Uncovering small membrane proteins in pathogenic bacteria: Regulatory functions and therapeutic potential

Preeti Garai, Anne Blanc-potard

To cite this version:

Uncovering small membrane proteins in pathogenic bacteria: regulatory functions and therapeutic potential

Preeti Garai$ and Anne Blanc-Potard*

1Laboratory of Pathogen-Host Interactions, Université de Montpellier, CNRS-UMR5235, Montpellier, France

$Present address: Department of Microbiology and Immunology, B2620, Biological Sciences Building, Medical College of Wisconsin, Milwaukee, USA

*Corresponding author: E-mail: anne.blanc-potard@umontpellier.fr
Abstract

Bacterial small proteins (below 50 amino acids) encoded by small open reading frames (sORFs) are recognized as an emerging class of functional molecules that have been largely overlooked in the past. While some were uncovered serendipitously, global approaches have recently been developed to detect these sORFs. A large portion of small proteins appears to be hydrophobic and located in the bacterial membrane. In the present review, we describe functional small hydrophobic proteins discovered in pathogenic bacteria and report recent advances in the discovery of additional ones. Small membrane proteins contribute to bacterial adaptation to changing environments and often appear to be implicated in negative feedback regulation loops by modulating the function or stability of larger membrane proteins. A subset of these proteins belongs to toxin-antitoxin modules. We highlight features of characterized hydrophobic small proteins that may pave the way for identification of the functional small proteins among novel sORFs discovered. Besides providing new insights into bacterial pathogenesis, identification of naturally-occurring small hydrophobic proteins of pathogenic bacteria can lead to new therapeutic interventions, as recently shown with the development of synthetic peptides derived from natural small proteins that display antibacterial or antivirulence properties.

Keywords: Small protein, pathogenic bacteria, membrane protein, regulation, antibacterial molecule
Introduction

In bacteria, small non-coding RNAs (sRNAs) have been recognized for decades for their regulatory role in a wide range of stress responses, including the response to host cells in the context of pathogenic bacteria. More recently, increasing evidence has shown that some sRNAs annotated as non-coding RNAs could in fact encode small proteins that were previously not considered in both eukaryotes (Chen, Brunner et al. 2020) and prokaryotes (Hemm, Weaver et al. 2020; Orr, Mao et al. 2020). Identification of small proteins that directly result from the translation of small open reading frames (sORFs) rather than cleavage of larger protein precursors adds to the pool of functional molecules. We use the term small protein for proteins that are below 50-amino acids in size, while they are also named in other reports as “very small proteins”, “microproteins”, “miniproteins” or “peptides”.

Small ORFs were historically missed in genome annotations (because of the minimal ORF length cutoff commonly used) and proteomic studies. Several translated sORFs from bacteria were discovered by serendipity. In addition, multiple global approaches, including bioinformatics prediction, mass spectrometry and ribosome profiling, have been used to identify bacterial small proteins (Sousa and Farkas 2018; Orr, Mao et al. 2020). These global studies have been performed in the model organism Escherichia coli (Hemm, Paul et al. 2010), in few pathogenic bacteria (see below) and more recently in bacteria of the microbiota (Sberro, Fremin et al. 2019). A few bacterial small proteins associated with diverse functions have been characterized and include toxins, chaperones, stabilizing factors of larger proteins and regulators (Storz, Wolf et al. 2014; Duval and Cossart 2017; Orr, Mao et al. 2020).

A substantial fraction of bacterial small proteins are hydrophobic with a single alpha-helix transmembrane domain and are predicted to localize in bacterial membrane, specifically the inner membrane in the case of Gram negative bacteria (Alix and Blanc-Potard 2009; Fontaine, Fuchs et al. 2011). In the present review, we focus on small membrane-associated
proteins identified in pathogenic bacteria (listed in Table 1) and their potential use in developing therapeutic strategies. These proteins appear to belong to two main functional classes: i) regulators, which modulate cellular functions like transport or antimicrobial resistance through direct interaction with other larger proteins, and ii) toxins, which are part of toxin-antitoxin modules and have antibacterial properties. Increasing evidence indicates that small membrane proteins can interfere with bacterial virulence and be used for antivirulence strategies. Moreover, common features of these small proteins will be highlighted here, which may facilitate the identification of new functional proteins from global studies carried out on pathogenic bacteria.

**Small hydrophobic proteins regulating membrane proteins in response to cation levels**

We describe here membrane-associated small proteins found in pathogenic bacteria that regulate larger membrane proteins, including virulence factors, involved in the adaptation of bacteria to changing concentration of cations in the environment (Table 1 and Table 2). These small proteins, which were mostly identified by serendipity, are regulated by two-component systems (TCS) (Table 2). As described below, implication of small proteins in negative regulation of virulence proteins can provide a basis for novel therapeutic perspectives.

**MgtR and MgtS: regulators of protein stability under Mg^{2+} limitation**

MgtR and MgtS, both present in *Salmonella enterica* serovar Typhimurium (*S. Typhimurium*), are two small hydrophobic proteins regulated by the PhoQP TCS, which is induced by Mg^{2+} limitation (Groisman 2001).

MgtR is a 30-amino acid long hydrophobic protein found to regulate MgtC in *S. Typhimurium* (Alix and Blanc-Potard 2008). MgtR is an inner membrane protein, with its N-
terminus in the cytoplasm and C-terminus in the periplasm (N\textsubscript{(in)}-C\textsubscript{(out)} orientation), encoded in the \textit{mgtCBR} operon, downstream to the gene coding for the magnesium transporter MgtB (Soncini, Garcia Vescovi et al. 1996; Groisman 2001). MgtR interacts with the inner membrane protein MgtC and promotes its degradation by FtsH protease, thus providing a negative feedback regulation (Alix and Blanc-Potard 2008). An Ala coil motif in the transmembrane domain of MgtR is essential for the interaction between MgtR and MgtC (Alix and Blanc-Potard 2008). The level of MgtB protein is not affected by MgtR, as shown by using \textit{mgtR} deletion mutant or \textit{mgtR} overexpressing strains (Alix and Blanc-Potard 2008; Olvera, Garai et al. 2019), but deletion of the \textit{mgtR} gene slightly enhanced expression of MgtA, another Mg\textsuperscript{2+} transporter of \textit{Salmonella} regulated by PhoP, indicative of a negative regulation by MgtR (Choi, Lee et al. 2012).

MgtC is a virulence factor that plays an important role in the adaptation to Mg\textsuperscript{2+} deprivation and in the intramacrophage survival of \textit{S. Typhimurium} (Blanc-Potard and Groisman 1997). Although a \textit{Salmonella mgtR} mutant is not significantly attenuated in macrophages (Alix and Blanc-Potard 2008), the overexpression of MgtR decreased the survival of \textit{Salmonella} inside macrophages, likely due to a decreased level of MgtC protein (Olvera, Garai et al. 2019). Hence, MgtR appeared as an antivirulence molecule. While MgtC has been described in other bacterial pathogens (Alix and Blanc-Potard 2007; Belon, Soscia et al. 2015), MgtR is naturally found only in \textit{Salmonella} (Alix and Blanc-Potard 2008). However, \textit{Salmonella} MgtR interacts with MgtC from other species and heterologous expression of \textit{mgtR} in mycobacteria and \textit{Pseudomonas aeruginosa} reduced intramacrophage survival, suggesting that the antivirulence strategy may apply to these bacteria as well (Jean-Francois, Dai et al. 2014; Belon, Soscia et al. 2015; Belon, Rosas Olvera et al. 2016).

A synthetic MgtR peptide was used to test the antivirulence property of MgtR when added exogenously to pathogenic bacteria. The synthetic MgtR peptide was successful in
reducing the pathogenicity of *S. Typhimurium* (Olvera, Garai et al. 2019). Interestingly, a derivative lacking the Ala-coil interacting motif was also efficient in reducing *Salmonella* intramacrophage survival and lowering not only the level of MgtC protein but also of other inner membrane proteins. This suggested that the effect of synthetic MgtR is less specific than that of endogenously produced MgtR. The growth of *S. Typhimurium* was not affected by synthetic MgtR, indicating a mode of action that differs from classical antimicrobial peptides (Olvera, Garai et al. 2019). Synthetic MgtR peptide was also found to reduce the level of MgtC protein and the survival of *P. aeruginosa* inside macrophages without bactericidal activity (Moussouni, Nogaret et al. 2019), supporting the therapeutic implication towards various pathogens.

The 31-amino acid long inner membrane protein MgtS (formerly denoted YneM) was first discovered in *E. coli*, where it acts to increase intracellular Mg\(^{2+}\) levels and maintain cell integrity upon Mg\(^{2+}\) depletion (Wang, Yin et al. 2017). MgtS interacts with MgtA to increase the levels of this P-type ATPase Mg\(^{2+}\) transporter under Mg\(^{2+}\)-limiting conditions. Overexpression of MgtS also leads to induction of the PhoRB regulon because MgtS can form a complex with a second protein, PitA, a cation-phosphate symporter (Yin, Wu Orr et al. 2019). Interestingly, MgtS is also present in *S. Typhimurium* (Table 1), but its role has not been studied. An effect of *Salmonella* MgtS on MgtA stabilization would antagonize the role of MgtR. Moreover, it would be of interest to test whether *Salmonella* MgtS can also stabilize MgtC or MgtB, which are not present in *E. coli*.

**MgrB**: modulator of Mg\(^{2+}\)-responsive signaling pathway

MgrB is a 47-amino acid long inner membrane protein, with an N\(_{\text{in}}\)-C\(_{\text{out}}\) orientation, first identified in *E. coli* and also found in enterobacterial pathogens including *Salmonella enterica* (designated as *yobG*), *Shigella flexneri*, *Yersinia pestis* and *Klebsiella pneumoniae*
(Lippa and Goulian 2009). Its transcription is activated by phosphorylated PhoP under magnesium limitation (Lippa and Goulian 2009).

MgrB negatively regulates the PhoQ sensor kinase activity by interacting with its periplasmic domain, thus preventing any further phosphorylation of PhoP and therefore $mgrB$ expression (Lippa and Goulian 2009). This MgrB-mediated negative feedback serves as a mode of a partial adaptation to achieve a steady state level of expression of genes induced by PhoP (Salazar, Podgornaia et al. 2016). Interestingly, this modulation of function of PhoP, mediated by MgrB, was observed even under high levels of magnesium, a condition that reduces PhoP phosphorylation (Lippa and Goulian 2009). Intriguingly, another small membrane protein of 65-amino acids, SafA (formerly known as B1500), also binds to PhoQ to promote its kinase activity and therefore opposes the action of MgrB (Eguchi, Itou et al. 2007; Eguchi, Ishii et al. 2012). MgrB has been proposed to mediate a redox sensitivity in the periplasm, which plays a critical role in the regulation of the PhoQP circuit (Lippa and Goulian 2012). Disulfide bonding by one or more MgrB cysteine residues seems important for the repressive effect of MgrB on PhoQ and a model has been proposed where DsbA affects PhoQP signaling by oxidizing C28 and C39 of MgrB (Lippa and Goulian 2012). In *E. coli*, MgrB is also required for survival under acid stress (Hobbs, Astarita et al. 2010), possibly due to a direct interaction between MgrB and IraM, leading to accumulation of RpoS and induction of acid resistance mediated by GadABC and HdeAB systems (Xu, Li et al. 2019).

In *S. Typhimurium*, where PhoQP plays a major role in virulence, MgrB acts as a negative regulator of the expression of the effector *steA* by inhibiting PhoP-dependent activation of *steA* (Cardenal-Munoz and Ramos-Morales 2013). In *K. pneumoniae* clinical isolates, inactivation of the $mgrB$ gene is responsible for acquired colistin resistance by upregulating PhoQP activity, which, in turn, activates the PmrBA system responsible of
modification of the lipopolysaccharide polymyxin target (Cannatelli, D'Andrea et al. 2013; Cannatelli, Giani et al. 2014).

PmrR: regulator of LPS in response to Fe\textsuperscript{3+} levels

PmrR is a 30-amino acid long hydrophobic protein located in the inner membrane of S. Typhimurium (Kato, Chen et al. 2012). PmrR was identified as a PmrA-activated sORF located immediately downstream to the \textit{pmrAB} operon encoding the PmrBA TCS (Kato, Chen et al. 2012). It was also independently identified by a bioinformatics approach (Alix and Blanc-Potard 2009). PmrBA, which responds to the presence of Fe\textsuperscript{3+} ions in the environment, is also responsible for modifying LPS in \textit{Salmonella} during the shift from Fe\textsuperscript{3+} poor to Fe\textsuperscript{3+} rich environment in order to prevent toxicity by Fe\textsuperscript{3+} ions (Wosten, Kox et al. 2000; Kato, Chen et al. 2012). PmrR interacts with and inhibits the activity of LpxT, an inner membrane enzyme responsible for increasing the LPS negative charge. This prevents the generation of 1-PP lipid A and decreases the negative charge of LPS, which further lowers Fe\textsuperscript{3+} binding to the bacterial cell (Kato, Chen et al. 2012). Because Fe\textsuperscript{3+} is an activating ligand for the sensor PmrB, transcription of PmrA-dependent LPS-modifying genes is reduced by PmrR. This mechanism acts as a negative feedback regulation for the expression of PmrA and enables bacteria to regulate their cell surface modifications based on the availability of an inducing signal.

The importance of LPS in virulence of \textit{Salmonella}, including intracellular replication of the pathogen, is well established (Nagy, Danino et al. 2006). Hence PmrR-dependent modification of LPS could be exploited for the development of antivirulence therapeutic agents against \textit{Salmonella} or other related pathogens, like \textit{K. pneumoniae}, which are known to contain PmrR (Kato, Chen et al. 2012).
KdpF: stabilizer of a K⁺ transporter

KdpF is a 30-amino acid long membrane protein with an N\textsubscript{(in)}-C\textsubscript{(out)} orientation, which is part of the P-type K⁺ transporter complex KdpFABC (Huang, Pedersen et al. 2017). Expression of \textit{kdpFABC} is induced by the KdpDE TCS under K⁺ limitation (Greie and Altendorf 2007; Gannoun-Zaki, Alibaud et al. 2013). KdpF, mainly studied in \textit{E. coli}, but present and annotated in other bacterial genomes, is proposed to be a subunit involved in the stabilization of the KdpABC complex \textit{in vitro}, possibly acting as a lipid-like peptide for the holoenzyme (Greie and Altendorf 2007). However, the role of KdpF \textit{in vivo} remains elusive because it is not essential for the K⁺-stimulated ATPase activity of the Kdp complex.

Interestingly, overexpression of KdpF in \textit{Mycobacterium bovis} BCG reduced intramacrophage growth and altered cording morphology (Gannoun-Zaki, Alibaud et al. 2013). This lower replication rate within macrophages is linked to an increased bacterial susceptibility to nitrosative stress (Rosas Olvera, Vives et al. 2017). Accordingly, \textit{Mycobacterium} KdpF can interact with membrane proteins involved in nitrosative stress detoxification or induced by nitrosative stress. Exogenous addition of KdpF synthetic peptide correlated with reduced level of these interacting target proteins (Rosas Olvera, Vives et al. 2017). Moreover, exogenous KdpF peptide presented similar biological effects as the endogenously expressed KdpF including nitrosative stress susceptibility and reduced intramacrophage replication rate for \textit{M. bovis} BCG, highlighting KdpF as a potent molecule with antivirulence properties. However, this is not common to all intracellular pathogens because, in contrast to mycobacteria, \textit{kdpF} overexpression in \textit{S. Typhimurium} increases intramacrophage multiplication, which may be related to opposite roles reported for the KdpDE TCS in the virulence of these two pathogens (Gannoun-Zaki, Belon et al. 2014).
Small membrane proteins implicated in adaptation to environmental stresses and antibiotic resistance

This section deals with additional small regulatory proteins (Table 1) that interact with larger proteins and participate in the coping strategies of pathogens during environmental stresses like oxidative stress or are involved in the development of resistance to antibiotics. These proteins appear to be beneficial for bacterial pathogens and although they have not been implicated in therapeutic strategies so far, one can propose that preventing interaction with their protein targets using mimicking peptides (Stone and Deber 2017) may reduce bacterial virulence.

CydX: cytochrome oxidase component involved in oxidative stress adaptation

CydX was first identified in *E. coli* during the biochemical study of the cytochrome *bd*-I oxidase (Miller and Gennis 1983) and its gene was identified as a sORF in the *cydAB* operon, previously named *ybgT* (VanOrsdel, Bhatt et al. 2013). CydX is a 37-amino acid long protein shown to be part of the oxidase complex by cryo-EM (Safarian, Hahn et al. 2019), required for the cytochrome *bd* oxidase activity and proposed to be involved in the folding and stabilization of the complex. CydX homologues are found in various bacteria, with a high degree of sequence variation in the C-terminal region (Allen, Brenner et al. 2014). The CydAB oxidase is known to play a role in pathogen survival within hosts and functional studies to specifically address the contribution of CydX have been carried out in two bacterial pathogens, *Brucella abortus* and *S. Typhimurium*. A *B. abortus* *cydX* mutant showed increased sensitivity to H$_2$O$_2$ and decreased acid tolerance (Sun, de Jong et al. 2012). *S. typhimurium* CydX plays an essential role in cytochrome *bd* function under stress conditions, including exposure to nitric oxide, and is required for bacterial replication within macrophages (Duc, Kang et al. 2020).
Prli42: activator of stressosome for adaptation to oxidative stress

Prli42 is a 31-amino acid long tail-anchored membrane protein identified in *Listeria monocytogenes* through a global approach of N-terminal proteogenomics (Impens, Rolhion et al. 2017) (see below). Prli42 relays oxidative stress signals to the stressosome to activate the general sigma factor B (σ^B) regulon mediated stress-sensing pathway. In contrast to the other small regulatory proteins described here, Prli42 directly interacts with a cytosolic component of the stressosome, RsbR, and not a membrane protein. This interaction involves the basic amino acids K4, K5 and K8 of the N-terminus of Prli42 and the acidic residues E109, E110 and E113 of N-terminus of RsbR (Impens, Rolhion et al. 2017). Cryo-electron microscopy revealed that this interaction enables RsbR to sense oxidative stress in the external environment leading to the activation of the stressosome. This is followed by a phosphorylation signaling cascade to activate σ^B and express stress response genes required for the survival of *L. monocytogenes* under oxidative stress (Williams, Redzej et al. 2019). The contribution of Prli42 in adaptation to oxidative stress explains its role in the survival of *L. monocytogenes* inside macrophages (Impens, Rolhion et al. 2017).

AcrZ: Regulator of drug efflux pump

AcrZ, a 49-amino acid long protein (previously known as YbhT) found in the inner membrane of *E. coli*, modulates the efflux activity of AcrAB-TolC drug efflux pump, which is required for resistance against several antibiotics (Hobbs, Yin et al. 2012). Accordingly, a mutation of AcrZ leads to sensitivity towards many antibiotics including chloramphenicol and tetracyclin (Hobbs, Yin et al. 2012). Transcription of *acrZ* is coregulated with *acrAB* (Hobbs, Yin et al. 2012). Cryoelectron microscopy showed that AcrZ interacts with the inner membrane.
protein AcrB (Wang, Fan et al. 2017) and AcrZ cooperate with lipids to allosterically modulate AcrB activity by inducing conformational changes in the drug-binding pocket (Du, Neuberger et al. 2020). AcrZ is highly conserved across enteric bacteria including enteropathogens like *Salmonella* and *Klebsiella* and therefore holds potential in the context of studying multi-drug resistance development by these pathogens.

**Small hydrophobic proteins as part of type I toxin-antitoxin modules**

Chromosomally encoded toxin-antitoxin (TA) systems are abundantly present in bacteria. TA loci encode a stable toxin whose ectopic overexpression either kills cells or confers growth stasis, and an unstable antitoxin. In type I TA systems, the antitoxin is a small antisense RNA that base-pairs with the toxin-encoding mRNA (Brantl 2012; Brielle, Pinel-Marie et al. 2016). We describe here two examples of transmembrane type I toxic peptides present in *S. Typhimurium* and *Staphylococcus aureus* (Table 1). Many additional toxic hydrophobic small proteins are known in *E. coli* and shared by enteropathogens (Fozo, Hemm et al. 2008; Hemm, Weaver et al. 2020). In terms of biological activity, these toxins are related to bacteriocins, which are the subset of antimicrobial peptides produced by bacteria, consisting of small amphipathic peptides that interact with bacterial membranes leading to cell death (Yang, Lin et al. 2014). However, bacteriocins are secreted, whereas toxins of type I TA systems usually remain membrane associated.

**TisB: role in persistence and antibiotic tolerance**

TisB, an SOS-induced membrane associated small protein identified in *E. coli*, is part of TisB/IstR1 TA system, where IstR is a non-coding RNA that inhibits the translation of the TisB toxin mRNA (Vogel, Argaman et al. 2004; Wagner and Unoson 2012). In addition to
slow ing bacterial growth down, TisB also plays a role in the formation of persister cells (Dorr, Vulic et al. 2010; Wagner and Unoson 2012), a subset of a bacterial population that enters a dormant state. Persister cells become refractory to the action of antibiotics and bacteria producing TisB toxin are tolerant to multiple antibiotics (Dorr, Vulic et al. 2010). TisB was shown to decrease proton motive force and ATP levels (Unoson and Wagner 2008), consistent with its role in formation of dormant cells. TisB forms an anion channel in artificial membranes (Gurnev, Ortenberg et al. 2012), thus functions as a typical antimicrobial peptide. The pattern of charges on the amphiphilic TisB helix suggested the presence of antiparallel dimers assembled via a ladder of salt bridges in the form of electrostatic charge-zipper (Steinbrecher, Prock et al. 2012). A computational-experimental approach found TisB to be inserted upright in the membrane as a tetramer (antiparallel dimer-of-dimers) (Schneider, Wadhwani et al. 2019).

Interestingly, TisB is also found in Salmonella (Table 1) but its role has not been investigated. A Salmonella mutant lacking the regulatory sRNA istR showed small, but reproducible, reduction in fitness during competitive infection in mice (Santiviago, Reynolds et al. 2009). Considering the role of TisB in E. coli, a deletion of tisB in S. Typhimurium may reduce persistence frequency in macrophages and play a role in antibiotic tolerance. Moreover, synthetic peptides derived from TisB may display antimicrobial potential against Salmonella.

PepA1, PepG1 and PepG2: role in toxicity towards bacteria and host cells

Several functional type I toxin-antitoxin (TA) modules with membrane peptide toxins are expressed by the human pathogen S. aureus. PepA1 (part of the SprA1/SprA1AS module) localizes in the membrane and triggers S. aureus death (Sayed, Nonin-Lecomte et al. 2012). Its solution structure was solved as a long bent, interrupted helix. When inserted into the S. aureus membrane, the PepA1 conformation switches to a continuous helix, presumably forming pores.
to alter membrane integrity. PepA1 expression is favored in an acidic environment and during oxidative stress, which may play a role in the survival of *S. aureus* after internalization into host immune cells (Sayed, Nonin-Lecomte et al. 2012). *S. aureus* also expresses a related toxin PepA2 (part of the SprA2/SprA2AS module), not redundant with PepA1, which internally triggers bacterial death (Germain-Amiot, Augagneur et al. 2019). Although PepA2 does not affect bacteria when it is present in the extracellular medium, it is highly toxic to other host cells such as polymorphonuclear neutrophils and erythrocytes.

Due to its antibacterial activity, PepA1 has been a starting point for the development of novel therapeutics. Exogenously provided PepA1 has activity against both Gram-positive and Gram-negative bacteria (Solecki, Mosbah et al. 2015). However, PepA1 was significantly hemolytic, thus limiting its use as an antibacterial agent. To overcome these limitations, PepA1 was converted into a cyclic heptapeptide derivative with no toxicity to human cells, but sharp antibacterial activity. Mechanistically, linear and helical PepA1 derivatives form pores at the bacterial and erythrocyte surfaces, while the cyclic peptide induces bacterial envelope reorganization with insignificant action on the erythrocytes (Solecki, Mosbah et al. 2015). Other optimized cyclic heptapeptides with unnatural amino acids that strengthen dynamic association with bacterial lipid bilayers are effective against both MRSA and *P. aeruginosa* in severe sepsis and skin infection models, respectively, thus confirming that they are promising lead candidates for new antibacterial agents (Nicolas, Bordeau et al. 2019).

*S. aureus* also produces an unconventional type I TA system composed of the cis-antisense RNA SprF1 and the stable SprG1 RNA, which encode two toxins (PepG1 and PepG2, of 44 and 31 amino acids, respectively) that accumulate at the membrane and trigger *S. aureus* death (Pinel-Marie et al., 2014). These two toxins can be secreted out by the pathogen to lyse human host cells, a process performed more efficiently by PepG1. Both secreted toxins also inactivate Gram-negative and -positive bacteria, with PepG2 being more effective (Pinel-Marie,
Brielle et al. 2014). Further studies are required to better characterize the role, kinetics of expression and processing of these two toxins during infection, as well as to evaluate their therapeutic potential using synthetic peptides.

**Global analyses leading to identification of sORFs in pathogenic bacteria**

The technologies to identify sORFs in bacteria have recently been improved (Miravet-Verde, Ferrar et al. 2019; Weaver, Mohammad et al. 2019), which can allow analysis of the small proteome of pathogenic bacteria and extend the identification of novel small proteins. As detailed below, at present, global analysis leading to identification of sORFs in bacterial pathogenic species has been carried out in *S. Typhimurium, L. monocytogenes, Mycoplasma pneumoniae* and *S. aureus*.

For *S. Typhimurium*, a combination of *in silico* softwares allowed the identification of 14 putative sORFs below 50 amino acids with a single trans-membrane domain (Alix and Blanc-Potard 2009). Nine of them were conserved in *E. coli*, whereas others were *Salmonella*-specific. MgtR and PmrR, which were identified independently in other studies (Alix and Blanc-Potard 2008; Kato, Chen et al. 2012), were also recovered. Later, an experimental global study based on ribosome profiling identified translated sORFs in *S. Typhimurium* (Baek, Lee et al. 2017). 113 small proteins below 50 amino acids were identified. Expression of 25 small proteins was validated in several conditions, including infection-relevant conditions of low Mg$^{2+}$ or pH 5.8. Interestingly, few sORFs were specifically expressed in infection-relevant conditions and hence are good candidates for playing a role during infection. This study however did not specify the proportion of predicted hydrophobic proteins, nor investigated the cellular location of small proteins. In addition, a mass spectrometry-based proteomics approach used to explore the RpoS proteome (alternative sigma factor involved in stress response) allowed the identification of several small proteins. This included small hydrophobic proteins...
identified in other studies and present in *E. coli* (YncL and YohP), as well as 4 uncharacterized ORFs smaller than 50 amino acids controlled by RpoS (Lago, Monteil et al. 2017).

A genome-wide N-terminomic approach allowed unambiguous detection of bacterial translation initiation sites (TISs) in *Listeria* and led to the discovery of 19 internal translation initiation sites and 6 small proteins, including Prli42 (see above) (Impens, Rolhion et al. 2017).

The method of N-terminal COFRADIC (COmbined FRActional DIagonal Chromatography) isolates N-terminal peptides via two sequential chromatographic separations. In prokaryotes, N-terminal peptides are formylated and this modification can be exploited to identify TISs upon treatment with a peptide deformylase inhibitor.

A combination of novel bioinformatics tool (RanSEPs) with omics approaches allowed the description of 109 small ORFomes in *M. pneumoniae* (Miravet-Verde, Ferrar et al. 2019). RanSEPs is a random forest-based tool for the prediction of small open reading frame-encoded proteins (SEPs). The functional study of SEPs highlighted an enrichment in the membrane, translation, metabolism and nucleotide-binding categories (Miravet-Verde, Ferrar et al. 2019).

Ribosome profiling was also used in *S. aureus* and examining the read coverage across the entire genome allowed the identification of 19 novel ORFs of less than 70-amino acids in length (Davis, Gohara et al. 2014). Most of them contained weak Shine-Dalgarno sequences and although some initiated with an AUG start codon, a vast majority of the newly identified ORFs initiated with alternative codons. Using an *in vitro* translation system with T7 promoter-driven templates carrying the native SD and initiation codons, the protein products of 10 ORFs was detected. Among them, PepA1 peptide, which was found independently (Sayed, Nonin-Lecomte et al. 2012), was recovered as Orf9.

**Uncovering novel functional hydrophobic small proteins in pathogenic bacteria**
With the increasing number of global analyses to uncover sORFs, the next challenge will be to define which ORFs are expressed, the required conditions for their expression and their function. Based on the amino acid sequence of known ORFs, one can distinguish some features that may help developing tools to identify additional functional sORFs (Table 1). Firstly, the size of these small proteins is often around 30-amino acids and few of these proteins are around 45-50-amino acids. Secondly, their amino acid composition displays some common features. Toxic peptides from TA systems have a C-terminal extremity enriched with positively charged lysine residues, especially as last amino-acid (Table 1). Positively charged residues (lysine and arginine) adjacent to the trans-membrane domain favor an N\textsubscript{(out)}-C\textsubscript{(in)} orientation, according to the “positive inside” rule (von Heijne 1992). The positive inside rule applies for single transmembrane domain proteins, as shown by the mutational analysis of conserved positive residues close to the transmembrane domain (Fontaine, Fuchs et al. 2011).

In contrast to toxic peptides, regulatory small membrane proteins are usually found with an N\textsubscript{(in)}-C\textsubscript{(out)} orientation (see above) and are not enriched with positively charged residues in the C-terminal part (Table 1). In addition, with the exception of CydX, we noticed that these proteins harbor aromatic residues at the C-terminal extremity and/or at the end of the trans-membrane domain (Table 1). This may have an impact on the membrane localization of the protein, as shown by the translocation of some outer membrane proteins (Soprova, Sauri et al. 2010). Tryptophan residues in the tail of small membrane protein may also contribute to the stabilization of important structural features, as known for bacteriocins whose activity is dependent on a tryptophan residue that helps in folding back of the C-terminal tail into the central helical part (Fimland, Eijsink et al. 2002). Moreover, tryptophan residues located in the transmembrane helices, which localize preferentially to the interfacial regions of phospholipid bilayers, can influence interactions between transmembrane proteins (White and von Heijne 2005). Accordingly, PmrR’s interaction with its partner LpxT is dependent on a single
tryptophan residue at the position 25 (W25) (Kato et al., 2012). In addition, the W29 residue of MgtS plays a role in the regulatory activity of MgtS towards MgtA (Wang, Yin et al. 2017).

Taking the size and amino acid sequence near the C-terminus into account may thus help to infer sORFs of interest among those identified with predicted transmembrane domain in global studies and to extend the number of small membrane proteins playing a role in pathogenic bacteria.

Concluding remarks

With the development of new approaches to identify small ORFs, the number of putative small proteins discovered in pathogenic bacteria is increasing and the current knowledge on functional hydrophobic sORFs provides insights into prediction of additional ORFs of interest. In functional studies, deletion of sORFs can have mild or unclear phenotypes, whereas overexpression of sORFs can provide better insights into biological relevance (keeping in mind that overproduction of such proteins can also drive non-specific effects on membrane lipids or proteins). Functionally characterized hydrophobic small proteins contribute to the adaptation of bacteria to changing environment and stress conditions, and are often implicated in negative feedback regulation loops through direct interaction with diverse membrane proteins (Figure 1). Negative regulation of virulence factors, such as MgtC, supports that a tiny regulation is required for optimal pathogenesis and opens the way for developing antivirulence strategies (Figure 2).

The increasing characterization of small proteins involved in the regulation of bacterial virulence, antibiotic resistance or bacterial toxicity provides new perspectives for therapeutic strategies against bacterial pathogens (Figure 1), as shown recently with the use of synthetic
peptides derived from the natural small hydrophobic proteins PepA1, MgtR and KdpF. Derivatives of toxic peptides, as PepA1, induce membrane permeability, leading to bacterial death. On the other hand, synthetic peptides derived from regulatory small proteins MgtR and KdpF act by lowering bacterial virulence (Figure 2). The therapeutic potential of synthetic transmembrane domains is gaining interest as they can disrupt the folding of membrane proteins or compete for protein-protein interactions in membranes (Stone and Deber 2017). Chemical modifications, such as cyclization or use of non-natural amino acids, are required to improve membrane insertion and stability of synthetic hydrophobic peptides (Stone and Deber 2017). Tests of such chemically modified peptides in suitable animal models, as performed for cyclic pseudopeptides derived from PepA1 (Nicolas, Bordeau et al. 2019), will then be performed to validate their efficiency. Strategies to reduce bacterial growth or virulence may also be derived from other known membrane proteins such as TisB or MgrB, which have been identified, but not yet studied, in pathogenic bacteria. Discovery of additional small membrane proteins in pathogenic bacteria will further increase the number of novel antibacterial or antivirulence peptidic molecules available for therapeutic uses.

Acknowledgments

We thank anonymous referees for their comments on a previous version of the manuscript, as well as Kenneth Brockman and Brandon Wills (Medical College of Wisconsin) for manuscript editing. Research in ABBP laboratory is supported by CNRS, Vaincre La Mucoviscidose and Association Grégory Lemarchal.

References


<table>
<thead>
<tr>
<th>Regulatory small proteins</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MgtR</strong> (S. Typhimurium) (30 aa)</td>
<td>MNRPDKIIALIFLLISLLVLCLALWQIVF</td>
</tr>
<tr>
<td><strong>MgtS</strong> (S. Typhimurium) (31 aa)</td>
<td>MLGSINLFIIVLGIILFGFLAAWFSHKWD</td>
</tr>
<tr>
<td><strong>PmrR</strong> (S. Typhimurium) (29 aa)</td>
<td>MKSYIYKSLTLCSVLIVSSFIYVWVTY</td>
</tr>
<tr>
<td><strong>MgrB</strong> (S. Typhimurium) (47 aa)</td>
<td>MKKKFRWVVGLGIVVVCLLWQAQVFNIMCDQDVQFFSGICAINKFIPW</td>
</tr>
<tr>
<td><strong>KdpF</strong> (S. Typhimurium) (29 aa)</td>
<td>MSAGVITGIVLFLLLGYLVYALINAEEF</td>
</tr>
<tr>
<td><strong>KdpF</strong> (M. tuberculosis) (30 aa)</td>
<td>MTTVDNIVGLVIAVALMAFLFAALLEPEKF</td>
</tr>
<tr>
<td><strong>CydX</strong> (S. Typhimurium) (37 aa)</td>
<td>MWYFAWILGTALLCAFGIITALALEHVEAGKTGQEES</td>
</tr>
<tr>
<td><strong>Prli42</strong> (L. monocytogenes) (31 aa)</td>
<td>MTNKKVVRVVLMLIAIVLSSVLTGVLML</td>
</tr>
<tr>
<td><strong>AcrZ</strong> (S. Typhimurium) (49 aa)</td>
<td>MLELLKSLVFAVIMVPVVMALIGLVLGEVFNIFSGIGQKDQSRQQHK</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Toxic hydrophobic small proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TisB</strong> (S. Typhimurium) (29 aa)</td>
</tr>
<tr>
<td><strong>PepA1</strong> (S. aureus) (30 aa)</td>
</tr>
<tr>
<td><strong>PepA2</strong> (S. aureus) (35 aa)</td>
</tr>
<tr>
<td><strong>PepG1</strong> (S. aureus) (44 aa)</td>
</tr>
<tr>
<td><strong>PepG2</strong> (S. aureus) (31 aa)</td>
</tr>
</tbody>
</table>

For each small protein, the number of amino-acids (aa) is indicated in parenthesis and the predicted TM domain based on TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) and TMPred (https://embnet.vital-it.ch/software/TMPRED_form.html) servers is underlined.
Aromatic residues in the C-terminus are in green. Positively charged residues in the C-terminus are in red.
Table 2. Characterized small membrane proteins regulating larger membrane proteins in response to cation levels

<table>
<thead>
<tr>
<th>Small membrane protein</th>
<th>Regulation by TCS</th>
<th>Target</th>
<th>Function</th>
<th>Therapeutic potential upon exogenous addition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MgtR</td>
<td>PhoQP</td>
<td>MgtC, MgtA</td>
<td>Destabilization of MgtC</td>
<td>Antivirulence action (reduced intramacrophage survival)</td>
</tr>
<tr>
<td>MgtS</td>
<td>PhoQP</td>
<td>MgtA, PitA</td>
<td>Stabilization of MgtA Modulation of PitA activity</td>
<td>Unknown</td>
</tr>
<tr>
<td>PmrR</td>
<td>PmrBA</td>
<td>LpxT</td>
<td>Inhibition of LpxT activity</td>
<td>Antivirulence potential (LPS modification)</td>
</tr>
<tr>
<td>MgrB</td>
<td>PhoQP</td>
<td>PhoQ</td>
<td>Reduction of PhoQP activity</td>
<td>Antivirulence potential (lower PhoP phosphorylation)</td>
</tr>
<tr>
<td>KdpF</td>
<td>KdpDE</td>
<td>KdpABC</td>
<td>Stabilization of KdpABC</td>
<td>Antivirulence action (increased susceptibility to nitrosative stress)</td>
</tr>
</tbody>
</table>
Figure 1. Small membrane proteins in pathogenic bacteria: from function as regulators or toxins to therapeutic potential. Regulatory small membrane proteins are generally expressed in response to particular environmental cues in order to regulate specific functions by interacting with their target protein within the membrane. Such interactions can either result in stabilization or destabilization of the target protein to help the pathogen regulate its virulence and adapt to certain environmental stress. Toxins are part of toxin-antitoxin modules, which cause cell toxicity by protein accumulation and pore formation in the membrane, and also contribute in the formation of persister cells. Therapeutic interventions against bacterial pathogens are developed using synthetic peptides derived from natural small membrane proteins known to destabilize virulence factors or display antibacterial action.

Figure 2. Model of action of the regulatory small protein MgtR when endogenously expressed or exogenously added in the context of therapeutic intervention. MgtR is regulated by the PhoQP TCS, which is activated under magnesium limitation (left panel). A direct interaction between MgtR and the MgtC virulence factor negatively regulates MgtC (left panel), which provides the basis for an antivirulence strategy. Addition to pathogenic bacteria of synthetic MgtR peptide has an antivirulence action by reducing the amount of MgtC protein (right panel) and possibly other membrane proteins (shown in blue).