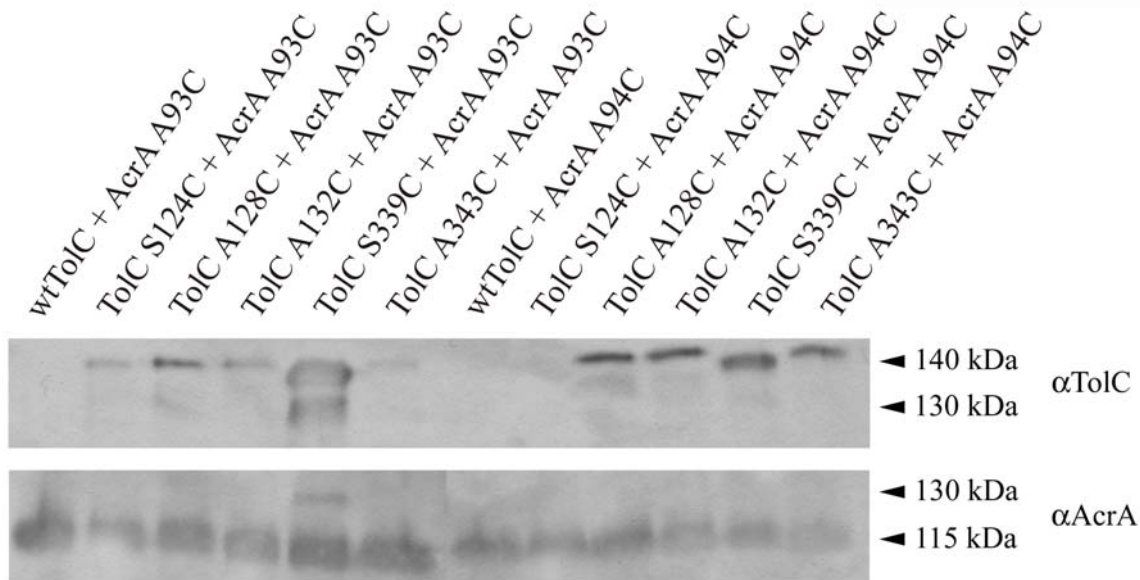
Figure 4.7: Western-Blot analysis of membrane fractions resuspended in 8M urea of DC14TC cells with pASK-TolC or one of the five cysteine mutant pASK-TolC plasmids and pBADacrA_{A93C}B-cat or pBADacrA_{A94C}B-cat, respectively. Samples were mixed with non-reducing loading-buffer and incubated at 100°C for 10 minutes prior to 6% SDS-PAGE. TolC and AcrA were detected with α-TolC and α-AcrA antibodies (Pineda Antibody Services, Berlin), respectively.

To prove this assumption the resuspended membrane probe of $TolC_{S339C}$ + $AcrA_{A93C}$ was purified using Strep-tactin sepharose (IBA). Purified TolC was examined for copurification of AcrA in Western-blot analysis. In fact AcrA is detectable in the elution of the $TolC_{S339C}$ + $AcrA_{A93C}$ probe while there is no copurified AcrA detectable for the equally treated control strains with wild-type TolC + $AcrA_{A93C}$ and $TolC_{S124C}$ + $AcrA_{A94C}$ (see Figure 4.8). Additionally, analysing the elution samples under non-reducing conditions the complex band at 130 kDa is again detectable for $TolC_{S339C}$ + $AcrA_{A93C}$ with both the αAcrA and the αTolC antibodies. This result strongly supports our assumptions about the AcrA/TolC interaction model on molecular level.
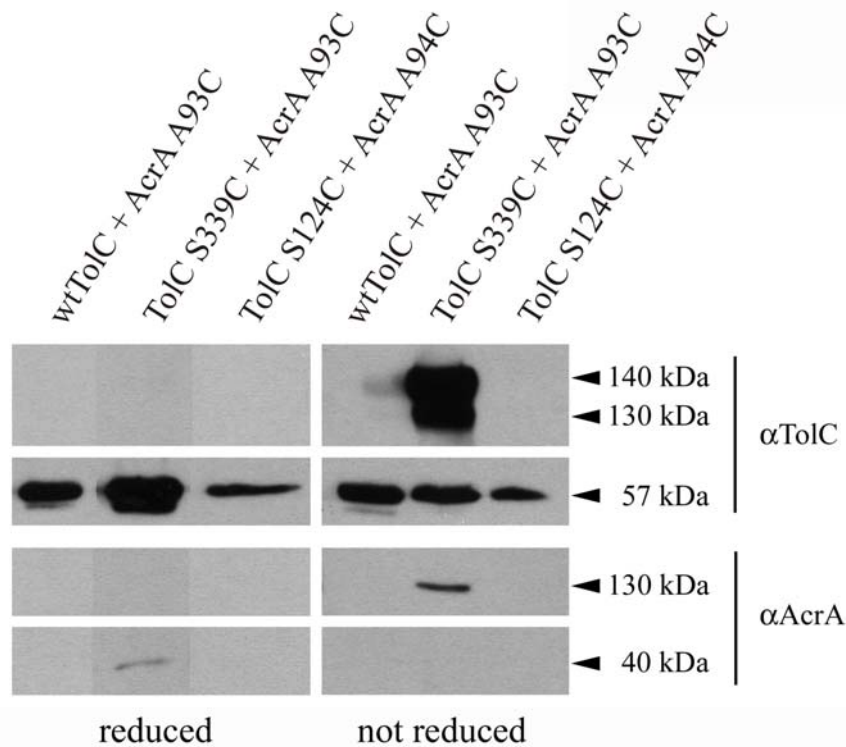
Figure 4.8: Western-Blot analysis of elution probes of strep-tag purified membrane extractions of DC14TC cells with pASK-TolC + pBADacrA$_{A93C}$B-cat, pASK-TolC$_{S339C}$ + pBADacrA$_{A93C}$B-cat or pASK-TolC$_{S124C}$ + pBADacrA$_{A94C}$B-cat. Samples were mixed with reducing or non-reducing loading-buffer and incubated at 100°C for 10 minutes prior to 7% SDS-PAGE. TolC and AcrA were detected with α-TolC and α-AcrA antibodies (Pineda Antibody Services, Berlin) respectively.

4.5. Conclusion

In the present study we demonstrate that the interaction model for TolC and AcrA posted here could be supported by *in vivo* measurements. We could show that the residues K383 and R390 in TolC are likely important contact sites to the hairpin domain of AcrA and more precisely to the residues aspartate 125 and glutamate 118, respectively. Our experiments, however, revealed that there probably does not exist one distinct and universal interaction site for each interacting protein. This is implicated by the result that the CH1 mutation has no effect on the functionality of the TolC/AcrAB multidrug efflux pump. A high flexibility of the interaction domain of TolC, however, seems to be very reasonable, because TolC has to be able to cooperate with several different inner membrane complexes (Koronakis *et al.*, 2004). Anyway, it is known that also the inner membrane protein AcrB interacts directly with TolC

independent of AcrA (Touze *et al.*, 2004; Tamura *et al.*, 2005). In addition, the results of isothermal titration calorimetry experiments performed by Touze and coworkers indicate the existence of several non-cooperative binding sites between AcrA and TolC. Furthermore, our GP1 mutation turned out that single residues do not just act as interaction sites but could be simultaneously important for the correct folding or assembly and also for the stability of the TolC trimer. In this context the results of mutations in the α/β domain of AcrA suppressing the effects of mutations in TolC (Gerken and Misra, 2004) can also be interpreted as an indirect influence, underlining how complex the assembly of the efflux pump is.

We could not yet prove our hypothesis of six AcrA monomers per one fully arranged TolC/AcrAB efflux pump, but the results presented here could be interpreted as a hint supporting this model. While the GP1-CH1 complementation implicates interaction of the AcrA hairpin domain with the modeled interaction site of TolC formed by helices H7/H8/H3, the disulfide cross-linking experiments point towards the possible H3/H4/H7 interaction site of TolC. Considering both interactions as necessary the assembly of the efflux pump would require two interacting AcrA molecules per TolC monomer. By site directed disulfide cross-linking, however, it was possible to determine the relative position of AcrA to TolC as postulated before.

Including the information of several other investigations, we propose the following model for the assembling of efflux pumps. The first step is the head-to-tail interaction of AcrB and TolC. As mentioned it was shown by cross-linking of AcrB and TolC that this interaction is independent of AcrA (Touze *et al.*, 2004). More precisely, we believe that the first contact between the transporter and the tunnel is limited to the interaction between the tip of the outer coiled coil of the tunnel and the tip of the N-terminal periplasmic loop of the transporter. As proven by Tamura and coworkers (2005) we propose that salt bridges formed between R143 of TolC and D256 of AcrB provide contact. They can be formed without conformational changes in both proteins. The distance between the arginine residues in TolC and the aspartate residues in AcrB is 38 Å and 41 Å, respectively.

The next step of the efflux pump assembling is the binding of the adaptor protein. This interaction stabilizes the drug efflux complex in a way that it can be purified as complex as shown for AcrAB/TolC and MexAB/OprM (Mokhonov et al., 2004; Tikhonova and Zgurskaya, 2004). It is known that AcrA and AcrB form a complex independent of TolC and that the C-terminus of AcrA is responsible for the binding to AcrB (Touze *et al.*, 2004; Zgurskaya and Nikaido, 2000a). Also, binding of AcrA to TolC is independent of the presence of the transporter (Husain *et al.*, 2004; Touze *et al.*, 2004), which means that no

conformational change in TolC or AcrA induced by interaction with the transporter is necessary. Based on our experimental data, we assume that two AcrA protomers bind to TolC, possibly in two sequential steps. The first AcrA binds to the H3/H4/H7 interaction site with its hairpin domain eventually without conformational changes in TolC. Binding of the second AcrA hairpin to the H7/H8/H4 binding site is associated with the outwards movement of the inner coiled coil formed by H7 and H8 resulting in the opening of the tunnel. For the purified TolC protein it could be shown that the open state is not stable and needs to be stabilized (Andersen *et al.*, 2002b). Therefore, beside interaction with the hairpin of the adaptor protein, a second binding site between the channel and the transporter could be established, which further stabilizes the open conformation of the tunnel. Site directed disulfide cross-linking of AcrB and TolC revealed that residues located at the turn of the TolC inner coiled coil formed by H7 and H8, interact with a periplasmic loop from the C-terminal half of AcrB (Tamura *et al.*, 2005). It is very likely that a salt bridge between R367 of TolC and D795 of AcrB is involved in this interaction. The distance between the arginine residues in TolC and the aspartate residues in AcrB in the resting state of both proteins is 10 Å and 21 Å, respectively, implying that the outwards movement of the inner coiled coil is prerequisite for this interaction. Interestingly, R367 of TolC is shown to be part of the circular network, which keeps the tunnel in the closed conformation (Andersen *et al.*, 2002b). Breaking the network renders Arg367 free for interaction with the transporter.

Applying this to the MexAB/OprM efflux pump, it is striking that the second interaction site between the channel and the transporter is different to the AcrAB/TolC system. The polarity of the residues forming the salt bridge is opposed. In OprM there is at the corresponding positions a negatively charged residue (D409), in MexB a positive charge (K794). This difference could also contribute to the fact that TolC and OprM could not be exchanged between the two systems without loosing function (Srikumar *et al.*, 1998; Tikhonova *et al.*, 2002).

Further experiments, however, will be necessary to elucidate the interaction of the pump components in more detail. But certainly this progress in establishing a reliable model will help to examine the cooperation of TolC and the AcrAB complex in a more targeted way in the future. The exact knowledge of function mechanisms of this and related multidrug resistance pumps will be very helpful to break this important defense of Gram-negative pathogens.

Chapter 5

Summary

5.1. Summary

Multidrug efflux pumps have serious impact on both intrinsic and acquired antibiotic resistance of clinically relevant Gram-negative pathogens. They are tripartite complexes consisting of an energized inner membrane transporter, a periplasmic adaptor protein and an outer membrane channel-tunnel protein of the TolC family, bridging the periplasmic space and transporting substances across the outer membrane. Analogous channel-tunnel-dependent export-systems are the type I protein secretion complexes, which are able to secrete proteins directly across the inner and outer membrane from the cytoplasm to the external medium without periplasmic intermediates. Within the last few years, the structures of all three components of the AcrAB/TolC multidrug efflux pump of *E. coli* could be solved, answering many questions, but the exact mechanism of the transport process is still not understood.

In the first part of this work, the very diverse adaptor protein family was ordered by bioinformatical sequence analysis and grouped in seven families and 32 subfamilies, identifiable by generated, specific sequence signatures. This new classification of the adaptor protein family will help to annotate genes of newly sequenced genomes correctly. The alignment of the consensus sequences of the different subfamilies showed that the seven families divide into two groups distinguishable by the C-terminus and the length of the hairpin domain. Interestingly, these characteristics can be clearly correlated with the topology of the corresponding inner membrane transporter. Adaptor proteins with an extra C-terminal domain and short hairpin domains form complexes with transporters that are characterized by periplasmic extensions, whereas adaptor proteins with a short C-terminus and long hairpin domains pair with transporters lacking periplasmic domains. Based on these findings,

structural models of prototypical export-systems of these two distinct groups were designed. One is a multidrug efflux pump with an RND transporter with periplasmic extension, whereas the second is a type I protein secretion apparatus with an ABC transporter without periplasmic extension. The models visualize the function of the hairpin domain. One function, common to both types of adaptor proteins, is to provide contact with the channel-tunnel. In the case of the efflux pumps, the hairpin length of the adaptor proteins of 50 to 60 Å is sufficient, because the periplasmic exit of the transporter and the tunnel entrance of the channel-tunnel protein are already in close proximity. A second function, exclusively provided by adaptor proteins interacting with transporters without periplasmic extension, is to form a conduit like assembly bridging the gap between the inner membrane transporter and the tunnel entrance. This model is supported by the finding, that all adaptor proteins of this group have a predicted hairpin length of at least 90 Å.

The second part of the work focuses on HI1462, the only functional TolC homologue of *H. influenzae*. It was expressed in *E. coli*, purified and reconstituted in planar lipid membranes. It exhibited a single-channel conductance of 43 pS in 1 M KCl and is the first known anion-selective channel-tunnel protein. A structural model of HI1462 was constructed, which explained its biophysical characteristics. An arginine residue was identified, which lines the tunnel entrance. The dominant role of this residue for anion-selectivity was proven by a mutant, where the arginine residue was changed to alanine. The selectivity of this mutant was reversed to cation-selectivity. Additionally, this mutant showed an increased single-channel conductance of 282 pS in 1 M KCl. This increase could be explained with the existence of a circular network of salt bridges formed by the arginine and a glutamate residue of the adjacent monomer. These salt bridges presumably keep the tunnel entrance in a tightly closed conformation. Therefore, the disruption of the network by the mutation led to an opening of the tunnel entrance, explaining its higher single-channel conductance. In *in vivo* experiments it could be demonstrated, that HI1462 is able to substitute TolC in the hemolysin secretion apparatus, whereas this was not possible in the AcrAB/TolC multidrug efflux pump. This implies, that interaction between the inner membrane complexes of drug efflux pumps and their corresponding channel-tunnels is more specific than the interaction of the inner and outer components of type I secretion systems.

In the third part of the work, the interaction of adaptor proteins with channel-tunnels in multidrug efflux pumps was studied on the AcrAB/TolC system of *E. coli*. This interaction is assumed to be important for the iris-like opening of the tunnel entrance enabling transport of substances. Based on the known structures of the components, the interaction between AcrA

and TolC was modeled and two charged residues on the tunnel-surface of TolC were identified as possible contact points for two countercharges on the hairpin domain of AcrA. Experiments with charge exchange mutants of TolC and AcrA demonstrated the importance of these residues for the correct function of the AcrAB/TolC efflux pump. Additionally, several residues in TolC and AcrA were changed to cysteine, and the mutants were investigated in disulfide cross-linking experiments. Cross-linking was observed for one certain combination of cysteine mutants of TolC and AcrA, concordantly with the interaction model. Taken together, the results of the charge exchange mutants and the disulfide cross-linking further support our proposed model of a fully arranged efflux pump.

5.2. Zusammenfassung

Efflux-Pumpen spielen sowohl für intrinsische als auch erworbene Resistenz von klinisch relevanten pathogenen Gram-negativen Bakterien eine wichtige Rolle. Sie setzen sich aus einem energetisierten Innenmembran-Transporter, einem periplasmatischen Adaptor-Protein und einem „channel-tunnel"-Protein aus der TolC-Familie in der äußeren Membran zusammen. Die aus diesen drei Komponenten bestehenden Komplexe durchspannen das Periplasma und transportieren Substanzen über die Außenmembran. Weitere channel-tunnel-abhängige Exportsysteme sind die Typ I Sekretions-Apparate, die Proteine ohne periplasmatische Zwischenstufen aus dem Cytoplasma direkt über Innen- und Außenmembran ins externe Medium transportieren. In den letzten Jahren konnten die Strukturen aller drei Komponenten der AcrAB/TolC Efflux-Pumpe von *E. coli* gelöst und viele Fragen beantwortet werden. Der genaue Mechanismus des Transportprozesses ist jedoch immer noch nicht verstanden.

Im ersten Teil der Arbeit wurde mittels bioinformatischer Sequenzanalyse die mannigfaltige Familie der Adaptor-Proteine geordnet und in sieben Familien und 32 Unterfamilien gruppiert, die über spezifische, selbst generierte Sequenzmotive identifiziert werden können. Diese neue Klassifizierung der Adaptor-Protein-Familie hilft, Gene in neu sequenzierten Genomen zukünftig richtig zuzuordnen. Der Vergleich der Konsensus-Sequenzen der verschiedenen Unterfamilien ergab, dass sich die sieben Familien in zwei Gruppen aufspalten, die sich im C-Terminus und der Länge der sogenannten „hairpin"-Domäne unterscheiden. Interessanterweise lassen sich diese Eigenschaften eindeutig mit der Topologie der zugehörigen Innenmembran-Transporter korrelieren. Adaptor-Proteine mit einer zusätzlichen C-terminalen Domäne und kurzen „hairpin"-Domänen bilden Komplexe mit Transportern, die periplasmatische Domänen aufweisen. Adaptor-Proteine mit kurzem C-Terminus und langen „hairpin"-Domänen gehören hingegen zu Transportern ohne periplasmatische Domänen. Auf der Grundlage dieses Befundes wurden Strukturmodelle von prototypischen Exportsystemen dieser zwei Gruppen erstellt. Das erste Modell ist eine Efflux-Pumpe mit einem RND-Transporter mit periplasmatischer Domäne, während das zweite einen Typ I Sekretions-Apparat mit einem ABC-Transporter ohne periplasmatische Erweiterung darstellt. In diesen Modellen ist die Funktion der „hairpin"-Domäne der Adaptor-Proteine erkennbar. Bei beiden prototypischen Exportsystemen ist die Funktion der „hairpin"-Domäne, den Kontakt mit dem „channel-tunnel" herzustellen. Im Falle der Efflux-Pumpen ist die „hairpin"-Länge der

Adaptor-Proteine mit 50 bis 60 Å ausreichend, da sich der periplasmatische Ausgang des Transporters und der Tunneleingang bereits in unmittelbarer Nähe zueinander befinden. Eine zweite Funktion erfüllen nur Adaptor-Proteine, die mit Transportern ohne periplasmatische Domäne interagieren. Sie bilden eine röhrenartige Struktur, die die Lücke zwischen dem Innenmembran-Transporter und dem Tunneleingang überbrückt. Dieses Modell wird von der Tatsache gestützt, dass alle Adaptor-Proteine dieser Gruppe eine vorhergesagte „hairpin"-Länge von mindestens 90 Å aufweisen.

Der zweite Teil der Arbeit beschäftigt sich mit HI1462, dem einzigen funktionellen TolC-homologen Protein von *H. influenzae*. Dieses wurde in *E. coli* exprimiert, aufgereinigt und in künstlichen Lipidmembranen rekonstituiert. HI1462 weist eine Einzelkanal-Leitfähigkeit von 43 pS in 1 M KCl auf und ist das erste beschriebene anionenselektive „channel-tunnel"-Protein. Ein Strukturmodell von HI1462 wurde erstellt, das die biophysikalischen Eigenschaften erklärt. Ein Arginin-Rest am Tunneleingang konnte als Ursache für die Anionenselektivität ausgemacht und dessen bestimmende Rolle durch eine Mutante nachgewiesen werden. In dieser Mutante wurde der Arginin-Rest gegen Alanin ausgetauscht und dadurch die Anionen- zu einer Kationenselektivität umgekehrt. Darüber hinaus zeigte diese Mutante eine erhöhte Einzelkanal-Leitfähigkeit von 282 pS in 1 M KCl. Dieser Anstieg konnte durch das Vorhandensein eines zirkulären Netzwerks von Salzbrücken zwischen dem Arginin- und einem Glutamat-Rest des benachbarten Monomers erklärt werden. Es ist zu vermuten, dass diese Salzbrücken den Tunneleingang in einer eng geschlossenen Konformation halten. Die Unterbindung dieses Netzwerks durch die Mutation führte daher zu einer Öffnung des Tunneleingangs, was die höhere Einzelkanal-Leitfähigkeit erklärt. Durch Experimente *in vivo* konnte gezeigt werden, dass HI1462 in der Lage ist, TolC im Hämolysin-Sekretions-Apparat von *E. coli* zu ersetzen, während eine funktionelle Substitution von TolC in der AcrAB/TolC Efflux-Pumpe nicht möglich war. Das deutet darauf hin, dass die Interaktion zwischen Innenmembran-Komplexen von Efflux-Pumpen und den zugehörigen „channel-tunnel"-Proteinen spezifischer ist als die Interaktion zwischen den Innen- und Außenmembran-Komponenten der Typ I Sekretions-Systeme.

Im dritten Teil der Arbeit wurde die Interaktion von Adaptor-Proteinen mit „channel-tunnel"-Proteinen in Efflux-Pumpen am Beispiel des AcrAB/TolC-Systems von *E. coli* untersucht. Es wird angenommen, dass diese Interaktion für das Iris-artige Öffnen des Tunnel-Eingangs wichtig ist, um den Transport von Substanzen zu ermöglichen. Basierend auf den bekannten Strukturen der Pumpenkomponenten wurde ein Modell der Interaktion zwischen AcrA und TolC erstellt und zwei geladene Reste auf der Tunneloberfläche von TolC als mögliche

Kontaktstellen für zwei Gegenladungen auf der „hairpin"-Domäne von AcrA identifiziert. Versuche mit Ladungsaustausch-Mutanten von TolC und AcrA zeigten die Bedeutung dieser Reste für die korrekte Funktion der AcrAB/TolC Efflux Pumpe. Darüber hinaus wurden mehrere Reste in TolC und AcrA gegen Cystein ausgetauscht und die Mutanten in Disulfid-cross-linking-Versuchen untersucht. In Übereinstimmung mit dem Interaktionsmodell konnte bei einer bestimmten Kombination von Cystein-Mutanten ein solches cross-linking beobachtet werden. Zusammengenommen unterstützen die Ergebnisse der Ladungsaustausch-Mutanten und des Disulfid-cross-linkings unser Modell einer komplett zusammengebauten Efflux-Pumpe.

Chapter 6

Appendix

6.1. References

1. Aendekerk,S., B.Ghysels, P.Cornelis, and C.Baysse. 2002. Characterization of a new efflux pump, MexGHI-OpmD, from *Pseudomonas aeruginosa* that confers resistance to vanadium. *Microbiology.* **148:**2371-2381.

2. Akama,H., T.Matsuura, S.Kashiwagi, H.Yoneyama, S.Narita, T.Tsukihara, A.Nakagawa, and T.Nakae. 2004a. Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in *Pseudomonas aeruginosa*. *J Biol Chem.* **279:**25939-25942.

3. Akama,H., M.Kanemaki, M.Yoshimura, T.Tsukihara, T.Kashiwagi, H.Yoneyama, S.Narita, A.Nakagawa, and T.Nakae. 2004b. Crystal structure of the drug discharge outer membrane protein, OprM, of *Pseudomonas aeruginosa*: dual modes of membrane anchoring and occluded cavity end. *J Biol Chem.* **279:**52816-52819.

4. Akatsuka,H., R.Binet, E.Kawai, C.Wandersman, and K.Omori. 1997. Lipase secretion by bacterial hybrid ATP-binding cassette exporters: molecular recognition of the LipBCD, PrtDEF, and HasDEF exporters. *J Bacteriol.* **179:**4754-4760.

5. Andersen,C., C.Hughes, and V.Koronakis. 2000. Chunnel vision. Export and efflux through bacterial channel-tunnels. *EMBO Rep.* **1:**313-318.

6. Andersen,C., C.Hughes, and V.Koronakis. 2001. Protein export and drug efflux through bacterial channel- tunnels. *Curr Opin Cell Biol.* **13:**412-416.

7. Andersen,C., C.Hughes, and V.Koronakis. 2002a. Electrophysiological behavior of the TolC channel-tunnel in planar lipid bilayers. *J Membr Biol.* **185:**83-92.

8. Andersen,C., E.Koronakis, E.Bokma, J.Eswaran, D.Humphreys, C.Hughes, and V.Koronakis. 2002b. Transition to the open state of the TolC periplasmic tunnel entrance. *Proc Natl Acad Sci U S A.* **99:**11103-11108.

9. Andersen,C., E.Koronakis, C.Hughes, and V.Koronakis. 2002c. An aspartate ring at the TolC tunnel entrance determines ion selectivity and presents a target for blocking by large cations. *Mol Microbiol.* **44:**1131-1139.

10. Andersen,C. 2003. Channel-tunnels: outer membrane components of type I secretion systems and multidrug efflux pumps of Gram-negative bacteria. *Rev Physiol Biochem Pharmacol.* **147:**122-165.

11. Aono,R., N.Tsukagoshi and M.Yamamoto. 1998. Involvement of outer membrane protein TolC, a possible member of the mar-sox regulon, in maintenance and improvement of organic solvent tolerance of *Escherichia coli* K-12. *J Bacteriol.* **180:**938-44.

12. Apweiler,R., T.K.Attwood, A.Bairoch, A.Bateman, E.Birney, M.Biswas, P.Bucher, L.Cerutti, F.Corpet, M.D.Croning, R.Durbin, L.Falquet, W.Fleischmann, J.Gouzy, H.Hermjakob, N.Hulo, I.Jonassen, D.Kahn, A.Kanapin, Y.Karavidopoulou, R.Lopez, B.Marx, N.J.Mulder, T.M.Oinn, M.Pagni, F.Servant, C.J.Sigrist and E.M.Zdobnov. 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* **29:**37-40.

13. Apweiler,R., A.Bairoch, C.H.Wu, W.C.Barker, B.Boeckmann, S.Ferro, E.Gasteiger, H.Huang, R.Lopez, M.Magrane, M.J.Martin, D.A.Natale, C.O'Donovan, N.Redaschi and L.S.Yeh. 2004. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **32:**D115-D119.

14. Attwood,T.K., P.Bradley, D.R.Flower, A.Gaulton, N.Maudling, A.L.Mitchell, G.Moulton, A.Nordle, K.Paine, P.Taylor, A.Uddin and C.Zygouri. 2003. PRINTS and its automatic supplement, prePRINTS. *Nucleic Acids Res.* **31:**400-402.

15. Awram,P. and J.Smit. 1998. The *Caulobacter crescentus* paracrystalline S-layer protein is secreted by an ABC transporter (type I) secretion apparatus. *J Bacteriol.* **180:**3062-9.

16. Azpiroz,M.F., E.Rodriguez and M.Lavina. 2001. The structure, function, and origin of the microcin H47 ATP-binding cassette exporter indicate its relatedness to that of colicin V. *Antimicrob Agents Chemother.* **45:**969-72.

17. Bahloul,A., J.Meury, R.Kern, J.Garwood, S.Guha and M.Kohiyama. 1996. Co-ordination between membrane oriC sequestration factors and a chromosome partitioning protein, TolC (MukA). *Mol Microbiol.* **22:**275-82.

18. Balakrishnan,L., C.Hughes, and V.Koronakis. 2001. Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *J Mol Biol.* **313:**501-510.

19. Baranova,N. and H.Nikaido. 2002. The baeSR two-component regulatory system activates transcription of the yegMNOB (mdtABCD) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. *J Bacteriol.* **184:**4168-4176.

20. Bateman, A., L.Coin, R.Durbin, R.D.Finn, V.Hollich, S.Griffiths-Jones, A.Khanna, M.Marshall, S.Moxon, E.L.Sonnhammer, D.J.Studholme, C.Yeats and S.R.Eddy. 2004. The Pfam protein families database. *Nucleic Acids Res.* **32:**D138-D141.

21. Benz,R., K.Janko, W.Boos, and P.Lauger. 1978. Formation of large, ion-permeable membrane channels by the matrix protein (porin) of *Escherichia coli*. *Biochim Biophys Acta.* **511:**305-319.

22. Benz,R., A.Schmid, and R.E.Hancock. 1985. Ion selectivity of gram-negative bacterial porins. *J Bacteriol.* **162:**722-727.

23. Benz,R., E.Maier and I.Gentschev. 1993. TolC of *Escherichia coli* functions as an outer membrane channel. *Zentralbl Bakteriol.* **278:**187-96.

24. Beveridge,T.J. 2001. Use of the gram stain in microbiology. *Biotech Histochem.* **76:**111-8.

25. Binet,R. and C.Wandersman. 1996. Cloning of the *Serratia marcescens* hasF gene encoding the Has ABC exporter outer membrane component: a TolC analogue. *Mol Microbiol.* **22:**265-273.

26. Blattner,F.R., G.Plunkett, C.A.Bloch, N.T.Perna, V.Burland, M.Riley, J.ColladoVides, J.D.Glasner, C.K.Rode, G.F.Mayhew, J.Gregor, N.W.Davis, H.A.Kirkpatrick, M.A.Goeden, D.J.Rose, B.Mau, and Y.Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science.* **277:**1453-1474.

27. Blum,H., H.Beier, and H.J.Gross. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis.* **8:**93-99.

28. Borges-Walmsley,M.I., K.S.McKeegan, and A.R.Walmsley. 2003. Structure and function of efflux pumps that confer resistance to drugs. *Biochem J.* **376:**313-338.

29. Brown,M.H., I.T.Paulsen and R.A.Skurray. 1999. The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol Microbiol.* **31:**394-5.

30. Bugg,T.D., G.D.Wright, S.Dutka-Malen, M.Arthur, P.Courvalin and C.T.Walsh. 1991 Molecular basis for vancomycin resistance in *Enterococcus faecium* BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. *Biochemistry.* **30:**10408-15.

31. Bush,K., G.A.Jacoby and A.A.Medeiros. 1995. A functional classification scheme for beta-lactamases and its correlation with molecular structure. *Antimicrob Agents Chemother.* **39:**1211-33.

32. Calladine,C.R., A.Sharff, and B.Luisi. 2001. How to untwist an alpha-helix: Structural principles of an alpha-helical barrel. *J Mol Biol.* **305:**603-618.

33. Cao,Z. and P.E.Klebba. 2002. Mechanisms of colicin binding and transport through outer membrane porins. *Biochimie.* **84:**399-412.

34. Chang,A.C. and S.N.Cohen. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol.* **134:**1141-56.

35. Chang,G. and C.B.Roth. 2001. Structure of MsbA from *E. coli*: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters. *Science.* **293:**1793-800.

36. Chen,C.M., T.K.Misra, S.Silver, and B.P.Rosen. 1986. Nucleotide sequence of the structural genes for an anion pump. The plasmid-encoded arsenical resistance operon. *J Biol Chem.* **261:**15030-15038.

37. Cheng,L.W. and O.Schneewind. 2000. Type III machines of Gram-negative bacteria: delivering the goods. *Trends Microbiol.* **8:**214-20.

38. Chopra,I. 1998. Protein synthesis as a target for antibacterial drugs: current status and future opportunities. *Expert Opin Investig Drugs.* **7:**1237-44.

39. Chuanchuen,R., C.T.Narasaki, and H.P.Schweizer. 2002. The MexJK efflux pump of *Pseudomonas aeruginosa* requires OprM for antibiotic efflux but not for efflux of triclosan. *J Bacteriol.* **184:**5036-5044.

40. Cohen,M.L. 1992. Epidemiology of drug resistance: implications for a post-antimicrobial era. *Science.* **257:**1050-5.

41. Colmer,J.A., J.A.Fralick, and A.N.Hamood. 1998. Isolation and characterization of a putative multidrug resistance pump from *Vibrio cholerae. Mol Microbiol.* **27:**63-72.

42. Cornelis,G.R. 2002. The *Yersinia* Ysc-Yop 'type III' weaponry. *Nat Rev Mol Cell Biol.* **3:**742-52.

43. Corpet,F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16:**10881-10890.

44. Corpet,F., F.Servant, J.Gouzy and D.Kahn. 2000. ProDom and ProDom-CG: tools for protein domain analysis and whole genome comparisons. *Nucleic Acids Res.* **28:**267-269.

45. Crick,F.H.C. 1953. The packing of α-helices – simple coiled coils. *Acta Crystallogr.* 6:689-697.

46. Culliton,B.J. 1992. Drug-resistant TB may bring epidemic. *Nature.* **356:**473.

47. da Silva,F.G., Y.Shen, C.Dardick, S.Burdman, R.C.Yadav, A.L.de Leon, and P.C.Ronald. 2004. Bacterial genes involved in type I secretion and sulfation are required to elicit the rice Xa21-mediated innate immune response. *Mol Plant Microbe Interact.* **17:**593-601.

48. Datsenko,K.A. and B.L.Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* **97:**6640-6645.

49. Davies,J. 1994. Inactivation of antibiotics and the dissemination of resistance genes. *Science.* **264:**375-82.

50. Davies,J. 1996. Bacteria on the rampage. *Nature.* **383:**219-20.

51. Dean,C.R., S.Narayan, D.M.Daigle, J.L.Dzink-Fox, X.Puyang, K.R.Bracken, K.E.Dean, B.Weidmann, Z.Yuan, R.Jain and N.S.Ryder. 2005. Role of the AcrAB-TolC efflux pump in determining susceptibility of *Haemophilus influenzae* to the novel peptide deformylase inhibitor LBM415. *Antimicrob Agents Chemother.* **49:**3129-35.

52. Delepelaire,P. 2004. Type I secretion in gram-negative bacteria. *Biochim Biophys Acta.* **1694:**149-61.

53. Dinh,T., I.T.Paulsen and M.H.Saier,Jr. 1994. A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of gram-negative bacteria. *J Bacteriol.* **176:**3825-31.

54. Dorman,C.J., A.S.Lynch, N.N.Bhriain and C.F.Higgins. 1989. DNA supercoiling in *Escherichia coli*: topA mutations can be suppressed by DNA amplifications involving the tolC locus. *Mol Microbiol.* **3:**531-40.

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

56. Duong,F., C.Soscia, A.Lazdunski and M.Murgier. 1994. The *Pseudomonas fluorescens* lipase has a C-terminal secretion signal and is secreted by a three-component bacterial ABC-exporter system. *Mol Microbiol.* **11:**1117-26.

57. Eswaran,J., C.Hughes, and V.Koronakis. 2003. Locking TolC entrance helices to prevent protein translocation by the bacterial type I export apparatus. *J Mol Biol.* **327:**309-315.

58. Eswaran,J., E.Koronakis, M.K.Higgins, C.Hughes, and V.Koronakis. 2004. Three's company: component structures bring a closer view of tripartite drug efflux pumps. *Curr Opin Struct Biol.* **14:**741-747.

59. Etz,H., D.B.Minh, C.Schellack, E.Nagy and A.Meinke. 2001. Bacterial phage receptors, versatile tools for display of polypeptides on the cell surface. *J Bacteriol.* **183:**6924-35.

60. Falquet, L., M.Pagni, P.Bucher, N.Hulo, C.J.Sigrist, K.Hofmann and A.Bairoch. 2002. The PROSITE database, its status in 2002. *Nucleic Acids Res.* **30:**235-238.

61. Federici,L., D.Du, F.Walas, H.Matsumura, J.Fernandez-Recio, K.S.McKeegan, M.I.Borges-Walmsley, B.F.Luisi, and A.R.Walmsley. 2005. The crystal structure of the outer membrane protein VceC from the bacterial pathogen *Vibrio cholerae* at 1.8 A resolution. *J Biol Chem.* **280:**15307-15314.

62. Fekkes,P. and A.J.Driessen. 1999. Protein targeting to the bacterial cytoplasmic membrane. *Microbiol Mol Biol Rev.* **63:**161-73.

63. Fernandez-Recio,J., F.Walas, L.Federici, P.J.Venkatesh, V.N.Bavro, R.N.Miguel, K.Mizuguchi, and B.Luisi. 2004. A model of a transmembrane drug-efflux pump from Gram-negative bacteria. *FEBS Lett.* **578:**5-9.

64. Fiedler,G., M.Arnold, S.Hannus, and I.Maldener. 1998. The DevBCA exporter is essential for envelope formation in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol.* **27:**1193-1202.

65. Finnie,C., N.M.Hartley, K.C.Findlay and J.A.Downie. 1997. The *Rhizobium leguminosarum* prsDE genes are required for secretion of several proteins, some of which influence nodulation, symbiotic nitrogen fixation and exopolysaccharide modification. *Mol Microbiol.* **25:**135-46.

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

67. Franke,S., G.Grass, C.Rensing, and D.H.Nies. 2003. Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *J Bacteriol.* **185:**3804-3812.

68. Funkhouser,A., M.C.Steinhoff, and J.Ward. 1991. *Haemophilus influenzae* disease and immunization in developing countries. *Rev Infect Dis.* **13** Suppl 6:S542-S554.

69. Garrido,M.C., M.Herrero, R.Kolter and F.Moreno. 1988. The export of the DNA replication inhibitor Microcin B17 provides immunity for the host cell. *EMBO J.* **7:**1853-62.

70. Gerken,H. and R.Misra. 2004. Genetic evidence for functional interactions between TolC and AcrA proteins of a major antibiotic efflux pump of *Escherichia coli*. *Mol Microbiol.* **54:**620-31.

71. German,G.J. and R.Misra. 2001. The TolC protein of *Escherichia coli* serves as a cell-surface receptor for the newly characterized TLS bacteriophage. *J Mol Biol.* **308:**579-85.

72. Glaser,P., H.Sakamoto, J.Bellalou, A.Ullmann and A.Danchin. 1988. Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of *Bordetella pertussis. EMBO J.* **7:**3997-4004.

73. Gonzalez-Pasayo,R. and E.Martinez-Romero. 2000. Multiresistance genes of *Rhizobium etli* CFN42. *Mol Plant Microbe Interact.* **13:**572-577.

74. Griffith,J.K., M.E.Baker, D.A.Rouch, M.G.Page, R.A.Skurray, I.T.Paulsen, K.F.Chater, S.A.Baldwin and P.J.Henderson. 1992. Membrane transport proteins: implications of sequence comparisons. *Curr Opin Cell Biol.* **4:**684-95.

75. Grkovic,S., M.H.Brown, and R.A.Skurray. 2002. Regulation of bacterial drug export systems. *Microbiol Mol Biol Rev.* **66:**671-701.

76. Gross,R. 1995. Domain structure of the outer membrane transporter protein CyaE of *Bordetella pertussis. Mol Microbiol.* **17:**1219-20.

77. Haft,D.H., B.J.Loftus, D.L.Richardson, F.Yang, J.A.Eisen, I.T.Paulsen and O.White. 2001. TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* **29:**1-43.

78. Hassan,M.E.T., D.van der Lelie, D.Springael, U.Romling, N.Ahmed, and M.Mergeay. 1999. Identification of a gene cluster, czr, involved in cadmium and zinc resistance in *Pseudomonas aeruginosa. Gene.* **238:**417-425.

79. Havarstein,L.S., H.Holo and I.F.Nes. 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:**2383-9.

80. Hayes,J.D. and C.R.Wolf. 1990. Molecular mechanisms of drug resistance. *Biochem J.* **272:**281-95.

81. Heidelberg,J.F., J.A.Eisen, W.C.Nelson, R.A.Clayton, M.L.Gwinn, R.J.Dodson, D.H.Haft, E.K.Hickey, J.D.Peterson, L.Umayam, S.R.Gill, K.E.Nelson, T.D.Read, H.Tettelin, D.Richardson, M.D.Ermolaeva, J.Vamathevan, S.Bass, H.Qin, I.Dragoi, P.Sellers, L.McDonald, T.Utterback, R.D.Fleishmann, W.C.Nierman, O.White, S.L.Salzberg, H.O.Smith, R.R.Colwell, J:J:Mekalanos, J.C.Venter, and C.M.Fraser. 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae. Nature.* **406:**477-483.

82. Higgins,C.F. 1992. ABC transporters: from microorganisms to man. *Annu Rev Cell Biol.* **8:**67-113.

83. Higgins,M.K., E.Bokma, E.Koronakis, C.Hughes, and V.Koronakis. 2004. Structure of the periplasmic component of a bacterial drug efflux pump. *Proc Natl Acad Sci U S A.* **101:**9994-9999.

84. Hinsa,S.M., M.Espinosa-Urgel, J.L.Ramos, and G.A.O'Toole. 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:**905-918.

85. Hiraga,S., H.Niki, T.Ogura, C.Ichinose, H.Mori, B.Ezaki and A.Jaffe. 1989. Chromosome partitioning in *Escherichia coli*: novel mutants producing anucleate cells. *J Bacteriol.* **171:**1496-505.

86. Hirakata,Y., R.Srikumar, K.Poole, N.Gotoh, T.Suematsu, S.Kohno, S.Kamihira, R.E.Hancock and D.P.Speert. 2002. Multidrug efflux systems play an important role in the invasiveness of *Pseudomonas aeruginosa*. *J Exp Med.* **196:**109-18.

87. Holland,I.B. and M.A.Blight. 1999. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J Mol Biol.* **293:**381-399.

88. Hueck,C.J. 1998. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev.* **62:**379-433.

89. Hui,F.M., L.Zhou, and D.A.Morrison. 1995. Competence for genetic transformation in *Streptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcin A secretion genes. *Gene.* **153:**25-31.

90. Husain,F., M.Humbard and R.Misra. 2004. Interaction between the TolC and AcrA proteins of a multidrug efflux system of *Escherichia coli*. *J Bacteriol.* **186:**8533-6.

91. Hwang,J.W., X.T.Zhong, and P.C.Tai. 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. *J Bacteriol.* **179:**6264-6270.

92. Jacobi,S. and K.Heuner. 2003. Description of a putative type I secretion system in *Legionella pneumophila*. *Int J Med Microbiol.* **293:**349-358.

93. Jellen-Ritter,A.S. and W.V.Kern. 2001. Enhanced expression of the multidrug efflux pumps AcrAB and AcrEF associated with insertion element transposition in *Escherichia coli* mutants selected with a fluoroquinolone. *Antimicob Agents Chemother.* **45:**1467-1472.

94. Johnson,J.M. and G.M.Church. 1999. Alignment and structure prediction of divergent protein families: Periplasmic and outer membrane proteins of bacterial efflux pumps. *J Mol Biol.* **287:**695-715.

95. Johnson,S., J.E.Deane and S.M.Lea. 2005. The type III needle and the damage done. *Curr Opin Struct Biol.* **15:**700-7.

96. Kalinowski,J., B.Bathe, D.Bartels, N.Bischoff, M.Bott, A.Burkovski, N.Dusch, L.Eggeling, B.J.Eikmanns, L.Gaigalat, A.Goesmann, M.Hartmann, K.Huthmacher, R.Kramer, B.Linke, A.C.McHardy, F.Meyer, B.Mockel, W.Pfefferle, A.Puhler, D.A.Rey, C.Ruckert, O.Rupp, H.Sahm, V.F.Wendisch, I.Wiegrabe, and A.Tauch. 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J Biotechnol.* **104:**5-25.

97. Kaneko,T., Y.Nakamura, C.P.Wolk, T.Kuritz, S.Sasamoto, A.Watanabe, M.Iriguchi, A.Ishikawa, K.Kawashima, T.Kimura, Y.Kishida, M.Kohara, M.Matsumoto, A.Matsuno, A.Muraki, N.Nakazaki, S.Shimpo, M.Sugimoto, M.Takazawa, M.Yamada, M.Yasuda, and S.Tabata. 2001. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res.* **8:**205-213.

98.   Kazama,H., H.Hamashima, M.Sasatsu and T.Arai. 1998. Distribution of the antiseptic-resistance gene qacE delta 1 in gram-positive bacteria. *FEMS Microbiol Lett.* **165:**295-9.

99.   Klyachko,K.A. and A.A.Neyfakh. 1998. Paradoxical enhancement of the activity of a bacterial multidrug transporter caused by substitutions of a conserved residue. *J Bacteriol.* **180:**2817-21.

100. Kobayashi,N., K.Nishino and A.Yamaguchi. 2001. Novel macrolide-specific ABC-type efflux transporter in *Escherichia coli*. *J Bacteriol.* **183:**5639-44.

101. Kobayashi,N., K.Nishino, T.Hirata, and A.Yamaguchi. 2003. Membrane topology of ABC-type macrolide antibiotic exporter MacB in *Escherichia coli*. *FEBS Lett.* **546:**241-246.

102. Koebnik,R., K.P.Locher and P.Van Gelder. 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol.* **37:**239-53.

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

104. Koronakis,V., A.Sharff, E.Koronakis, B.Luisi, and C.Hughes. 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature.* **405:**914-919.

105. Koronakis,V., C.Andersen and C.Hughes. 2001. Channel-tunnels. *Curr Opin Struct Biol.* **11:**403-7.

106. Koronakis,V., J.Eswaran, and C.Hughes. 2004. Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annu Rev Biochem.* **73:**467-489.

107. Kranitz,L., H.Benabdelhak, C.Horn, M.A.Blight, I.B.Holland and L.Schmitt. 2002. Crystallization and preliminary X-ray analysis of the ATP-binding domain of the ABC transporter haemolysin B from *Escherichia coli*. *Acta Crystallogr D Biol Crystallogr.* **58:**539-41.

108. Laemmli,U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227:**680-685.

109. Lagos,R., J.E.Villanueva and O.Monasterio. 1999. Identification and properties of the genes encoding microcin E492 and its immunity protein. *J Bacteriol.* **181:**212-7.

110. Letoffe,S., J.M.Ghigo, and C.Wandersman. 1994. Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *J Bacteriol.* **176:**5372-5377.

111. Letunic,I., L.Goodstadt, N.J.Dickens, T.Doerks, J.Schultz, R.Mott, F.Ciccarelli, R.R.Copley, C.P.Ponting and P.Bork. 2002. Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Res.* **30:**42-244.

112. Li,X.Z., H.Nikaido and K.Poole. 1995. Role of mexA-mexB-oprM in antibiotic efflux in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother.* **39:**1948-53.

113. Li,X.Z. and K.Poole. 2001. Mutational analysis of the OprM outer membrane component of the MexA-MexB-OprM multidrug efflux system of *Pseudomonas aeruginosa*. *J Bacteriol.* **183:**12-27.

114. Locher,K.P., A.T.Lee and D.C.Rees. 2002. The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science.* **296:**1091-8.

115. Lomovskaya,O. and K.Lewis. 1992. Emr, an *Escherichia coli* locus for multidrug resistance. *Proc Natl Acad Sci U S A.* **89:**8938-42.

116. Lucas,C.E., K.E.Hagman, J.C.Levin, D.C.Stein and W.M.Shafer. 1995. Importance of lipooligosaccharide structure in determining gonococcal resistance to hydrophobic antimicrobial agents resulting from the mtr efflux system. *Mol Microbiol.* **16:**1001-9.

117. Ma,D., D.N.Cook, M.Alberti, N.G.Pon, H.Nikaido, and J.E.Hearst. 1995. Genes acrA and acrB encode a stress-induced efflux system of *Escherichia coli. Mol Microbiol.* **16:**45-55.

118. Makino,K., K.Oshima, K.Kurokawa, K.Yokoyama, T.Uda, K.Tagomori, Y.Iijima, M.Najima, M.Nakano, A.Yamashita, Y.Kubota, S.Kimura, T.Yasunaga, T.Honda, H.Shinagawa, M.Hattori, and T.Iida. 2003. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V cholerae. Lancet.* **361:**743-749.

119. Manting,E.H. and A.J.Driessen. 2000. *Escherichia coli* translocase: the unravelling of a molecular machine. *Mol Microbiol.* **37:**226-38.

120. Marger,M.D. and M.H.Saier,Jr. 1993. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. *Trends Biochem Sci.* **18:**13-20.

121. Maxwell,A. 1997. DNA gyrase as a drug target. *Trends Microbiol.* **5:**102-9.

122. Mikolosko,J., K.Bobyk, H.I.Zgurskaya and P.Ghosh. 2006 Conformational flexibility in the multidrug efflux system protein AcrA. *Structure.* **14:**577-87.

123. Misra,R. and P.R.Reeves. 1987. Role of micF in the tolC-mediated regulation of OmpF, a major outer membrane protein of *Escherichia coli* K-12. *J Bacteriol.* **169:**4722-30.

124. Moellering,R.C.Jr. 1998. Problems with antimicrobial resistance in gram-positive cocci. *Clin Infect Dis.* **26:**1177-8.

125. Mokhonov,V.V., E.I.Mokhonova, H.Akama and T.Nakae. 2004. Role of the membrane fusion protein in the assembly of resistance-nodulation-cell division multidrug efflux pump in *Pseudomonas aeruginosa. Biochem Biophys Res Commun.* **322:**483-9.

126. Moore,R.A., D.DeShazer, S.Reckseidler, A.Weissman and D.E.Woods. 1999. Efflux-mediated aminoglycoside and macrolide resistance in *Burkholderia pseudomallei. Antimicrob Agents Chemother.* **43:**465-70.

127. Morita,Y., K.Kodama, S.Shiota, T.Mine, A.Kataoka, T.Mizushima and T.Tsuchiya. 1998. NorM, a putative multidrug efflux protein, of *Vibrio parahaemolyticus* and its homolog in *Escherichia coli. Antimicrob Agents Chemother.* **42:**1778-82.

128. Morona,R. and P.Reeves. 1982. The tolC locus of *Escherichia coli* affects the expression of three major outer membrane proteins. *J Bacteriol.* **150:**1016-23.

129. Mulder,N.J., R.Apweiler, T.K.Attwood, A.Bairoch, D.Barrell, A.Bateman, D.Binns, M.Biswas, P.Bradley, P.Bork, P.Bucher, R.R.Copley, E.Courcelle, U.Das, R.Durbin, L.Falquet, W.Fleischmann, S.Griffiths-Jones, D.Haft, N.Harte, N.Hulo, D.Kahn, A.Kanapin, M.Krestyaninova, R.Lopez, I.Letunic,

D.Lonsdale, V.Silventoinen, S.E.Orchard, M.Pagni, D.Peyruc, C.P.Ponting, J.D.Selengut, F.Servant, C.J.Sigrist, R.Vaughan and E.M.Zdobnov. 2003. The InterPro Database, 2003 brings increased coverage and new features. *Nucleic Acids Res.* **31:**315-318.

130. Murakami,S., R.Nakashima, E.Yamashita, and A.Yamaguchi. 2002. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature.* **419:**587-593.

131. Muth,T.R. and S.Schuldiner. 2000. A membrane-embedded glutamate is required for ligand binding to the multidrug transporter EmrE. *EMBO J.* **19:**234-40.

132. Nagel de Zwaig,R. and S.E.Luria. 1967. Genetics and physiology of colicin-tolerant mutants of *Escherichia coli*. *J Bacteriol.* **94:**1112-23.

133. Narita,S., S.Eda, E.Yoshihara, and T.Nakae. 2003. Linkage of the efflux-pump expression level with substrate extrusion rate in the MexAB-OprM efflux pump of *Pseudomonas aeruginosa. Biochem Biophys Res Commun.* **308:**922-926.

134. Neu,H.C. 1992. The crisis in antibiotic resistance. *Science.* **257:**1064-73.

135. Neuwald,A.F., J.S.Liu, D.J.Lipman, and C.E.Lawrence. 1997. Extracting protein alignment models from the sequence database. *Nucleic Acids Res.* **25:**1665-1677.

136. Ng,E.Y., M.Trucksis and D.C.Hooper. 1994. Quinolone resistance mediated by norA: physiologic characterization and relationship to flqB, a quinolone resistance locus on the *Staphylococcus aureus* chromosome. *Antimicrob Agents Chemother.* **38:**1345-55.

137. Nies,D.H., A.Nies, L.Chu and S.Silver. 1989. Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system from *Alcaligenes eutrophus. Proc Natl Acad Sci U S A.* **86:**7351-5.

138. Nikaido,H. and M.Vaara. 1985. Molecular basis of bacterial outer membrane permeability. *Microbiol Rev.* **49:**1-32.

139. Nikaido,H. 1989. Outer membrane barrier as a mechanism of antimicrobial resistance. *Antimicrob Agents Chemother.* **33:**1831-6.

140. Nikaido,H. 1994. Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science.* **264:**382-388.

141. Nikaido,H. 1996. Multidrug efflux pumps of gram-negative bacteria. *J Bacteriol.* **178:**5853-9.

142. Nishino,K. and A.Yamaguchi. 2001. Overexpression of the response regulator evgA of the two-component signal transduction system modulates multidrug resistance conferred by multidrug resistance transporters. *J Bacteriol.* **183:**1455-8.

143. Paulsen,I.T. and R.A.Skurray. 1993. Topology, structure and evolution of two families of proteins involved in antibiotic and antiseptic resistance in eukaryotes and prokaryotes--an analysis. *Gene.* **124:**1-11.

144. Paulsen,I.T., L.Nguyen, M.K.Sliwinski, R.Rabus, and M.H.Saier,Jr. 2000. Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *J Mol Biol.* **301:**75-100.

145. Peterson,J.D., L.A.Umayam, T.Dickinson, E.K.Hickey and O.White. 2001. The Comprehensive Microbial Resource. *Nucleic Acids Res.* **29:**123-125.

146. Pilsl,H. and V.Braun. 1995. Novel colicin 10: assignment of four domains to TonB- and TolC-dependent uptake via the Tsx receptor and to pore formation. *Mol Microbiol.* **16:**57-67.

147. Pimenta,A.L., J.Young, I.B.Holland, and M.A.Blight. 1999. Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E coli* haemolysin translocator. *Mol Gen Genet.* **261:**122-132.

148. Poole,K., K.Tetro, Q.Zhao, S.Neshat, D.E.Heinrichs and N.Bianco. 1996. Expression of the multidrug resistance operon mexA-mexB-oprM in *Pseudomonas aeruginosa*: mexR encodes a regulator of operon expression. *Antimicrob Agents Chemother.* **40:**2021-8.

149. Poole,K. 2005. Efflux-mediated antimicrobial resistance. *J Antimicrob Chemother.* **56:**20-51.

150. Prilipov,A., P.S.Phale, P.VanGelder, J.P.Rosenbusch, and R.Koebnik. 1998. Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from *E. coli*. *FEMS Micobiol Lett.* **163:**65-72.

151. Putman,M., H.W.vanVeen and W.N.Konings. 2000. Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev.* **64:**672-93.

152. Rasmussen,B.A. and K.Bush. 1997. Carbapenem-hydrolyzing beta-lactamases. *Antimicrob Agents Chemother.* **41:**223-32.

153. Ren,Q., K.H.Kang and I.T.Paulsen. 2004. TransportDB: a relational database of cellular membrane transport systems. *Nucleic Acids Res.* **32:**D284-D288.

154. Saier,M.H.Jr., R.Tam, A.Reizer and J.Reizer. 1994. Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport. *Mol Microbiol.* **11:**841-7.

155. Saier,M.H.Jr., I.T.Paulsen, M.K.Sliwinski, S.S.Pao, R.A.Skurray and H.Nikaido. 1998. Evolutionary origins of multidrug and drug-specific efflux pumps in bacteria. *FASEB J.* **12:**265-74.

156. Saier,M.H.Jr., J.T.Beatty, A.Goffeau, K.T.Harley, W.H.Heijne, S.C.Huang, D.L.Jack, P.S.Jahn, K.Lew, J.Liu, S.S.Pao, I.T.Paulsen, T.T.Tseng and P.S.Virk. 1999. The major facilitator superfamily. *J Mol Microbiol Biotechnol.* **1:**257-79.

157. Saier,M.H.Jr. 2000. A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev.* **64:**354-411.

158. Sambrook,J., E.F.Fritsch and T.Maniatis. 1988. Molecular cloning – A Laboratory Manual, 2nd edn. *Cold Spring Harbor Laboratory Press.*

159. Sanchez,L., W.Pan, M.Vinas, and H.Nikaido. 1997. The acrAB homolog of *Haemophilus influenzae* codes for a functional multidrug efflux pump. *J Bacteriol.* **179:**6855-6857.

160. Schmitt,L., H.Benabdelhak, M.A.Blight, I.B.Holland, and M.T.Stubbs. 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:**333-342.

161. Schulein,R., I.Gentschev, H.J.Mollenkopf, and W.Goebel. 1992. A topological model for the hemolysin translocator protein HlyD. *Mol Gen Genet.* **234:**155-163.

162. Shaw,K.J., P.N.Rather, R.S.Hare and G.H.Miller. 1993. Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. *Microbiol Rev.* **57:**138-63.

163. Spratt,B.G. and K.D.Cromie. 1988. Penicillin-binding proteins of gram-negative bacteria. *Rev Infect Dis.* **10:**699-711.

164. Spratt,B.G. 1994. Resistance to antibiotics mediated by target alterations. *Science.* **264:**388-93.

165. Srikumar,R., T.Kon, N.Gotoh and K.Poole. 1998. Expression of *Pseudomonas aeruginosa* multidrug efflux pumps MexA-MexB-OprM and MexC-MexD-OprJ in a multidrug-sensitive *Escherichia coli* strain. *Antimicrob Agents Chemother.* **42:**65-71.

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

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

168. Stover,C.K., X.Q.Pham, A.L.Erwin, S.D.Mizoguchi, P.Warrener, M.J.Hickey, F.S.Brinkman, W.O.Hufnagle, D.J.Kowalik, M.Lagrou, R.L.Garber, L.Goltry, E.Tolentino, S.Westbrock-Wadman, Y.Yuan, L.L.Brody, S.N.Coulter, K.R.Folger, A.Kas, K.Larbig, R.Lim, K.Smith, D.Spencer, G.K.Wong, Z.Wu, I.T.Paulsen, J.Reizer, M.H.Saier, R.E.Hancock, S.Lory, and M.V.Olson. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature.* **406:**959-964.

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

170. Tamura,N., S.Murakami, Y.Oyama, M.Ishiguro and A.Yamaguchi. 2005. Direct interaction of multidrug efflux transporter AcrB and outer membrane channel TolC detected via site-directed disulfide cross-linking. *Biochemistry.* **44:**11115-21.

171. Tate,C.G., E.R.Kunji, M.Lebendiker and S.Schuldiner. 2001. The projection structure of EmrE, a proton-linked multidrug transporter from *Escherichia coli*, at 7 A resolution. *EMBO J.* **20:**77-81.

172. Tatusov, R.L., E.V.Koonin and D.J.Lipman. 1997. A genomic perspective on protein families. *Science.* **278:**631-637.

173. Tatusov, R.L., D.A.Natale, I.V.Garkavtsev, T.A.Tatusova, U.T.Shankavaram, B.S.Rao, B.Kiryutin, M.Y.Galperin, N.D.Fedorova and E.V.Koonin. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* **29:**22-28.

174. Thanabalu,T., E.Koronakis, C.Hughes, and V.Koronakis. 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.

175. Thanassi,D.G., G.S.Suh and H.Nikaido. 1995. Role of outer membrane barrier in efflux-mediated tetracycline resistance of *Escherichia coli*. *J Bacteriol*. **177:**998-1007.

176. Thanassi,D.G. and S.J.Hultgren. 2000. Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr Opin Cell Biol*. **12:**420-30.

177. Thompson,S.A., O.L.Shedd, K.C.Ray, M.H.Beins, J.P.Jorgensen and Blaser MJ. 1998. *Campylobacter fetus* surface layer proteins are transported by a type I secretion system. *J Bacteriol*. **180:**6450-8.

178. Tikhonova,E.B., Q.Wang, and H.I.Zgurskaya. 2002. Chimeric analysis of the multicomponent multidrug efflux transporters from gram-negative bacteria. *J Bacteriol*. **184:**6499-6507.

179. Tikhonova,E.B. and H.I.Zgurskaya. 2004. AcrA, AcrB, and TolC of *Escherichia coli* form a stable intermembrane multidrug efflux complex. *J Biol Chem*. **279:**32116-24.

180. Tomb,J.F., O.White, A.R.Kerlavage, R.A.Clayton, G.G.Sutton, R.D.Fleischmann, K.A.Ketchum, H.P.Klenk, S.Gill, B.A.Dougherty, K.Nelson, J.Quackenbush, L.Zhou, E.F.Kirkness, S.Peterson, B.Loftus, D.Richardson, R.Dodson, H.G.Khalak, A.Glodek, K.McKenney, L.M.Fitzegerald, N.Lee, M.D.Adams, E.K.Hickey, D.E.Berg, J.D.Gocayne, T.R.Utterback, J.D.Peterson, J.M.Kelley, M.D.Cotton, J.M.Weidman, C.Fujii, C.Bowman, L.Watthey, E.Wallin, W.S.Hayes, M.Borodovsky, P.D.Karp, H.O.Smith, C.M.Fraser, and J.C.Venter. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*. **388:**539-547.

181. Touze,T., J.Eswaran, E.Bokma, E.Koronakis, C.Hughes and V.Koronakis. 2004. Interactions underlying assembly of the *Escherichia coli* AcrAB-TolC multidrug efflux system. *Mol Microbiol*. **53:**697-706.

182. Towbin,H., T.Staehelin and J.Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A*. **76:**4350-4.

183. Trepod,C.M. and J.E.Mott. 2004. Identification of the *Haemophilus influenzae* tolC gene by susceptibility profiles of insertionally inactivated efflux pump mutants. *Antimicrob Agents Chemother*. **48:**1416-1418.

184. Tseng,T.T., K.S.Gratwick, J.Kollman, D.Park, D.H.Nies, A.Goffeau and M.H.Saier,Jr. 1999. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J Mol Microbiol Biotechnol*. **1:**107-25.

185. Turk,D.C. 1984. The pathogenicity of *Haemophilus influenzae*. *J Med Microbiol*. **18:**1-16.

186. Van Dyk,T.K., L.J.Templeton, K.A.Cantera, P.L.Sharpe, and F.S.Sariaslani. 2004. Characterization of the *Escherichia coli* AaeAB efflux pump: a metabolic relief valve? *J Bacteriol*. **186:**7196-7204.

187. Walsh,C.T., S.L.Fisher, I.S.Park, M.Prahalad and Z.Wu. 1996. Bacterial resistance to vancomycin: five genes and one missing hydrogen bond tell the story. *Chem Biol*. **3:**21-8.

188. Walsh,C. 2000. Molecular mechanisms that confer antibacterial drug resistance.*Nature*. **406:**775-81.

189. Wandersman,C. and P.Delepelaire. 1990. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc Natl Acad Sci U S A*. **87:**4776-4780.

190. Wang,H.W., Y.J.Lu, L.J.Li, S.Liu, D.N.Wang, and S.Sui. 2000. Trimeric ring-like structure of ArsA ATPase. *FEBS Lett*. **469:**105-110.

191. Weisblum,B. 1995. Erythromycin resistance by ribosome modification. *Antimicrob Agents Chemother.* **39:**577-85.

192. Williams,D.H. 1996. The glycopeptide story--how to kill the deadly 'superbugs'. *Nat Prod Rep.* **13:**469-77.

193. Wipat,A., S.C.Brignell, B.J.Guy, M.Rose, P.T.Emmerson, and C.R.Harwood. 1998. The yvsA-yvqA (293 degrees-289 degrees) region of the *Bacillus subtilis* chromosome containing genes involved in metal ion uptake and a putative sigma factor. *Microbiology.* **144**:1593-1600.

194. Wong,K.K., F.S.Brinkman, R.S.Benz, and R.E.Hancock. 2001. Evaluation of a structural model of *Pseudomonas aeruginosa* outer membrane protein OprM, an efflux component involved in intrinsic antibiotic resistance. *J Bacteriol.* **183:**367-374.

195. Wu,C.H., L.S.Yeh, H.Huang, L.Arminski, J.Castro-Alvear, Y.Chen, Z.Hu, P.Kourtesis, R.S.Ledley, B.E.Suzek, C.R.Vinayaka, J.Zhang, and W.C.Barker. 2003. The Protein Information Resource. *Nucleic Acids Res.* **31:**345-347.

196. Yerushalmi,H. and S.Schuldiner. 2000. An essential glutamyl residue in EmrE, a multidrug antiporter from *Escherichia coli*. *J Biol Chem.* **275:**5264-9.

197. Yoneyama,H., H.Maseda, H.Kamiguchi and T.Nakae. 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:**4628-34.

198. Yoshida,H., M.Bogaki, M.Nakamura and S.Nakamura. 1990. Quinolone resistance-determining region in the DNA gyrase gyrA gene of *Escherichia coli. Antimicrob Agents Chemother.* **34:**1271-2.

199. Zgurskaya,H.I. and H.Nikaido. 1999. AcrA is a highly asymmetric protein capable of spanning the periplasm. *J Mol Biol.* **285:**409-420.

200. Zgurskaya,H.I. and H.Nikaido. 2000a. Cross-linked complex between oligomeric periplasmic lipoprotein AcrA and the inner-membrane-associated multidrug efflux pump AcrB from *Escherichia coli*. *J Bacteriol.* **182:**4264-4267.

201. Zgurskaya,H.I. and H.Nikaido. 2000b. Multidrug resistance mechanisms: drug efflux across two membranes. *Mol Microbiol.* **37:**219-225.

202. Zhang,L., X.Z.Li and K.Poole. 2001. SmeDEF multidrug efflux pump contributes to intrinsic multidrug resistance in *Stenotrophomonas maltophilia*. *Antimicrob Agents Chemother.* **45:**3497-503.

6.2. Curriculum vitae

Name: Johann Georg Polleichtner

Address: Dettelbachergasse 9,
97070 Würzburg

Date and place of birth: 14.01.1977   Aschaffenburg/Germany

Marital status: married

Schooling:

1983-1987	Elementary school Laufach
1987-1996	Kronberg-Gymnasium Aschaffenburg
06/1996	Abitur

Course of Study:

11/1996-12/2001	Study of biology at the Julius-Maximilians-Universität of Würzburg Major subject: Biotechnology Minor subjects: Cell- and Developmental Biology; Genetics
01/2000-06/2000	Study at the Université Paul-Sabatier, Toulouse as participant of the SOKRATES/ERASMUS program
12/2001	Diploma-thesis at the department of Biotechnology in the laboratory of Prof. Dr. R. Benz: "Study of the pore-forming domain of the adenylate cyclase toxin (ACT) of *Bordetella pertussis*"
01/2002	Start of Ph.D.-thesis

Chapter 6 - Appendix

6.3.   Acknowledgements