From Department of Microbiology, Tumor and Cell Biology
Karolinska Institutet, Stockholm, Sweden

Regulation of biofilm formation in
Salmonella typhimurium and
Escherichia coli NISSLE 1917

Cláudia Monteiro

Stockholm 2011
Cover illustration: Cells of *Salmonella enterica* serovar Typhimurium producing the extracellular matrix component cellulose.

All previously published papers were reproduced with permission from the publishers

Published by Karolinska University Press

Printed by Larseric Digital Print AB

© Cláudia Monteiro, 2011
ISBN 978-91-7457-184-4
ABSTRACT

Bacteria have the ability to grow in cell communities designated biofilms. This mode of growth is widespread and offers numerous advantages to the bacteria in terms of survival, persistence and propagation. Bacteria have developed different ways of building up a biofilm. Complex regulatory mechanisms control this sophisticated mode of growth in response to environmental conditions.

This thesis focuses on the regulation of biofilm formation by the food-borne pathogen *Salmonella enterica* serovar Typhimurium and the probiotic strain *Escherichia coli* Nissle 1917.

Commonly, species of the family of Enterobacteriaceae produce the biofilm extracellular matrix components cellulose and curli fimbriae at low temperature. The expression of cellulose and curli is activated by the transcriptional regulator CsgD. In this work, we demonstrated an altered pattern of biofilm regulation in *E. coli* Nissle 1917 (Paper I). Biogenesis of curli fimbriae was activated by CsgD at low temperature, while cellulose production at 28°C and 37°C did not require CsgD nor the di-guanylate cyclase AdrA. Cellulose production was, however, still dependent on the second messenger c-di-GMP. This regulatory pattern of cellulose and curli fimbriae production has been conserved in *E. coli* Nissle 1917 clonal isolates for more than 80 years implying biological significance. Production of cellulose mediated adhesion of *E. coli* Nissle 1917 to the gastrointestinal epithelial cell line HT-29 and to the mouse epithelium in vivo, thus possibly playing a role in colonization of the gut.

A characteristic of biofilm formation is cell heterogeneity. In *S. typhimurium*, expression of the master regulator CsgD was bistable during biofilm development (Paper II). Bistability led to task distribution, whereby the subpopulation of cells, which expressed high amounts of CsgD, was associated with microcolony formation and the production of cellulose.

CsgD expression is tightly regulated and responds to a variety of environmental conditions such as nutrient starvation and oxygen tension. Several global regulators contribute directly or indirectly to CsgD regulation. In this work, we identified novel factors involved in the complex CsgD regulation. Two lytic transglycosylases, MltE and MltC redundantly activated CsgD and rdar morphotype expression (Manuscript III). The absence of these two lytic transglycosylases could be partially compensated by the second messenger c-di-GMP. The chaperone Hfq and two Hfq dependent sRNAs, ArcZ and RyeB, also activated rdar morphotype expression by controlling the
expression of CsgD (Manuscript IV). We demonstrated that ArcZ is a key regulator of biofilm formation. In addition, ArcZ played a role in the transition between sessility and motility and was involved in the timing of type I versus curli fimbriae surface attachment.
LIST OF PUBLICATIONS

This thesis is based on the following publications, which are referred to in the text by the Roman numerals:


Publications by the author, which are not included in the thesis:
# TABLE OF CONTENTS

1 Introduction .......................................................................................................................... 1  
   1.1 Microorganisms ............................................................................................................. 1  
       1.1.1 *Salmonella typhimurium* ................................................................................... 1  
       1.1.2 *Escherichia coli* Nissle 1917 ........................................................................... 2  
   1.2 Biofilms .......................................................................................................................... 4  
       1.2.1 Detrimental and beneficial biofilms ..................................................................... 4  
       1.2.2 Biofilm formation ................................................................................................. 5  
       1.2.3 The biofilm matrix ................................................................................................. 5  
       1.2.4 The rdar morphotype ........................................................................................... 7  
           1.2.4.1 Regulation of rdar morphotype expression ................................................. 8  
       1.2.5 Biofilm on abiotic surfaces .................................................................................... 13  
       1.2.6 Biofilm in the intestine ........................................................................................ 15  
   1.3 Heterogeneity within biofilms ....................................................................................... 17  
       1.3.1 The phenomenon of bistability ............................................................................ 18  
   1.4 Lytic transglycosylases ................................................................................................. 19  
       1.4.1 Functions of lytic transglycosylases ................................................................... 19  
   1.5 Small RNAs .................................................................................................................. 21  
       1.5.1 The RNA chaperone Hfq ..................................................................................... 21  
       1.5.2 Small RNAs in *S. typhimurium* ....................................................................... 22  
       1.5.3 The sRNA ArcZ in *S. typhimurium* ................................................................. 23  
2 Aims ...................................................................................................................................... 24  
3 Methodology ....................................................................................................................... 25  
   3.1 Molecular biology methods ......................................................................................... 25  
   3.2 Approaches used in the study of biofilms ...................................................................... 26  
   3.3 Gene expression studies ............................................................................................... 27  
   3.4 Infection biology models ............................................................................................. 28  
4 Results and discussion ........................................................................................................ 29  
   4.1 Paper I ........................................................................................................................... 29  
   4.2 Paper II .......................................................................................................................... 32  
   4.3 Manuscript III .............................................................................................................. 35  
   4.4 Manuscript IV .............................................................................................................. 38  
5 Conclusions and future perspective .................................................................................. 41  
6 Acknowledgements ............................................................................................................ 43  
7 References ........................................................................................................................... 45
# LIST OF ABBREVIATION

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. typhimurium</td>
<td><em>Salmonella enterica</em> serovar Typhimurium</td>
</tr>
<tr>
<td>S. enterica</td>
<td><em>Salmonella enterica</em></td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em> Nissle 1917</td>
</tr>
<tr>
<td>E. coli Nissle 1917</td>
<td><em>Escherichia coli</em> Nissle 1917</td>
</tr>
<tr>
<td>P. aeruginosa</td>
<td><em>Pseudomonas aeruginosa</em></td>
</tr>
<tr>
<td>IL-8</td>
<td>Interleukine-8</td>
</tr>
<tr>
<td>rdar</td>
<td>red, dry and rough</td>
</tr>
<tr>
<td>pdar</td>
<td>pink, dry and rough</td>
</tr>
<tr>
<td>bdar</td>
<td>brown, dry and rough</td>
</tr>
<tr>
<td>saw</td>
<td>smooth and white</td>
</tr>
<tr>
<td>csg</td>
<td>Curli subunit gene</td>
</tr>
<tr>
<td>bcs</td>
<td>Bacterial cellulose synthase</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
<tr>
<td>IL-6</td>
<td>Interleukine-6</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Tumor necrosis factor-alpha</td>
</tr>
<tr>
<td>AdrA</td>
<td>AgfD regulated protein</td>
</tr>
<tr>
<td>c-di-GMP</td>
<td>Cyclic diguanosine monophosphate</td>
</tr>
<tr>
<td>sRNAs</td>
<td>Small RNAs</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence Activated Cell Sorting</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative real time PCR</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

1.1 Microorganisms

1.1.1 Salmonella typhimurium

*Salmonella enterica* serovar Typhimurium (*S. typhimurium*) is a gram negative, facultative anaerobe, facultative intracellular pathogen. It belongs to genus *Salmonella*, species *Salmonella enterica* (*S. enterica*). The species *S. enterica* is further subdivided into serovars according to their antigenic specificity. For clinical purposes *S. enterica* serovars are divided in two groups. The first group causes enteric fever in humans and includes serovars *Salmonella typhi* and *Salmonella paratyphi* A, B and C. The second group is referred to as non-typhoidal *Salmonella* and causes gastroenteritis in humans. It includes the serovars *S. typhimurium* and *S. enteritidis*. In humans, *S. typhimurium* infection is usually acquired via food derived from infected livestock, eggs or contaminated plants (Berger *et al*., 2010). The infection causes gastroenteritis, which lasts 3-5 days (Ina *et al*., 2003) or a more severe enterocolitis that can last 2-3 weeks. In healthy individuals the infection is usually self-limiting but cases of bacteremia and localized systemic infections are observed in immuno-compromised patients (Hohmann, 2001, Mastroeni, 2003). The mechanisms of pathology are associated to the adhesion and invasion of the intestinal epithelium by *S. typhimurium* and consequent immune response by the host, which is mainly characterized by a massive neutrophil recruitment (Harris *et al*., 1972, McGovern & Slavutin, 1979, Wick, 2004). The main reservoirs of *Salmonella* are intestinal tracts of live stock and wild animals. *Salmonella* is released into the environment by the feces and can survive outside the host for extended periods of time (Winfield & Groisman, 2003). Contamination of drinking and irrigation water occurs, which can subsequently lead to contamination of vegetables. Vegetable contamination can also occur directly through feces. In addition, improper handling of food products and lack of maintenance of industrial facilities are also cause of *Salmonella* outbreaks (Rhen, 2007).

*S. typhimurium* and other *Salmonella* serovars were some of the first bacterial organisms to be sequenced (McClelland *et al*., 2001). A variety of tools is available to genetically manipulate these microorganisms (Datsenko & Wanner, 2000, Hsiao & Zhu, 2009). *S. typhimurium* serves as a model organism to study typhoid fever in mice and more recently the streptomycin pretreated mice became a mouse model for *S. typhimurium* colitis (Hapfelmeier & Hardt, 2005). These infection models have the
advantage that both host and bacteria can be genetically manipulated, which allows the study of specific signaling pathways of *S. typhimurium*-host interaction. In conclusion, *S. typhimurium* is a very attractive model organism to study.

1.1.2 *Escherichia coli* Nissle 1917

*Escherichia coli* (*E. coli*) is a gram-negative, facultative anaerobic bacterium, which belongs to genus *Escherichia*. This species contains strains that can be classified as pathogens, commensals and even probiotics, according to their effects on humans. Probiotic microorganisms are defined as “live organisms which when administrated in adequate amounts confer a health benefit on the host” (Guarner, 2006). *Escherichia coli* Nissle 1917 (*E. coli* Nissle 1917) was isolated during the World War I from a soldier who, in contrast to all his comrades, was not affected by a gastroenteritis outbreak (Nissle, 1919). Probiotics can confer protective effects in different ways, some possible explanations include: suppression of growth and colonization of pathogenic microorganisms, contribution to strengthening the epithelial barrier and immunomodulation (Lebeer et al., 2010).

*E. coli* Nissle 1917 is sold as a medicine and many studies have proven the efficiency of the treatment in gastrointestinal diseases (Schultz, 2008, Henker et al., 2008). *E. coli* Nissle 1917 is a relatively safe organism for humans, since it lacks most virulence factors (Grozdanov et al., 2004, Gronbach et al., 2010) and is a good colonizer of the intestine (Barth et al., 2009). *E. coli* Nissle 1917 possesses genetic information which mediates fitness properties such as iron uptake systems and microcins, which give a competitive advantage against other enterobacteria (Grosse et al., 2006, Valdebenito et al., 2006, Patzer et al., 2003). *E. coli* Nissle 1917 has been demonstrated to inhibit the growth of pathogenic bacteria and fungi in vitro, like *Salmonella enteritidis*, *Shigella dysenteriae*, *E. coli*, *Proteus vulgaris* and *Candida albicans* (Sonnenborn & Greinwald, 1990). *E. coli* Nissle 1917 is able to prevent invasion of epithelial cells by *Salmonella* and adherent-invasive *E. coli* strains isolated from patients with Crohn’s disease (Altenhoefer et al., 2004, Boudeau et al., 2003).

*E. coli* Nissle 1917 has anti-inflammatory and immunostimulatory properties. It induces IL-8 response and the expression of other chemokines in host cells (Lammers et al., 2002, Ukena et al., 2005). *E. coli* Nissle 1917 displays a distinct immune response through the induction of antimicrobial peptides. Thereby, expression of flagella of *E. coli* Nissle 1917 induces the production of the antimicrobial peptide β-defensin 2 in gastrointestinal epithelial cells (Wehkamp et al., 2004, Schlee et al.,...
2007). *E. coli* Nissle 1917 inhibits the cell cycle and cell multiplication of peripheral T-cells, which can be interpreted as an anti-inflammatory effect (Sturm *et al.*, 2005). Recently, *E. coli* Nissle 1917 was also shown to activate dendritic cells and their cytokine synthesis and proved to have a protective effect against dust allergy, effects that were in part mediated through Toll-like receptor 4 (Adam *et al.*, 2010). The antagonistic effects of *E. coli* Nissle 1917 have also been demonstrated in vivo, as *E. coli* Nissle 1917 protected gnotobiotic piglets against the virulent strain *S. typhimurium* (Trebichavsky *et al.*, 2010).
1.2 Biofilms

In a diversity of environments bacteria form biofilms, communities of cells embedded into a self produced extracellular matrix adherent to each other and/or to surfaces and interfaces (Costerton et al., 1995). This mode of growth, recently described as the most successful form of life on earth (Flemming & Wingender, 2010), is the most prevalent mode of growth of microorganisms including bacteria.

1.2.1 Detrimental and beneficial biofilms

Biofilm formation offers advantages for the bacteria, promoting survival and propagation in many different environments. From the human perspective, biofilm formation is usually seen as harmful, as up to 80% of human bacterial infections are caused by biofilm forming bacteria (Costerton et al., 1999). However, in many cases biofilm formation is beneficial to humans.

Bacteria grown as biofilms have increased resistance to antimicrobial agents, e.g. antibiotics, phagocytic cells and other factors of the immune system (Mah & O'Toole, 2001). Chronic infections such as recurrent urinary tract infections caused by E. coli (Anderson et al., 2004), cystic fibrosis lung infection caused by Pseudomonas aeruginosa (P. aeruginosa) (Bjarnsholt et al., 2009) and infections associated with indwelling medical devices such as prosthesis and catheters (Lynch & Abbanat, 2010) are diseases typically caused by biofilm forming bacteria. On artificial devices, more than one species can usually cause the infection. Prosthetic valve endocarditis is associated with infections by species such as Staphylococcus, Streptococcus spp. and gram-negative bacteria, while urinary-tract catheters can be colonized predominantly by P. aeruginosa, Enterococcus faecalis, E. coli and Proteus mirabilis (Lynch & Abbanat, 2010).

Biofilm formation is also responsible for problems in industry, especially when materials are in contact with water. In drinking water reservoirs and distribution pipes pathogenic microorganisms like Legionella pneumophila and E. coli can be a threat to human health (Flemming, 2002, Juhna et al., 2007). Biofilms are also a major concern in the development and spread of antibiotic resistance, due to elevated mutation rates and higher exchange of genetic elements (Anderson & O'Toole, 2008, Hoiby et al., 2010). Biofilms can, however, also be beneficial. Biofilms of commensal bacteria in the human gastrointestinal tract are important for health maintenance, as alterations of the gut microbiota can lead to disease (Sekirov et al., 2010). The application of
microbial biofilms in industrial processes such as bioremediation (Singh et al., 2006), fermentation (Kunduru & Pomotto, 1996), biofuel production (Wang & Chen, 2009) and even generation of electricity in microbial fuel cells (Rabaey et al., 2007) is increasing.

1.2.2 Biofilm formation

Biofilm formation has been studied and described mainly in P. aeruginosa biofilms, where five stages of development have been identified (Sauer et al., 2002, O'Toole & Kolter, 1998, Watnick & Kolter, 2000). These sequential five stages are: initial attachment, irreversible attachment, microcolony formation, maturation and finally dispersion. In a first phase, bacteria attach loosely to a surface and may even migrate on the surface. Once the loose attachment becomes irreversible, cell division leads to formation of microcolonies. At this stage, the bacteria produce extracellular matrix components that help them to attach to the surface and to each other. Subsequently, a three dimensional structure is formed, which comprises a mature biofilm with a complex architecture containing live and dead cells plus a substantial amount of extracellular material. Within the mature biofilm, channels are formed, through which water, nutrients and other products can circulate. From a mature biofilm, cells or blocks of cells can actively or passively disperse into the environment, ensuring the propagation of the bacteria. To the dispersion of the cells contribute secreted enzymes that degrade the biofilm extracellular matrix components, as well as the stimulation of regulatory networks that allow a sessile bacteria to become motile again (Gjermansen et al., 2006, Wang et al., 2007). Consequently, biofilms are not static, as the dispersal stage ensures a continuous construction and renovation of the mature biofilm and seeding of new biofilms (Watnick & Kolter, 2000).

1.2.3 The biofilm matrix

It is estimated that microorganisms constitute less than 10% of the dry mass in a biofilm, while 90% is extracellular material (Flemming & Wingender, 2010). Extracellular matrix can include polysaccharides, proteins, nucleic acids and lipids. The composition of the matrix depends on the microorganism and other variants such as environmental conditions and the surface. The three dimensional structure created by the microorganisms and the extracellular matrix provides mechanical support and allows the creation of a special environment where other components, such as enzymes
and nutrients are maintained. The extracellular matrix components can themselves serve as nutrient source when degraded by enzymes. The matrix provides an important advantage for the microorganisms, as it protects from desiccation, oxidization, biocides, antibiotics, metallic cations and ultraviolet radiation (Branda et al., 2005, Flemming & Wingender, 2010).

Table 1. Biofilm matrix components in *Salmonella enterica* and *Escherichia coli*

<table>
<thead>
<tr>
<th>Biofilm component</th>
<th>Species</th>
<th>Relevance for biofilm formation in <em>Salmonella</em> and <em>E. coli</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Extracellular Polysaccharides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cellulose</td>
<td><em>S. typhimurium</em>&lt;br&gt;<em>E. coli</em>&lt;br&gt;<em>G. xylinus</em>&lt;br&gt;<em>K. pneumonia</em>&lt;br&gt;<em>E. coli</em>&lt;br&gt;<em>Bordeletella</em> spp.&lt;br&gt;<em>Actinobacillus pleuropneumonia</em>&lt;br&gt;<em>E. coli</em>&lt;br&gt;<em>Salmonella</em> spp.&lt;br&gt;<em>Shigella</em> spp.&lt;br&gt;<em>Enterobacter</em> spp.</td>
<td>Cell-cell interactions</td>
</tr>
<tr>
<td>PGA polysaccharide</td>
<td></td>
<td>Adhesion, microcolony formation</td>
</tr>
<tr>
<td>Colanic acid</td>
<td><em>S. typhimurium</em>&lt;br&gt;<em>Salmonella</em> spp.&lt;br&gt;<em>Shigella</em> spp.&lt;br&gt;<em>Enterobacter</em> spp.</td>
<td>Microcolony formation, biofilm maturation</td>
</tr>
<tr>
<td>Capsule</td>
<td><em>S. typhimurium</em>&lt;br&gt;<em>S. enteritidis</em></td>
<td>Biofilm formation, protection against dessication</td>
</tr>
<tr>
<td>Proteinaceous components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curli fimbriae</td>
<td><em>S. typhimurium</em>&lt;br&gt;<em>E. coli</em>&lt;br&gt;<em>Citrobacter</em> spp.&lt;br&gt;<em>E. sakazakii</em>&lt;br&gt;<em>Salmonella</em> spp.&lt;br&gt;<em>K. pneumonia</em>&lt;br&gt;<em>E. agglomerans</em>&lt;br&gt;<em>E. coli</em></td>
<td>Adhesion, microcolony formation, biofilm maturation</td>
</tr>
<tr>
<td>Type 1 fimbriae</td>
<td></td>
<td>Adhesion</td>
</tr>
<tr>
<td>Mat fimbriae</td>
<td></td>
<td>Early stages of biofilm development</td>
</tr>
<tr>
<td>BapA</td>
<td><em>S. typhimurium</em></td>
<td>Biofilm formation</td>
</tr>
<tr>
<td>Antigen 43</td>
<td><em>E. coli</em></td>
<td>Microcolony formation</td>
</tr>
<tr>
<td>Conjugative pili</td>
<td><em>E. coli</em></td>
<td>Biofilm maturation</td>
</tr>
</tbody>
</table>

This table was made based on the following publications: (Van Houdt & Michiels, 2005, Romling, 2002, Zogaj et al., 2003, Rehm, 2010, Kaplan et al., 2004, Parise et al., 2007, Gibson et al., 2006, Hahtela et al., 1985, Latasa et al., 2005, Lehti et al., 2010)

Extracellular matrix components perform diverse tasks when building up a biofilm, consequently having distinct biological significance.

Cellulose is produced by a variety of pathogens and environmental bacteria as indicated in table 1. In soil bacteria, such as in the family of *Rhizobiacea* cellulose is required for firm adherence and aggregation of bacteria to plant roots, which leads to a survival advantage for the bacteria (Smit et al., 1992, Romling, 2002). Several species of the family of *Enterobacteriacea* produce cellulose. Cellulose production prevented curli mediated invasion of commensal *E. coli* strains and *S. typhimurium* into gastrointestinal
epithelial cells (Wang et al., 2006, Lamprokostopoulou et al., 2010). In Paper I of this thesis we show that cellulose production is required for adhesion of the probiotic E. coli Nissle 1917 to gastrointestinal epithelial cells and to the intestinal lining. In parallel, others showed that curli and cellulose act in concert during cell adherence and biofilm formation in enterohaemorrhagic and enteropathogenic E. coli (Saldana et al., 2009). Enteric bacteria such as E. coli and Salmonella produce the extracellular matrix component curli fimbriae, which has several distinct roles in pathogenesis. Curli fimbriae have been implicated in attachment and invasion of host cells. E. coli strains show higher adherence to epithelial cells when curli are expressed (Wang et al., 2006, Kikuchi et al., 2005). Moreover, expression of curli fimbriae mediates invasion of E. coli and Salmonella strains into host cells (Gophna et al., 2001, Wang et al., 2006, Lamprokostopoulou et al., 2010). Curli fimbriae were also implicated in activation of the immune system and interaction with host proteins. Curli fimbriae are recognized by TLR2/TLR1 (Tukel et al., 2010) and consequently activate the immune system resulting in induction of IL-8, IL-6 and TNF-α (Bian et al., 2000, Lamprokostopoulou et al., 2010, Wang et al., 2006). Interaction of curli fimbriae with host proteins such as plasminogen was proposed to facilitate bacterial dissemination through the host (Bian et al., 2000, Sjobring et al., 1994).

1.2.4 The rdar morphotype

In the laboratory, biofilm formation can be studied using various model systems. Formation of rugose (also called rough or wrinkled) colonies on agar plates containing absorptive dyes, pellicle formation at the air-liquid interface of a standing culture, attachment to surfaces in steady-state culture and biofilm formation under continuous flow conditions in a flow-cell used in combination with confocal laser scanning microscopy are all used as biofilm models (Branda et al., 2005).

Species of the family of Enterobacteriaceae, such as S. typhimurium and E. coli produce the biofilm extracellular components cellulose and curli fimbriae (Bokranz et al., 2005, Romling, 2005, Zogaj et al., 2001, Zogaj et al., 2003). The expression of these two components leads to a characteristic morphotype on agar plates containing the diazo Congo Red dye. This morphotype was designated rdar due to the red, dry and rough aspect of the colonies. Congo Red has characteristic spectrophotometric properties. The absorbance spectrum of Congo Red differs upon binding cellulose or curli, allowing the detection of the two components when present alone or in combination. Therefore,
strains producing cellulose appear pink dry and rough (pdar), while production of curli fimbriae results in brown dry and rough (bdar) colonies. When no matrix components are expressed the phenotype is smooth and white (saw). The different Congo Red morphotypes constitute a valuable tool to study regulatory networks underlying the expression of the two extracellular matrix components cellulose and curli fimbriae (Romling, 2005).

Figure 1. Colony morphology prototypes of S. typhimurium UMR1 and isogenic mutants. From the left to the right are examples of rdar, pdar, bdar and saw morphotypes. UMR1 expresses a highly regulated rdar morphotype at ambient temperatures.

1.2.4.1 Regulation of rdar morphotype expression

CsgD, a master regulator of rdar morphotype expression

CsgD (curli subunit gene D), previously called AgfD (aggregative fimbriae D), is a master regulator of rdar morphotype expression. CsgD is a response regulator with a DNA-binding output domain. According to its C-terminal DNA-binding helix-turn-helix domain, CsgD is a member of the LuxR superfamily. CsgD belongs to the FixJ family, based on the homology of the N-terminal receiver domain. FixJ family members are response regulators which are part of two component systems (Romling, 2005, Goulian, 2010). Usually response regulators of two component systems are activated by phosphorylation of a conserved aspartate in the receiver domain by a cognate sensor kinase. However, no sensor kinase has been identified for CsgD. Interestingly, recent data suggests that the response regulator CsgD is active in its unphosphorylated form under biofilm formation conditions (Zakikhany et al., 2010). CsgD can be phosphorylated by acetyl phosphate in vitro, however, whether and when
CsgD is phosphorylated in vivo is not clear. In vivo studies with CsgD mutants mimicking phosphorylation showed no biofilm activation (Zakikhany et al., 2010). CsgD is a component of the divergently transcribed csgDEFG-csgBAC operon, which is required for the biosynthesis of curli fimbriae. The csgBAC operon encodes the structural subunits for curli fimbriae formation; CsgA is the major subunit and CsgB is a nucleator protein facilitating CsgD polymerization (Hammar et al., 1995, Hammar et al., 1996). CsgC is hardly expressed, but CsgC might be involved in facilitating assembly of the extracellular CsgA (Collinson et al., 1996, Hammar et al., 1995, Gibson et al., 2007). The csgDEFG operon encodes three accessory proteins required for curli production, secretion and assembly of CsgA and CsgB (Epstein et al., 2009, Robinson et al., 2006, Chapman et al., 2002).

CsgD also activates cellulose biosynthesis. The two divergently transcribed operons bcsABZC-bcsEFG are responsible for cellulose biosynthesis (Wong et al., 1990, Solano et al., 2002). BcsA and BcsB form the cellulose synthase complex where BcsA is the cellulose synthase (Saxena et al., 1991, Zogaj et al., 2001, Romling, 2002). In Gluconacetobacter xylinus, the synthase complex is located in the cytoplasmic membrane, implying a transmembrane complex over the cytoplasmic and outer membrane to secrete the cellulose (Kimura et al., 2001). Cellulose-synthesizing multienzyme complexes organized in a row, can be observed by electron microscopy (Kimura et al., 2001).

CsgD positively regulates transcription of the csgBAC operon and indirectly activates cellulose biosynthesis. Thereby, CsgD directly binds to the promoter and activates transcription of the divergently transcribed csgBAC operon (Romling et al., 2000, Hammar et al., 1995, Zakikhany et al., 2010). In addition, CsgD also binds to the promoter of the adrA gene (Zakikhany et al., 2010). AdrA (AgfD regulated protein A) is a diguanylate cyclase responsible for the production of the second messenger c-di-GMP. C-di-GMP presumably binds to the PilZ domain of BcsA, the cellulose synthase and consequently activates the synthesis of cellulose (Romling et al., 1998b, Romling et al., 2000). However, also CsgD-independent cellulose expression occurs in S. typhimurium (Garcia et al., 2004, Solano et al., 2002, Da Re & Ghigo, 2006).

In S. enteritidis, CsgD also activates the expression of the O-antigen capsule and BapA, a large secreted surface protein. The yih operons encode genes required for O-antigen capsule assembly, translocation and regulation. CsgD downregulates a repressor of the yih assembly operon. Moreover, the expression of the four biofilm components, curli
fimbriae, cellulose, O-antigen capsule and BapA is coregulated in *S. enteritidis* (Gibson *et al.*, 2006, Latasa *et al.*, 2005). CsgD is often referred to as the master regulator of biofilm formation. The divergently transcribed operons *csgDEFG-csgBAC* are found in several enterobacterial species such as *Salmonella* spp., *E. coli*, *Shigella* spp., *Enterobacter* spp., *Citrobacter* spp. and *Raoultella ornithinolytica*, demonstrating the importance of these fimbrial operons and the regulator CsgD (Zogaj *et al.*, 2003). In addition, there are a few other microorganisms that harbor CsgD homologues such as *Vibrio* species, which actually contain more than one copy of the CsgD paralogues per genome. *Vibrio vulnificus* and *Vibrio parahaemolyticus* contain three copies while *Vibrio cholerae* has two copies. VpsT, a CsgD homologue in *Vibrio cholerae* is involved in biofilm formation and motility by directly sensing c-di-GMP (Casper-Lindley & Yildiz, 2004, Krasteva *et al.*, 2010).

CsgD expression is highly regulated. The complex network responsible for CsgD regulation has been mainly studied in *S. typhimurium* and *E. coli*. CsgD expression is highly affected by environmental conditions, whereby temperature, pH, osmolarity, oxygen tension and nutrient availability play a crucial role (Gerstel & Romling, 2001, Romling *et al.*, 1998b, Prigent-Combaret *et al.*, 2001). In the model organisms *S. typhimurium* UMR1 and *E. coli* K-12, maximal expression of CsgD occurs at stationary phase when nutrients are scarce, with a 370-fold increase in transcription in stationary phase compared to logarithmic phase (Gerstel & Romling, 2001). Fecal *E. coli* isolates show a great variability in production of the extracellular matrix components curli fimbriae and cellulose (Bian *et al.*, 2000, Bokranz *et al.*, 2005). In contrast, almost all *S. typhimurium* strains express the rdar morphotype at 28°C, but not at 37°C (Gerstel & Romling, 2001).

Several global regulators have been identified which regulate CsgD expression in response to environmental conditions. The stationary phase sigma factor RpoS is responsible for transcription initiation (Robbe-Saule *et al.*, 2006, Gualdi *et al.*, 2007, Ogasawara *et al.*, 2010a). In addition, more than 10 regulatory factors were identified to directly control the *csgD* promoter positively or negatively (Ogasawara *et al.*, 2010b, Gerstel & Romling, 2003, Ishihama, 2010). For example, the response regulator OmpR is essential for *csgD* transcription in *S. typhimurium* and *E. coli* (Romling *et al.*, 1998a, Gerstel & Romling, 2001, Prigent-Combaret *et al.*, 2001). On the other hand, RstA activates *csgD* transcription under acidic conditions (Ogasawara *et al.*, 2007). In contrast, CpxR, the response regulator of the CpxR/CpxA two-component system,
represses $csgD$ transcription in response to high osmolarity (Prigent-Combaret et al., 2001). There is cooperation between the positive factors OmpR-IHF and RstA-IHF and the negative factors CpxR-H-NS (Ogasawara et al., 2010a).

CsgD regulation occurs also at the post-transcription level by two redundant small RNAs, OmrA and OmrB. Downregulation of CsgD expression by OmrA and OmrB occurs within the 147 bp long $csgD$ 5’-UTR far upstream of the ribosome binding site by a direct antisense interaction (Holmqvist et al., 2010).

Figure 2. Factors involved in regulation of $csgD$, the master regulator gene of biofilm formation. Transcription factors that directly regulate $csgD$ are indicated by white circles, nucleoid proteins directly affecting $csgD$ transcription are indicated by grey circles.

**The stress sigma factor RpoS**

The stationary phase sigma factor RpoS controls entrance into the stationary phase promoting stress resistance (Navarro Llorens et al., 2010). Regulation of RpoS is complex and occurs at various levels: transcription, translation, stability/degradation and activity (Navarro Llorens et al., 2010).

RpoS was shown to be required for transcriptional activation of the stationary phase inducible $csgD$ promoter (Romling et al., 1998a). Thereby, RpoS directly binds to the $csgD$ promoter to activate transcription (Robbe-Saule et al., 2006).

Transcription of CsgD and development of rdar morphotype expression by RpoS requires the Crl protein (Robbe-Saule et al., 2006). The effect of Crl is most pronounced at low levels of RpoS. When RpoS is expressed in higher amounts from a
plasmid, the rdar morphotype is restored in a crl deletion mutant. Crl interacts directly with RpoS in vitro and facilitates the formation of the RNA polymerase-RpoS complex, the holoenzyme $\sigma^S$, thus enhancing transcription of the csgD and adrA promoters (Robbe-Saule et al., 2006, Bougdour et al., 2004).

On the other hand, CsgD expression is also indirectly regulated by RpoS. RpoS activates expression of the transcriptional regulator MlrA, the expression of which is required for csgD transcription (Brown et al., 2001, Romling et al., 1998a, Ogasawara et al., 2010b).

RpoS and CsgD expression are also interconnected. Not only does RpoS regulate CsgD, but also CsgD positively controls RpoS, through activation of iraP, which encodes a factor involved in RpoS stabilization. Thus RpoS and CsgD are involved in a positive feedback loop, which might have implications in coordination of biofilm formation with other biological responses regulated by RpoS (Gualdi et al., 2007).

**Figure 3.** Representative diagram of the CsgD/RpoS autoactivation loop. See the text for details.

**The secondary messenger c-di-GMP**

The second messenger c-di-GMP is another positive regulator of CsgD expression. The discovery of c-di-GMP dates back to 1987, when c-di-GMP was identified as an allostERIC activator of the cellulose synthase complex in *Gluconacetobacter xylinum* (Ross et al., 1987). The c-di-GMP molecule was only a few years ago recognized as a general secondary messenger in bacteria (Jenal, 2004). Many phenotypes are affected by c-di-GMP, among others biofilm formation (Romling, 2005), motility (Simm et al.,
2004, Wolfe & Visick, 2008), the production of extracellular polysaccharides (Simm et al., 2005), multicellular behaviour (Romling, 2005) and bacterial-host interactions (Lamprokostopoulou et al., 2010). C-di-GMP has also been involved in virulence of several pathogens including S. typhimurium (Tamayo et al., 2008, Hisert et al., 2005, Lamprokostopoulou et al., 2010, Tamayo et al., 2007). C-di-GMP positively regulates expression of extracellular matrix components and consequently biofilm formation. On the other hand, c-di-GMP negatively regulates motility (Tamayo et al., 2007). C-di-GMP is synthesized by diguanylate cyclases (GGDEF domain containing proteins), while it is degraded by EAL and HD-GYP domain proteins that function as phosphodiesterases (Christen et al., 2005, Ryan et al., 2006, Simm et al., 2004, Tamayo et al., 2005). The second messenger c-di-GMP is almost ubiquitous among bacteria (Romling et al., 2005), but the numbers of diguanylate cyclases and phosphodiesterases can be highly variable among bacterial species (Romling et al., 2005). S. typhimurium and E. coli K-12 harbour twenty and thirty one GGDEF/EAL proteins, respectively. The high number of proteins involved in c-di-GMP metabolism suggests some redundancy in function, although the presence of different sensory domains in the N-terminus indicates that these proteins might respond to different signals (Romling, 2005). Experimental evidence demonstrated the existence of task distribution (Kader et al., 2006) with the existence of different pools of c-di-GMP in the cell. In fact, asymmetric distribution of c-di-GMP in the cell was recently reported (Christen et al., 2010). In S. typhimurium, the c-di-GMP pool which controls CsgD expression is influenced by at least six proteins: two GGDEF-EAL domain proteins, STM2123 and STM3388, which function as diguanylate cyclases, and three EAL domains proteins, STM3611, STM1827 and STM4264 plus one GGDEF-EAL domain protein, STM1703, which function as phosphodiesterases (Kader et al., 2006, Simm et al., 2007). In E. coli K-12, the pair YdaM/YciR, a GGDEF and a GGDEF-EAL protein, respectively, is dedicated to csgD transcription control, while YegE and YedQ, two GGDEF domain proteins and YhjH, an EAL domain protein, regulate flagellar activity and positively modulate CsgD transcription (Pesavento et al., 2008).

1.2.5 Biofilm on abiotic surfaces

Microorganisms have the ability to interact with abiotic surfaces. As referred to in chapter 1.2.1, biofilms on abiotic surfaces can be beneficial or can have deleterious effects. For Salmonella the most well known example is the contamination of
equipment used in food-processing (Cliver, 2006). Several components required for
dar morphotype expression such as cellulose, curli fimbriae and the large surface
protein BapA are also involved in adherence to glass and polystyrene surfaces.
Adherence in this case is regulated by CsgD, which is involved in activation of
expression of these extracellular matrix components (Latasa et al., 2005, Romling et al.,
In general, attachment of microorganisms to a surface depends on a variety of
structures on the cell surface and outer membrane proteins (Goulter et al., 2009).
In E. coli flagella-mediated swimming motility is essential for biofilm formation as it
allows the bacteria to overcome repulsive forces and to spread along the surface (Pratt
& Kolter, 1998, Genevaux et al., 1996, Wood et al., 2006). Type 1 fimbriae and curli
fimbriae contribute to irreversible attachment to surfaces, both in Salmonella and E.
coli (Austin et al., 1998, Pratt & Kolter, 1998, Cookson et al., 2002, Orndorff et al.,
2004, Vidal et al., 1998, Uhlich et al., 2006). Outer membrane proteins were shown to
cover approximately 50% of the E. coli cell surface (Lin et al., 2002). The importance
of outer membrane proteins such as OmpX, BtuB and OmpW for adhesion to abiotic
surfaces has been demonstrated (Otto & Hermansson, 2004, Rivas et al.,
2008). Altered growth conditions modulate the attachment to surfaces as, for example,
temperature and salt can affect the expression of extracellular matrix components curli
fimbriae and cellulose (Romling, 2005).
Physicochemical factors of the bacteria, the nature of the surface and culture conditions
affect attachment (Goulter et al., 2009).
Bacterial surface charge is one factor that determines the interaction with the surface.
The majority of bacterial cells are negatively charged due to the presence of carboxyl
and phosphate groups on the outer cell membrane, nevertheless variability in surface
charge is observed between species and strains (Ukuku & Fett, 2002, Rivas et al.,
2005). However, some studies demonstrated that charge influences surface attachment,
while in other studies this association was not found (Ukuku & Fett, 2002, Rivas et al.,
2007).
Hydrophobicity also seems to positively contribute to surface attachment, although data
do not demonstrate a clear association between hydrophobicity and the attachment
ability. This lack of correlation is probably due to differences in the experimental
procedures, surfaces and bacterial isolates used in the different studies (Liu et al., 2004,
Rivas et al., 2007, Rivas et al., 2005, Goulter et al., 2009).
In addition to the bacterial properties, the surface properties also have to be considered in order to fully understand the complex adhesion process to abiotic surfaces. Surfaces can be hydrophilic, such as stainless steel and glass, or hydrophobic, such as Teflon and polystyrene. In addition, surface roughness has been shown to be an important property affecting bacterial attachment (Oliveira et al., 2006).

In conclusion, there is an enormous complexity of bacteria-abiotic surface interactions. Correlation of the physicochemical properties of the bacterial cells and surfaces with attachment requires further characterization.

1.2.6 Biofilm in the intestine

The intestinal microflora is mainly developed during infancy, when the number and diversity of microorganisms increases until up to one year. However, the microbiota composition evolves continuously during life. Bacteroides and Firmicutes dominate the adult intestinal lumen (Eckburg et al., 2005). It is estimated that 40,000 bacterial species inhabit the human intestine (Frank & Pace, 2008) and recent studies demonstrate that the majority of the microorganisms are uncultivated or unknown species (Eckburg et al., 2005). The microbial composition varies in different parts of the intestine (Frank et al., 2007) and bacteria can be found as individual cells, microcolonies or in large communities (Macfarlane & Dillon, 2007).

The intestinal epithelial layer is separated from the lumen by the mucus layer; within these three distinct spaces there is a heterogeneous distribution of the microflora. The epithelial surface and mucous layer is mainly colonized by Clostridium, Lactobacillus and Enterococcus, while the intestinal lumen contains a bigger variety of bacteria such as Bacteroides, Bifidobacteria, Streptococcus and Enterobacteriacea (Swidsinski et al., 2005). It was proposed that the intestinal microbiota forms a biofilm-like community. The retention of the microorganisms in a polysaccharide rich mucous layer benefits the host by promoting beneficial functions of the microbiota such as digestion and strengthening of the host defenses, which keeps pathogens away from the epithelial cells (Sonnenburg et al., 2004). In fact, all Bifidobacterium and Lactobacilli species isolated from human gut biopsies, tightly bound to the epithelium, were able to form biofilm (Fakhry et al., 2009). Medical treatment and change of diet can lead to disruption of the bacterial community of the intestine. Once unprotected, the intestinal lining can be colonized by pathogenic organisms. In a recent study, it was demonstrated that adherent-invasive E. coli colonized a high number of patients suffering from
Crohn’s disease (Rolhion & Darfeuille-Michaud, 2007). Microflora protects against epithelial injury (Rakoff-Nahoum et al., 2004), regulates fat storage (Caesar et al., 2010), stimulates angiogenesis (Stappenbeck et al., 2002) and protects from inflammatory intestinal diseases. These findings lead to a greater interest in probiotic microorganisms and their possible use as therapies for intestinal disease or diseases related to disturbance of the intestinal microbiota.

*S. typhimurium* is also found in the mucus layer in the gut, while only a fraction of the bacteria invades the epithelium (Santos et al., 2009). This suggests that *S. typhimurium* resides in a biofilm stage in the gut. Indirect evidence for CsgD expression in the intestine exist, since csgD expression was actually detected in the small intestine of mice (White et al., 2008). In a cell culture model, invasion of the gastrointestinal epithelial cell line HT-29 by *S. typhimurium* at high c-di-GMP levels is mainly inhibited by CsgD and cellulose (Lamprokostopoulou et al., 2010).

The extracellular matrix component curli fimbriae was shown to play a role in the pathogenesis of *E. coli* and *Salmonella*. Curli expression is involved in eliciting IL-8 expression in macrophages during sepsis in *E. coli* (Bian et al., 2000) and deletion of csgBA in *S. typhimurium* leads to reduced intestinal inflammation in a calf model (Tukel et al., 2005). Although csgBA expression was not detected in the mouse gut in one study (White et al., 2008), other studies identified CsgA as a pathogen-associated molecular pattern (PAMP) that signals through Toll-like receptors 1 and 2 (TLR1/TLR2), which interact in order to recognize curli fibrils (Tukel et al., 2005, Tukel et al., 2010).
1.3 Heterogeneity within biofilms

In nature, biofilms are usually not formed by one species, but exist mainly as complex consortia of microorganisms with a variety of different species. The dental plaque, probably one of the most complex biofilms, contains hundreds of species (ten Cate, 2006). Little is known about the regulation and the role of each species in these kind of biofilms.

Heterogeneity in biofilms is, however, not uniquely related to different species. Single species biofilms are also heterogeneous as different cell types perform defined and specialized functions (Costerton et al., 1995). Some authors even like to compare biofilms to multicellular organisms, where bacteria exhibit cooperative and unselfish behavior (Shapiro, 1998). The existence of subpopulations of cells is of great advantage to the bacteria increasing its chances of survival under various growth and stress conditions (Stoodley et al., 2002). Subpopulations of cells can exhibit different phenotypes and even genotypes. Several mechanisms that contribute to phenotypic heterogeneity have been identified: microscale chemical gradients, adaptation to local environment conditions, stochastic gene expression and finally genotypic variation that occurs through mutation and selection (Stewart & Franklin, 2008). Growth in high density communities leads to the creation of unique niches due to nutrient gradients, waste products and signaling compounds. Examples of microscale chemical gradients are oxygen, nutrients and waste products. In *S. epidermidis* biofilms, the existence of spatial patterns of DNA replication and protein activity due to different oxygen and nutrient concentrations was demonstrated (Rani et al., 2007). Moreover expression of the master regulator CsgD is highly sensitive to oxygen tension (Gerstel & Romling, 2001).

The local environments can also affect the physiology of the cells, whereby the contact with a surface is maybe the most prominent change compared to the swimming organism. The transition between planktonic cells and adherent cells can be observed when attachment to a surface occurs. In *E. coli* the expression of genes regulated by the Cpx two-component system was induced during initial adhesion to abiotic surfaces. This response required NlpE, an outer membrane lipoprotein (Otto & Silhavy, 2002).

Genetic variation through mutation is associated with biofilm heterogeneity. In *P. aeruginosa* biofilms, during chronic lung infection in cystic fibrosis patients, extensive genetic alterations occur. These mutations play a role in *P. aeruginosa* genome evolution and adaptation during chronic respiratory infection (Mena et al., 2008).
Stochastic variation in cell physiology happens in *E. coli* during exponential growth, where two subgroups of cells exist, the motile cells and the non-growing cells. The non-growing cells are able to survive antibiotic treatment since antibiotics only kill growing cells. This subpopulation is designated the persisters (Balaban *et al.*, 2004).

### 1.3.1 The phenomenon of bistability

Conventional thinking considered genetically identical cells growing under the same conditions identical with respect to pattern of gene expression and protein synthesis. However, expression analysis of fluorescent reporter proteins on the single cell level demonstrated that bacteria can actually exist in distinct physiological states (Veening *et al.*, 2005, Grantcharova *et al.*, 2010). Bistability originates from unimodal noise in the expression of a master regulatory gene that passed a threshold, thus giving rise to an elevated gene expression. Two mechanisms were proposed to explain this phenomenon: first, positive autoregulation that responds in a non-linear way; second, the existence of a pair of mutually repressing repressors (Dubnau & Losick, 2006, Graumann, 2006, Veening *et al.*, 2008). Bistability in biofilm formation was described for *Bacillus subtilis*, where production of the extracellular matrix components exopolysaccharide and the protein TasA occurs only in a subpopulation of cells. SinI, an antirepressor of the SinR repressor which controls the operons responsible for the expression of the extracellular matrix components is exclusively expressed in this subpopulation (Chai *et al.*, 2008). In this thesis (Paper II) another example of bistability in biofilms is described. We found that in *S. typhimurium* biofilms, the expression of the master regulator of biofilm formation, CsgD, is bistable. Only a subpopulation of cells expresses CsgD and is engaged in the production of the extracellular matrix component cellulose (Grantcharova *et al.*, 2010).
1.4 Lytic transglycosylases

Lytic transglycosylases are enzymes ubiquitous in bacteria, which are involved in the degradation of the bacterial cell wall peptidoglycan. Lytic transglycosylases catalyse the cleavage of the β-1,4-glycosidic bond between N-acetylmuramic acid and N-acetylglucosamine with the concomitant formation of anhydromuropeptide monomers (Holtje et al., 1975). Together with other peptidoglycan hydrolases such as amidases, lytic transglycosylases are responsible for the coordinated regulation of cell wall degradation events. Four different families of lytic transglycosylases can be distinguished based on consensus motifs, however this classification does not imply specific functions (Blackburn & Clarke, 2001). The apparent redundancy in lytic transglycosylases in bacterial genomes makes the assignment of functions complex.

E. coli harbours seven lytic transglycosylases, the soluble Slt70 and six outer membrane anchored lipoproteins, MltA-F (Scheurwater & Clarke, 2008). Although Slt is the principal lytic transglycosylase responsible for the turnover of peptidoglycan (Park & Uehara, 2008), an effect on cell morphology and cleavage of the peptidoglycan septum is only observed when six lytic transglycosylases are deleted (Heidrich et al., 2002).

1.4.1 Functions of lytic transglycosylases

The main function of lytic transglycosylases is the biosynthesis and recycling of the peptidoglycan sacculus constituting the bacterial cell wall, during cell growth and division (Scheurwater & Clarke, 2008, Holtje, 1998). In E. coli, half of the cell wall is broken down in one generation (Goodell, 1985). The subsequently produced anhydromuropeptide monomers are reused directly or transported into the cytoplasm through the permease AmpG (Lindquist et al., 1993), where they can induce β-lactamase (Jacobs et al., 1994). Together with other peptidoglycan degrading enzymes, lytic transglycosylases separate the daughter cells through cleavage of the septum during cell division (Heidrich et al., 2002).

Another important function of lytic transglycosylases is the insertion of macromolecular cell wall spanning complexes (Koraimann, 2003). Space in the peptidoglycan wall is created by lytic transglycosylases allowing the establishment of macromolecular secretion systems. As an example, biogenesis of the type III secretion system in Pseudomonas syringae and the formation of flagella and pili in Caulobacter
*crescentus* has been shown to require lytic transglycosylases (Oh *et al.*, 2007, Viollier & Shapiro, 2003).

In addition to its basic functions, lytic transglycosylases contribute to bacterial-host interaction. Anhydromuropeptide monomers are recognized by eukaryotic pattern recognition receptors (Girardin *et al.*, 2003, Viala *et al.*, 2004). Bacteria as diverse as *Bordetella pertussis, Neisseria gonorrhoeae* and *Vibrio fischeri* release significant amounts of anhydromuropeptides which are required for virulence and symbiosis (Melly *et al.*, 1984, Koropatnick *et al.*, 2004, Goldman *et al.*, 1982).
1.5 Small RNAs

Research in the field of bacterial small RNAs (sRNAs) has been increasing during the last years, while *S. typhimurium* became a model organism for the study of sRNA mediated regulation (Vogel, 2009). At least 70 sRNAs have been identified in *Salmonella*. These sRNAs are typically 50-250 nucleotides in length and the majority works by an antisense mechanism. The sRNA interacts with the messenger RNA (mRNA) at or close to the translation start sites thus affecting mRNA stability and protein translation (Vogel, 2009). sRNAs that regulate gene expression by base-pairing with mRNA are divided into two classes, the *cis* and *trans* encoded base pairing sRNAs. *Cis* sRNAs are encoded by the antisense strand of their target mRNA, therefore having an extensive potential to base-pair with their target. *Trans* sRNAs are encoded at a different genomic location than their target, thus having a limited complementarity. In many cases *trans* sRNA-mRNA interaction requires Hfq, an RNA chaperone (Waters & Storz, 2009).

1.5.1 The RNA chaperone Hfq

In *E. coli*, Hfq is a very abundant protein. It is estimated that 50,000-60,000 copies exist per cell, corresponding to approximately 10,000 hexamers (Brennan & Link, 2007). Hfq has pleiotropic function, regulating approximately 20% of genes in the *S. typhimurium* genome (Ansong *et al.*, 2009). *Trans* sRNAs require Hfq in order to stabilize and facilitate the binding with the target mRNA. Hfq binds to A/U rich regions in single-stranded sRNAs and mRNAs (Brennan & Link, 2007). In *S. typhimurium*, deletion of *hfq* affects a number of phenotypes and leads to alterations in gene and protein expression. As an example, Hfq affects invasion of epithelial cells, survival in macrophages and motility (Sittka *et al.*, 2007). The expression of the stress sigma factor RpoS, which has been shown to be essential for *Salmonella* virulence in mice (Fang *et al.*, 1992), is severely impaired in the *hfq* deletion mutant (Brown & Elliott, 1996). In the absence of the Hfq chaperone, outer membrane proteins accumulate, which leads to sigma E mediated envelope stress response (Bang *et al.*, 2005, Figueroa-Bossi *et al.*, 2006, Sittka *et al.*, 2007, Bossi *et al.*, 2008).
1.5.2 Small RNAs in *S. typhimurium*

sRNAs adjust bacterial physiology and behavior in response to changing environmental conditions, regulating the expression of a wide variety of targets such as outer membrane proteins, two-component systems and sigma factors (Figueroa-Bossi et al., 2009, Mandin & Gottesman, 2009, Thompson et al., 2007). Regulation of outer membrane protein genes by sRNAs, has been extensively studied in *S. typhimurium* and involves the sRNAs, RybB, InvR, MicA and CyaR. RybB has a pleiotropic function, repressing the synthesis of all major porins (OmpA/C/D/F) and minor outer membrane proteins (Udekwu et al., 2005, Figueroa-Bossi et al., 2006, Bossi & Figueroa-Bossi, 2007, Sittka et al., 2007, Pfeiffer et al., 2007, Papenfort et al., 2006). Other sRNAs such as InvR, CsrB and CsrC are involved in mediating the cross talk between the core genome and the virulence regions thereby affecting the *Salmonella* pathogenicity island (SPI) related invasion phenotypes (Pfeiffer et al., 2007, Fortune et al., 2006). In addition to the effect on invasion, CsrB and CsrC have been implicated in regulation of motility and biofilm formation through the regulation of c-di-GMP metabolizing proteins (Jonas et al., 2010). One sRNA such as RybB is able to control several genes, thereby coordinating responses to certain environmental signals. On the other hand, distinct sRNAs can respond to different environmental signals resulting in one integrated response. In *E. coli*, RpoS is positively regulated by the sRNAs ArcZ, RprA, DsrA and negatively regulated by the sRNA OxyS. ArcZ, RprA and DsrA were shown to directly base pair with the transcript encoding RpoS, sequestering the ribosome binding site (Majdalani et al., 1998, Majdalani et al., 2002, Soper et al., 2010). Evidence suggests that RpoS integrates distinct stress signals through regulation by these sRNAs since they are expressed in response to different environmental stress conditions including stationary phase growth, low temperature, envelope stress and oxidative stress (Altuvia et al., 1997, Majdalani et al., 1998, Majdalani et al., 2002, Papenfort et al., 2009, Beisel & Storz, 2010). RpoS is highly conserved between *E. coli* and *Salmonella*, however, differences in regulation by sRNAs occur. Although DsrA and RprA contribute to RpoS regulation in *Salmonella*, RpoS expression is only slightly affected by these two sRNAs when compared to *E. coli*. Direct binding of DsrA and RprA to the *rpoS* mRNA was not shown until now (Jones et al., 2006). ArcZ was reported to affect RpoS expression in *Salmonella* but a direct interaction was not demonstrated (Papenfort et al., 2009). Moreover, there is experimental evidence that, even between *Salmonella* isolates, sRNA expression and consequently function can
differ (Vogel, 2009). These evidences demonstrate the complexity of post-transcriptional regulation by sRNAs and the flexibility of the regulatory systems that evolve in different organisms.

1.5.3 The sRNA ArcZ in S. typhimurium

ArcZ, previously named SraH and RyhA, was first discovered in E. coli. ArcZ is a conserved sRNA among Enterobacteriaceae (Argaman et al., 2001, Wassarman et al., 2001) and is abundantly present as a 50 nt species derived from a 120 nt processed precursor transcript (Papenfort et al., 2009, Argaman et al., 2001). Deep sequencing analysis revealed that ArcZ constituted 10% of the Hfq-bound Salmonella sRNAs (Sittka et al., 2008). Three targets of the sRNA ArcZ were identified in S. typhimurium: sdaC, involved in serine uptake; tpx, involved in oxidative stress and STM3216, a chemotaxis protein. ArcZ represses its target mRNAs through a mechanism of post-transcriptional regulation, more specifically, sdaCB and STM3216 mRNAs are regulated by sequestration of the ribosome binding site while tpx mRNA is targeted in the coding sequence (Papenfort et al., 2009).
2 AIMS

Bacterial species belonging to the family of Enterobacteriaceae display the multicellular rdar morphotype, characterized by the production of the two extracellular matrix components cellulose and curli fimbriae. The master regulator CsgD commonly regulates this multicellular behavior. The global aim of this thesis was to analyse the regulation of biofilm formation in two strains, which display rdar morphotype expression, *S. typhimurium* and *E. coli* Nissle 1917.

Specific aims were:

- To elucidate the mechanisms of regulation of the temperature independent rdar morphotype expression displayed by the probiotic strain *E. coli* Nissle 1917. In this context to characterize cellulose production and the biological role of cellulose in *E. coli* Nissle 1917. *(Paper I)*

- Analyze expression of the master regulator CsgD at the single cell level in different biofilm models. *(Paper II)*

- Identify novel components of the regulatory network leading to expression of the master regulator CsgD and study their mechanisms of action. *(Manuscript III and IV)*
3 Methodology

The most relevant methods used in the papers and manuscripts of this thesis are described in this chapter. Detailed information about protocols is provided in the Material and Methods section of each paper.

3.1 Molecular biology methods

Construction of deletion mutants

Construction of chromosomal deletion mutants in *S. typhimurium* and *E. coli* Nissle 1917 was done to investigate the function of a gene and the corresponding protein. The deletion mutants were constructed using the Datsenko and Wanner method (Datsenko & Wanner, 2000). This procedure is based on the expression of the Red system from phage λ and allows effective homologous recombination between DNA stretches as short as 40nt using a linear DNA fragment. In this way, a gene can be replaced by an antibiotic resistance gene using a PCR product. In total, three genes constitute the Red system γ, β and exo, coding for Gam, Bet and Exo respectively. Gam inhibits the host RecBCD exonuclease V, so that the electroporated PCR product is not cleaved. The recombination is promoted by Bet and Exo.

In a first step, a PCR product is generated using primers with 40nt extensions homologous to the first and last 40nt of the gene of interest. The template plasmid pKD3 Cm® or pKD4 Km®, which carries the antibiotic resistance gene is flanked by FRT (FLP recognition target) sites. The generated PCR product is subsequently transformed by electroporation into the target strain containing the λ Red recombinase. Recombinant clones are selected for their antibiotic resistance and subsequently grown at 42°C to promote the loss of the temperature sensitive plasmid carrying the λ Red recombinase. The presence of the FRT sites flanking the antibiotic gene allow the removal of the resistance cassette when necessary.

Reporter fusions

Translational reporter fusions are usually used to study expression of a specific protein. Joining two or more genes, which separately encode for two distinct proteins, by removal of the stop and start codons, creates fusion proteins. Translation of the two merged genes originates a single polypeptide, which keeps the functional properties of the two individual proteins. In paper II the *csgD* gene without the stop codon was fused
in frame to the open reading frame of the Green Fluorescent Protein (GFP) without the start codon. This allowed visualization of the fusion protein in individual bacterial cells by fluorescence microscopy and quantification of expression by Fluorescence Activated Cell Sorting (FACS). In manuscript IV, sRNA-target interactions are studied using gene fusions to GFP. The use of a constitutive promoter fused to the leader sequence of a gene allows the identification of the sRNA that binds to the 5’UTR of the target gene and thus contribute to translational control. For that purpose a fragment containing the leader sequence plus start codon of the csgD gene was fused to GFP on a low copy vector. Combined with another plasmid with the sRNA cloned, the effects of a particular sRNA can be studied (Urban & Vogel, 2007).

3.2 Approaches used in the study of biofilms

Figure 4. Approaches used in this thesis to study biofilm formation. (a) High cell density colonies grown on agar containing the Congo Red dye. (b) Biofilm formed on polystyrene and (c) glass abiotic surfaces. Staining with the dye crystal violet allows visualization and quantification of the biofilm. (d) The pellicle, biofilm formation at the air-liquid interface. (e) Example of biofilm formation using the flow-cell model, as analyzed by confocal laser scanning microscopy.
3.3 Gene expression studies

Western-blot

Western Blot is a technique used to detect the expression of a certain protein of interest or an attached tag based on their ability to bind to specific antibodies. In a first step, the proteins mixture in the sample is separated by gel electrophoresis according to size and subsequently transferred to a membrane. For our work, after incubation with specific antibodies, chemiluminescent detection was used to visualize the protein. The polyclonal anti-CsgD peptide antibody was used to detect CsgD protein by Western Blot as described before (Romling et al., 2000).

Fluorescence-activated cell sorting (FACS)

FACS is a specialized type of flow cytometry. This method allows the separation of cells according to their physical and chemical characteristics from a heterogeneous mixture. It is based on light scattering and detection of fluorescence signals emitted from individual cells. This allows separation of individual cells according to size and expression of proteins or genes. The strength of the FACS technique is a fast, objective and quantitative analysis. In addition, cells with different features can be separated and recovered for further analysis. In paper II this technique was used in order to quantify the expression of the CsgD-GFP fusion protein in individual cells, which allows quantification of CsgD-GFP expression. Consequently, monomodal and bimodal CsgD expression was observed.

Quantitative real time RT-PCR

Quantitative real time polymerase chain reaction (qRT-PCR) is a technique used to simultaneously amplify and quantify a targeted DNA molecule (VanGuilder et al., 2008, Freeman et al., 1999). This method is based on the principles of PCR, with the addition that the amplified DNA is quantified in each amplification cycle as it accumulates. Two approaches are used to detect the DNA products: (1) fluorescent dyes which intercalate into double-stranded DNA, such as SYBR-Green, an approach that was used in manuscript III and IV of this thesis and (2) fluorescent modified DNA oligos such as taqman probes. In this thesis, qRT-PCR was used to estimate mRNA levels. For that purpose reverse transcription of the mRNA into cDNA was performed prior to qRT-PCR. The result was analysed in terms of relative quantification of gene
expression, which determines the change in expression of a specific target gene in a test sample in relation to the same gene in a calibrator sample (VanGuilder et al., 2008).

3.4 Infection biology models

In paper I, we investigated the impact of cellulose expression by *E. coli* Nissle 1917 in the interaction with the intestinal mucosa. For that purpose, two relevant infection models were chosen: (1) the gastrointestinal cell-line HT-29 for in vitro studies and (2) the ileal-loop model for in vivo studies in BALB/C mice.

**HT-29 cell line**

The HT-29 cell line is a human colon carcinoma cell-line, which offers a convenient and widely used system to study bacteria-epithelial cell interactions. This cell line was used to investigate adherence and invasion of bacteria into host cells. In addition, the production of the pro-inflammatory cytokine IL-8 stimulated upon co-incubation with bacteria was assessed.

**Ileal-Loop model in BALB/c mice**

BALB/c mice are widely used for infection studies and mouse colonization with *E. coli* Nissle 1917 has been conducted using this animal model (Adam et al., 2010, Arribas et al., 2009). The ileal-loop model offers the possibility to study, in vivo, the initial interactions between the bacteria and the intestinal mucosa (Jones et al., 1994). An ileal-loop in the intestine is made by creating a defined space in the small intestine through closing with a knot. This technique allows the study of direct interactions between bacteria and the epithelium. Using this method, the number of animals per study can be reduced, as several loops can be done in one mouse, allowing the study of more than one bacterial strain or mutants of the same strain in the same mouse.
4 Results and Discussion

4.1 Paper I

Characterization of cellulose production in *Escherichia coli* Nissle 1917 and its biological consequences

*E. coli* isolates show a great variability in the production of the extracellular matrix components curli fimbriae and cellulose (Bian et al., 2000, Bokranz et al., 2005, Da Re & Ghigo, 2006). In this study, we show that the probiotic *E. coli* strain Nissle 1917 is able to produce these two extracellular matrix components. However, while cellulose expression was observed at 28°C and 37°C, curli fimbriae were only produced at 28°C. Cellulose production by *E. coli* Nissle 1917 was investigated in more detail by transmission electron microscopy where extensive fibrillar structures were observed.

In *S. typhimurium* and in the commensal strain *E. coli* TOB1, the transcriptional regulator CsgD regulates the expression of the extracellular matrix components cellulose and curli fimbriae (Romling et al., 1998b, Bokranz et al., 2005). Previous studies showed that curli fimbriae biosynthesis is directly regulated by CsgD. Cellulose expression is indirectly regulated by CsgD via the diguanylate cyclase AdrA that produces c-di-GMP, the allosteric activator of the cellulose synthase (Romling et al., 2000, Simm et al., 2004). In contrast, our results in *E. coli* Nissle 1917 demonstrate that CsgD activates the production of curli fimbriae at 28°C, but biosynthesis of cellulose at 28°C and 37°C does not require CsgD. In addition, cellulose biosynthesis by *E. coli* Nissle 1917 does also not require the diguanylate cyclase AdrA. In conclusion, while curli fimbriae biosynthesis follows the regulatory pathway observed in many Salmonella and *E. coli* strains, cellulose biosynthesis exhibits an unconventional pattern of regulation.

*E. coli* Nissle 1917 was isolated during the First World War from humans (Nissle, 1919) and recently from swine herds in Germany (Nissle, 1919, Kleta et al., 2006). We demonstrated that the pattern of production of the extracellular matrix components curli fimbriae and cellulose is conserved in *E. coli* isolates of the Nissle clonal type after 90 years. Moreover, the unconventional regulatory pathway for cellulose biosynthesis seems to be retained in these isolates. Investigation of colony morphology of the deletion mutants *csgD, adrA* and *bcsA* genes of the recent isolated Nissle 1917 clone members also showed that the regulatory pattern of rdar morphotype expression is conserved as cellulose expression was independent of CsgD and AdrA.
In *S. typhimurium* and *E. coli*, activation of the cellulose synthase can occur independently of AdrA, which is transcriptionally activated by CsgD (Garcia *et al.*, 2004, Da Re & Ghigo, 2006). In these cases the diguanylate cyclase STM1987/YedQ was required. Therefore we investigated the role of this protein in cellulose biosynthesis in *E. coli* Nissle 1917. Surprisingly our findings demonstrate that in *E. coli* Nissle 1917 YedQ had evolved into a new protein that we named YedQ1. In YedQ1 a new domain with no obvious relation to c-di-GMP metabolism replaced the GGDEF C-terminal domain of YedQ.

Although AdrA and YedQ1 do not participate in cellulose biosynthesis, the second messenger c-di-GMP is required for cellulose biosynthesis in *E. coli* Nissle 1917. Expression of phosphodiesterases YhjH and STM1827 lead to downregulation of the rdar morphotype suggesting that cellulose biosynthesis in *E. coli* Nissle 1917 is dependent on c-di-GMP. Stimulation of the *E. coli* Nissle 1917 cellulose synthase by c-di-GMP was also verified by an in vitro assay for cellulose synthesis (Saxena & Brown, 1995). Since c-di-GMP is required for activation of the cellulose synthase by directly binding to the PilZ domain (Ryjenkov *et al.*, 2006), we speculate that one of the other GGDEF or GGDEF/EAL proteins might be responsible for c-di-GMP synthesis in Nissle 1917. Alternatively, perhaps due to a redundant effect of other diguanylate cyclase the effect of AdrA could be hidden.

Biofilm formation is variable in *E. coli*, however, the pattern of expression of cellulose in *E. coli* Nissle 1917 and its clonal isolates was conserved for more than 80 years. Consequently, we hypothesized that the constitutive expression of cellulose could provide an important advantage for *E. coli* Nissle 1917 persistence. This led us to further characterize the role of cellulose production in the interaction of *E. coli* Nissle 1917 with the host. Using the HT-29 cell line we confirmed previous observations that *E. coli* Nissle 1917 significantly adheres to epithelial cells (Lasaro *et al.*, 2009). We also demonstrated that a deletion mutant of the cellulose synthase dramatically decreased *E. coli* Nissle 1917 adherence. In vivo studies using the ileal loop model mice, confirmed the importance of cellulose in adherence to the intestinal epithelial lining. The requirement of cellulose in adhesion to HT-29 cells in *E. coli* Nissle 1917 is in contrast to the role of cellulose in the commensal strain *E. coli* TOB1 where cellulose inhibits adherence (Wang *et al.*, 2006). This shows that the role of cellulose in bacterial-host interactions is dependent on the strain background in *E. coli*. A probable cause of these differences is the coexpression with other extracellular matrix components. *E. coli* TOB1 coexpresses cellulose and curli fimbriae at 37°C in contrast
to strain *E. coli* Nissle 1917 which produces cellulose alone. Parallel work by other group demonstrated that synchronized production of curli and cellulose enhances adhesion to HT-29 cells in enterohaemorrhagic and enteropathogenic *E. coli* strains (Saldana *et al*., 2009). *E. coli* Nissle 1917 is a good colonizer of the intestine (Barth *et al*., 2009), a good biofilm former (Hancock *et al*., 2010) and expresses other types of fimbriae such as type 1 fimbriae and F1C fimbriae (Pere *et al*., 1985). F1C fimbriae were shown to be required for biofilm formation and adherence to epithelial cells (Lasaro *et al*., 2009). We can speculate that the efficiency of *E. coli* Nissle 1917 to colonize the intestine is associated to the expression of various extracellular matrix components. These properties could help Nissle 1917 to outcompete other pathogenic strains, which might be important for the probiotic effects of this strain.

The potential of *E. coli* Nissle 1917 to invade intestinal epithelial cells was described to be very low (Boudeau *et al*., 2003), a result that was confirmed in this study. Cellulose production inhibited invasion since, in the absence of cellulose, *E. coli* Nissle 1917 was significantly more invasive. A similar role of cellulose in invasion was found in the commensal strain *E. coli* TOB1 (Wang *et al*., 2006).

*E. coli* Nissle 1917 induces production of the pro-inflammatory cytokine IL-8 in epithelial cells (Lammers *et al*., 2002). IL-8 production was reduced in the cellulose deficient mutant, which might be a consequence of the reduced adherence in comparison with the wild type. Interestingly, in *E. coli* TOB1 the absence of cellulose lead to an increase of adherence and IL-8 response (Wang *et al*., 2006). This evidence strongly supports our assumption that reduced adherence is the cause of reduced IL-8 response by *E. coli* Nissle 1917.
4.2 Paper II

**Bistable Expression of CsgD in Biofilm Development of *Salmonella enterica* Serovar Typhimurium**

On the population level, CsgD expression is low in the logarithmic phase with a maximum of expression in stationary phase after 24h (Simm *et al.*, 2009). In this study, we investigated the pattern of expression of CsgD, the master regulator of biofilm formation in *S. typhimurium* at the single cell level. For that purpose a CsgD-GFP fusion protein was created at the native chromosomal locus of *csgD* under its natural promoter. The CsgD-GFP fusion protein was functional, as the characteristic rdar morphotype, indicative for the presence of the extracellular matrix components cellulose and curli fimbriae was produced. On the molecular level, expression of the CsgD-controlled genes *csgB* and *adrA* reached wild type levels in the CsgD-GFP fusion strain. The CsgD-GFP fusion protein was stable as detection of CsgD or GFP using anti-CsgD and anti-GFP antibodies did not reveal major degradation. The pattern of CsgD-GFP expression reached a maximum after 24h of growth (Simm *et al.*, 2009) similar as the CsgD expression in the wild type.

Fluorescence microscopy observations revealed that, when expressed, the CsgD-GFP fusion protein was homogeneously distributed in the cytoplasm of the cells. Only in a small percentage of cells CsgD-GFP was localized at the cell pole close to the membrane. CsgD expression was unequally distributed in the population. At early time points, most of the cells did not express CsgD, while a few cells showed high CsgD expression. A gradual increase in the numbers of cells expressing high levels of CsgD-GFP over time was observed, reaching a maximum after 24h of growth. This is in agreement with the results obtained to detect the fusion protein over time by Western blot analysis.

It was also observed that only a portion of the cell population was expressing CsgD-GFP, which therefore showed cell-to-cell variation.

CsgD has been shown to control multicellular behavior by regulating expression of the extracellular matrix components cellulose and curli fimbriae (Romling *et al.*, 1998b). We therefore investigated whether CsgD expression was associated with the formation of cell communities.

Using different biofilm models, such as colony growth on agar plates, biofilm formation in steady-state liquid culture and in flow-cell chambers, differential CsgD-GFP expression could be observed. Cellulose was mainly produced by aggregates of
cells, which expressed high levels of CsgD-GFP. To quantify the observed variability at the single cell level we analyzed CsgD-GFP expression by FACS analysis. Biofilm formation is associated with the formation of cell clumps due to the production of extracellular matrix components, therefore the use of wild type strains would not lead to accurate numbers (Romling et al., 2000). Consequently deletion mutants bcsA csgBA, were constructed in order to facilitate quantitative analysis of single cells. FACS analysis of UMR1 revealed differential expression of CsgD-GFP, with the formation of two distinct subpopulations of cells after 24h of growth, which showed a pattern of on/off CsgD-GFP expression. In conclusion, this part of the study demonstrates that expression of the master regulator CsgD is bimodal which leads to the establishment of phenotypic diversity during biofilm development.

Three fold enhanced CsgD-GFP expression obtained through a csgD promoter mutation, which leads to CsgD expression independent of the RpoS sigma factor (Romling et al., 1998a), abolished the bistable expression of CsgD-GFP and led to constitutive expression. This fact indicates that the bistable expression of CsgD requires wild type promoter expression levels of CsgD and possibly RpoS dependence and is generated at the level of transcriptional regulation. Two mechanisms were proposed to explain the phenomenon of bistability: positive feedback, and the existence of a pair of mutually repressing repressors (Dubnau & Losick, 2006, Graumann, 2006, Veening et al., 2008). The bistable expression of CsgD could be caused by a positive feedback loop. In E. coli a positive feedback loop for CsgD expression was described. CsgD positively regulates IraP, which is required for RpoS stabilization. RpoS in turn positively regulates the csgD promoter (Gualdi et al., 2007, Romling et al., 1998b). However, autoactivation of CsgD has not been observed in S. typhimurium, therefore an upstream origin for the bistable expression has to be considered, possibly at the level of RpoS or another regulator.

Elevated CsgD-GFP levels due to an increase in c-di-GMP concentrations in the cell also lead to monophasic expression of CsgD-GFP, which shows that c-di-GMP mediated upregulation of CsgD expression is able to overcome the bistability effect. The occurrence of bistability and consequent phenotypic variation during biofilm development might be an important advantage for the bacteria in terms of management of resources. Possibly only a subpopulation of cells needs to produce extracellular matrix components, which creates an environment from which all the cell community benefits. Indeed, it has been shown that cells can be part of the biofilm without producing the extracellular matrix components (Latasa et al., 2005). Additionally, the
bistable CsgD expression allows the existence of a subpopulation of cells that is not
dedicated to biofilm formation and might have other tasks in the biofilm community
such as motility. Recently, White and co-authors observed a global metabolic shift
associated with multicellular development in *S. typhimurium*. They speculated that
CsgD might be the signal for aggregation of *S. typhimurium* and compared the rdar
morphotype of *S. typhimurium* with the developmental process of sporulation in
*Bacillus spp* and *Myxococcus* (White *et al.*, 2010).
4.3 Manuscript III

Role of Lytic Transglycosylases in the Rdar Morphotype Expression in Salmonella enterica serovar Typhimurium

In this study we identified lytic transglycosylases associated with the development of multicellular behavior in S. typhimurium. Previous studies suggested the involvement of the lytic transglycosylase MltE in biofilm formation in E. coli (Niba et al., 2007). In addition the mltE and ycgR genes are arranged in tandem and transcribed in opposite directions. The ycgR gene codes for the c-di-GMP binding protein YcgR, which is involved in regulation of motility (Ryjenkov et al., 2006). Consequently, as the rdar morphotype is regulated by c-di-GMP (Romling, 2005) and genes adjacenty located on the chromosome are often functionally related, this motivated us to investigate the potential role of MltE in rdar morphotype expression.

Nine lytic transglycosylases were identified in S. typhimurium. Seven belong to family I, while MltA and MltB belong to family II and III respectively. In comparison to E. coli, S. typhimurium has two additional lytic transglycosylases of family I, (Scheurwater & Clarke, 2008).

Investigation of combinations of deletion mutants revealed that two lytic transglycosylases, MltE and MltC, are specifically involved in rdar morphotype expression. However, the effects are only observed when both lytic transglycosylases, MltE and MltC are deleted, demonstrating redundancy of function. MltE and MltC have different domain structures, yet, with 38% identity in their transglycosylase SLT domains, MltE and MltC show the highest identity between two transglycosylase SLT domains in S. typhimurium. Nevertheless, while MltE expressed from a plasmid fully complemented the mltE mltC double mutant, MltC expressed from the same plasmid could only partially complement the double mutant.

The low expression of MltC, in contrast to MltE expression from the plasmid, might explain partial complementation. Alternatively, the permease encoded by the nupG gene located downstream of mltC, in the same predicted operon, could be specifically required for MltC activity. In E. coli MltC was demonstrated to be part of an operon together with NupG (Gifford & Wallace, 1999). Alternatively, MltC may require the activity of MltE, which is endo-acting in contrast to all other investigated lytic transglycosylases, including MltC, which are exoenzymes (Kraft et al., 1998).

The double mutant mltE mltC showed a significant reduction of CsgD expression, approximately 30% of wild type UMR1. Determination of csgD mRNA levels
demonstrated that regulation of CsgD by MltE and MltC occurs at the transcriptional and translational level.

Using plasmids expressing MltE and MltC with mutations in the catalytic motif of the transglycosylase SLT domains, we demonstrated that the catalytic activity of MltE and MltC was in fact responsible for the regulatory effects of MltE and MltC on rdar morphotype expression. The mutated proteins were not able to complement the deletion mutants. In addition, a zymogram assay demonstrated the loss of hydrolytic activity of the mutated lytic transglycosylases. These results establish for the first time, a communication pathway between the cell wall turnover and the regulation of biofilm components.

In order to get insights in the regulation of CsgD by the lytic transglycosylases we tested if the stress sigma factor RpoS and the second messenger c-di-GMP signaling pathways were involved in the regulation. RpoS was previously shown to be required for CsgD expression (Romling et al., 1998a). However, investigation of the role of MltE and MltC in a strain with RpoS-independent CsgD expression, demonstrated that CsgD regulation by MltE and MltC occurs independently of RpoS. To test the influence of the c-di-GMP signaling in regulation of CsgD by the lytic transglycosylases MltE and MltC, the effects of elevated c-di-GMP levels in the double mutant mltE mltC were tested. The results demonstrate that c-di-GMP can partially compensate the effect of MltE and MltC on rdar morphotype and CsgD expression.

In E. coli, deletion of more than one transglycosylase altered cell morphology and permeability (Heidrich et al., 2002). Therefore we analyzed cell morphology and permeability as phenotypes potentially altered in the mltE mltC double mutant. Interestingly, the double mutant displayed formation of long chains of cells when grown at 28°C in media without salt. Transmission electron microscopy revealed that chain formation is due to the impairment of the cleavage of the murein septum. Alterations in cell envelope permeability were also observed in the mltE mltC double mutant compared to the wild type UMR1, as the mltE mltC double mutant does not grow on MacConkey agar plates. Indicating sensitivity to the bile acids present in these media. Nonetheless, our results also demonstrate that the membrane integrity is only partially lost, as the double mutant is still resistant to vancomycin.

Regulation of CsgD expression by the lytic transglycosylases MltE and MltC could occur through signaling from the periplasmic space to the cytoplasm. The anhydromuropeptide monomers, products of peptidoglycan degradation are transported to the cytoplasm by the permease AmpG (Jacobs et al., 1994), where they could serve
as a signal of cell wall turn over. However, deletion of AmpG and other transporters for peptidoglycan degradation products did not have any effect on rdar morphotype expression. Thus, cytosolic anhydromuropeptides or other peptidoglycan degradation products do not seem to be involved in the signaling. An additional possibility could be the sensing of extracytoplasmic stress, which has been shown to downregulate CsgD and rdar morphotype expression (Vianney et al., 2005) (Prigent-Combaret et al., 2001). The main function of lytic transglycosylases is the biosynthesis and recycling of the peptidoglycan sacculus in cell growth and division. The Rcs phosphorelay signaling, a key signaling pathway, senses perturbations in peptidoglycan turnover (Laubacher & Ades, 2008) (Callewaert et al., 2009) (Huang et al., 2006). Therefore signaling through the Rcs system could be involved in regulation of the rdar morphotype by lytic transglycosylases. Finally, since we observed an alteration in permeability of the outer membrane in the double mutant mltE mltC we hypothesize that altered membrane permeability might lead to a different milieu in the periplasmic space, which could alter signaling. Other conditions that lead to a loss of outer membrane integrity are the deletion of the peptidoglycan binding lipoproteins, some outer membrane proteins and the Tol system (Lloubes et al., 2001) (Cascales et al., 2002). Deletion of Tol system has been shown to downregulate curli-mediated biofilm formation (Vianney et al., 2005).

It is not the first time that an autolysin has been associated with biofilm formation. As referred above, the MltE was identified to promote biofilm formation in a screen of E. coli mutants (Niba et al., 2007). In gram-positive bacteria, interference with turnover of the peptidoglycan network alters surface adherence and biofilm formation (Heilmann et al., 1997) (Ahn & Burne, 2006) (Dubrac et al., 2007). Peptidoglycan turnover might be required for alteration of surface charge or exposure of adhesins, as the peptidoglycan network of gram-positive bacteria is in direct contact with the environment. In contrast, in gram-negative bacteria, the outer membrane covers the peptidoglycan layer and is tightly connected with it. Therefore, the mechanisms through which lytic transglycosylases affect biofilm formation in gram-negative and gram-positive bacteria might be different.
4.4 Manuscript IV

Hfq dependent small RNAs ArcZ and RyeB contribute to biofilm development

Hfq has been shown to regulate a variety of phenotypes (Sittka et al., 2007, Papenfort et al., 2008). In this study we found that the chaperone Hfq has a high impact on biofilm formation in S. typhimurium. When grown on Congo Red plates the hfq deletion mutant displays a saw morphotype, indicating that the extracellular matrix components cellulose and curli fimbriae are not expressed. In addition, Hfq was required for biofilm formation on hydrophobic and hydrophilic surfaces.

The biofilm phenotypes tested were shown to require expression of cellulose and curli fimbriae, which are regulated by CsgD, the master regulator of rdar morphotype expression (Romling, 2005). Using real-time RT-PCR to analyze csgD mRNA levels and Western blot to analyze CsgD protein levels, we demonstrated that Hfq regulates biofilm formation through the transcriptional activator CsgD.

The sigma factor RpoS plays a major role in CsgD and rdar morphotype expression (Romling et al., 1998a), while Hfq was shown to be required for RpoS expression (Sittka et al., 2008, Fang et al., 1992). We therefore investigated the role of RpoS in the Hfq dependent rdar morphotype expression. RpoS levels were dramatically decreased in the hfq mutant, but cross-complementation of the hfq mutant with RpoS expressed from a plasmid was not sufficient to restore CsgD and rdar morphotype expression despite RpoS expression. Besides RpoS, additional factors seem to be required to restore CsgD and rdar morphotype expression in the hfq mutant. Therefore we investigated the roles of sRNAs.

Trans sRNAs require Hfq to facilitate and stabilize the binding to the target mRNA (Brennan & Link, 2007). In a screening of a library of 38 sRNA deletion mutants, we identified two Hfq dependent sRNAs that affected rdar morphotype expression. Deletion of arcZ led to a highly significant downregulated rdar morphotype expression. Although csgD mRNA levels were not significantly altered in the arcZ mutant, CsgD protein levels were dramatically decreased, demonstrating that regulation of the rdar morphotype by ArcZ occurs at the post-transcriptional level. Moreover, RpoS levels were also decreased in the arcZ deletion mutant whereas ArcZ expressed from a plasmid was able to restore RpoS expression as well as CsgD and rdar morphotype expression. ArcZ expressed from a plasmid was also able to restore rdar morphotype
and CsgD expression in the *hfq* deletion mutant. Thus ArcZ demonstrates to be a key sRNA in regulation of expression of the rdar morphotype. Although RpoS levels are decreased in the *arcZ* deletion mutant and ArcZ has recently been demonstrated to positively contribute to RpoS translation (Soper et al., 2010), RpoS levels in the *arcZ* mutant were sufficient for expression of other RpoS dependent phenotypes, as the catalase activity which is RpoS dependent was not affected in the *arcZ* mutant. We therefore concluded that ArcZ regulates CsgD independently of RpoS. In addition, RpoS expressed from a plasmid in the *arcZ* mutant did not restore CsgD expression, indicating that RpoS alone is not sufficient to overcome the effect of ArcZ sRNA on CsgD and rdar morphotype expression. As *csgD* mRNA levels are hardly affected in the *arcZ* mutant, ArcZ might regulate CsgD expression post-transcriptionally or post-translationally. However, a direct regulation of CsgD expression by the ArcZ sRNA cannot be excluded at this time.

The second sRNA identified in the screening for rdar morphotype expression was RyeB. Deletion of *ryeB* led to a slightly downregulated rdar morphotype compared to UMR1, which demonstrates a minor influence when compared to the effect of the ArcZ sRNA. *csgD* mRNA levels as well as CsgD protein levels are similarly decreased in the *ryeB* mutant indicating that regulation occurs at the transcriptional level. Pointing to that RyeB indirectly affects CsgD transcription. Interestingly, RpoS levels are increased in the *ryeB* mutant, thus RyeB has an opposing regulatory effect on CsgD and RpoS.

ArcZ sRNA was also involved in motility regulation, as swimming motility was enhanced in the *arcZ* mutant. The effect of ArcZ on swimming motility has been reported before, where overexpression of ArcZ was shown to inhibit motility (Papenfort et al., 2009). As such ArcZ seems to regulate a switch between sessility and motility.

Studies of the biofilm phenotype on hydrophobic and hydrophilic surfaces revealed a peculiar phenotype for the *arcZ* mutant. While reduced pellicle formation was observed, adherence to the tube walls did not follow the expected pattern of downregulation of biofilm formation. We speculated that another biofilm component could play a role. Recently, it was demonstrated that *fimA*, which encodes the major subunit of type 1 fimbriae, is upregulated in the *arcZ* mutant (Papenfort et al., 2009). As type 1 fimbriae mediate surface adherence, we suspected that type 1 fimbriae might compensate for the lack of curli fimbriae in the adherence phenotype. Biofilm formation on a surface required type 1 fimbriae and curli fimbriae (Pratt & Kolter, 1998, Romling et al., 1998b). Using combinations of the *arcZ*, *fimA* and *csgBA*
mutants, we saw a time-dependent contribution of both types of fimbriae to surface-adherence. While both type 1 and curli fimbriae participated in the adherence process during the first 24hs, only curli fimbriae seem to be involved at later time points. In addition, in the arcZ deletion mutant adherence is mainly dependent on type 1 fimbriae during the first 24h and dependent on curli fimbriae at later time points. Consequently, ArcZ seems to regulate the timing of expression of type 1 fimbriae versus curli fimbriae.

**Figure 5.** Scheme of the regulatory network involving the sRNAs ArcZ and RyeB in expression of the extracellular matrix components cellulose and curli fimbriae through the master regulator CsgD. Full lines represent direct effects, discontinuous lines represent effects not proven to be direct.
5 Conclusions and future perspective

The work presented in this thesis contributes to better understand the regulation of biofilm formation in the probiotic bacterium *E. coli* Nissle 1917 and *S. typhimurium*. In paper I we found that production of the extracellular matrix components cellulose and curli fimbriae is conserved in Nissle 1917 clone members. In contrast to the majority of *E. coli* and *S. typhimurium* strains, in Nissle 1917 the master regulator CsgD activates the production of curli fimbriae, however CsgD is not required for cellulose biosynthesis. Neither the diguanylate cyclases AdrA nor YedQ were required for activation of the cellulose synthase, however, cellulose biosynthesis is still dependent on c-di-GMP. Cellulose production by the probiotic Nissle 1917 alters the interaction with the host in the ileal-loop mouse model. One might speculate that cellulose production possibly has a role in colonization and immunoregulation. This study reinforces the idea that the regulation of production of extracellular matrix components varies even within the same species. Nissle 1917 produces cellulose in a temperature independent way and mediates adherence to gastrointestinal cell and mouse mucosa. This indicates that cellulose might play an important role in colonization of different hosts. In contrast, curli fimbriae are produced in a temperature dependent way in *E. coli* Nissle 1917. Curli fimbriae are recognized by TLR2/TLR1 (Tukel et al., 2005, Tukel et al., 2010) and the fact that Nissle 1917 does not produce curli fimbriae at 37°C might be associated with the probiotic effect. As a future perspective, it would be important to understand the role of cellulose in long term colonization and biofilm formation in the gut.

In paper II, we found that biofilm formation in *S. typhimurium* involves a switch at the level of CsgD expression. Bistable expression of CsgD leads to a corresponding pattern of task distribution during biofilm formation where cells producing CsgD produced high amounts of cellulose. Our results also demonstrate that bistable regulation of CsgD is dependent on the presence of the wild type csgD promoter. Moreover, high levels of c-di-GMP lead to growth phase independent monophasic expression of CsgD. This study describes heterogeneity of biofilm communities and how bacteria cooperate in biofilms. While only one part of the population is dedicated to produce extracellular matrix components, the whole population can benefit from them. This type of behavior is recognized as being altruistic (Shapiro, 1998). Cooperation allows cells in the biofilm to perform other tasks or to be ready to answer to other challenges. In the future, it would be important to determine at which level bistability is generated and what are the
mechanisms involved. One can speculate that it could be at the level of CsgD expression or at the level of another upstream regulator.

In manuscript III and IV, we identified novel factors involved in the complex regulatory system of the master regulator CsgD (Gerstel et al., 2003, Ogasawara et al., 2010a, Jubelin et al., 2005, Kader et al., 2006, Simm et al., 2007, Brown et al., 2001, Holmqvist et al., 2010).

In manuscript III, two lytic transglycosylases, MltE and MltC that redundantly regulate rdar morphotype expression through CsgD were identified. This study also revealed that c-di-GMP can partially compensate for the absence of these two lytic transglycosylases.

In manuscript IV, we found that the chaperone Hfq and the Hfq dependent sRNAs, ArcZ and RyeB, regulate biofilm formation through CsgD. The sRNA ArcZ is the key regulator of the biofilm phenotype, as it can compensate for the absence of the chaperone Hfq. ArcZ regulates the transition between sessility and motility and also the timing of type 1 versus curli fimbriae surface-attachment.

These two studies could be developed further by determining the mechanisms through which the two lytic transglycosylases MltE and MltC and the two sRNAs, ArcZ and RyeB, affect expression of CsgD and the extracellular matrix components. Understanding these mechanisms will elucidate how the cell wall turnover affects pathways and how sRNAs affect the biofilm phenotype.
6 Acknowledgments

Many people accompanied me during this five years journey.
In particular, I would like to thank

My supervisor **Ute Römling** for offering me the opportunity to do my PhD in your group, for your expertise and enthusiasm for science. Thank you for helping me to achieve my goals and for keeping your office door always open for any question I had.

My cosupervisor **Sun** for the time spent in your lab and for having received me so well in Umeå.

Our collaborators, **Professor Jörg Vogel**, **Dr. Mark Gomelsky**, **Professor R.Malcolm Brown Jr**, **Dr. Annelie Brauner** and **Dr. Mikael Rhen** for your valuable input which pushed the projects forward.


The European Union for supporting my PhD studies with a Marie Curie Early Stage Research Training Fellowship (Sixth Framework Program, MEST-CT-2004-8475).

Present and former members in Ute Römling group: **Devrim** for helping me during my first weeks in Ute’s group; **Roger** for your friendship and numerous scientific suggestions; **Nina** for your friendship and enthusiasm for science; **Agaristi** for taking care of me when I arrived; for your friendship and all kinds of discussions, I enjoyed walking the PhD path in parallel with you; **Katherina, Verena, Uwe, Irfan, Nedaa, Xiaoda, Kader, Eugenie, Mathias, Kristin, Heili, Astrid, Soazig** and **Srisuda** for your friendship, for all the discussions about science, politics, life and other things, for the fun we had in the lab and during our lunch time.
My students **Karina, Romy** and **Charleyne**, for your motivation and contribution to my projects. I learned a lot supervising you.

My IMO-train colleagues **Agaristi, Kristina, Sönke, Jorrit, Aurel, Nicolas** and **Katharina**, for all the conversations and seminars we had together.

All other people at MTC, who were part of my daily life for the last five years, for being nice and happy people.

All other friends who made my life happier in Stockholm, **Joana, Oliver, Maroof, Iulian, Alina, Alessandro, Kristina, Christian, Ricardo, Sofia, Lina, Diogo, Vasco, Levi, Gonçalo, Tiago, Pedro, Isabel, Nina, Volodya, Katherina, Stefan, Agaristi, Hamid** and any other people who I might have forgotten to mention.

À minha nova família **Ilda e António** por todo o apoio e por me receberem sempre tão bem em vossa casa.

À minha família e principalmente aos meus pais, **Maria da Conceição e António** por toda a confiança e apoio que me deram quando decidiram vir para a Suécia fazer este doutoramento e por terem uns “miminhos” à minha espera sempre que regressava a casa.

Ao **Pedro**, por teres passado estes cinco anos ao meu lado, por me teres ouvido e motivado quando as coisas não corriam tão bem, por teres partilhado comigo muitos bons momentos e acima de tudo pelo teu amor e amizade.
7 References


organization of the cellulose synthase operon in Acetobacter xylinum. Proc Natl Acad Sci USA 87: 8130-8134.


