Molecular analysis of transcription factors in uropathogenic \textit{E. coli} adhesin operons

Stina Lindberg

Department of Molecular Biology
Laboratory for Molecular Infection Medicine Sweden (MIMS)

Umeå University
Umeå 2009
Till min familj

Great things are done by a series of small things brought together

- Vincent van Gogh
# TABLE OF CONTENTS

**ABSTRACT** ....................................................................................................................... 7

**PAPERS IN THIS THESIS** ............................................................................................. 8

1. **INTRODUCTION** ........................................................................................................ 9

   *Escherichia coli* ............................................................................................................. 9

   Pathogenicity islands (PAIs) ...................................................................................... 10

ADAPTATION TO ENVIRONMENTAL CHANGES ................................................................ 11

REGULATION OF GENE EXPRESSION ............................................................................... 11

   Transcriptional regulation .......................................................................................... 12

   The nucleoid associated proteins .............................................................................. 13

   Silencing by H-NS ....................................................................................................... 13

   Anti-silencing ............................................................................................................. 14

   Transcriptional activation ........................................................................................ 15

   CRP-cAMP ............................................................................................................... 16

PROTEIN – DNA INTERACTIONS ....................................................................................... 17

   Major groove binding proteins ................................................................................. 17

   Minor groove binding proteins ................................................................................. 17

BACTERIAL PROTEIN SECRETION AND TRANSLOCATION SYSTEMS ................................. 18

   The general secretion pathway (GSP) ...................................................................... 18

   Type I secretion ......................................................................................................... 19

   Outer membrane vesicles .......................................................................................... 19

UROPATHOGENIC E. COLI (UPEC) ................................................................................. 21

   Virulence factors of UPEC ....................................................................................... 21

   Fimbriae ....................................................................................................................... 21

   P-fimbriae ................................................................................................................... 22

   Regulation of the pap operon .................................................................................... 23

   The PapB family of transcriptional factors .................................................................. 24

   Type 1 fimbriae ........................................................................................................... 25

   F1C-fimbriae ............................................................................................................... 26

   Toxins ........................................................................................................................... 27

   α-haemolysin, Hly .................................................................................................... 27

CROSS-TALK BETWEEN DIFFERENT FIMBRIAL SYSTEMS ............................................. 28

SECRETION OF SURFACE ASSOCIATED ORGANELLES ................................................... 29

   The chaperone/usher pathway .................................................................................. 29

2. **AIMS OF THIS THESIS** ............................................................................................ 31
3. RESULTS AND DISCUSSION ............................................................... 32

PAPER I ........................................................................................................ 32
  Regulatory Interactions among Adhesin Gene Systems of Uropathogenic Escherichia coli .......................................................... 32

PAPER II ...................................................................................................... 37
  Purification, crystallization and preliminary data analysis of FocB: a transcription factor in the regulation of fimbrial adhesin expression in uropathogenic E. coli .... 37

PAPER III ...................................................................................................... 39
  Structure of FocB: a member of a family of transcription factors regulating fimbrial adhesin expression in uropathogenic E. coli .................. 39

PAPER IV ...................................................................................................... 43
  Interaction between the PapI protein and the alpha subunit of Escherichia coli RNA polymerase ................................................................. 43

PAPER V ...................................................................................................... 46
  Release of the type I secreted haemolysin via outer membrane vesicles from Escherichia coli ................................................................. 46

PAPER VI ...................................................................................................... 49
  Detection of functional P fimbriae proteins in E. coli outer membrane vesicles .... 49

4. CONCLUDING REMARKS ................................................................. 52

ACKNOWLEDGEMENTS ............................................................................. 53

REFERENCES .............................................................................................. 55
ABSTRACT

The main causative agent of human urinary tract infections is the uropathogenic *Escherichia coli* (UPEC) pathotype. It may cause disease due to its ability to express a number of bacterial virulence factors. Fimbrial adhesins are particularly important for the initial establishment of infection in the urinary tract. The fimbriae are hair-like structures protruding from the bacterial cell and by attaching to specific receptors in the urinary tract they mediate adherence to different cell types, allowing the bacteria to resist the shear forces from urine flow. The UPEC strains generally carry multiple determinants for fimbrial adhesins. Previous studies have indicated that there is a co-regulation between different fimbrial genes and one factor that has been implicated in this is the PapB protein, acting as a transcriptional regulator of P-fimbrial expression. The PapB protein can be regarded as the prototype of a family of fimbrial regulators that show high homology between different fimbrial operons. One homolog is FocB, regulator of F1C fimbriae.

In this study, the role of the FocB protein in the regulation of F1C fimbriae as well as in the co-regulation with other fimbrial genes was investigated. It was observed that FocB binds to DNA, similarly to PapB, in an oligomeric fashion and that PapB and FocB can form hetero-oligomeric complexes, which appear to have a repressive role in the regulation of the F1C fimbriae. In addition, the FocB protein also had a repressive effect on transcription of the *fim* operon, which encodes the Type 1 fimbriae. For further analysis of FocB *in vitro*, we developed efficient procedures for purification of the protein and established conditions for its crystal formation with the aim to conduct X-ray diffraction studies. By the hanging-drop vapour-diffusion method, we obtained crystals that in the X-ray analysis diffracted sufficiently well to allow modelling of a high resolution structure of FocB. The structural model was considered in relation to the DNA binding properties of the protein. The FocB analysis represents the first structural model of this family of transcriptional factors. This model should aid in further understanding of the roles and functions of these proteins in the regulation of the UPEC fimbrial operons.

The complexity of the system, with multiple factors involved in the regulation of fimbrial operons, was revealed in earlier studies of the PapI protein showing that PapI activates transcription of the *pap* operon as a part of a complex with the global regulator Lrp. However, PapI itself did not appear to bind to DNA and its mode of action has remained unclear. By genetic analyses and *in vitro* studies we show that PapI may interact also with the α subunit of the RNA polymerase. This finding indicates that PapI might directly interact with the transcriptional apparatus and thus aid in the activation of *pap* expression.

Bacteria are frequently releasing outer membrane vesicles (OMVs) from their surface. We studied the release of the haemolysin toxin from *E. coli* in connection with formation of OMVs and found that the toxin was tightly associated with the vesicles in an active form. By overproduction of the PapB or PapI regulators in order to maximise the population of bacteria expressing fimbriae, we could detect P fimbriae proteins associated with OMVs that displayed specific adhesion to receptor-coated beads. This suggests a possible scenario in which the vesicles can function as directed vehicles of bacterial virulence factors.
PAPERS IN THIS THESIS

I. Regulatory Interactions among Adhesin Gene Systems of Uropathogenic Escherichia coli
Stina Lindberg, Yan Xia, Berit Sondén, Mikael Göransson, Jörg Hacker, and Bernt Eric Uhlin.

II. Purification, crystallization and preliminary data analysis of FocB: a transcription factor in the regulation of fimbrial adhesin expression in uropathogenic E. coli.
Submitted

III. Structure of FocB: a member of a family of transcription factors regulating fimbrial adhesin expression in uropathogenic E. coli.
Ulrika Hultdin, Stina Lindberg, Christin Grundström, Shenghua Huang, Bernt Eric Uhlin, and A. Elisabeth Sauer-Eriksson.
Manuscript

IV. Interaction between the PapI protein and the alpha subunit of Escherichia coli RNA polymerase
Jurate Straseviciene, Hyun-Sook Park, Stina Lindberg, and Bernt Eric Uhlin.
Manuscript

V. Release of the type I secreted haemolysin via outer membrane vesicles from Escherichia coli
Carlos Balsalobre, Jose Manuel Silván, Stina Berglund(*), Yoshimitsu Mizunoe, Bernt Eric Uhlin, and Sun Nyunt Wai.
Molecular Microbiology, 2006, 59 (1), p. 99–112

VI. Detection of functional P fimbriae proteins in E. coli outer membrane vesicles
Manuscript

(*) now Lindberg
1. INTRODUCTION

Escherichia coli

Escherichia coli is a common inhabitant of the intestinal tract of humans. It is a rod-shaped, motile, facultative anaerobic Gram-negative bacterium that in most cases does not cause infection in man. However, there is a large spectrum of different E. coli species and there are many pathotypes that do cause infection with very diverse symptoms, from mild diarrhoea to life-threatening systemic diseases.

The E. coli species is often divided into three major groups, the non-pathogenic commensal (i), the intestinal pathogenic (ii) and the extra-intestinal pathogenic (iii) (78). The pathogenic groups can, in turn, be divided into several sub-groups, or pathotypes (41), dependent on the nature of their mode of infection (Table 1).

<table>
<thead>
<tr>
<th>Intestinal pathogenic E. coli (IPEC) (cause enteric/diarrhoeal diseases)</th>
<th>Enteropathogenic E. coli</th>
<th>Enterohaemorrhagic E. coli</th>
<th>EPEC</th>
<th>Enterotoxigenic E. coli</th>
<th>EHEC</th>
<th>Enteroinvasive E. coli</th>
<th>ETEC</th>
<th>Enteraggregative E. coli</th>
<th>EAEC</th>
<th>Diffusely adherent E. coli</th>
<th>DAEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraintestinal pathogenic E. coli (ExPEC) (cause UTI and meningitis)</td>
<td>Uropathogenic E. coli</td>
<td>Meningitis-associated E. coli</td>
<td>UPEC</td>
<td>MNEC (NMEC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Summary of different pathotypes of E. coli. The commonly used abbreviations are shown in the right column. Summarised from Kaper et al. (41)

The extraintestinal bacterial strains often reside in the intestinal tract, but there they usually do not cause infection. However, upon migration to other sites of the human body the infection can take place. The reason for the diversity in the pattern
of diseases caused by the *E. coli* species can to a great extent be explained by the acquisition of extramobile DNA elements harboring different virulence factors. These DNA elements have been named Pathogenicity Islands (PAIs) (31).

**Pathogenicity islands (PAIs)**

Pathogenicity islands are mobile genetic elements in the size of 10-200 kb. The DNA of the PAI often differs from the core genome in terms of G/C content and it is flanked by small direct repeats and transfer RNA (tRNA) genes. Association of the PAI to genes of insert elements (IS), transposases or integrases are also often found (Fig.1). This is an indication of acquisition through horizontal gene transfer, where the tRNA genes often serve as target. (32)

![Pathogenicity Island Diagram](image)

**Fig.1** Schematic view of a pathogenicity island. The bold line represents parts of the core genome, light grey boxes represent PAI-specific genes. The arrows indicate the presence of direct repeats at the ends of the PAI. Abbreviations: DR, direct repeats; *int*, integrase gene; *vir*, virulence-associated gene; *mob*, mobility gene. *mob* genes encode proteins involved in mobility of the genome. (32)

The genes in the PAIs may also be virulence-associated and encode a variety of different virulence factors, such as adherence factors, toxins, secretion systems and iron uptake systems. The PAIs are not unique to the *E. coli* species, but are found in a variety of pathogenic bacterial species, both Gram-
negative and Gram-positive, such as *L. monocytogenes*, *Yersinia* spp., *S. typhimurium*, *S. aureus* and *V. cholerae*, just to mention a few.

As the pathogenicity islands are mobile elements, the insertion of a PAI into a strain is not necessarily permanent. Deletion frequencies in the range of one in $10^{-5} – 10^{-6}$ have been observed in the uropathogenic isolate 536 (12, 62). However, the excision rate differs between different PAIs and some seem to be more prone to deletion than others. It has been shown that certain PAIs are more sensitive than others, in terms of deletion frequencies, to environmental stimuli such as temperature, nutrient availability and cell density (62).

**Adaptation to environmental changes**

In order for bacteria to be successful in surviving in different environments, such as for example the transition from free-living to the gut of an animal/human host, it is crucial to be able to sense, and adapt to, the environmental changes. There is great diversity in the different milieus that the bacteria can encounter and the change can be rapid, from one second to another.

We can imagine a bacterium that resides in a pond where the nutrient availability and osmolarity is low, the temperature is around 25°C and the pH is neutral. At one point a person starts drinking the pond water, and suddenly the bacterium is on its way down to the gastrointestinal tract. The temperature is raised to 37°C, the osmolarity increases and the pH drops dramatically. Further on, the bacterium needs to handle exposure to bile salts and the differences in availability and composition of nutrients, all within a couple of minutes.

The bacterial species has evolved a number of mechanisms to be able to handle these environmental challenges.

**Regulation of gene expression**

Facing all the different alterations in the environment, the bacterium can adapt by changing the expression of its housekeeping and virulence genes. Common environmental signals
that cause these changes are temperature, iron, osmolarity, Ca\(^{2+}\), carbon source, etc. The adaptations to these signals can occur at different levels.

Acquisition of new genetic material from surrounding bacteria (for example PAIs) or duplication of the same gene on the chromosome is one adaptation. The expression of a gene can also be altered by other gene rearrangements. One example is the invertible promoter of the *fim* operon, which uses the recombinases FimB and FimE to switch the promoter fragment in front of the structural genes, resulting in either “ON” or “OFF” expression of the fimbriae (see a more detailed explanation below). This type of regulation is called phase variation.

Another type of gene rearrangement is antigenic variation. In this case the sequence of a gene is altered so that the gene product is no longer recognised by the host’s immune defence. This approach is used by, for example, *Streptococcus pyogenes* and the change in the hypervariable N-terminal region of the M-protein (37).

Transcriptional regulation of gene expression is the most complex and most intensely studied area of gene regulation and the regulation can be very diverse. A brief description of the different regulatory control mechanisms relevant to this work will be presented below.

**Transcriptional regulation**

Gene transcription is the synthesis of RNA from a DNA template. This process is dependent on an RNA polymerase, which in *E. coli* consists of two \( \alpha \) subunits, one \( \beta \) subunit and one \( \beta' \) subunit building up the core enzyme, and one sigma (\( \sigma \)) subunit that together with the core enzyme builds up the holoenzyme.

For transcription to occur, the gene is dependent on a promoter sequence that is present in the -10 to -35 region upstream of the transcriptional start. The nucleotide sequence of the promoter determines its strength and the type of \( \sigma \) subunit connected to the RNAP determines the specificity with which promoter the RNAP can bind. Upstream of the -35 region is situated the upstream (UP) element. This is the recognition site for the C-terminal part of the RNAP \( \alpha \) subunit and it contributes
strongly to the efficiency of the transcriptional initiation (76). There are a relatively small number of bacterial proteins that act as global regulators of gene transcription by controlling a large portion of the *E. coli* chromosome.

**The nucleoid-associated proteins**

Bacterial chromosomal DNA is organised into a compact structure called a nucleoid. In order to keep this structure intact, a number of factors are important, such as the nucleoid-associated proteins. The major nucleoid proteins are the FIS (factor for inversion stimulation), HU (heat unstable protein), IHF (integration host factor protein) and H-NS (histone-like nucleoid structuring). The most abundant nucleoid protein during exponential growth is the FIS protein. It recognises a poorly conserved 15bp binding site but has non-specific affinity for DNA as well. When FIS is at its maximal expression, it induces strong compaction of the DNA (82). IHF binds DNA with great specificity and induces strong bending of the DNA (2). It has a clear consensus sequence for binding, but can also bind DNA nonspecifically and be substituted by HU. The HU protein binds preferentially to supercoiled DNA (85) and shows little sequence specificity, although it tends to bind to structural distortions in DNA, such as nicks and gaps (7, 40). The H-NS protein will be described in more detailed below.

The expression pattern of the different nucleoid proteins differs with the bacterial growth phase and the role of these proteins are different in terms of compaction and as antagonists of compaction. This provides a means to locally modulate the structure of the nucleoid, which is dependent on growth conditions and is important for transition between transcriptionally active and repressed chromatin architectures (see Dame, 2005 for review (17)).

**Silencing by H-NS**

H-NS is a small (about 15 kDa), abundant protein in *E. coli* with preferential binding to T/A-rich, curved DNA. H-NS binds to DNA as a dimer or higher order oligomer and the initial binding region can function as nucleation sites from which the protein can polymerise along the DNA (3). H-NS in itself does not seem to
induce very strong DNA bending, but it has a tendency to bind DNA that is intrinsically bent. The silencing effect by H-NS was first described in the temperature-regulated pap operon in E. coli (30). A mutation in the drdX locus induced transcription of the fimbrial genes in a non-temperature-dependent manner. Further analyses of the mutation revealed that the gene encodes the H-NS protein. The results also implied that the interactions between H-NS and the regulatory region of the pap operon are different at high and low temperatures (30).

Promoter regions of virulence genes often have the properties of being both curved and T/A-rich, which make them a good target for silencing by H-NS. As reviewed by Dorman (20), the acquisition of horizontally transferred virulence genes requires effective transcriptional silencing, to avoid the risk of expressing a newly transferred gene at an unfavourable time point. It has been shown that H-NS binds to newly acquired foreign DNA of low G/C content and that H-NS silences expression by restricting the access of RNA polymerase to the DNA (58, 65).

Anti-silencing

In order for transcription to occur, it is necessary that the binding of H-NS and other nucleoid-associated proteins can be modulated to relieve gene repression and give access to anti-repressors. The binding of H-NS to DNA is not very strong, which means that it can be displaced without the help of other proteins. This has been studied in the virF promoter region of S. flexneri (74), where H-NS has two binding sites distant from each other. At lower temperature, the DNA is bent and H-NS can keep the DNA in a closed loop. As the temperature rises, the bending of the DNA decreases and the distance between the H-NS binding sites gets longer, so that the H-NS no longer can keep the structure closed. This enables other factors to gain access to the DNA and initiate transcription (74).

This mechanism has also been seen in the pap operon in E. coli, where expression is repressed by H-NS at 25°C, but as the temperature rises, the repression is relieved and the operon is transcribed (30). In a drdX- strain with the CRP binding site deleted in the pap UAS, the expression of the pap operon reached levels similar to wt, indicating a role for the cAMP-CRP complex as
an anti-silencer rather than activator of *pap* transcription (25). Similar results were obtained when the gene or primary binding site for the PapB DNA binding protein was mutated in combination with the *hns* gene mutation.

Another example of anti-silencing is where H-NS is outcompeted by other regulatory factors. In *Vibrio vulnificus* the *rtxA1* operon, encoding the RTX toxin, is repressed by H-NS. The expression of the *rtxA1* operon is dependent on the transcriptional regulator HlyU (56). It was shown by competitive footprinting and gel shift analyses that HlyU does not act as an activator, but rather as an anti-repressor, that alleviates the repression of H-NS by competition at two AT-rich sites upstream of the *rtxA1* operon promoter (57).

**Transcriptional activation**

Proteins that are activators of transcription usually have their binding sites close to the promoter. They often make contact with the RNA polymerase, through contact with the CTD-domain of the α-subunit (Fig. 2) (reviewed by Ebright and Busby, 1995 (22)).

![Fig. 2](image)

*(Adapted from Ebright and Busby, 1995)*

**Fig. 2** A promoter containing an activator site that functions through contact with the α-CTD of the RNA polymerase (e.g. the *lac* promoter). The α-CTD makes specific protein-protein contacts (black circle) with the activator (22).
**CRP-cAMP**

A well-studied activator that functions through contact with the RNAP is the CAP (catabolite gene activator protein) or CRP (cAMP receptor protein). CRP only functions with cAMP bound (CRP-cAMP) and when the cAMP levels in the cell are high, it will bind to CRP, which in turn binds its activator site. The CRP-cAMP complex bends the DNA and makes contact with the $\alpha$-C-terminal domain ($\alpha$-CTD) of the RNA polymerase (83, 105).

There are two classes of promoters that have been described where CRP acts as a sole regulator. In class I promoters (i.e. *lac*) the CRP protein makes contact with the $\alpha$-CTD of the RNA polymerase and stabilises its binding to the promoter, allowing transcription to occur (21, 64). In class II promoters (i.e., *gal*), the CRP protein makes contact with both the $\alpha$-CTD and the $\alpha$-N-terminal domain of the RNA polymerase to activate transcription (67). By analysis of mutant variants of the $\alpha$-CTD domain, specific determinants have been identified to be important for CRP-dependent activation of the *lac* promoter (reviewed by Busby and Ebright, 1999 (14)).

One promoter in the case of virulence-associated proteins that has been shown to depend on the CRP-cAMP for activation is the $P_{papBA}$, driving the expression of the *pap* operon, encoding P-fimbriae (29). Studies have shown that the role of CRP in the activation of the $papBA$ promoter share many characteristics with activation of class I promoters, even though the CRP site is positioned considerably further away from the $P_{papBA}$ (215.5 bp) than in the classical class I *lac* promoter (-61.5) (102).
Protein – DNA interactions

The structure of the B-DNA, with the two strands wrapped around each other forming a double helix, forms two grooves between the two strands – the major and minor grooves. The width of the major groove is 11.6 Å and minor groove is 6.0 Å, but it also varies with the sequence. Non-alternating runs of A:T base pairs tend to have a narrower width, in the range of 3-4 Å, compared to the average width of 6 Å (reviewed by Neidle, 2001 (66)). The characteristics of the DNA, as well as the sequence, are important for the accessibility of different regulatory proteins.

Major groove binding proteins

A common structure of a DNA binding protein is the helix-turn-helix motif that recognises and binds specific sequences of the DNA. The recognition helix is the second helix of the motif and when it binds to DNA, it lies in the major groove. The amino acids in helix 2 can contact and form hydrogen bonds with specific bases in the DNA (90).

A typical example of a helix-turn-helix protein is the CRP protein of *E. coli*. It was shown by structural studies that the C-terminal helix binds in the major groove of B-DNA while the N-terminal helix binds to cAMP (61). This type of helix-turn-helix motif was later on found to be a general trait of DNA binding proteins (90).

Minor groove binding proteins

While the typical regulatory proteins bind to specific sequences of the DNA in the major groove, the minor groove binding proteins are to a great extent proteins that contribute to the structural organisation of the DNA, building up, for example, the nucleoid (10). An important task for the minor groove DNA binding proteins is to bend or kink the DNA, allowing distant sites to come closer to each other, making protein interactions possible for DNA compaction, as well as for the control of transcriptional regulation. This is can be done through a mechanism called intercalation, where the protein side chain is inserted between the bases in the minor groove, inducing the bend or kink (101).
Proteins in prokaryotic systems that have been described as being important for bending of the DNA are, among others, the *E. coli* purine repressor protein PurR (84), the integration host factor IHF (recognises specific sequences) (75) (91), HU (non-specific recognition) (93) and H-NS (89).

**Bacterial protein secretion and translocation systems**

Another important and essential feature of the bacterial cell in terms of adaptation to environmental changes is the ability to translocate and secrete proteins. The bacteria commonly obtain nutrients by secreting enzymes that degrade macromolecules to smaller constituents that in turn can be taken up, actively or by passive diffusion, by the bacterial cell. Components of surface-exposed organelles and lipid structures need to be transferred across the bacterial inner and outer membranes (IM and OM) from the cytosol where they are produced. Outer membrane proteins need to pass the inner membrane and be inserted into the OM. Secretory signal molecules and other effectors and toxins need to be secreted to exert their effect and misfolded or damaged cell products need to be degraded and/or disposed of to avoid potentially damaging effects on cellular functions.

The bacteria have evolved a number of different mechanisms to complete these tasks, and I will describe briefly below some of these mechanisms that are relevant to this work.

**The general secretion pathway (GSP)**

The GSP is the mechanism by which proteins are translocated through the cytoplasmic membrane. The proteins transported by this system have an 18-30 amino acid long N-terminal signal peptide (or leader peptide), that is cleaved off during transport through the cytoplasmic membrane. The system transporting the protein through the membrane is the Sec (secretory) translocase. Two pathways can be used for transport to the Sec translocase, either via the SecB cytoplasmic chaperone or the SRP (signal recognition particle). In both cases, the protein will be kept in an unfolded state and brought to the membrane-
associated SecA, from which it is secreted through the SecYEG transmembrane pore (Reviewed by de Keyzer et al., 2003 (18)).

However, for many proteins secretion is not complete with transport across the cytoplasmic membrane. There are several ways for the protein to be exported subsequently across the outer membrane, the so-called terminal branches. Three major terminal branches exist; the chaperon/usher pathway (i), used by many fimbrial systems (96) (see a more detailed description below), (ii) the type II secretion mechanism that is used by e.g., the cholera toxin (77) and (iii) the type IV secretion system used by e.g., the *H. pylori* CagA protein (5).

**Type I secretion**

RTX toxins (repeat in toxin) are a group of toxins produced by a range of Gram-negative bacteria that are commonly secreted to the extracellular space via the type I secretion system. The type I secretion system is built up of three proteins, spanning the cell envelope. First, there is a transmembrane domain (TMD) that is localised in the cytoplasmic membrane. The TMD is built up of a dimer and recognises the signal sequence of the substrate. A membrane fusion protein (MFP) spans the periplasmic space, connecting the inner and outer membranes, and an outer membrane protein of the TolC class, acting as the link to the extracellular space (19).

This system has been extensively studied with the *E. coli* haemolysin secretion as a model. The haemolysin (HlyA) has an N-terminal signal sequence and the protein is fully folded prior to secretion. HlyA is directed to the transmembrane domain HlyB, which transports the toxin, by hydrolysis of ATP, through the cytoplasmic membrane and periplasm via HlyD and out to the extracellular space via the outer membrane protein TolC (48).

**Outer membrane vesicles**

Many Gram-negative bacteria have been shown to frequently release small vesicles from the outer membrane of the cells (OMVs). The vesicles are spherical in shape and in the size range of 50–250 nm in diameter. Naturally, the composition of the OMVs is very similar to that of the outer membrane, that is, consisting of LPS, proteins and phospholipids, but it also includes
some periplasmic material that is enclosed during the process of vesicle formation (reviewed by Beveridge, 1999 (9)). However, the vesicles have been suggested to have a more specific function in terms of sorting certain proteins into the vesicles as a mechanism of delivery into the extracellular space (1, 8, 42, 49, 98).

Wai et al. (2003) showed that vesicles secreted from *E. coli* cells contained the toxin cytolysin A (ClyA), but more interestingly, that assembly of ClyA into the vesicles was needed for the formation of its active oligomeric form. Furthermore, it was eight times more toxic to eukaryotic cells when delivered by OMVs than in its monomeric form (98). It has also been described that the vesicle content can be modified and delivered to host cells. Kesty and Kuehn (2004) showed that periplasmic green fluorescent protein (GFP) could be incorporated into the OMVs, and thus be used to track the course of the OMVs after release (44). By overexpression of the *Yersinia enterocolitica* adhesin/invasion protein Ail, they also showed that Ail could be incorporated into the OMVs, and in turn, alter the adhesion and internalisation properties of the OMVs with respect to mammalian cells (44). Kuehn et al. (2004) also showed that Enterotoxigenic *E. coli* can target the delivery of OMV-bound LT toxin to host cells via the toxin receptor, and subsequently become internalised through lipid rafts (45).

As OMVs contain high concentrations of many surface antigens, such as outer membrane proteins and LPS, which are highly immunogenic, OMVs have the potential of being used for vaccine development. It was recently shown that immunisation with OMVs from *V. cholerae* induced strong antibody responses in mice as well as long-lasting protection against *V. cholerae* infection (81). Schild et al. also demonstrated that a specific immune response could be induced by OMVs harbouring the heterologously expressed PhoA, suggesting that OMVs could be used as antigen-delivery vehicles by packaging antigens from different pathogens into OMVs (80).

The role of outer membrane vesicles in the secretion of proteins as opposed to other protein secretion mechanisms is, however, not fully understood. It is possible that their role in secretion in some cases is minor, but that they play a larger role in other cases (60).
Uropathogenic *E. coli* (UPEC)

Uropathogenic *E. coli* belongs to the group of extraintestinal pathogenic *E. coli* and accounts for about 80% of urinary tract infections in individuals <55 years old (16). The large intestine often serves as a reservoir for the UPEC, and given an opportunity, the bacteria can enter the urinary tract and cause infection (110). The most common type of infections of the urinary tract are urethritis and cystitis (uncomplicated infection of the urethra or bladder), and acute pyelonephritis (infection of the kidney, usually more severe). The risk of acquiring a bloodstream infection is also higher in the case of pyelonephritis, given that the kidney is a highly vascularised organ.

There are several factors that frequently are associated with clinical isolates recovered from patients with urinary tract infections, such as different adhesins including P-fimbriae and different toxins. These virulence factors will be described in more detail below.

**Virulence factors of UPEC**

**Fimbriae**

Fimbriae (also called pili) are hair-like proteinaceous structures protruding from the surface of the bacterial cell. They are built up of major and minor subunits and have a tip fibrillum with an adhesin at the distal end of the fibrillum that mediates the specific interaction with a receptor (38, 50). The fimbriae can be several microns in length and have a diameter of about 10 nm (Fig. 3).

Uropathogenic *E. coli* has the ability to express several different types of fimbriae, with different adhesive properties. This ability enables them to adhere to different receptor structures, thus increasing the possibility for establishment in different environments.
Fig. 3 AFM micrograph of a F1C-fimbriated *E. coli* bacterium. The enlarged section shows the tip fibrillum, which carries the adhesin at the distal end. Scale 2x2 µm.

*P fimbriae*

The *pap* gene cluster resides within pathogenicity islands that have been acquired through horizontal gene transfer. The bacteria with P fimbriae are often associated with acute pyelonephritis and the ability to express such fimbriae has also been shown to be important for early establishment of *E. coli* in the human urinary tract (106). The adhesin of the P fimbriae is called PapG and the receptor is (Gal(α1 → 4)βGal)-containing glycolipids (53). There are different variants of the P fimbriae that only differ in receptor binding properties. The different PapG variants (PapGI, PapGII and PapGIII (or PrsG - *pap*-related sequence)) have different isoreceptors, giving them different specificities for binding (reviewed by Lane and Mobley, 2007 (52)). The PapGI adhesin binds preferentially to the P-blood group antigen present on human erythrocytes and uroepithelial cells, whereas the PapGIII (PrsG) adhesin binds the Forssman antigen present on sheep erythrocytes and in the human renal pelvis (53, 59). The different variants of *pap* are thought to be the result of gene duplications (4) and they are often present within the same bacterial strain.
Regulation of the pap operon

The *pap* gene cluster, defined by the cloning and genetic analysis of genes required for production of the P fimbriae, consists of nine structural genes (*papA-G*) and two regulatory genes *papI* and *papB*. The *papBA-G* genes are transcribed as a polycistronic operon and the *papI* gene is transcribed from a separate promoter in the opposite direction from the other genes (Fig. 4). Recent studies have shown that there are also additional genes in the promoter distal region of the major operon (86, 87). The transcription of the *papBA-G* operon is under phase variation, thus there is either an ON or OFF status of transcription.

A DNA binding protein with a regulatory role is PapB, which influences transcription of both the *papAB-G* genes and the *papI* operon by binding to a specific DNA sequence in the region between the two divergent *papI* and *papB* promoters (PapB site 1) (Fig. 4) (see a more detailed description below). PapB also has an auto-regulatory role where it represses its own transcription and this was suggested to be due to its binding to the two lower-affinity sites PapB site 2 and site 3 (24).

![Fig. 4](image_url) Genetic organisation of the regulatory region of the *pap* operon. The binding sites for Lrp, Crp and PapB are indicated with purple circles, green and black boxes respectively. Straight arrows are indicating factors important for the transcriptional regulation, which are described in more detail in the text.
PapI is another transcriptional activator that, in complex with the Lrp protein, binds the Lrp binding sites 4-6 allowing activation of transcription. Dependent on the methylation state of the DNA at the regulatory region, Lrp acts as a repressor, by binding Lrp sites 1-3 (39, 68, 99).

The PapI protein is an 8.8-kDa positive regulator of the pap operon (15). It was suggested to be a winged-helix-turn-helix (wHTH) protein with some non-specific DNA binding ability, however it is found in a complex together with the Lrp protein which binds strongly to DNA (34, 39, 43, 68). The PapI protein is highly conserved among other fimbrial operons. However, extensive sequencing and alignment of papI genes from 54 UPEC isolates revealed 10 different PapI variants (97). The varying residues were localised to the proposed Lrp binding domain of the protein, whereas the DNA binding domains were conserved (43). Totsika et al. proposed that the PapI variants have evolved to prevent cross-activation of other related fimbrial operons within the same isolate (97).

Other important players in the regulation of the pap-operon are the CRP-cAMP protein, which is essential for papI and papBA transcription by RNAP contact (15, 102). The DAM-methylase methylates DNA and thereby alters the affinity of Lrp-PapI to the GATC_distal binding sites (13, 34, 68). The H-NS protein is a silencer of pap-transcription at low temperatures and it was also suggested to be acting as a DNA methylation blocker (30, 103).

The PapB family of transcription factors

The PapB protein is an 11-kDa DNA binding protein. It binds to DNA in an oligomeric fashion with 8-10 monomers. It has specificity for TTT/AAA triplets with 9 bp repeats and it was proposed to be a minor groove binding protein (107). An amino acid alignment of PapB homologous proteins from different fimbrial operons reveals high similarity (Fig. 5). It also shows that there are certain regions within the proteins that are conserved. These regions were shown to be important for the ability of the PapB protein to oligomerise and bind DNA (109).
Fig. 5 Protein sequence alignment of PapB homologous proteins. Conserved aa residues are boxed. Areas important for DNA binding are marked with a black circle and areas important for oligomerisation are marked with a striped oval (109).

**Type 1 fimbriae**

The genes encoding Type 1 fimbriae (fim) are located on the core chromosome, which means that they are present in common laboratory strains as well as in pathogenic isolates. Type 1 fimbriae can bind to both abiotic and biotic surfaces (73) and has been shown to be involved in attachment, invasion and establishment of biofilms in the bladder (63, 104). The adhesin of the Type 1 fimbriae is called FimH and mediates adherence to mannose-containing receptors, and the attachment can be blocked by soluble α-D-mannoside (73).

The fim gene cluster consists of one polycistronic operon containing the structural genes (fimA-H) and two monocistronic operons containing the genes for the two recombinases FimB and FimE. The regulation of Type 1 fimbriae is phase variable and under the control of FimB and FimE, which are responsible for changing the direction of the promoter that is driving the expression of the structural genes. By changing the direction of the invertible 314 bp DNA fragment where the promoter is situated, the expression of the operon will turn either “ON” or “OFF” (27, 46) (Fig. 6).
Fig. 6 Illustration of the *fim* switch at the promoter of the *fim* operon. The FimB recombinase is mainly responsible for the switch of the invertible element in OFF → ON direction, whereas FimE turns the switch from ON → OFF.

The frequency of the “switch” is dependent on temperature, with the fastest switch from OFF-ON occurring between 37 and 41°C. Under normal laboratory conditions, the promoter is mostly in OFF position and only a sub-population expresses the fimbriae on the bacterial surface (26).

Other regulators that are important in the regulation of the *fim* operon are the DNA binding factors H-NS (heat-stable nucleoid structuring protein), IHF (integration host factor) and Lrp (leucine-responsive protein) (23, 28, 70).

**F1C fimbriae**

The genes responsible for the expression of the F1C fimbriae are the *foc* gene cluster. It is highly homologous to the *sfa* gene cluster (72) which encodes S fimbriae. The receptor for the F1C fimbriae has been identified as galactosylceramide and globotriaosylceramide. The galactosylceramide is present in all tissues within the ascending urinary tract except for the urethra, while the globotriaosylceramide is specifically expressed in renal tissue (6). The transcription factors of the *foc* operon, FocI and FocB share high amino acid homology with the *pap* transcription factors PapI and PapB.
Toxins

$\alpha$-haemolysin, Hly

The $\alpha$-haemolysin is an exotoxin that is frequently associated with uropathogenic *E. coli*. It belongs to the group of RTX toxins, which all contain a tandem duplication of nine amino acids. This group of toxins acts by creating pores in the eukaryotic cell membrane. Pore formation is dependent on the binding of Ca$^{2+}$. The toxin is secreted from the bacterial cell by the Type 1 secretion system. It is secreted directly to the extracellular space through a pore consisting of HlyB, HlyD and TolC, without any periplasmic intermediates (94). The synthesis, activation and secretion of the $\alpha$-haemolysin are determined by the *hlyCABD* operon (100). The 110-kDa HlyA protein is produced in its inactive form, but activated in the cytoplasm by HlyC, an fatty acid acyltransferase (33).
Cross-talk between different fimbrial systems

Uropathogenic *E. coli* strains often carry several different fimbrial operons within the genome. For example, the UPEC strain CFT073 carries genes encoding 12 putative fimbriae, including the *fim* operon encoding for Type 1 fimbriae, two *pap* operons encoding P fimbriae and a *foc* operon encoding F1C fimbriae (Welch *et al.*, 2002). Several of these operons may be cryptic and contain pseudogenes, but the possibility for multiple fimbrial expression is clearly present.

Also, the UPEC strain J96 has been shown to carry at least four different fimbrial operons, including *fim, pap, prs* and *foc* (11, 59, 92). However, it has been shown in an *E. coli* pyelonephritis strain, which carries the genes for P, Type 1 and F1C fimbriae, only a sub-fraction (about 3-9%) of the bacteria express more than one fimbrial type simultaneously on the bacterial surface, and that there is rapid phase variation between fimbriated and afimbriated phenotypes (69). This gives a basis for the possibility of cross-regulation between different fimbrial operons.

There have been several reports on fimbrial cross-talk in different uropathogenic isolates. Xia *et al.*, as well as Holden *et al.*, showed that there is cross-talk between the P and Type 1 fimbriae, where the PapB regulator turns off Type 1 expression by affecting the *fim* switch (36, 108). This has also been shown to be the case for FocB, the regulator of F1C fimbriae, which affected Type 1 expression in a similar manner (55).

The presence of Type 1 fimbriae has also been shown to turn off the expression of the Ag43 surface protein, an aggregative factor involved in bacterial aggregation and biofilm formation in *E. coli* (79). The effect of Type 1 fimbrial expression on P fimbriae *in vivo* in a murine urinary tract model has also been reported (88). Cross-talk between fimbriae and flagella has been described in the uropathogen *P. mirabilis* (MR/P fimbriae) (54) and between S fimbriae and Type 1 fimbriae and motility in *E. coli* (86). Cross-regulation between different *pap* paralogues as well as between fimbrial types with homologous regulatory proteins has also been described (55, 97).
Secretion of surface-associated organelles

The translocation of fimbrial subunits to the outer membrane of the bacterial cell and the assembly of the fimbriae has been described extensively with the chaperone/usher secretion pathway of P fimbriae as model system (see review by Thanassi and Hultgren, 2000 (95)). The organisation and functions of the subunits of the different fimbrial operons are often very similar (Fig. 7). More than 30 adhesive structures of Gram-negative pathogenic bacteria are known to use a chaperone/usher pathway for their secretion (96). It has also been shown that the chaperone/usher system is interchangeable between Type 1 and F1C fimbriae, as long as both the usher and the chaperone from the same fimbrial type work together (47). The amino acid identity is 68% between FocC and FimC and 58% between FocD and FimD.

![Fig. 7 Genetic organisation of different fimbrial operones. Boxes with the same color represent genes with similar function. The originating strain is indicated. Modified from Sjöström et al., 2009 (87).]

The chaperone/usher pathway

The fimbrial subunits of the pap operon are translocated in an unfolded state over the cytoplasmic membrane via the general secretion system (sec) pathway. But in order to be completely released into the periplasmic space and be folded properly, the chaperones PapD and DsbA, a periplasmic disulphide isomerase, is needed. In the absence of PapD, the proteins will be...
misfolded and degraded by the protease DegP (Fig.8). The misfolded proteins in the periplasm are sensed by a two-component system CpxA-CpxR, in which the CpxA is the sensory kinase and CpxR the DNA binding response-regulator, activating the transcription of $degP$ (Reviewed by Thanasi et al, 1998 (96)). The folded pilus subunits will be directed to the channel-formed PapC usher in the outer membrane, where subunit-subunit interactions will lead to the translocation of a linear fiber across the membrane. Once outside the cell, the fimbriae will twist into its final shape.

Fig.8 Schematic picture of the chaperone/usher secretion pathway exemplified by the secretion of P fimbrial subunits. The different steps are described in the text. (Modified from Thanassi and Hultgren, 2000 (95))
2. AIMS OF THIS THESIS

- To investigate the specific interactions between the transcription factors of the *pap* and *foc* operons in uropathogenic *Escherichia coli* (UPEC), as well as the role of the PapB family of transcription factors in the regulation of the different fimbrial systems.

- To assess the functional properties of the FocB protein, in relation to its role as a DNA-binding protein, using a structure based approach.

- To study possible interactions between the PapI transcription factor and RNA polymerase in the context of *pap* transcription.

- To study how outer membrane vesicles may carry UPEC virulence factors and to explore the possibility of enhancing the presence of functional fimbriae proteins associated with OMVs.
3. RESULTS AND DISCUSSION

Paper I

*Regulatory interactions among adhesin gene systems of uropathogenic* *Escherichia coli*

FocB is a transcription factor affecting the expression of the *foc* operon. This work was aimed towards characterising FocB in terms of its DNA-binding and oligomerisation properties. We also studied the interactions between FocB and PapB, a homologous protein in the *pap* operon, and the effect of these proteins on F1C fimbrial expression.

*The FocB protein binds both foc and pap regulatory regions*

In order to investigate the DNA-binding properties of FocB to the regulatory region of the *pap* and *foc* operons, the proteins were purified and subjected to *in vitro* DNA mobility shift analyses. We used DNA fragments of approximately 400 bp, corresponding to the *pap* and *foc* intercistronic region, respectively, where PapB has been shown to bind to a 52 bp sequence (107) in the *pap* DNA (site 1) (Paper I, Fig. 1).

The results of the gel mobility shift analyses showed that FocB bound efficiently to the *foc* DNA, whereas binding to the corresponding *pap* DNA was somewhat less efficient. The binding of FocB to *foc* DNA was also observed at the single molecule level by atomic force microscopy (AFM) imaging. Measurements of the protein binding to the DNA indicated that FocB binds in an oligomeric fashion and in the region corresponding to *pap* site 1 (Fig. 9).
To further assess the binding properties of FocB and PapB to foc and pap DNA, DNA fragments containing both DNA binding sites 1 and 2 were used in gel mobility shift analyses. The results indicated that the binding of FocB to both foc sites 1 and 2 was more efficient than PapB binding to the same DNA (Paper I, Fig. 4A-C). No binding of PapB to foc site 2 did occurred within the concentrations tested.

### Overexpression of PapB decreases the expression of F1C fimbriae

The ability of PapB to bind the foc regulatory DNA region prompted us to investigate the effect of PapB on the expression of F1C fimbriae. A strain harbouring the cosmid pBSN50, encoding the complete foc operon, was used. The presence of FocA, the major fimbrial subunit of F1C, was determined by Western blotting and the phenotype of cell expressing whole fimbriae expression was examined by AFM imaging. Previous studies have identified regions of the PapB protein that are important for its ability to bind DNA or to oligomerise (Fig. 5) and selected mutants were used in order to determine their effect on F1C expression.

The results showed that by expressing wt PapB in trans, the expression of F1C was abolished, whereas by using a PapB mutant deficient in oligomerisation, the effect on expression of F1C was modest (Paper I, Fig. 2). However, when a PapB mutant deficient in DNA binding was introduced, F1C expression remained repressed, indicating that hetero-oligomeric complexes had formed.

---

**Fig. 9** AFM images showing binding of FocB to foc DNA. Panel (a), only DNA, panels (b-d) DNA incubated with 3.2 µM FocB. Arrows indicate the regions where the protein is binding to the DNA.
between PapB and FocB where PapB could titrate out the activating FocB complexes.

**Oligomerisation of FocB and the formation of PapB/FocB hetero-oligomers**

Previous studies have shown that the PapB protein binds to DNA, and that it does so in an oligomeric fashion (107, 109). To investigate if FocB has the same property, we performed dimerisation/oligomerisation experiments both in vivo and in vitro.

We used a system based on the λ cI repressor. λ cI has two domains, the N-terminal DNA-binding domain and the C-terminal dimerisation domain. The λ cI protein can only repress transcription if it works as a dimer bound to two separate operator sites. We constructed a chimeric protein where FocB replaced the C-terminal dimerisation domain. If FocB can dimerise, the chimeric protein can function like the wild type and thus repress transcription of the downstream reporter genes. The results showed that the chimeric protein clearly could repress transcription of the reporter genes thus suggesting the formation of FocB dimers in vivo (Paper I, Table 3).

The effect of PapB on the expression of F1C fimbriae intrigued us, prompting us to test whether PapB and FocB proteins could interact in the regulation of foc operon expression. This was initially tested by incubating different concentrations of FocB and PapB together with foc DNA in a gel mobility shift assay. The results showed that the binding of PapB to foc DNA gave a considerable sharper band shift if FocB was present. This indicates that heteromeric PapB/FocB complexes may be formed and may alter the DNA interactions (Paper I, Fig. 4D).

Furthermore, the hetero-oligomerisation ability was tested by cross-linking purified PapB and FocB proteins in vitro as well as in a bacterial strain overexpressing the two proteins in vivo. The results from both of these experiments indicated the presence of FocB dimers and higher order oligomers, as well as the formation of hetero-oligomeric complexes of FocB/PapB (Paper I, Fig. 5).
The role of the FocB/PapB hetero-oligomers was also tested in vivo in the clinical uropathogenic isolate J96. The expression of FocA was determined by Western immunoblotting while the levels of PapB were gradually increased by induction from an inducible lac promoter. As the levels of PapB increased, the FocA levels decreased (Paper I, Fig. 6).

Taken together, these results strongly suggest that the FocB protein forms dimers and oligomers in vivo and that the FocB and PapB proteins can form hetero-oligomeric complexes that in turn could have an impact on the regulation of the different fimbrial operons. The results also indicate that PapB can sequester FocB and titer out the activating FocB. The increased ability of hetero-oligomeric FocB/PapB complexes to bind foc DNA also suggests a role in inhibition of transcription at the promoter region of foc.

FocB affects the expression of both pap and fim operons

The PapB protein was earlier shown to inhibit the transcription of Type 1 fimbriae, by affecting the fim switch. PapB could increase the ON \( \rightarrow \) OFF transition by increasing FimE expression as well as blocking the activity of FimB at the fim switch thus preventing switching on fim expression (35, 108).

We wanted to investigate if FocB also can affect the pap and fim operons and we did this by using reporter systems of both operons. Using a reporter fusion of papBA::lacZ and its isogenic papB1A::lacZ together with a plasmid encoding FocB, we observed that FocB could restore the papA transcription to 90% in the papB mutant.

The effect of FocB on fim expression was tested in two different reporter fusions. The first one was a fimA::lacZ fusion with both recombinases intact upstream of the fimA promoter region. The second was a “switch locked on” fimA::lacZ fusion strain where both recombinases were deleted, thus the effect seen will be due to effects exerted on the fimA promoter. In the “switch locked on” strain, no effect was observed upon expression of the FocB protein, whereas more than a 50% reduction of the \( \beta \)-galactosidase level was seen in the wt fimBE fusion (Paper I, Fig. 7B).
From these data we concluded that FocB can complement a PapB mutant and also inhibit Type 1 fimbriae expression. The latter effect was due to FocB having an effect on recombinases similar to that of PapB and thus both proteins may play a role in regulatory cross-talk with the \textit{fim} determinant.

\textit{Conclusions}

Taken together, the experiments presented in this paper suggested to us that there is a hierarchy in the cross-regulation between the \textit{pap}, \textit{fim} and \textit{foc} operons. FocB could stimulate the expression of \textit{pap}, whereas PapB did not alone affect \textit{foc} expression. However, together with FocB, PapB had a repressive effect on the \textit{foc} operon, presumably by the formation of heterooligomeric complexes. In turn, both PapB and FocB showed a repressive effect on the expression of the \textit{fim} operon, encoding for the Type 1 fimbriae (Fig. 10).

\textbf{Fig. 10} Summary of the observed regulatory crosstalk between the \textit{pap}, \textit{foc} and \textit{fim} operons in uropathogenic isolate J96.
Paper II

*Purification, crystallization and preliminary data analysis of FocB: a transcription factor in the regulation of fimbrial adhesin expression in uropathogenic E. coli.*

The PapB family of transcription factors is clearly important in the regulation of – and cross-regulation between – fimbrial operons. However, the mechanism behind the regulation is still in many respects not known. To further understand how the FocB protein is acting at the fimbrial promoter region, we decided to take a structural approach. We developed efficient procedures for purification of the FocB protein and established conditions for its crystal formation with the aim of conducting X-ray diffraction studies.

**Purification and crystallisation of the FocB protein**

To obtain a pure preparation of the FocB protein at high concentration, we cloned the *focB* gene into an inducible vector carrying a 6x His-tag. The protein was subsequently expressed and affinity purified with Ni-NTA agarose and gel filtration. In the initial trials aiming at purifying a protein that had the ability to form crystals for X-ray diffraction studies, we used a protein that carried an N-terminal 6x histidine tag. However, the crystals stopped growing at a very small size and the optimisation of the crystallisation conditions did not improve the crystals.

NMR analysis of the His-FocB protein revealed that a part of the protein was unfolded, presumably the His-tag. The NMR spectra also indicated that the protein tended to be more stable at higher temperatures (Paper II, Fig. 1).

In order to remove the His-tag, the FocB protein was cloned into a vector harbouring a restriction site for the TEV (tobacco etch virus) protease between the His-tag and the *focB* gene. The protein was purified and subsequently the His-tag could be cleaved off. The non-tagged FocB protein had the ability to form crystals of sufficient dimensions in a short time period. The protein was crystallised by the hanging-drop vapour-diffusion method and
crystals grew after one week at 22°C. X-ray diffraction data from these crystals were collected to a resolution of 2.0Å.

**Further analysis of FocB**

The data set obtained from the X-ray diffraction studies of the FocB protein suggested that there were two molecules in the asymmetrical unit. A secondary structure prediction suggested an all-helical structure of seven helices and a CD spectrum indicated an α-helical structure as well (Paper II, Fig. 2).

Furthermore, *in vitro* cross-linking studies were performed in order to investigate the oligomerisation capability of the FocB protein. The results indicated that the dimeric form dominated (close to 40%), although a smaller fraction appeared as trimers (5-10%) (Paper II, Fig. 4). Gel-filtration clearly indicated that the protein is oligomeric and forms dimers or eventually trimers. At this point, we believe that the dimeric form more likely represents the native protein, in accordance with most other prokaryotic transcription factors.
Paper III

Structure of FocB: a member of a family of transcription factors regulating fimbrial adhesin expression in uropathogenic E. coli

The experiments presented in Paper II set the basis for solving the structure of the FocB protein. Further studies, with the help of a structural model of FocB, were aimed at increasing the understanding of the function of PapB/FocB transcription factors in relation to the DNA binding properties of the protein.

Structure of FocB

We solved the structure of FocB at 1.4Å resolution. The R-factor was 0.2026 and Rfree was 0.2296. The structural model revealed a homodimer where each chain consists of five α-helices of different lengths connected by loops. The first four helices and a part of the fifth build up the body of the protein. Helices α4 and α5 of each chain build up a helix-turn-helix (HTH) motif. The crystal packing contacts suggest two possibilities for the homodimeric FocB interactions (Fig. 11).

Fig. 11 Ribbon diagram of the two alternative FocB homodimers. (A) shows dimer “Interface-I”, where chain A is shown in dark green and chain B is shown in light green. (B) shows dimer “Interface-II”, where chain A is shown in dark blue and chain B is shown in light blue. Helices are marked α1-α5. Arg61 and Cys65 are marked in red.
Both dimer alternatives are suggested to be stable in solution, according to computational analyses. However, calculations of the free energy of assembly dissociation and the solvation energy effects indicate that Interface-II is the most probable dimer variant.

Earlier studies of the PapB protein, regulator of P fimbriae transcription and a FocB homolog, revealed amino acids important for DNA binding and oligomerisation of the protein (109). The residues important for DNA binding, Arg61 and Cys65, were positioned in helix $\alpha_4$ of the FocB protein (Fig. 11). In dimer Interface-I, the Arg61 and Cys65 side chains are located in the cleft between the two helix $\alpha_4$ chains. In dimer Interface-II, the side chains are separated and directed outwards from the molecule (Fig. 12, in red). Residues suggested to be important for oligomerisation are shown for Interface-II dimer (Fig. 12). Leu34 and Leu36 are marked in yellow and Tyr74, Phe75 and Ser76 are marked in green. However, how the FocB protein oligomerises along the DNA as well as the mode of interaction with the DNA helix are still unclear.

**Fig. 12** Filled structure model of FocB Interface-II dimer. Conserved residues within the PapB/FocB family of transcription factors that were suggested to be important for DNA binding are shown in red and residues important for oligomerisation are shown in yellow and green.
**DNA binding properties of FocB**

The HTH motif is a common structure of DNA-binding proteins that interact with the major groove of DNA. However, helix $\alpha_5$, which would constitute the recognition helix, is blocked in the dimeric form of FocB. As shown in Fig. 11, Arg61 and Cys65 are differentially located within the two Interface variants and they are both located in helix $\alpha_4$, which precedes the recognition helix $\alpha_5$. Database searches for structural similarities of FocB to other known DNA-binding proteins generated a number of homologues. However, the majority have a preference for binding the major groove of DNA using the classical HTH recognition helix.

By competitive gel mobility shift assays using the minor and major groove binding drugs distamycin and methyl green, respectively, we wanted to assess how FocB could interact with DNA. The results showed that FocB was outcompeted by distamycin, but not by methyl green (Paper III, Fig. 4), similar to what has been shown for PapB (107). This indicates that FocB binds in the minor groove of the DNA.

A common trait of minor groove binding proteins is that they in many cases induce bending of the DNA. However, we have no indications that FocB has that ability. FocB binds in the intercistronic region between the focI and focB genes, a region that has high A/T content and is intrinsically bent (Sondén et. al. unpublished data). The nucleoid-structuring protein H-NS binds in the same region and acts as a repressor of fimbrial transcription, as well as binding the minor groove of the DNA. There is a possibility that FocB has a role in alleviating the repression exerted by H-NS, thus opening the way for fimbrial expression. It could also have a role in stabilising the curvature of the DNA when H-NS is not present.

Taken together, further studies are needed to understand the DNA-binding properties of FocB and an obvious step is to crystallise the FocB protein together with its target DNA. However, the conclusion that FocB is a minor groove binding protein is in good agreement with the structural model.
Conclusions

The work presented in Papers II and III resulted in a crystal structure solution of the FocB protein, creating the first structural model of the PapB/FocB family of transcription factors. The model revealed a homodimer with a classical helix-turn-helix motif common for DNA-binding proteins. The competition assay with the minor groove binding drug distamycin and the apparent block of helix V (the major groove recognition helix) in the dimeric structural model suggest a role for FocB as an oligomeric minor groove binding protein. Furthermore, this is in accordance with the previous mutational analysis of the DNA-binding and oligomerisation properties of the homologous PapB protein.
Paper IV

*Interaction between the PapI protein and the alpha subunit of Escherichia coli RNA polymerase*

Transcriptional regulation of P fimbriae is complex and involves a number of factors. The PapI protein is one of these factors and, similar to PapB, PapI is conserved in a large number of different fimbriae systems. The role of PapI has been studied earlier and it was found to act together with the global regulator Lrp (for references, see the Introduction) as an activator of pap transcription. However, its molecular mechanisms and modes of action are not well understood. In this study, we wanted to investigate the role of PapI in relation to the transcriptional apparatus, specifically the finding that it might interact with the α-CTD of RNA polymerase.

*Enhanced expression of the RNAP α-subunit inhibits pap transcription*

In initial studies, we observed that elevated levels of the RNAP α-subunit inhibited pap transcription. To investigate this finding further, we tested different mutant alleles of the α-subunit. Since the α-subunit is an essential component of RNA polymerase and is required for *E. coli* growth, most mutations would be impossible to generate in the chromosomal locus (*rpoA*) of the bacterium. Instead, many studies have been performed with diploid strain constructs, i.e., mutant forms of the protein are overproduced from plasmids in trans to the wild type α-subunit. Any “dominant” effect(s) by individual mutant proteins can then be scored. By introducing plasmids containing wt and mutant α with alanine substitutions of different residues into an *E. coli* strain with a papBA::lacZ fusion, we found that expression of pap was repressed in the strains carrying the 260A and 265A mutations (Paper IV, Fig. 1). This indicated that these residues might function as contacts between α and some regulatory component involved in pap expression.
CRP is important for the activation of pap transcription, and interactions between CRP and the α-subunit of RNAP have been studied in great detail (14). Both substitutions at positions 265 and 287 cause severe defects in CRP-dependent transcription, while a change at position 260 does not seem to have an effect in that context. The α mutations were subsequently introduced into the UPEC E. coli isolate J96, to study the effect on PapA expression and CRP-dependent transcription. The altered levels of α did not have an effect on CRP expression, as determined by Western blot analyses. The levels of PapA were slightly reduced in the transformants expressing low, uninduced levels, of the 260A and 265A mutants (Paper IV, Fig. 2A). However, stronger effects were observed after inducing α production from a plasmid (Paper IV, Fig. 2B). The effects seen with the J96 transformants were less distinct than what we observed in the papBA::lacZ operon fusion strain constructs. J96 is a clinical isolate that carries gene clusters encoding several different fimbrial systems, including at least two highly homologous pap operons (pap and prs). We monitored PapA expression only at the protein level in J96. Presumably, it will be of interest in further studies to look at the effect(s) at the transcriptional level of the individual fimbriae operons in this strain.

**Interactions between PapI and α-subunit of RNAP**

We were interested in investigating the roles of some of the pap operon regulatory proteins of the in the α-induced repression of pap. We introduced plasmids expressing CRP or PapB and PapI into the papBA::lacZ fusion strain overexpressing α. Additional expression of CRP showed no effect, whereas co-expression of PapB and PapI restored the pap expression to about one third of the level of the control strain (Paper IV, Fig. 3A). The results suggested that α may interact with the pap transcriptional factors, *i.e.*, PapB and/or PapI. However, there is also a possibility that the intercistronic DNA region between papB and papI on the introduced plasmid is involved directly and thus contributes to the effect.

To further investigate the possible interaction between α and PapI, we created a number of amino acid substitutions in
PapI. We selected the mutations based on residues that are conserved among PapI-related fimbrial transcription factors. We introduced the mutant and wt PapI variants into the \textit{papBA::lacZ} fusion strain harbouring the \(\alpha\)-overexpressing plasmid. In particular one mutant, I9A-L10A, seemed less affected by the \(\alpha\) overexpression, compared to the strain containing wt PapI, indicating that the mutation affects the possible interaction with \(\alpha\). However, it is still unclear if PapI mutants form unstable complexes, or perhaps have enhanced interaction with Lrp, leading to increased \textit{pap} transcription. We also introduced the \(\alpha\)-mutants into a \textit{papBA::lacZ} fusion strain that expresses \textit{pap} independently of PapI (DL2121). Interestingly, the expression of \textit{pap} was less sensitive to the presence of mutant \(\alpha\), indicating that the repression of \textit{pap} induced by these mutants is dependent of PapI.

To study the suggested direct interactions between PapI and \(\alpha\) \textit{in vitro}, we constructed a PapI-GST fusion protein and used it in pull-down assay experiments with purified \(\alpha\)-subunit. By binding the GST-PapI to sepharose beads and mixing with \(\alpha\), we observed that the \(\alpha\)-protein eluted together with PapI, after washing away unbound components. This suggests that PapI and \(\alpha\) may interact specifically \textit{in vitro}.

\textit{Conclusions}

The role of PapI in the transcription of \textit{pap} has mainly been studied with respect to its interactions with Lrp. It is a key factor in the phase variation of the \textit{pap} operon and is required for the transition from the phase OFF \(\rightarrow\) ON state, by interacting with Lrp. In this study, we propose a new role for PapI in \textit{pap} expression. It appears that PapI is involved in alternative complexes and our results suggest that it may interact with the \(\alpha\) subunit of RNAP. Elevated levels of \(\alpha\) repressed expression of \textit{pap}. This may occur by sequestering PapI and thus reducing the levels of PapI-Lrp complex needed for \textit{pap} transcription.
Paper V

**Release of the type I secreted haemolysin via outer membrane vesicles from Escherichia coli**

It has previously been observed that the α-haemolysin (HlyA) of *E. coli* can be associated with the outer membrane of the cell, after secretion via the Type 1 secretion system (71). Recent studies show that different toxins associate with outer membrane vesicles (OMVs). The aim of this work was to investigate whether α-haemolysin is associated with OMVs that are released from *E. coli* cells.

**Association of α-haemolysin with OMVs**

In order to find out where the HlyA is located after secretion, we used a laboratory strain expressing the α-haemolysin from a plasmid and the empty vector as a control. OMVs from both strains were purified. By SDS-PAGE and immunoblot analysis using α-HlyA anti-serum, we identified the HlyA protein in both whole cells and OMV fractions from the α-haemolysin-producing strain (Paper V, Fig. 1). By electron microscopy we also observed that the strain producing the α-haemolysin released vesicles that varied in size, compared to the smaller uniform vesicles produced from the vector control strain. We also found that a majority of the HlyA was indeed associated with OMVs (~66 %) compared to the free soluble fraction in the supernatant (~34 %).

To determine if the HlyA was physically associated with the OMVs or if it only co-purified, we performed a series of dissociation tests (Paper V, Fig. 2). However, only after treatment with the non-ionic detergent Triton X-100, the HlyA was completely dissociated from the vesicles as a result of vesicle membrane disruption. We concluded from these experiments that the HlyA is tightly associated with the OMVs released from the bacteria.
The OMV-associated α-haemolysin is active

To further investigate the role of haemolysin-associated OMVs, we wanted to test whether the HlyA was cytolytically active. This was tested by measuring the release of haemoglobin after incubation with horse red blood cells. Haemolysis of the red blood cells was clearly observed after incubation with OMVs from the α-haemolysin-producing strain, but no such observation was made with the corresponding vector control (Paper V, Fig. 4A). The haemolytic activity of OMVs was also completely dependent on Ca$^{2+}$, which is in accordance with the Ca$^{2+}$ dependency for activity of HlyA.

The activity of α-haemolysin was also tested by inducing cell detachment of a HeLa cell monolayer. Incubation with whole cells, filtered supernatants (including OMVs) or OMV preparations from the α-haemolysin-producing strain showed a 50% level of detachment after 90 min, compared to the vector control strain (Paper V, Fig. 4B). Thus, these results show that the α-haemolysin associated with the OMVs was active on both lytic (red blood cells) and non-lytic (HeLa cells) assay systems.

Altered morphology and protein content of α-haemolysin-containing OMVs

The observed differences in size of the OMVs isolated from the α-haemolysin-producing strain made us hypothesise that the larger vesicles might contain α-haemolysin. To test this hypothesis, we performed a density gradient centrifugation of the OMV preparations and analysed the protein content of the different fractions obtained.

As revealed by Western blot analyses, the OMV marker OmpA, an outer membrane protein, showed a peak in two fractions No. 9 and No. 15, indicating two different populations of OMVs (Paper V, Fig. 7). The lower density fraction of the two (No. 9) also contained the HlyA protein as well as the TolC protein, a part of the Type 1 secretion machinery required for extracellular transport of HlyA. As expected, the lower-density fractions also showed the highest haemolytic activity (Paper V, Fig. 7C). By electron microscopy we could also observe that fraction No. 9 had a larger
portion of large-sized OMVs than OMVs from fraction No. 15. We concluded that these low density OMVs contained the α-haemolysin as they also had the highest haemolytic activity.

**OMV-associated α-haemolysin from other E. coli isolates**

As the data presented so far was restricted to the presence of α-haemolysin in OMVs produced from a laboratory strain, we were interested in knowing how the α-haemolysin is localised in extraintestinal *E. coli* clinical isolates. We studied four different strains from the *E. coli* collection of reference (ECOR) that showed different levels of HlyA expression (51). OMVs were isolated from each of the strains and analysed for the presence of HlyA as well as for their haemolytic activity. Also, the levels of OMV-associated α-haemolysin were estimated. The levels were shown to be highly dependent on the strain, varying from a few percent to over 30% in the case of strain ER60. By atomic force microscopy (AFM) of OMVs from strain ER60, we could also observe the presence of both large (>150 nm) and smaller OMVs (Fig. 13). Taken together, the results from the natural isolates show results similar to what was shown for the *E. coli* K-12 strain studies, and we conclude that the association of α-haemolysin with OMVs is something that can occur in environmental as well as clinical *E. coli* strains.

![Fig. 13 AFM micrograph of OMVs produced from the clinical isolate ER60. The insert highlights an example of the larger vesicles in the preparation. Bar equals 100 nm.](image)
Paper VI

Detection of functional P fimbriae proteins in E. coli outer membrane vesicles

The findings presented in Paper V, that the α-haemolysin can associate with OMVs after secretion to the extracellular milieu, as well as the lack of knowledge about the role for these OMVs after budding off from the bacterial cell surface, prompted us to investigate the possibility of adhesive properties of the E. coli OMVs.

Association of P-fimbrial proteins to OMVs

To investigate the possibility of adhesive properties of OMVs from E. coli, we transformed the laboratory strain MC1061 with the plasmid pBEU74 containing a variant of the pap operon that does not express papA, which encodes the major subunit. The resulting strain expressed the adhesin on the surface, but not on the fimbrial rod. This strain was used for the vesicle preparations to avoid breakage of the fimbrial structures during preparation, causing a loss of possible adhesive properties of the OMVs. The adhesive phenotype was confirmed by haemagglutination assays using the MC1061/pBEU74 strain. The corresponding vector control did not agglutinate human red blood cells.

Characterisation of the OMV preparations revealed the presence of the TolC outer membrane protein in the OMVs from all strains. Furthermore, the P-fimbrial usher protein PapC was detected in whole cells from MC1061/pBEU74, as well as in the OMV preparation from the same strain (Paper VI, Fig. 1). This suggested to us that the OMVs can carry the P-fimbrial adhesin.

To see if the presence of P-fimbrial proteins could be detected in natural isolates, we prepared OMVs from the uropathogenic strain J96. However, regulation of the P fimbriae involves phase variation, which makes the population heterogeneous in its fimbriation and only a sub-fraction of the bacteria are fimbriated. By introducing and expressing the papI or papB gene from a plasmid in J96, the majority of the population will be fimbriated, enabling the study of OMVs specifically from P-
fimbriated bacteria. By immunoblot analysis of the OMVs from the different strains using PapC anti-serum, we detected the PapC protein in OMVs from the strain overexpressing PapI (Paper VI, Fig. 2). This strain also expressed the highest level of PapA protein, which builds up the fimbrial structure.

By density gradient centrifugation, we further analysed the OMVs and we observed that the PapC protein was only detected in certain fractions of the vesicle preparations (Paper VI, Fig. 4). The outer membrane protein OmpA was, however, detected in all fractions.

**Adhesive properties of OMVs**

The presence of PapC in the OMV fractions of both MC1061 and J96 led us to investigate the adhesive properties of the vesicles. By fluorescence labelling of OMVs from MC1061/pPAN5 expressing a short and more rigid mutant variant of the P fimbriae, and its vector control, we visualised their binding properties to latex beads coated with the P-fimbrial galabioside receptor.

We could observe attachment of the OMVs from the P fimbriae-expressing strain, whereas little or no attachment was seen by OMVs from the vector control strain. Furthermore, the attachment could be blocked by the addition of soluble galabiose, the receptor for the P-fimbrial adhesin (Paper VI, Fig. 5).

**Conclusions**

Based on the experiments presented in this paper we suggest that functional fimbriae proteins may be associated with OMVs. We also propose that they may have adhesive properties, mediated by P fimbriae proteins exposed on the surface of the vesicles, and the observed galabiose-specific adhesion indicates that the OMVs are carrying the PapG adhesin protein exposed on the surface as well. Taken together with the results presented in Paper V, it is exciting to speculate about the role of OMVs serving as vehicles for toxins that could be directed to a specific target cell by means of fimbrial adhesins.
As shown by our results with genetically manipulated strains, it is feasible to obtain optimized adherent OMVs by the use of cloned fimbrial gene systems and expression of selected fimbrial transcription factors. This would be desirable *e.g.* for the use of OMVs in vaccine preparations.
4. CONCLUDING REMARKS

The P fimbriae of uropathogenic *E. coli* are under the control of a complex regulatory network. The presence of several different fimbrial operons in the same bacterium makes the regulatory system even more intricate. Coordinating expression of the different fimbrial types at the right time and place is probably important for establishment of the bacteria in the urinary tract. Many studies have shown that there is cross-talk between different fimbrial types that ensures their coordinated expression, and we are just beginning to understand the complexity of this cross-regulation. In this work, I have focused on transcription factors that are involved in the regulation of fimbrial operons, in particular the PapB/FocB family.

The findings indicate that the different PapB-like transcription factors may form heterodimeric/oligomeric complexes that can have an influence on the fimbrial cross-talk. The FocB protein also represses the expression of Type 1 fimbriae. (Paper I)

The crystal structure of FocB was solved. This is the first member of the PapB/FocB family of fimbrial transcription factors to be crystallised. The protein appeared in dimeric form and the results indicate that it binds in the minor groove of the DNA. (Paper II and III)

Investigations of the PapI transcriptional regulator revealed that it may interact with the \( \alpha \)-CTD of RNA polymerase, a feature that needs to be considered in the context of *pap* regulation. (Paper IV)

In the study of outer membrane vesicles, the \( \alpha \)-haemolysin was shown to associate with the OMVs after secretion via the Type 1 secretion machinery. Further analysis of OMVs from a uropathogenic isolate revealed that P-fimbrial proteins could be associated with the OMVs and suggests a possible role for OMVs in directed delivery of bacterial virulence factors to target cells. (Paper V and VI)
ACKNOWLEDGEMENTS

My Ph.D. studies at the Department of Molecular Biology have finally come to an end! Many years have passed and to thank every single one of you who has made this time such a memorable period of my life, would require a thesis of its own. So, to ensure that none of you feels neglected, I would like to say THANK YOU! It has been a true pleasure! I also want to give a special thanks to all co-authors and collaborators contributing to the publications and manuscripts in this thesis.

There are, however, a number of people that I would like to thank personally. First I would like to thank both former and present members of the BEU group, students included, but especially Bernt Eric, my supervisor – thanks for giving me the opportunity to work in your group. I have learned a lot and I’ve developed a lot of independence 😊! Thank you Sun for teaching me all the Japanese I know and for the many nice gatherings at the “BEU-SNW residence” with great food! Monica, for all your support and help – both when it comes to labwork and personal matters – may we never stop having our downtown summer lunches! Annika, you’re one of a kind – who else but you would put spaghetti through your ears or run with plastic bags over your feet when it’s raining 😊. It would never have been the same without you. Berit “dun-di-run”, for making me understand that I’m not the only one who doesn’t know what to do when I grow up, and for joining me at the ceramics course. Ulrika, for nice collaborations and all the chit chats about how to solve the really big issues in life, and Liz for still wanting to make nice crystals with us. Anna Å, for making me want to be as organised as you are (I don’t think I will ever make it though, but I’ll do my best...), Jurate, for your hospitality and sense for nice food! Claudia, my favourite German with your lovely laughter, Connie for always speaking from your heart, Hyun-Sook for your kindness, Ting the little lab-squirrel – I don’t think I ever saw you without a smile on your face, Ryoma for joining us at the morning “fikas” and bringing nice cakes (thank you Ryoma’s wife! (?) and chocolate, Carlos – I miss your Britney Spears covers, Barbro for your sneezing, Tianyan for being such a nice and positive person, Takahiko, you’re the most Japanese I have ever met, Pramod keep smiling, Karolis – so what about the pictures from Japan?? Jörgen, not all dirty jokes go well with lunch...😊, Edmund for being Edmund, Teresa, Jonas and Sakura – I can’t believe that I mistook you for “Sakamura, the new Japanese student in group JJ” 😊, Maria N you are a good friend and a good spirit, I wish you all the best. Lisandro (or should I call you Andy?), Stefan, Jenny,
Sara, Emma for making such delicious cakes, Olena, etc. The SvB group; Sven, Marie, Pär, Christer, Marija for loving my cat, Yngve – I wish you good luck!, Ingela, Lelle, Mari, Elin, Betty, etc. And thanks to Marita, Maria W, Berit N, Helena and Ethel, for being the solid rocks at the department, to Johnny – I especially like your chocolate deliveries, to Sequence-Siv, Media and Marienella – if it were up to me, I would give you a big raise!


This research was supported by project and equipment grants from The Swedish Research Council, The Wallenberg Foundations, The Faculty of Medicine, Umeå University and the Swedish Foundation for International Cooperation in Research and Higher Education (STINT). The work was performed within the Umeå Centre for Microbial Research (UCMR) and was carried out in the framework of the European Virtual Institute for Functional Genomics of Bacterial Pathogens (CEE LSHB-CT-2005-512061) and the ERA-NET project “Deciphering the intersection of commensal and extraintestinal pathogenic E. coli.”

Ett steg i sänder!
Sedan ett steg till!
Det blir härligt att se vad som händer när man går åt det håll man själv vill!

- Kent Andersson
REFERENCES


34. **Hernday, A. D., B. A. Braaten, and D. A. Low.** 2003. The mechanism by which DNA adenine methylase and PapI activate the *pap* epigenetic switch. Mol Cell 12:947-957.


