REGULATION OF THE UROPATHOGENIC *ESCHERICHIA COLI* *tos* OPERON AND ITS IMPLICATIONS FOR AN EXPANSION OF MICROBIAL RECIPROCAL REGULATION BETWEEN ADHERENCE- AND MOTILITY-RELATED GENES

by

Michael David Engstrom

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in the University of Michigan
2016

Doctoral Committee:

Professor Harry L. T. Mobley, Chair
Associate Professor Matthew Chapman
Professor Victor DiRita, Michigan State University
Associate Professor Maria Sandkvist
Professor Michele Swanson

"Our imagination is stretched to the utmost, not, as in fiction, to imagine things which are not really there, but just to comprehend those things which 'are' there."

Richard P. Feynman

To my family.

ACKNOWLEDGEMENTS

I will begin by first expressing my deepest thank you to my dissertation advisor and mentor, Frederick G. Novy Distinguished University Professor Harry L. T. Mobley. Indeed, to have received training from an individual held to such a high esteem and respect by his fellow colleagues and leaders in field of molecular pathogenesis was truly a great honor. In addition, Harry's optimism, sense of humor, generosity, interest in my ideas and work, commendation during my successes, and encouragement and patience during my failures made working with him a privilege and exceptionally well suited to my continued growth as an individual and in my career in the field of microbiology.

I also wish to express much gratitude to the other members of my dissertation committee: Dr. Matthew Chapman, Dr. Victor DiRita, Dr. Maria Sandkvist, and Dr. Michele Swanson. It was always a privilege to meet with such a group of eminent leaders in the field of microbiology at the same time. Furthermore, it was an honor to have this group of scientists take an interest in my work and ideas, provide their own valuable ideas, experiences, and direction to improve my work, and give suggestions and encouragement for my continued growth as individual and for my future career.

I am also grateful to members of the Mobley laboratory and other members of the Department of Microbiology and Immunology. Both Dr. Christopher Alteri and Dr.

David Friedman have always been great sources for discussion of my work, especially owing to their vast knowledge of microbiology. Even when we did not agree with each other, I always felt that our conversations improved my work and caused me to view it work from another, previously unimagined point of view. While all members of the Mobley laboratory have made my journey through graduate school enjoyable, I wish to single out my fellow graduate students Dr. Ariel Brumbaugh, Lauren Cooper, Courtney Luterbach, and Dr. Daniel Reiss for making it especially enjoyable. In addition, I have always enjoyed and appreciated the opportunity to have lengthy discussions with Dr. Chelsie Armbruster, Dr. Sébastien Crépin, Dr. Laura Mike, and Dr. Sargurunathan Subashchandrabose on the nature of physical law, how it applies to microbiology, and how it specifically applies to my work. It was also a great honor to have known and been able to discuss these same topics, first, with Dr. A. Cody Springman. Likewise, her dedication to her work and strength while battling leukemia was truly an inspiration, and she will be greatly missed. Additionally, I must also thank Stephanie Himpsl, who always kept the Mobley laboratory running smoothly and ensured that I always had the necessary resources available to perform my experiments. Likewise, I must also thank Heidi Thompson, who has always provided encouragement, ensured that all of my academic matters were taken care of, and rapidly diffused any issues (many of which I am sure that I potentiated) that arose while I was graduate student. Sheryl Smith, who always provided support in scheduling meetings with Harry, making travel arrangements for conferences, and did numerous other things behind the scenes, also deserves many thanks.

Finally, I wish to thank my parents, Kim and Cindy Engstrom, and my younger brother Christopher Engstrom. Without their constant and selfless dedication, support, and sacrifice, I can confidently say that I would not be the person I am today and would not have been able to advance this far in my career. In particular, their dedication to their work, hardworking mentality, and strong morals have served as an inspiration and something to emulate. In addition, it was my family that originally fostered my interest in science, for which I shall be forever grateful.

Michael David Engstrom
University of Michigan
February 2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ACP	acyl carrier protein
asialo-GM_1[GgO_4Cer]	gangliotetraosylceramide
asialo-GM_2[GgO_3Cer]	gangliotriaosylceramide
BIg	bacterial Ig-like fold
bp	basepair
c-di-GMP	cyclic-di-GMP
C-terminus	carboxy-terminus
CFU	colony forming unit
CNF1	cytotoxic necrotizing factor 1
DIG	digoxigenin
EMSA	electrophoretic mobility shift assay
ExPEC	extraintestinal pathogenic *Escherichia coli*
fimS	Promoter element containing P_{fimA}
Fis	factor for inversion
FISSEQ	fluorescent *in situ* sequencing
GalCer2	β1-linked galactosylceramide 2
GASP	growth advantage at stationary phase
GI	gastrointestinal
GlcCer	glucosylceramide
H-NS	heat-stable nucleoid structuring protein
HlyA	α-hemolysin
HU	histone-like protein from *Escherichia coli* strain U93

ICE	integrative and conjugative elements
IHF	integration host factor
ITC	isothermal calorimetry
IVIAT	*in vivo*-induced antigen technology
kb	kilobase
LB	lysogeny broth
Lrp	leucine-responsive regulatory protein
N-terminus	amino-terminus
NAP	nucleoid-associated protein
nLc_4Cer	paragloboside
NS-left	nonstructured-left
NS-right	nonstructured-right
ORF	open reading frame
Ori	origin
PAI	pathogenicity island
Pap	pyelonephritis associated pilus
Pic	protein involved in colonization
ppGpp	guanosine tetraphosphate
RACE	rapid amplification of cDNA ends
rdar	red, dry, and rough
RTX	repeats-in-toxin
Sat	secreted autotransporter toxin
SELEX	systematic evolution of ligands by exponential enrichment
SeqFISH	sequential *in situ* hybridization
SPATE	serine protease autotransporter of *Enterobacteriaceae*
SPI	*Salmonella* pathogenicity island
Ter	terminus
TosA	type one secretion protein A
Tsh	temperature-sensitive hemagglutinin

UPEC	uropathogenic *Escherichia coli*
UTI	urinary tract infection
Vat	vacuolating autotransporter toxin
YESTA	yeast extract and tryptone agar

ABSTRACT

Urinary tract infections (UTIs) are common bacterial infections, affecting nearly half of all women. The vast majority of UTIs are caused by a heterogeneous collection of extraintestinal pathogenic *Escherichia coli* (ExPEC) strains referred to as uropathogenic *E. coli* (UPEC). Unlike some pathogens, UPEC do not have a fixed set of virulence and fitness factors associated with uropathogenesis. Thus, biological covariance especially predominates in uropathogenesis, and UPEC use a variety of virulence and fitness factors to perform essentially the same functions (colonize the urinary tract and establish infection). There are also a number of genetic networks connecting UPEC virulence and fitness genes, and gaining any traction understanding UPEC virulence genes and virulence gene regulation forces one to gain a better understanding of these networks. One of these networks is the reciprocal regulation network connecting *E. coli* motility- and adhesin-encoding genes. It is widely believed that when adhesin genes are expressed, motility genes are repressed, and when motility genes are expressed, adhesin genes are repressed.

The UPEC strain CFT073 genome includes the *tos* operon, which encodes the regulatory, secretion, and adherence machinery of a repeats-in-toxin (RTX) nonfimbrial adhesin, TosA (for *t*ype *o*ne *s*ecretion protein A). This nonfimbrial adhesin promotes

UPEC adherence to host cells derived from the upper urinary tract and is expressed primarily *in vivo*. However, little else was known about why the *tos* operon is preferentially repressed *in vitro*. I hypothesized that regulators encoded in the *tos* operon, regulators encoded elsewhere in the *E. coli* genome, and environmental conditions encountered in the urinary tract mediate *tos* operon regulation. It is also my hypothesis that reciprocal regulation of adhesin and motility genes is an integral part of *tos* operon regulation, and regulators encoded in the *tos* operon are participants in this regulatory network. Using a variety of *in vitro* approaches, I identified that TosR, a member of the PapB family, is a *tos* operon dual positive and negative regulator. In addition, I found that the *tos* operon promoter, P_{tos}, is located upstream of *tosR*, and there are at least two TosR binding sites in the vicinity of P_{tos}. Nucleoid-associated proteins H-NS and Lrp, both associated with adherence and motility reciprocal regulation, also serve a function in negative and positive regulation of the *tos* operon, respectively. High levels of leucine inhibit positive regulation of the *tos* operon. Therefore, the *tos* operon may also be responsive to environmental conditions encountered in the urinary tract (low leucine). TosEF, encoded by two genes in the *tos* operon, were found to suppress motility by inhibiting FliC production, and TosR additionally inhibits P fimbria production while promoting curli and/or cellulose production. Thus, *tos* operon regulation is also coupled to adherence- and motility-related gene reciprocal regulation. Our work has both explained *tos* operon regulation and expanded current knowledge of reciprocal regulation to now include RTX nonfimbrial adhesin genes.

CHAPTER 1

INTRODUCTION

Urinary tract infections

The burden to public health

Nearly half of all women will experience at least one urinary tract infection (UTI) in their lifetime (1), and almost a quarter of all women experiencing one UTI will experience recurrent UTIs (2). Coincidental with the prevalence of UTIs, these infections result in numerous visits to healthcare providers resulting in annual direct costs of at least $3.5 billion (3). This figure, however, would greatly increase if indirect costs of UTI were included (4-6). In addition, with the rise of antibiotic resistance among organisms causing UTIs (6-9), both direct and indirect costs will surely increase as these infections become refractory to classical treatments.

Sites of infection and identification

The urinary tract itself can be divided into the upper and lower urinary tract. The upper urinary tract consists of the kidneys and ureters, and the lower urinary tract consists of the bladder and urethra. UTIs occur in either division of the urinary tract, but most

frequently UTIs occur when bacteria colonizing the GI (gastrointestinal) tract access the urethra and ascend to the bladder (1, 10, 11). Bladder colonization by an uropathogen usually results in a self-limiting infection, cystitis (12, 13). However, bacteria can ascend to the kidneys leading to a more severe infection, acute pyelonephritis (12-15). It is also possible for bacteria to cross the kidney tubules and capillary endothelium into the bloodstream, leading to bacteremia and potentially fatal urosepsis (12-15).

Signs and symptoms of UTI vary among individuals, especially when considering the mental state, but in the case of lower UTIs these might include urinary urgency, pain associated with voiding urine, low-grade fever, and noticing changes in urine characteristics (*i.e.,* noticing blood, strong odors, and urine turbidity) (16, 17). In the case of upper UTIs, additional signs and symptoms might include high-grade fever, malaise, flank pain, nausea and vomiting, and renal impairment (17-19). In addition to a history of these signs and symptoms, diagnosis of a UTI often occurs by microscopic observation of neutrophils and bacteria in the urine, detection of nitrites and leukocyte esterases with a urinalysis dipstick, and positive urine cultures (lower limit $\geq 10^2$ CFU/mL [colony forming units per mL]) (16, 17).

Classification and causative organisms of UTIs

UTIs can be classified as complicated or uncomplicated. Complicated UTIs occur in association with urinary tract obstruction, presence of instrumentation (*e.g.,* a catheter), impaired voiding, metabolic abnormalities, and being immunocompromised (20, 21). In addition, a UTI during pregnancy is sometimes classified as a complicated UTI (21), but

debate about this classification remains ongoing. An uncomplicated UTI occurs in an otherwise healthy individual (*i.e.,* in the absence of the complicating features above). While numerous bacterial species cause UTIs, species often include *Escherichia coli*, *Proteus mirabilis*, *Klebsiella pneumoniae*, and *Staphylococcus saprophyticus* (22-24).

Uropathogenic *E. coli* (UPEC) and the plurality of virulence factors

UPEC are extraintestinal pathogenic *E. coli* (ExPEC)

The overwhelming majority of uncomplicated UTIs (≥80%) are caused by uropathogenic *E. coli* (UPEC) (1, 13, 22, 24, 25). As the urinary tract is not directly continuous with the gastrointestinal tract, by definition UPEC form a subgroup of extraintestinal pathogenic *E. coli* (ExPEC) (26, 27). However, of all ExPEC strains, many colonize sites outside of the urinary tract (26, 27); therefore, the UPEC-ExPEC equivalence is not generalizable (*i.e.,* not all ExPEC are UPEC). In principle, ExPEC can be found in any one of the five *E. coli* phylotypes (referred to as A, B1, B2, D, and E). Nevertheless, most known UPEC (and ExPEC) strains are part of the B2 and D lineages (25, 28-30). Indeed, among the most studied UPEC strains (CFT073, UTI89, and 536), all are part of the B2 lineage (30). UPEC strains CFT073 and 536 are pyelonephritis isolates, and UTI89 is a cystitis isolate (31).

UPEC diversity

While all UPEC are *E. coli*, there is substantial diversity among *E. coli* strains (32-36). Indeed, UPEC strain CFT073 has an extra 590,209 bp of chromosomal DNA, compared with the commensal *E. coli* strain MG1655 (36). One source of such diversity is horizontal gene transfer, whereby *E. coli* acquires foreign DNA through plasmids, other episomes (*e.g.*, ICEs [integrative and conjugative elements]), and phage-mediated transduction (37, 38). Many fitness or virulence factors are often encoded in genes harbored on pathogenicity islands (PAIs), which are predicted to be horizontally acquired genetic elements, 30-100 kbs (kilobases) long, with GC content typically differing from that of the *E. coli* backbone chromosome (39-43). Of note, but not surprising given the genetic diversity discussed above, a well-defined core complement of virulence and fitness factors does not exist among UPEC strains (32-36, 40). Indeed, some UPEC virulence or fitness factors are present among avirulent commensal *E. coli* strains (33-35, 40). Nevertheless, some virulence and fitness factors are outlined in **Figure 1-1** and include iron acquisition systems, secreted toxins, specific metabolism pathways, motility systems, fimbrial adhesins, and nonfimbrial adhesins (23, 31-36, 39, 40, 44, 45). Thus, it can be seen that gaining a complete picture of urovirulence and uropathogenesis requires one to focus on the combination of many distinct types of virulence factors and possibly complex regulatory networks connecting them. In this work, to remain sharply focused, only secreted toxins, metabolism, motility systems, fimbrial adhesins, and nonfimbrial adhesins will be discussed.

Secreted toxins

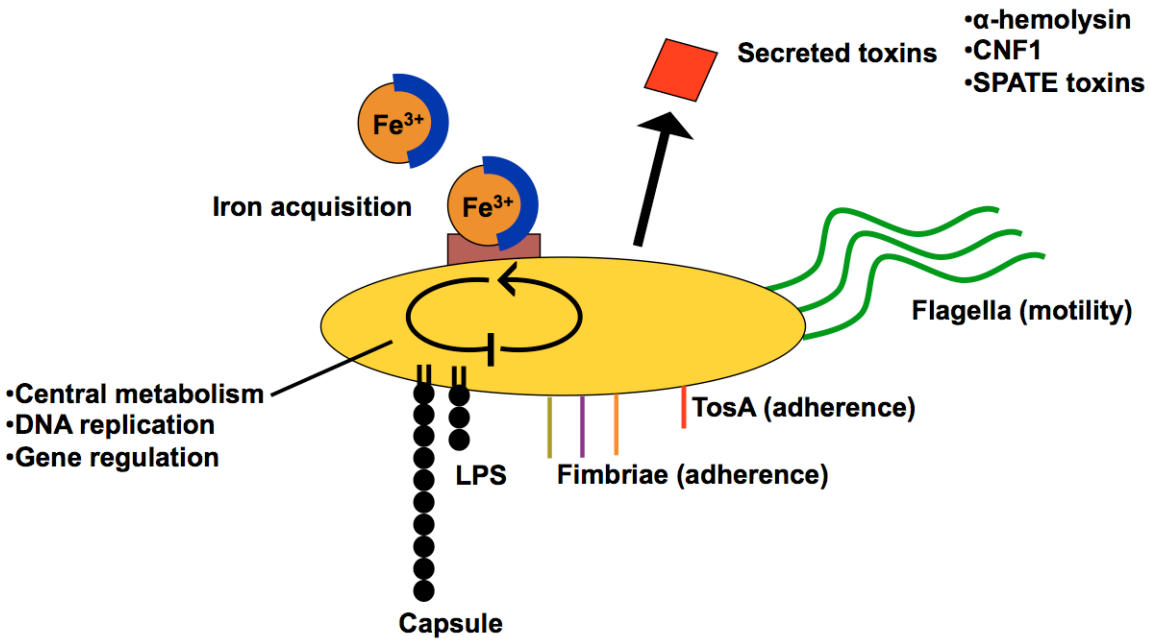

Figure 1-1. UPEC have a variety of virulence and fitness factors to promote UTI. Factors classically associated with UPEC virulence and fitness include flagella, adhesins, LPS (lipopolysaccharide), capsule, iron acquisition systems, and secreted toxins. Also depicted is a representation of systems not classically associated, or associated in nontrivial ways, with UPEC virulence. These systems include central metabolism, cell and DNA replication, and sophisticated gene regulation networks.

The hemolytic activity of some extraintestinal *E. coli* isolates is a well-known phenotype, associated with virulence, and is mediated through HlyA (α-hemolysin) (46, 47). Importantly, HlyA is the prototype of the RTX (repeats-in-toxin) family of secreted proteins (48). The RTX family itself harbors a diverse array of proteins with functions that include cell lysis, cell-to-cell adherence, lipase activity, and metalloprotease activity (49, 50). All RTX proteins are identified by a glycine- and aspartate-rich repeat (GGXGXD) near the C-terminus (carboxy-terminus) of the protein (48-51) and are unified by a single method of secretion (type 1 secretion) (49). Others have already speculated that the significance of these repeats is to coordinate the binding of divalent metal cations (particularly Ca^{2+}), which facilitates extracellular folding (49, 50). HlyA, in particular, contains 13 tandem repeats of the RTX sequence near its C-terminus, and extracellular Ca^{2+} is associated with HlyA activity (51-54).

The genes necessary for hemolytic activity and secretion of HlyA are harbored in the *hly* operon (*hlyCABD*) (55). HlyC, in conjunction with the ACP (acyl carrier protein), mediates acylation of HlyA (56-60). Acylation of HlyA (at K564 and K690) is necessary for host membrane binding and possibly HlyA oligomerization (59-62). HlyA is the RTX cytolysin encoded by the operon. In addition to erythrocyte lysis (53, 54), this protein is cytotoxic toward granulocytes (63, 64), monocytes, and to a lesser extent lymphocytes (64). At the same time, HlyA promotes lysis of renal proximal tubular cells (65) and promotes damage and hemorrhaging in the uroepithelium and bladder during experimental infection (66). In addition to cytolysis, HlyA may perturb host cell membrane signaling through Akt, a protein kinase. In particular, HlyA stimulates

dephosphorylation of Akt, which perturbs the host inflammatory response and impairs host cell survival (67). Others have shown that HlyA perturbs host cell proteolysis pathways, which affects host cell survival, the host cell cytoskeleton, and the inflammatory response (68).

HlyA secretion depends on HlyB, HlyD (both encoded in the *hly* operon), and TolC (encoded at another locus) (69-72); these three genes form the basis of a prototypical type 1 secretion system (69). The order of HlyA secretion events follows. HlyB dimers recognize a poorly characterized signal at the C-terminus of HlyA (69, 73, 74); at least some of the recognition of the HlyA secretion cargo, in addition, may be through an HlyB-HlyD complex (69, 75). Upon ATP binding, an HlyB conformational change occurs (69, 74, 75), and HlyA enters the HlyD periplasmic channel with limited exposure to the periplasm itself (69, 75-77); trimers of HlyD form a stable complex with the outer membrane pore protein, TolC (69, 75). While in the HlyD secretion channel, it is predicted that HlyA begins folding into its native conformation, partially assisted by HlyD itself (78). Subsequently, HlyA exits the channel through TolC, and the secretion system is ready for another round of secretion following ATP hydrolysis in the HlyB dimers (69). At some point, the proton motive force may contribute to HlyA secretion, but the precise mechanism of this contribution is poorly understood (69, 70).

Additional toxins found in some UPEC strains include CNF1 (cytotoxic necrotizing factor 1) and the SPATE (serine protease autotransporter of *Enterobacteriaceae*) toxins (23, 44). CNF1 causes cytopathic effects in a variety of cell types derived from the upper and lower urinary tract (79). During an experimental UTI,

CNF1 also promotes bladder inflammation and submucosal edema (66). Many of the CNF1-mediated host cell toxic effects proceed through reactive oxygen species-mediated damage to the uroepithelium; these reactive species are hypothesized to originate from neutrophils succumbing to actin cytoskeleton remodeling by CNF1 (80, 81).

Sat (secreted autotransporter toxin) is a SPATE family member and toxin that mediates vacuole formation in cultured cells derived from the upper and lower urinary tract (82, 83). Of note, kidney cells appear to be markedly susceptible to Sat, suggesting that Sat may have some function during upper UTI. Pic (protein involved in colonization) is another SPATE family member expressed during UTI and associated with around 15% and 31% of cystitis and pyelonephritis strains, respectively (84). This suggests that Pic serves some important function during UTI. Vat (vacuolating autotransporter toxin), similar to Sat, is another member of the SPATE family that mediates vacuolation of host-derived cells (85). Tsh (temperature-sensitive hemagglutinin), although considered a SPATE protein and exhibits some proteolytic activity (86), mediates *E. coli* adherence to erythrocytes, hemoglobin, and extracellular matrix proteins (87-89).

Metabolism

Although not traditionally classified as a mediator of virulence, UPEC metabolism in the urinary tract is a critical contributor to uropathogenesis (90, 91). Indeed, gluconeogenesis and the TCA cycle are required during UTI, but are predicted to be less important during colonization in more nutrient-rich sites, such as the GI tract (90,

91). Amino acid catabolism is important for UPEC growth in urine and the urinary tract (90, 91). It is of note that metabolic state (*i.e.,* growth in nutrient deplete or replete conditions) can be communicated to the UPEC cell in the form of DNA supercoiling and differential nucleoid geometry (*i.e.,* promoted by different nucleoid structuring proteins discussed below), both of which contribute to gene regulation (92-109). Thus, a single "metabol-nucleoid" phenomenon integrating metabolism, DNA supercoiling, and geometry of the nucleoid may be contributing to how genes are regulated in a variety of situations *E. coli* may encounter.

Motility

Motility is an important phenomenon during UTI. Indeed, nonmotile UPEC mutants are outcompeted by the parental wild-type strain in a murine model of ascending UTI (110, 111). For instance, nonmotile UPEC cells colonize the kidneys at low levels (112). In this same pursuit, imaging studies have shown that flagellar gene expression is coincident with ascension into the upper urinary tract and transient in nature (112). Intriguingly, however, flagellar genes are downregulated during an experimental UTI (113) and human infection (114). However, examining gene expression during experimental and human UTI, at present, only accounts for a population-level description of gene expression (*i.e.,* there is no temporal or spatial information in these data). Therefore, it is predictable that population heterogeneity exists for flagellar gene expression. Likewise, a kinetics study following bladder and kidney colonization during an experimental UTI found heterogeneity among sequence tagged lineages arriving in the

kidney (*i.e.,* not necessarily reflective of the abundance of constructs in the bladder) (115). This also supports a model of flagellar gene expression heterogeneity, as this organelle is important for kidney colonization.

Flagellar assembly occurs in a highly organized fashion. Indeed, three discrete flagellar gene classes (class I, class II, and class III) are responsible for flagellar gene expression and assembly (116-119). Class I genes encode the master flagellar operon regulators FlhD and FlhC (116-119). A multimeric complex of FlhDC ($FlhD_4C_2$) is responsible for activation of class two flagellar genes (116-119). Class II flagellar genes encode FliA (σ^{28}), FlgM (anti-σ^{28}), the flagellar basal body, and the flagellar hook (116-119). Upon basal body and hook assembly, FlgM is secreted from the cell, and FliA promotes the expression of class III flagellar genes (116-119). Class III genes encode proteins involved in hook-filament assembly, the flagellar filament (FliC), the flagellar cap (FliD), the flagellar motor proteins (MotAB), and several proteins involved in bacterial chemotaxis (116-119).

Flagellar-mediated motility, often coupled with chemotaxis, occurs when *E. coli* peritrichous flagella are bundled together through coherent counterclockwise flagellar and flagellar motor rotation (120, 121); noncoherent and clockwise rotation of flagella results in tumbling motion of the bacterium (120, 121). Methyl-accepting chemotaxis proteins, involved in the sensing of attractants and repellants, include Trg, Tar, Tsr, and Tap (122-126). Trg senses saccharides (122, 124); Tar and Tsr sense amino acids (122, 123, 126); Tap senses dipeptides (122, 125). It is of note, however, that the *trg* and *tap* are often not present or severely mutated in UPEC strains (127), which suggests some

common ancestry among UPEC strains and that these receptors are not necessary for UPEC lifestyles. Upon attractant binding to a chemotaxis sensor above, a phosphorylated CheB will mediate demethylation of one of these sensors and reduce the activity of the CheA kinase (120, 121). Reduced CheA kinase activity will result in less CheY phosphorylation (enhanced by CheZ dephosphorylation), and the flagellum and flagellar motor will remain in counterclockwise motion (120, 121). However, upon binding of a repellent (or attractant diffusion from the chemotaxis sensor) CheR promotes methylation of this sensor and continued kinase activity of CheA (120, 121). This in turn promotes CheY phosphorylation and clockwise rotation of the flagellum and flagellar motor (120, 121).

It is intriguing to speculate that flagellar-mediated motility, like flagellar assembly, is highly coordinated, especially with other UPEC lifestyles. Indeed, flagella are highly immunogenic and recognized by toll-like receptor 5 (128). Therefore, it would be of some advantage to *E. coli* if part of the cell population could swim (albeit be detectable by the host immune system) and another part of the population be nonmotile for a period of time (but refractory to some host immune system sensors) (129). In addition, as reviewed below and in **Figure 1-2**, adherence and motility are reciprocally regulated. Thus, this is another manner in which flagellar-mediated motility could be highly coordinated, especially with other UPEC lifestyles.

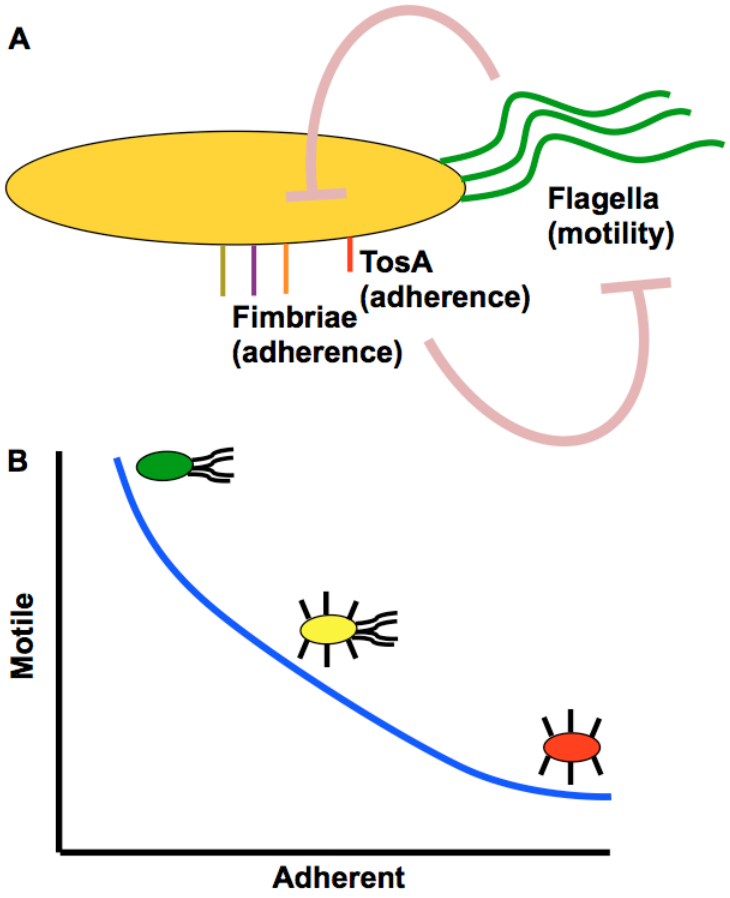

Figure 1-2. Motility and adherence are coordinated phenotypes. **(A)** A representation of the phenomenon of motility and adherence reciprocal regulation is depicted. Pink inhibition bars represent the possibility that some regulation could be indirect. **(B)** A predicted representation of motility and adherence reciprocal regulation as a curve is shown. Motility and adherence are abstracted to form orthogonal axes, and cellular behavior is abstracted to the indicated blue curve. Green cells with multiple flagella represent a regime where the cells are preferentially motile and less adherent; red cells showing multiple surface adhesins represent the regime where cells are preferentially adherent, and yellow cells with flagella and surface adhesins represent a regime where cells may transiently show limited preference between the other two regimes. This third regime could represent a transition period (*i.e.,* into another regime or back into a previous one).

c-di-GMP (cyclic-di-GMP) and motility

The universal bacterial second messenger c-di-GMP (cyclic-di-GMP) contributes to the coordination of flagellar-mediated motility (130-132). Enzymes with diguanylate cyclase activity, often containing a conserved GGDEF (glycine-glycine-aspartate-glutamate-phenylalanine) motif, mediate the formation of c-di-GMP from two GTP molecules; enzymes with phosphodiesterase activity, often containing a conserved EAL (glutamate-alanine-leucine) motif, mediate c-di-GMP degradation (130-132). At high c-di-GMP levels, flagella-mediated motility is reduced (130-132). This effect is mediated through both reducing flagellar levels (130-132) and decreasing flagellar motor rotation (133). However, regulation by c-di-GMP binding and metabolism enzymes is complex. For example, YdiV, containing a degenerate EAL domain, may instead serve a function during the repression of flagellar operons (134). In addition, any one bacterial species may harbor many genes encoding proteins with diguanylate cyclase and phosphodiesterase activities (131, 135, 136), furthering this regulatory complexity. Nevertheless, most diguanylate cyclases and phosphodiesterases are linked either directly or indirectly (such as through two-component systems) to a variety of environmental signals (130, 131, 137). The levels of c-di-GMP also rise upon entry into stationary phase (138), which suggests that some of these systems may be coupled to cell growth phase and reduced motility during stationary phase (139). The effect of c-di-GMP levels on adhesin operons will be discussed below.

Adhesins

UPEC adherence is an important virulence property during UTI, especially to overcome the flow of urine. It is of note that adherence, especially in the context of a biofilm, is a key component of microbial multicellularity (131, 140-142). Of all UPEC adhesins, fimbrial adhesins have been the most extensively studied. Some UPEC strains encode genes for the synthesis of as many as 11 distinct fimbrial adhesins (34, 36). Indeed, a plethora of diverse fimbrial operons is typical of UPEC strains, compared with nonpathogenic *E. coli* (34). UPEC strains encode an average of eight fimbrial types, compared to an average of three fimbrial types for commensal *E. coli* (34). Fimbriae, like flagella, are multimeric extracellular organelles and are assembled through a chaperone-usher pathway (143). Of the variety of fimbriae synthesized by UPEC, Type 1, P, F1C fimbriae, and biofilm mediated adherence will be reviewed below.

Type 1 fimbria

Type 1 fimbriae are associated with almost all *E. coli* strains (*i.e.,* no differential prevalence among pathogenic and nonpathogenic strains) (144). As with other fimbriae, the *fim* operon and locus harbors genes necessary for regulation of the *fim* operon, assembly of a chaperone-usher secretion system, and fimbrial assembly (145-147). Coupled with nucleoid structuring proteins, FimB and FimE are two recombinases involved in phase variation (flipping) of the *fimS* invertible element, containing P_{fimA}, at two inverted repeat sequences bracketing this element (148-152). FimB mediates the ON-to-OFF or OFF-to-ON inversions of *fimS*, and FimE preferentially mediates the ON-to-OFF inversion of *fimS* (148-150, 152). The involvement of nucleoid structure in

this process will be discussed below. In addition to FimBE, the chaperone-usher system (assembly machinery), major fimbrial subunit (stalk protein), and tip adhesin are all encoded by the remaining *fim* operon genes (*fimAICDFGH*) (**Figure 1-3A**) (143, 145-147, 153-157). Multimers of FimA comprise the majority of a Type 1 fimbria, and FimFGH form the tip adhesin complex (143, 153, 157). FimH is the major adherence element of Type 1 fimbria (153, 157). The FimH adhesin tip binds to a mannoside moiety found associated with a variety of surface glycoproteins and glycolipids (158-162). Some work has explored the selective pressure on the FimH adhesin tip and urovirulence. Greater urovirulence is seen for some FimH variants conferring stronger mannose binding (163). However, there does appear to be some fitness cost associated with such FimH variants at other sites in a host (163), which suggests that selective pressures and fitness costs are not always uniform at all sites in a host and must be balanced by ExPEC.

Type 1 fimbria is a well-established virulence factor, especially in the murine model of UTI (163-167). Given the near universal prevalence of Type 1 fimbriae (among *E. coli* strains), however, it seems more likely that the combinations of Type 1 fimbria with other virulence determinants are more important than Type 1 fimbria alone. Furthermore, during experimental murine UTI, UPEC strains appear to differ in their ability to control the *fimS* invertible promoter element (166, 168, 169). In particular, F11 (cystitis strain) favors the ON inversion, and the ON inversion is less favored by CFT073 (pyelonephritis strain) late in an experimental infection (166, 169). The importance of Type 1 fimbria during human infection is poorly understood; low expression of the *fim* operon is sometimes observed during human UTIs (114). However, as these expression

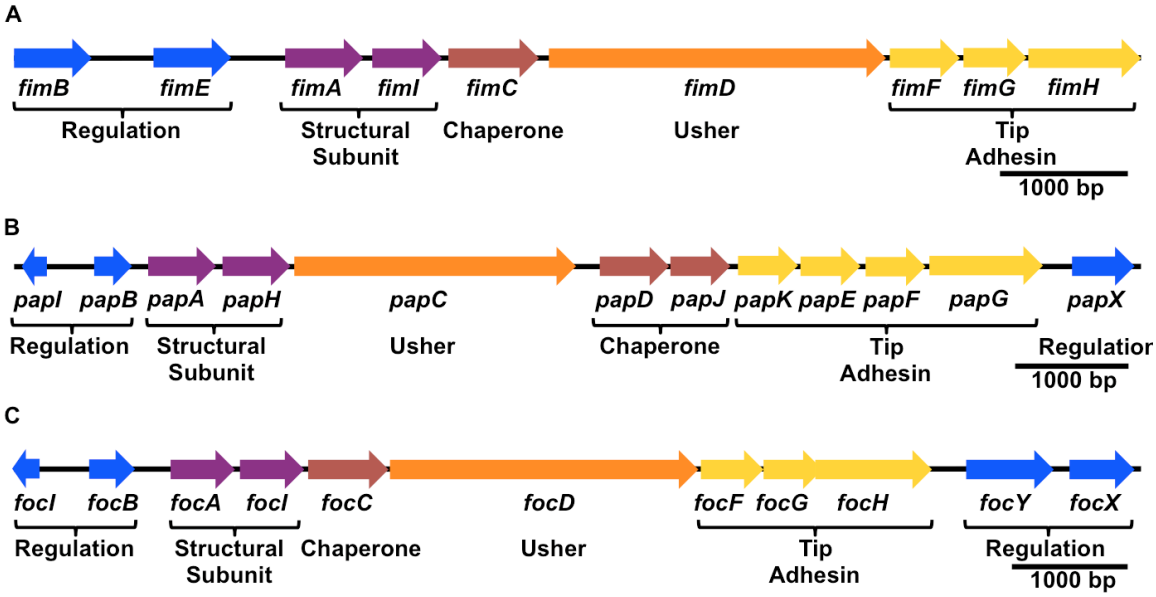

Figure 1-3. Representation of the structure of several adhesin operons. (**A**) The *fim operon* encoding the genes for assembly of Type 1 fimbria is represented to scale. The *fim* regulatory genes are depicted in blue, structural subunit genes are depicted in purple, the chaperone gene is represented in maroon, the usher is shown in orange, and the adhesin tip components are shown in yellow. (**B**) The *pap* operon genes necessary for assembly of P fimbria are represented to scale. As above, the color of each *pap* operon gene represents the function of the encoded protein. (**C**) Genes of the *foc* operon responsible for the assembly of F1C fimbria are represented to scale. The color of each gene in the *foc* operon, as above, represents the function of the encoded protein.

studies lack temporal-spatial data, it seems probable that subpopulations of *E. coli* expressing the *fim* operon may be found during infection.

P fimbria

The genes necessary to produce P fimbria were among the first cloned virulence factor genes (145). Unlike Type 1 fimbria, P fimbriae are epidemiologically associated with pathogenic *E. coli* strains (34, 170-177). As is the case for Type 1 fimbria, nearly all genes necessary for P fimbrial operon regulation, secretion, and assembly are encoded by the *pap* operon (*pap*, *p*yelonephritis *a*ssociated *p*ilus) (143, 178-181). PapI and PapB are two regulators involved in perturbing nucleoid structure to mediated *pap* operon negative and positive regulation (152, 179-181). PapB is hypothesized to overcome H-NS (heat-stable nucleoid structuring protein) silencing by binding near P_{papI} and mediating *papI* expression (positive *pap* regulation) (181). At sufficiently high levels, PapB binds to and represses P_{papBA} (negative *pap* regulation) (180). Thus, PapB control over the *pap* operon exhibits examples of both positive and negative feedback regulation (depending on the DNA binding context) (180, 181). PapB may mediate DNA binding through forming an oligomer on an AT-rich region in the vicinity of both P_{papI} and P_{papBA}, which does not contain inverted repeats often recognized by other regulators (182). Instead, it seems likely that PapB recognizes DNA structure over a specific DNA sequence (182). PapI, on the other hand, acts to control Lrp (leucine-responsive regulatory protein) binding to two GATC sites (Dam-mediated methylation sites) (152, 179, 183). Lrp is both a positive and negative regulator of the *pap* operon (152, 179,

183), and the specific mechanism of this regulation will be explored below. It is also intriguing to speculate that epigenetic regulation of the *pap* operon, through DNA methylation, could couple DNA replication to this process (184).

The remaining genes of the *pap* operon (*papAHCDJKEFG*) encode the chaperone-usher system, major fimbrial subunit, and tip adhesin (**Figure 1-3B**) (143, 178, 185). An additional gene, *papX*, regulates motility as discussed below (186-188). The major structural unit of P fimbria is composed of PapA multimers (143, 178, 185), and the P fimbrial adhesin tip is composed of PapK, PapE, and PapG subunits (143, 178, 185). The PapG tip mediates binding to a α-D-Gal-(1-4)-β-D-Gal moiety, which is also known as the glycosphingolipid-anchored P blood group antigen (189).

Although P fimbria promotes adherence to cells derived from the uroepithelium (190), demonstrating that *pap* operon mutants have virulence defects during experimental infection remains challenging. In independent infections using the murine model of UTI, a weeklong study failed to find that *pap* operon mutants were any less virulent than the UPEC parental strain (191). However, a subsequent signature-tagged mutagenesis screen and coinfection studies identified *pap* operon mutants as being significantly outcompeted by parental and other mutant strains, and this mutation could be complemented with a cloned *pap* operon (192). In addition, a CFT073 *pap* mutant has a reduced rate of kidney infection, compared with the parental CFT073 strain (193). There is also evidence to suggest that the *pap* operon is expressed during a human infection (194), which includes the production of antibodies against P fimbria in infected individuals (195). Thus, taken

together, P fimbria clearly serves a function during a UTI, but this function has been difficult to measure.

F1C fimbria

F1C fimbriae, like P fimbriae, are associated with UPEC strains, compared to nonpathogenic *E. coli* strains (196-199). As is the case for both Type 1 and P fimbriae, all activities and structures necessary for the regulation, secretion, and assembly of F1C fimbria are encoded by the *foc* operon (200-205). The *foc* operon is regulated in a manner similar to the *pap* operon (152, 200-205). FocB, a PapB homolog, controls expression of FocI at P_{focI}, and FocB negatively regulates P_{focBA} at high FocB levels (202, 203). FocI controls Lrp binding to GATC sites in P_{focBA}, which in turn mediates positive and negative regulation of the *foc* operon (152, 200-205). It is of note that FocB and PapB cross-regulate the *pap* and *foc* operons, respectively (203). However, A detailed description of fimbrial operon cross-regulation will be discussed below.

The remaining genes of the *foc* operon (*focAICDFGHY*) encode the chaperone-usher system, major fimbrial subunit, and tip adhesin (**Figure 1-3C**) (200-203, 205). An additional gene, *focX*, may contribute to motility repression (187). The major structural unit of F1C fimbria is composed of FocA multimers (200-203, 205), and the F1C fimbrial adhesin tip is composed of FocG and FocH subunits (200-203, 205). FocH mediates adherence to specific targets containing GlcCer (glucosylceramide), GalCer2 (β1-linked galactosylceramide 2), lactosylceramide, globotriaosylceramide, nLc_4Cer (paragloboside), lactotriaosylceramide, asialo-GM_2[GgO_3Cer] (gangliotriaosylceramide),

and asialo-GM_1[GgO_4Cer] (gangliotetraosylceramide) moieties (206). F1C fimbriae promote adherence to human kidney epithelial cells found in the distal tubules and collecting duct and promote adherence to vascular endothelial cells (207). In agreement with these observations, expression of F1C fimbriae during human infection has been demonstrated (197).

TosA, the first *E. coli* RTX nonfimbrial adhesin

A model adhesin for understanding UPEC adherence regulation is the RTX nonfimbrial adhesin, TosA (TosA, *t*ype *o*ne *s*ecretion protein A). *tos* operon (*tosRCBDAEF*) (**Figure 1-4A)** regulation will also serve as a model of reciprocal regulation between adhesin and flagellar operons. In addition to harboring *tosA*, the remaining genes of the *tos* operon (*tosRCBDEF*) encode a type 1 secretion system (TosCBD) and putative regulators (TosREF) (**Figure 1-4B**) (208). TosC, TosB, and TosD have homology with known type 1 secretion system components TolC, HlyB, and HlyD, respectively (208). TosR is a member of the PapB family, discussed above, and TosEF (originally annotated as *c0364* and *c0365*) are both members of the LuxR family (208). The *tos* operon itself is localized to pathogenicity island-*aspV* (PAI-*aspV*), harbored by UPEC strain CFT073 (39). At least one-in-four UPEC isolates harbor the *tos* operon, and the *tos* operon is a potential predictor of *E. coli* virulence (35).

The RTX adhesins, which include TosA, form a growing group of the RTX family (50). TosA is localized to the outer membrane and specifically mediates adherence to cells derived from the upper urinary tract (208). In addition to TosA,

A

B

Protein	MW (kDa)	Homolog	% AA Identity
TosR	11.7	PapB Family	45
TosC	46.8	TolC	44
TosB	79.0	Lss-like (HlyB)	62
TosD	46.9	HlyD	77
TosA	>250	OM adhesin-like protein	28
TosE	18.5	LuxR/Sigma 70	32
TosF	20.3	LuxR	22

Figure 1-4. The *tos* operon encodes regulator, secretion system, and adhesin homologs. (A) ORFs (open reading frames) are depicted to scale with colored arrows (representing the function of the encoded proteins) in the direction of transcription. Blue arrows represent putative regulator encoding genes, orange arrows represent genes encoding the structural proteins for a putative type I secretion system, and a red arrow represents the gene encoding the RTX nonfimbrial adhesin, TosA. **(B)** The predicted molecular weights (kDa), identified homologs, and percent amino acid identity with the identified homologs are noted for each protein encoded by a gene in the *tos* operon. Image is modified from (208).

membership of this family includes LapA (*Pseudomonas putida* and *Pseudomonas fluorescens*), LapF (*Pseudomonas putida* and *Pseudomonas fluorescens*), BapA (*Salmonella enterica*), SiiE (*Salmonella enterica*), FrhA (*Vibrio cholerae*), BpfA (*Shewanella oneidensis*), RtxA (*Legionella pneumophila*), and *Mp*AFP (*Marinomonas primoryensis*) (50, 209, 210).

Both LapA and LapF serve functions during biofilm development (211-213). In particular, LapA promotes early surface attachment events (including the switch to irreversible surface binding), and LapF promotes biofilm cell-to-cell interactions (*i.e.,* biofilm structure) (50, 211-213). Differential regulation by the nucleoid protein Fis (factor for inversion stimulation), of the genes encoding LapA and LapF, may also serve a function in modifying LapA- and LapF-mediated biofilm formation (214). BapA, like LapA and LapF, is associated with biofilm formation (215). *bapA* expression is coordinated with curli fiber synthesis, and BapA has an additional function during invasive infection (*e.g.,* liver and spleen infection) (215).

SiiE is a very large (>5,000 amino acids in length) RTX adhesin composed almost entirely of BIg (bacterial Ig-like) folds (50, 216). The *sii* operon is on SPI-4 (*Salmonella* pathogenicity island-4), and SiiE is important for cattle intestinal colonization (especially binding to and invading polarized epithelial cells) and is a mediator of inflammation in mice (217, 218). SiiE requires extracellular Ca^{2+} coordination within the BIg folds to maintain a rigid structure (216), which may contribute to SiiE-mediated adherence of *S. enterica* to host cells.

FrhA is found almost exclusively in *Vibrio* species and contains repeated cadherin-like domains (50, 219), which are known to contribute to eukaryotic cell-to-cell interactions (*i.e.,* adherence) (220). The FrhA RTX nonfimbrial adhesin contributes to hemagglutination of erythrocytes, epithelial cell adherence, biofilm formation, chitin binding, and intestinal colonization of infant mice (221). Interestingly, regulators associated with *Vibrio cholerae* flagellar gene expression contribute to positive regulation of *frhA* expression (221). *frhA* negative regulation may occur in response to low levels of intracellular c-di-GMP and suggests that even nonfimbrial adhesins might participate in coordinated adherent and motility phenotypes (221).

BpfA, like SiiE, contains internal BIg folds in addition to other repeated motifs (50). BpfA, like the other RTX nonfimbrial adhesins above, serves an important function during *S. baltica* biofilm formation (222). Biofilm formation mediated by BpfA, in turn, may be suppressed by an interaction with the inner membrane protein BpfD (222). High intracellular levels of c-di-GMP may promote BpfA dissociation from BpfD and promote biofilm formation (222).

RtxA, an RTX nonfimbrial adhesin from *L. pneumophila*, contains a von Willebrand factor type A domain (50, 223), which may facilitate binding to and invasion of amoebae and macrophage cells (50, 224-226). RtxA may also promote *L. pneumophila* intracellular survival after host cell invasion, and, as such, RtxA is a noted virulence factor during experimental murine lung infection (224).

To date, *Mp*AFP (an antifreeze protein), of *M. primoryensis*, is the largest known RTX nonfimbrial adhesin (>13,000 amino acids long) (209, 210). The large *Mp*AFP size

is again due to the presence of repetitive BIg folds, which coordinate Ca^{2+} to obtain adhesin rigidity (209, 210). In terms of adherence, *Mp*AFP binds ice (209, 210). This adherence phenomenon is predicted to mitigate damage caused by cell contact with ice and suspend *M. primoryensis* in the extracellular milieu, possibly where nutrient and oxygen levels are higher (209, 210). Like M*p*AFP and many of the other RTX adhesins above, TosA contains a number of internal repetitive motifs (208). In particular, TosA contains at least five tandem direct repeats 335 amino acids long (208).

Individual challenge and competition assays have revealed that TosA is both an important virulence and fitness factor, especially during an invasive infection (*i.e.,* it mediates kidney, liver, and spleen infection) (35, 39, 208, 227). In addition, *E. coli* CFT073 cells producing TosA enhanced lethality in a zebra fish model of bacteremia, compared with *E. coli* CFT073 not producing TosA (208). Furthermore, mice pre-vaccinated with TosA were protected from invasive infection of the spleen when challenged with *E. coli* CFT073 by the transurethral route (208).

TosA was previously identified through an IVIAT (*in vivo*-induced antigen technology) screen (227), which can be used to identify genes preferentially expressed *in vivo* (228). In agreement with the IVIAT finding, *tosA* and the remainder of the *tos* operon were repressed during growth under laboratory conditions (*i.e.,* in lysogeny broth [LB] at 37°C) (208, 227), while induced *in vivo* (as estimated from pooled voided urine from experimental murine UTIs) (208, 227). Little else was previously known about *tos* operon regulation, but considering that TosA is an adhesin, this may suggest that

regulation of the *tos* operon could be related to the same phenomenon governing motility and adherence reciprocal regulation.

c-di-GMP and adherence

The bacterial secondary messenger c-di-GMP contributes to adhesin regulation and biofilm formation (130-132, 142). *E. coli* biofilms are networks of cells, extracellular matrix proteins (often adhesins), DNA, and extracellular polysaccharides (140-142, 229-237). Biofilms, a form of multicellular behavior, are associated with microbial adherence to both biotic and abiotic surfaces (141, 142), and these structures have gained some notoriety for imparting an antibiotic-resistant phenotype, being the source of some persistent infections, and causing environmental problems (141, 142). High levels of c-di-GMP are related to higher levels of *E. coli* adherence and multicellularity (130-132, 142). Indeed, diguanylate cyclases YdaM, YegE, YedQ, and YeaP function to increase CsgD levels, which in turn increases curli fiber and cellulose levels (both components of *E. coli* biofilms) (131, 140, 142, 229-236). Supplemental to CsgD, diguanylate cyclase YaiC may contribute to increasing cellulose levels (131). Of note, many of these enzymes are under the regulatory of control of σ^S (131), which may explain the rise of c-di-GMP levels upon entry into stationary phase (138). As lower levels of c-di-GMP reduce adherence, enzymes with phosphodiesterase activity or degenerate EAL domains function by decreasing the levels of curli, cellulose, Type 1 fimbria, and P fimbria (131, 134, 238).

Reciprocal regulation between adhesin and motility gene expression

As noted above, adhesin and flagella expression are coordinated (*i.e.,* they are reciprocally regulated) (**Figure 1-2**). Intuitively, this coordination seems logical; it is improbable (and highly counterproductive) for a UPEC cell to both be maximally anchored to a specific site (nonmotile) and simultaneously swimming in the extracellular milieu (motile). A great deal of this coordinated behavior in *E. coli* is mediated through proteins encoded by genes at the end of adhesin operons (*e.g., papX* and *focX*) (186-188). Indeed, a *papX* mutation was previously identified as a suppressor mutation of a nonmotile *fimS* locked-ON mutation (188). PapX mediates a decrease in motility by binding in the vicinity of P_{flhDC} and repressing expression of *flhDC*, which encodes the master positive regulator of flagellar gene expression as described above (187, 188). FocX is virtually identical to PapX (187). Therefore, it is reasonable to hypothesize that FocX also contributes to decreased motility in a similar manner. In addition, regulators encoded by nearly all adhesin operons of another uropathogen, *Proteus mirabilis*, mediate decreased motility when expressed (186, 239). Perhaps the most studied of these motility regulators in *P. mirabilis* is MrpJ, a homolog of PapX (186, 239, 240).

To date, the manner in which flagella and flagellar operons decrease adhesin expression is unclear, but attractive mechanisms may include perturbing intracellular c-di-GMP levels, nucleoid-mediated regulation, and the activity of bundled flagella themselves. FliA, which controls expression of class III flagellar genes, may control expression of the gene encoding phosphodiesterase YhjH (131, 241). In turn, YhjH promotes c-di-GMP degradation and subsequent repression of curli fiber expression (131,

241). Some of the reciprocal regulation between adhesin and flagellar genes may be the result of switches between NAPs (nucleoid-associated proteins), which will be reviewed below. The activity of bundled flagella themselves could also mediate downregulation of adhesin operons. For instance, flagellar activity reduces fluid shear viscosity (242). In turn, fluid shear could be sensed by *E. coli* (243, 244). Indeed, fluid shear has already been implemented in the regulation of Intimin and Tir synthesis and activity (adhesin and receptor proteins of enteropathogenic and enterohemorrhagic *E. coli*) (244) and the enhanced adherence phenotype of Type 1 fimbria (245). In addition, P fimbria is resistant to shear forces (possibly from the flow of urine) due to fimbrial stalk coiling (246). High c-di-GMP levels suppress flagellar motor rotation (133), which could further drive cells toward an adherent state.

Another aspect of reciprocal regulation is the coordination between different adhesin operons. One example of this form of regulation is between PapB (P fimbria) and Type 1 fimbria (203, 247). Indeed, earlier work found that PapB represses the *fim* operon through binding in *fimS* (247). In a manner similar to PapB, FocB negatively regulates the *fim* operon (203, 248). As noted above, FocB and PapB cross-regulate each other (203). Cross-regulation of the *foc* and *pap* operons is achieved through specific binding of P_{focBA} and P_{papBA} by both PapB and FocB (203). Although the mechanism of cross-regulation can only be speculated to be through specific DNA sequence binding, a PrfB and SfaB (both PapB family members) may also mediate cross-regulation between the *prf* and *sfa* operons (249, 250). However, it is intriguing to note that this cross-regulation between PapB family members is only well described for family members

with high homology (203, 204, 249, 250), which suggests that this form of reciprocal regulation is not an unexpected phenomenon. Whether or not this reciprocal regulation is generalizable for additional PapB family members is unknown, but would confirm the existence of an adhesin operon expression hierarchy (251, 252). The precise function of nucleoid structure in this form of reciprocal regulation is still unclear.

E. coli nucleoid regulation

The *E. coli* nucleoid and associated proteins

To accommodate an entire chromosome within a small cellular volume, *E. coli* must compact its DNA. Indeed, if completely uncoiled, the circumference of the *E. coli* chromosome would extend around 1500 μm, compared to 1.6-3.9 μm for the *E. coli* cell itself (depending on growth condition) (94, 253). The problem of chromosomal compaction is further underscored by the fact that a rapidly dividing cell could initiate chromosome replication before cell division is complete (254-256). The compacted chromosome, including associated proteins and RNAs, is referred to as the *E. coli* nucleoid.

The *E. coli* nucleoid itself is a mosaic of both plectoneme (**Figure 1-5A**) and solenoid (**Figure 1-5B**) compacted regions (94, 257-262), with plectoneme compaction predominating (258-262). The core of the nucleoid is composed of densely packed DNA and associated NAPs; plectoneme DNA braids emanate from this core, and the cellular

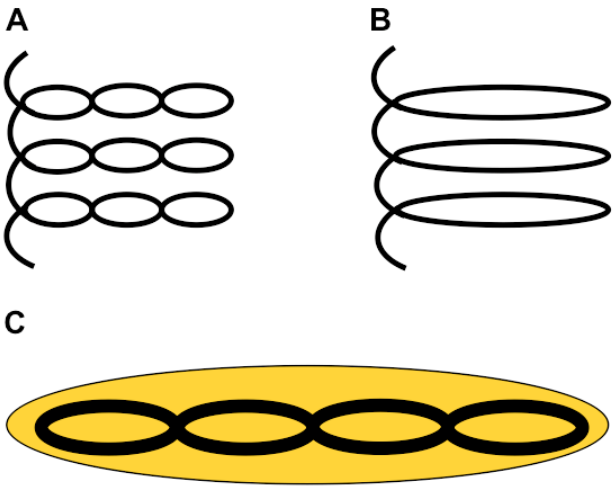

Figure 1-5. **Alternative forms of compacted DNA are represented. (A)** An example of the braided structure of plectoneme chromosome compaction is indicated. The compacted black line represents a double helix of DNA. **(B)** The looped structure of a chromosome compacted as a solenoid is shown. As above, the black line represents a double helix of DNA. **(C)** A representation of the writhed and coiled architecture of the *E. coli* chromosome is shown. Thick black lines represent compacted areas of the *E. coli* chromosome within the cell (yellow oval).

transcriptional and translational machinery are coincident with these DNA braids at the periphery of the cell interior (263, 264). The entire nucleoid itself is described as having a coiled and folded structure (94, 265-267) (**Figure 1-5C**). In addition, the *E. coli* nucleoid is further organized into at least six macrodomains (Ori [Origin], NS-right [Nonstructured-right], Right, Ter [Terminus], Left, and NS-left [Nonstructure-left]) (94, 268-271), which have defined boundaries and limited interactions between these domains.

Cell division, growth phase, and metabolism, especially energy generation, are two aspects of *E. coli* biology coupled to the nucleoid. For instance, the *E. coli* chromosome is generally in a negative (underwound) supercoiled state (257, 259, 272). During cell division, particularly DNA replication, DNA becomes more positively supercoiled through the activity of a DNA helicase (259, 273, 274). DNA gyrase subsequently adds negative supercoils (*i.e.,* the DNA becomes underwound) back to the chromosome (273-275). This process requires ATP hydrolysis (273, 275); thus, metabolism, particularly energy generation, could be coupled to nucleoid structure, especially the supercoiled state of the chromosome (92-94). Furthermore, these phenomena may be linked to the more prominent appearance of different NAPs during different growth phases (94, 95, 104-109) **Table 1-1**. While H-NS is prominent in exponential phase, it can be present in all growth phases (94, 95, 109); Lrp has higher prominence near middle to late exponential phase, but Lrp levels could drop at some point in stationary phase (94, 95, 104, 105, 109); IHF (integration host factor) α and β subunits appear to have prominence during the transition from exponential to stationary

Table 1-1. Nucleoid structuring proteins vary with growth phase

Regulator	Most prominent growth phase	References
CbpA	Stationary	(95, 109)
CbpB	Exponential and stationary	(95, 109)
DnaA	Exponential	(95, 109)
Dps	Stationary	(95, 109)
Fis	Lag and exponential	(94, 95, 109)
Hfq	Exponential	(95)
HUα	Lag* and exponential	(94, 95, 107-109)
HUβ	Exponential and stationary*	(94, 95, 107-109)
H-NS	Exponential	(94, 95, 109)
IciA	Exponential	(95)
IHFα	Stationary	(94, 95, 106, 109)
IHFβ	Stationary	(94, 95, 106, 109)
Lrp	Exponential and stationary	(94, 95, 104, 105, 109)
StpA	Exponential	(95, 109)

*Although this subunit is predominate at this phase, levels of HU are highest in exponential phase

phase (94, 95, 106, 109). In addition, Fis appears to be more prominent during lag and exponential phase (94, 95, 109). The expression of the genes encoding the subunits of HU (histone-like protein from *E. coli* strain U93) shows a complex regulatory profile. HUα subunits predominate during lag phase, HUαβ subunits are both present during exponential phase, and HUβ subunits are also present during stationary phase (94, 95, 107-109). All NAPs, however, may be expressed during any growth phase (*i.e.,* in addition to the point where levels are the highest). Nevertheless, the differential appearances of NAPs may approximate switches (especially in regulatory occupancy) where global gene expression profiles may be perturbed through varying levels (and possibly ratios) of regulating NAPs as described below. Most strains of *E. coli* encode the genes for at least 12 distinct NAPs including CbpA, CbpB, DnaA, Dps, Fis, Hfq, H-NS, HU, IciA, IHF, Lrp, and StpA (95). Hha and YdgT are two accessory NAPs supplementing gene regulation at AT-rich sequences (276). However, only the five key NAPs previously listed above will be reviewed in detail below including, IHF, HU, Fis, H-NS, and Lrp.

IHF and HU, two homologs (277), are both involved in gene regulation. IHF is often thought of as a DNA architectural catalyst, as it mediates significant bending and looping of bound DNA (277-281). Indeed, IHF-mediated DNA bending is predicted to promote FimB and FimE site-specific recombination (279, 282, 283), which will either promote or inhibit *fim* operon expression based on whether *fimS* invertible element containing P_{fimA} is in the ON or OFF orientation after recombination. DNA looping by IHF is also implicated in allowing RNA polymerase to contact distal transcription factors,

from specific promoters, to regulate gene expression (281, 284, 285). The IHF binding site itself is a chimera of AT-rich tracts at the 5' end and the sequence (A/T)ATCAANNNNTT(A/G) at the 3' end (286). In addition, IHF may be involved in the regulation of genes necessary during nutrient-limiting conditions, and rises in prominence during stationary phase (94, 95, 109, 287). Like IHF, HU mediates DNA bending and looping (288, 289), and HU also mediates RNA polymerase contact with distal transcription factors, from specific promoters (290, 291). Unlike IHF, however, HU does not bind to a specific sequence (292-294). It is intriguing to note that both IHF and HU have two subunits, which may show differential DNA binding activities based on dimerization state (106-108).

Fis is another NAP, which appears during lag and exponential phase and regulates gene expression. Fis, of all other NAPs discussed here, is the most abundant with nearly 60,000 copies per cell (95). Fis has some function in DNA bending and binds AT-rich sequences that may be more tractable toward bending (295-297). As noted above, Fis positively and negatively regulates RTX nonfimbrial adhesin genes *lapA* and *lapF*, respectively (214). It also regulates genes necessary for entry into exponential growth phase, which include metabolic pathways, replication, transcription, and translation machinery genes (298-302). As Fis constrains DNA supercoiling and promotes spatiotemporal coordination of supercoiled DNA (302, 303), this may be another mechanism where global gene expression is regulated by Fis. Fis, in addition, induces expression of *hns* (304), which may also contribute to the NAP predominance switching

phenomenon during different growth phases discussed above (*i.e.*, *hns* represses *fis*). At high Fis concentrations, however, *hns* may be repressed (304).

H-NS regulation

Perhaps one of the most studied NAPs is H-NS. Indeed, H-NS is one of the NAPs at the core of the nucleoid, and H-NS serves a key function in compaction of the nucleoid into a coiled structure (94, 305). As is the case with Fis above, H-NS appears to be involved in constraining DNA supercoiling and the spatiotemporal coordination of supercoiled DNA available to transcriptional machinery (305-308). Intriguingly, H-NS may regulate Fis (258), which further suggests that complex regulatory networks may exist among NAPs. H-NS itself binds AT-rich sequences (41-43, 94, 309). However, H-NS binding to DNA may be nucleated at a specific sequence (tCGATAAATT) and subsequently spread (*i.e.,* oligomerize to form a filament) from this site (258, 310). In the presence of this nucleating site, H-NS binding may even extend into lower affinity sequences (258). In addition, H-NS has affinity for curved DNA (311, 312), which is often associated with AT-rich sequences (313).

H-NS promotes plectoneme compaction of the nucleoid, and H-NS-mediated gene regulation is often through bridging two distinct DNA segments together (307, 314-316). Indeed, H-NS oligomers are hypothesized to form a right-handed helical filament coordinating this bridging (314). The favored mechanisms of H-NS-mediated repression include obstructing promoters from transcription factors and transcription machinery, RNA polymerase trapping, and constraining DNA supercoils (307). Occlusion occurs

when RNA polymerase and other relevant proteins are unable to bind sequences already occupied by H-NS, which is largely due to steric inhibition (307, 317). The occlusion phenomenon may be extended to include inhibition of the function of RNA polymerase (*e.g.,* open complex formation) (307, 318, 319). H-NS-mediated DNA looping may also contribute to RNA polymerase trapping. In this case, it is suggested that H-NS mediates repression through the prevention of RNA polymerase-mediated elongation of a transcript; thus, RNA polymerase is essentially trapped at these promoters, and the associated genes will not be transcribed (307, 310, 320-322). However, the force generated by a translocating RNA polymerase is sufficient to displace H-NS (307, 323, 324), which suggests that H-NS can only trap a stationary RNA polymerase. Intriguingly, it is possible that RNA polymerase itself may cooperate with H-NS binding to participate in the formation of RNA polymerase trapping complexes at these promoters (307, 322, 325). Both RNA polymerase trapping and RNA polymerase-mediated H-NS repression may be sensitive to the presence of alternative sigma factors (317, 326, 327). This may contribute to growth phase regulation switching discussed above (*i.e.,* σ^S may overcome H-NS occupancy at some promoters). In the case of constraining supercoiling, H-NS bridging between DNA strands is sufficient to maintain local supercoiling, even when there are double stranded DNA breaks at other sites (305-308). This suggests that H-NS DNA binding could locally overcome the activity of topoisomerases. Constrained supercoiling subsequently modulates activation and repression at promoters sensitive to supercoiling (305, 307, 328-330). In addition, H-NS constrained supercoiling may also contribute to RNA polymerase stalling and Rho factor-dependent termination of

transcription (322, 331, 332). This stalling activity may even contribute to post-transcriptional regulation of gene expression and protein synthesis, but this activity is still largely speculative.

As is the case with H-NS silencing above, there are multiple mechanisms to overcome this activity (307, 315). Several of these anti-silencing mechanisms include H-NS displacement by RNA polymerase, altered nucleoid structure by other NAPs, H-NS displacement by classical sequence or motif-specific transcription factors (including filament formation inhibition), filament and bridging antagonism by H-NS-like molecules, anti-silencing by environmental stimuli, such as pH, temperature, and osmolarity, and non-coding RNA-mediated anti-silencing (307, 315, 333). Although RNA polymerase may contribute to H-NS-mediated silencing as described above, it may be the case that displacement of H-NS by RNA polymerase may be mediated through altering promoter geometry (especially in complex with different sigma factors) (307, 317, 326, 327). Likewise, a translocating RNA polymerase may displace H-NS (307, 323, 324), which is another mechanism whereby H-NS-mediated silencing is overcome. As discussed above, the NAP content of the nucleoid varies with growth phase (94, 95, 104-109). In the case of anti-silencing by altered nucleoid structure, other NAPs may, thus, antagonize H-NS silencing activity (304, 307, 334, 335). In this sense, anti-silencing of H-NS is reminiscent of a switch between nucleoid states and underlying gene regulation (*i.e.*, the prominence of NAP regulation at a locus may switch). Sequence-specific or motif-specific regulators can also contribute to overcoming H-NS-mediated silencing (181, 307, 315, 336, 337). These regulators could compete with H-NS for DNA binding

sites (307, 315). In addition, sequence-specific regulators may overcome H-NS-mediated silencing through inhibiting the extension of an H-NS filament (307, 315, 338). In addition, these sequence-specific regulators may also alter the geometry of promoters and facilitate H-NS filament disengagement from the silenced DNA (307, 315). Some sequence specific-regulators are dispensable for positive gene regulation when H-NS is absent, but other regulators may still promote gene expression in other cases (307, 339). In H-NS-like molecule antagonism of H-NS silencing, it has been shown that truncated H-NS-like molecules may oligomerize with H-NS, but suppress DNA binding and bridge formation (307, 315). Likewise, pH, temperature, and osmolarity are all known to alter the structure of DNA (340-342). As such, H-NS binding may be altered by this different geometry (307, 315, 333). While, these changes in geometry usually result in the inhibition of H-NS silencing (307, 315, 333, 343), it has been suggested that H-NS silencing could be augmented by these environmental conditions (307). The DsrA RNA is a non-coding RNA that inhibits *hns* mRNA translation (307, 315, 344). DsrA activity, thus, lowers the cellular pool of H-NS, which in turn has the potential to contribute to nucleoid restructuring.

H-NS also allows *E. coli* to have regulatory activities that are not encoded for in the genome. These activities include regulating genes acquired by horizontal transfer (306, 307, 315) and silencing expression from intergenic and intragenic promoters (345), both cases where *E. coli* may not encode dedicated regulators for these recently acquired genes or sequences. Harboring foreign DNAs without regulation could potentially be detrimental to a host, as the host may expend energy producing unnecessary or toxic

proteins. PAIs, such as PAI-*aspV* harboring the *tos* operon, are often noted for containing genes that are AT-rich, compared with the rest of the host genome (40-43). As noted above, H-NS binds AT-rich sequences (41-43, 94, 309). Thus, it has been well documented that H-NS regulates AT-rich sequences acquired by horizontal transfer (41-43). Of particular note for UPEC, nearly all classical virulence factors (*i.e.,* HlyA, fimbriae, and iron uptake systems) are upregulated in an *hns* mutant (346), which further underscores the importance of understanding H-NS-mediated regulation. Specific functions for H-NS in the regulation of adhesins and motility genes will be reviewed below. H-NS also suppresses expression from promoters located inside of genes (intragenic) or between genes (intergenic), which may otherwise promote expression of silencing or nonfunctional RNAs (345). Therefore, this may be another mechanism whereby *E. coli* mitigates possible detrimental effects of having foreign DNA (*i.e.,* expressing unnecessary RNAs or RNAs that silence expression of necessary genes).

Lrp regulation

Lrp, like H-NS above, is a global regulator of gene expression (105, 152, 179, 183, 250, 296, 347-355). Indeed, Lrp regulates nearly 10% of all *E. coli* genes (105). The vast majority of these genes are most prominently regulated at the transition into stationary growth phase (intriguingly, when Lrp levels fall) (105). Lrp may wrap DNA in a manner reminiscent of a solenoid (356); DNA wrapping has also been observed for the Lrp homologs, using atomic force microscopy (357). It does not appear that a single Lrp-binding site consensus sequence exists. Nevertheless, using a classical DNase I

protection assay, the sequence $GN_{2\text{-}3}TTT$ was identified as a putative Lrp-binding sequence in the vicinity of P_{papI} and P_{papBA} (358). There is, in addition, some evidence that Lrp has affinity for other AT-rich sequences (359).

Lrp exists as an octamer, hexadecamer, or leucine-bound octamer (356, 360, 361). The Lrp octamers emerge as a consequence of being arranged as a tetramer of dimers (356). However, one of the interfaces between adjacent dimers does not mediate further oligomerization, which suggests this structure is essentially open (*i.e.*, not a closed ring) and that asymmetry exists in this octamer (356). Higher order multimers (*i.e.*, hexadecamers) of the Lrp octamers, nevertheless, are observed in the cellular Lrp pool (360, 361). In the presences of leucine, however, these hexadecamers disassociate to leucine-bound octamers or fail to form (360, 361). Furthermore, Lrp-mediated regulatory activities may be different for Lrp octamers, hexadecamers, and leucine-bound octamers (360-362). In addition, other amino acids, including alanine, histidine, isoleucine, methionine, serine, and threonine, can also modulate Lrp-mediated regulation through DNA binding (363), but the extent and consequences of these different binding behaviors have yet to be fully realized. DNA wrapping by Lrp may either obstruct transcriptional machinery from binding to relevant sequences or promote RNA polymerase interactions with Lrp itself or distal elements necessary for transcription (364-366).

Lrp is often studied as a regulator of genes involved in metabolism and is usually described as a feast or famine regulatory protein (152, 367). Indeed, Lrp levels usually rise in middle to late exponential phase, but could drop later in stationary phase (94, 95, 104, 105, 109). Thus, Lrp regulation of the starvation response is complex. Consistent

with its function as a feast or famine regulator, Lrp is regulated by the alarmone ppGpp (guanosine tetraphosphate), but this regulation may be complex and indirect (104). The alarmone ppGpp itself rises in response to nutrient stresses (368). Of the genes regulated by Lrp, the general trend is that genes encoding protein involved in anabolic (synthetic) pathways are positively regulated by Lrp and genes involved in catabolism (degradation) pathways are repressed (105, 152); genes encoding proteins involved in amino acid synthesis and degradation are especially noted for differential regulation (105, 152, 369). An *lrp* mutation is also one of the known GASP (growth advantage at stationary phase) mutations (370), which further underscores that regulation by Lrp is necessary to promote specific behaviors in response to nutrient limitation. In addition, adhesin and flagellar levels are governed by Lrp regulation (148, 152, 179, 202, 204, 279, 348, 352, 354, 355), but this regulation will be discussed below. Taken together, it seems reasonable to speculate that Lrp functions to start the switch into stationary phase, and as such controls the synthesis of proteins necessary for stationary phase survival, while resources and nutrients are still available.

H-NS and Lrp controlled reciprocal regulation (adherence-motility and multicellular-unicellular behavior)

As noted above, H-NS and Lrp both regulate adhesin operons (**Table 1-2**). For instance, H-NS is a negative regulator of nearly all adhesin operons described in this work (*e.g.*, *pap*, *fim*, and *foc*) (148, 179, 181, 202, 351, 354, 371-374). Lrp, on the other

Table 1-2. Summary of H-NS- and Lrp-mediated reciprocal regulation.

Regulator	Effect	Reference
H-NS	Dual regulator of motility	(375-378)
	Negative regulator of Type 1 fimbria	(148, 150, 374)
	Negative regulator of P fimbria	(181, 183, 354, 371, 372)
	Negative regulator of F1C fimbria	(202)
	Negative regulator of biofilm (cellulose- and/or curli-mediated)	(235, 326, 376, 379-381)
	Negative regulator of *S. enterica* multicellularity (type 6 secretion system-mediated)	(382)
Lrp	Negative regulator of motility	(355)
	Dual regulator of Type 1 fimbria	(148, 152, 279, 348, 351, 352, 383)
	Dual regulator of P fimbria	(152, 179, 183)
	Dual regulator of F1C fimbria	(152, 200-202, 204, 205)
	Positive regulator of biofilm (cellulose- and/or curli-mediated)	This work
	Positive regulator of the *P. mirabilis* Dienes phenomenon (type 6 secretion system-mediated multicellularity)	This work

hand, is a dual regulator (both positive and negative regulator) of these and other adhesin operons, depending on the DNA binding context (148, 152, 179, 202, 204, 279, 348, 352, 354). Thus, a switch between these two regulators, particularly gaining prominence in regulatory occupancy in the *E. coli* chromosome, could account for this differential regulation (**Figure 1-6**). In one specific example, H-NS is the presumptive negative regulator of the *pap* operon (**Figure 1-7A**) (183, 354, 371, 372). When the distal GATC (Dam methylation) site is methylated and the proximal GATC site is unmethylated, Lrp binds the proximal sites (Lrp-binding sites 1, 2, and 3), and the *pap* operon is OFF (**Figure 1-7B**) (152, 179, 183). When PapB displaces H-NS, the *papI* regulatory gene is expressed, and the distal GATC site is unmethylated, PapI acts to promote Lrp binding to the distal sites (Lrp-binding sites 4, 5, and 6) (152, 181, 183, 384) (**Figure 1-7C**). A similar regulation model is also predicted to exist for the *foc* operon (152, 200-205). Unfortunately, H-NS-mediated negative regulation of the *pap* operon can be complex. For instance, alleviation of H-NS negative regulation could be masked by the fact that Lrp is also a negative regulator of the *pap* operon. Nevertheless, it is accepted that H-NS can be a negative regulator of the *pap* operon, independent of Lrp (354). Thus, the prominence in regulation of the *pap* operon can switch from an H-NS state (negative regulation) to an Lrp state (dual regulation), which suggests that an H-NS and Lrp switch is involved in *pap* operon regulation. It is important to note, however, that H-NS and Lrp are both global regulators (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 370, 372, 375, 377, 378, 385-389). Therefore, this switch may only be an abstraction of the two NAPs and associated regulators perturbed by the changing ratio of

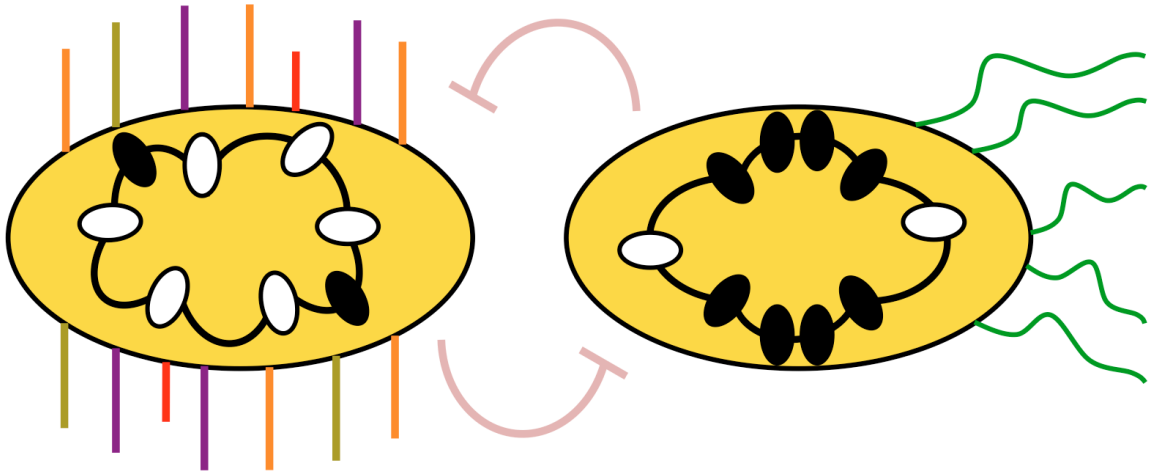

Figure 1-6. A switch between H-NS and Lrp may be involved in adherence and motility reciprocal regulation. A schematic representing a switch between occupancy of Lrp (white oval) and H-NS (black oval) in the *E. coli* chromosome (black circle) is noted. Correspondingly, if the switch is in the Lrp occupancy predominating state (possibly as seen in the switch from exponential and stationary phase), adherence is predicted to predominate over motility (adhesins, colored bars). When the switch is in the H-NS occupancy predominating state (possibly as seen in exponential phase), motility is predicted to predominate over adherence (flagella, green curves). Pink bars between the adherent and motile states represents the necessity that some of the reciprocal regulation of adherence- and motility-related genes is indirect.

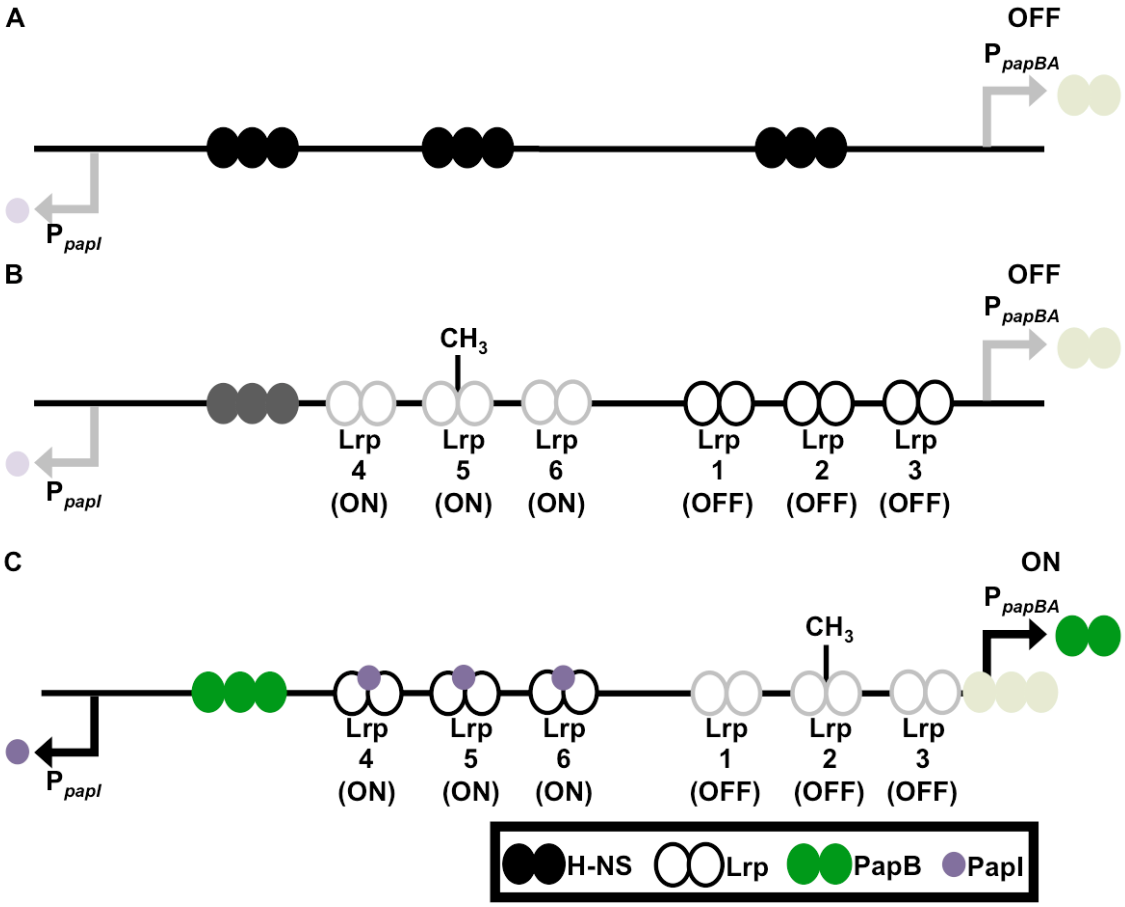

Figure 1-7. Regulation of the *pap* operon is dependent on NAP (nucleoid-associated protein) binding context and epigenetics. **(A)** H-NS, a negative regulator of the *pap* operon, binds to the locus represented in the figure and manipulates DNA methylation state. The *pap* operon-related promoters are, thus, OFF (gray arrows). **(B)** Lrp-binding site 5 is methylated, and Lrp will not bind the activating Lrp-binding sites (4, 5, and 6), denoted "ON" and indicated by faded ovals. Lrp-binding site 2 is unmethylated, and Lrp binds to repressive Lrp-binding sites (1, 2, and 3), denoted "OFF" (solid white ovals). The *pap* operon-related promoters are OFF (gray arrows), and H-NS may still bind to this locus (faded black ovals). **(C)** Lrp binding site 2 is methylated, and Lrp will not bind the Lrp-repressive binding sites, indicated by faded ovals. PapB (green ovals) displaces H-NS and promotes *papI* expression. In turn, PapI (purple circles) promotes Lrp binding to the unmethylated activation sites. The *pap* operon-related promoters are, thus, turned ON (black arrows). When PapB levels increase, PapB will bind to an area in the vicinity of P_{papBA}, which will turn the *pap* operon-related promoters OFF. This is represented by faded green ovals [*i.e.,* the *pap* operon is ON in **(C)**].

H-NS and Lrp at a variety of regulated sequences (**Figure 1-2**). This leads to the notion that such switches can be virtual (*i.e.*, they are indirect, instead of being a literal competition for binding sites at a specific locus).

Regulation of the *fim* operon includes a similar switch between H-NS, Lrp, and IHF (**Table 1-2**) (148, 150, 152, 279, 348, 351, 352, 374). In particular, H-NS negatively regulates *fimB*, which encodes the FimB recombinase modulating both OFF-to-ON and ON-to-OFF inversions of the *fimS* invertible element (148-150, 152). The *fimE* gene, which encodes the FimE recombinase modulating ON-to-OFF recombination, is also negatively regulated by H-NS (148-150, 152). Intriguingly, H-NS biases the *fimS* switch in the OFF state (**Figure 1-8A**), which suggests that H-NS regulation of the *fim* operon is complex (*i.e.,* a FimB or FimE recombination event to the OFF state coupled with transcription from the inverted P_{fimA} may be necessary for this H-NS-mediated biasing into the OFF state) (374). Lrp, in conjunction with IHF, positively and negatively regulate *fim* expression (148, 152, 279) (**Figure 1-8B)**. Lrp binding to the *fim* locus may weakly alter transcription of *fimB* (upregulate) and *fimE* (downregulate) (390). Lrp regulates the inversion of *fimS* (both ON-to-OFF and OFF-to-ON), through specific binding to sites within this element (148, 152, 279, 348, 351, 352, 383). Indeed, there are three Lrp binding sites in the *fimS* element, and when Lrp binds sites 1 and 2, site-specific recombination of the promoter is favored (383). As Lrp wraps or bends DNA, Lrp binding may facilitate site-specific recombination at the inverted repeats bracketing *fimS*, by bringing the inverted repeats into proximity with each other (279). In contrast, Lrp binding to sites 1, 2, and 3 inhibits site-specific recombination. In addition, branched

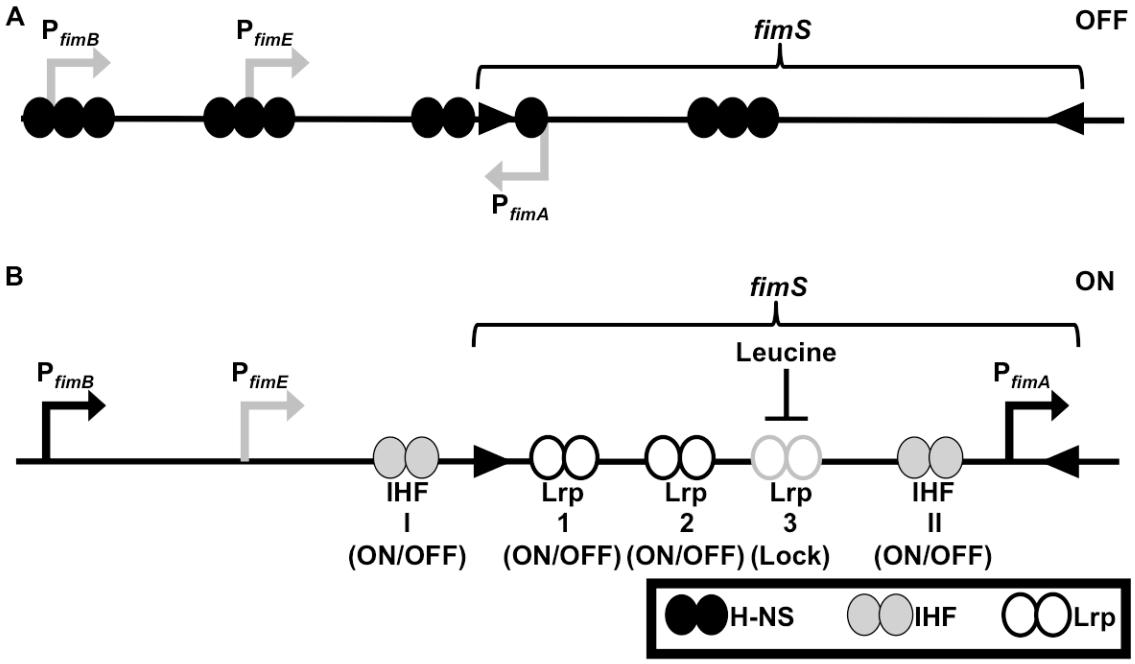

Figure 1-8. Like regulation of the *pap* operon, *fim* operon regulation is dependent on NAP binding context and inversion of the *fimS* promoter element. (A) H-NS binds and represses expression of the inverted repeat (depicted by inverted black arrowheads) recombinase genes *fimB* and *fimE* (gray arrows). H-NS binding in the vicinity of the *fimS* element also inhibits recombination of this element. As P_{fimA} faces toward the recombinases, the *fim* operon is OFF (gray arrow). **(B)** When Lrp binds to Lrp-binding sites 1 and 2, with IHF binding to IHF-binding sites I and II, recombination at the *fimS* element could be promoted, and P_{fimA} could be directed into the ON or OFF orientation, denoted "ON/OFF". In the case of the ON orientation, P_{fimA} faces away from the recombinase genes (black arrow). Also depicted is differential regulation of *fimB* and *fimE*; *fimB* is induced (black arrow) and *fimE* is repressed (gray arrow) in the depicted scenario. In addition, Lrp-binding site 3, sensitive to Lrp binding to leucine, inhibits recombination of *fimS*. As such, it may be thought of as locking the *fim* element into a particular orientation, denoted "Lock". Lrp binding to site 3 is depicted with a faded oval (*i.e.,* recombination will be able to proceed in this scenario).

chain amino acids seem to favor Lrp dissociation from site 3, which integrates environmental cues to Lrp-mediated regulation of the *fim* operon (383). IHF, the other *fim* operon dual regulator, binds one site in the vicinity of *fimS* and another site in *fimS* to promote DNA bending and FimB- and FimE-mediated recombination, perhaps in a manner similar to Lrp above (148, 279). Again, this shows how Lrp and H-NS, along with IHF, could be thought of as a switch vying for prominence in *fim* regulation.

Expression of flagellar genes appears to be linked to an H-NS and Lrp switch (**Figure 1-6** and **Table 1-2**). Indeed, an *hns* mutation substantially decreases motility, and an *lrp* mutation increases motility (355, 375, 376). It is interesting to note, however, that H-NS is a dual regulator of flagellar gene regulation (378). In particular, H-NS binding to a site after the transcriptional start site of P_{flhDC} mediates positive regulation of flagellar genes, whereas H-NS binding to additional sites in the vicinity of P_{flhDC} may contribute to downregulation of motility genes (378). In addition, H-NS negatively regulates *hdfR*, which encodes a known negative regulator of *flhDC* (376). H-NS also negatively regulates *rcsD*, which encodes a kinase regulating the activity of RcsB, another negative regulator of *flhDC* (376, 377). Likewise, H-NS represses *csgD*, a gene encoding the curli operon regulator (235, 376, 380). CsgD, in turn, decreases motility (376, 391). At present, it remains unclear why an *lrp* mutation promotes motility, but it has been suggested that lack of positive regulation of adhesin operons and associated motility regulators may explain this phenotype (355). Nevertheless, as both NAPs are involved in regulating motility, particularly H-NS both positively and negatively and Lrp negatively, elements of the switch discussed above are part of this regulation. Additional

regulators (associated with the underlying H-NS and Lrp-mediated regulation), however, could be part of this switch, underscoring the possible virtual nature of this switch.

The switch between multicellular (sometimes adherent) and unicellular (sometimes motile) lifestyles exhibits elements of H-NS and Lrp control (**Table 1-2**). H-NS, for instance, negatively regulates genes associated with curli biosynthesis (326, 376, 379, 381). H-NS, in addition, negatively regulates a number of enzymes with diguanylate cyclase activity (*i.e.,* mediating c-di-GMP formation) (376). Thus, H-NS may inhibit cellulose synthesis in a CsgD- and c-di-GMP-dependent manner (392-394). Cellulose and curli are both components of some adherent biofilms (140, 229-236). This negative regulation is relieved by an *hns* mutation in CFT073 (**Figure 1-9A**), a UPEC strain that ordinarily does not produce curli and cellulose (238). Furthermore, ectopic expression of *lrp* promotes synthesis of curli and/or cellulose (**Figure 1-9B**). Additionally, Type 1 fimbrial operon regulation includes a possible switch between H-NS and Lrp (148, 150, 279, 348, 351, 352, 374), and these fimbriae contribute to biofilm formation (395, 396). These examples further underscore that elements of the H-NS and Lrp switch may be associated with changes between multicellular and unicellular behaviors.

H-NS and Lrp also coordinate the production and activity of type 6 secretion systems (**Table 1-2**). In terms of further regulating a switch to multicellularity, H-NS silences the expression of the SPI-6 type 6 secretion system genes in *S. enterica* (382). This type 6 secretion system is hypothesized to inhibit competing bacteria and homogenize and synchronize the *S. enterica* population to promote infection (382), which

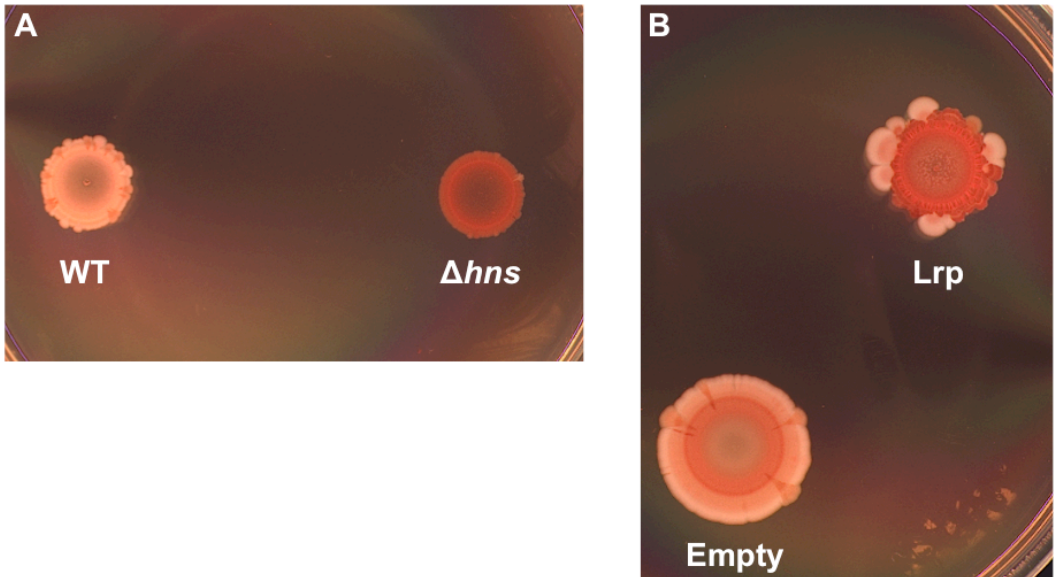

Figure 1-9. H-NS negatively regulates and Lrp positively regulates curli and/or cellulose production in *Escherichia coli* strain CFT073. **(A)** After 48 hours of incubation at 37 °C on a YESTA (yeast extract and tryptone agar) plate containing 40 μg/mL Congo red and 1 μg/mL Coomassie Brilliant Blue R-250, a UPEC strain CFT073 *hns* mutant binds the Congo red dye indicating the production of curli and/or cellulose; the UPEC CFT073 wild-type strain does not bind the Congo red dye. **(B)** After 48 hours of incubation at 37 °C on a YESTA plate containing Congo red, Coomassie Brilliant Blue R-250, ampicillin (100 μg/mL), and 10 mM L-arabinose, UPEC CFT073 harboring pBAD-*lrp* exhibits the rdar (red, dry, and rough) phenotype indicating the production of curli and/or cellulose; UPEC CFT073 harboring the empty pBAD plasmid does not exhibit the rdar phenotype.

can be thought of as a multicellular behavior. In the same pursuit, Lrp has some regulatory function governing the Dienes phenomenon in *P. mirabilis*, another uropathogen (**Figure 1-10**). The Dienes phenomenon occurs during swarming motility (a type of multicellular behavior) and is dependent on a type 6 secretion system (397). Lrp, in addition to mediating the Dienes phenomenon, is implicated in promoting swarming motility (398). Genes encoding energy and metabolic pathways also appear to be important for swarming motility (399-401), which possibly extends differential nucleoid regulation to this behavior.

Swarming motility itself, however, blurs the distinction between adherence and motility. For example, flagella themselves become numerous on the cell surface (*i.e.,* resembling a cell with numerous fimbriae on its surface, instead of a cell with a few peritrichous flagella) and mediate swarm cell interaction (*e.g.,* adherence) to each other during movement across a surface (402). Thus, swarming motility may be a unique blending of adherence and motility. Although not explored, it is also reasonable to speculate that the *P. mirabilis* type 6 secretion system, as was suggested for the *S. enterica* type 6 secretion system above, may have a function synchronizing and coordinating the cell population and communicating nucleoid state to neighboring cells.

E. coli encodes numerous NAPs, which change in prominence during the switches between *E. coli* growth phases. Thus, generalizing a switch governing adherence and motility reciprocal regulation, through H-NS and Lrp, may require the activity of additional regulators including other NAPS. This is especially true as both H-NS and Lrp negatively regulate *lrp* (403). Thus, an underlying mechanism may be necessary to

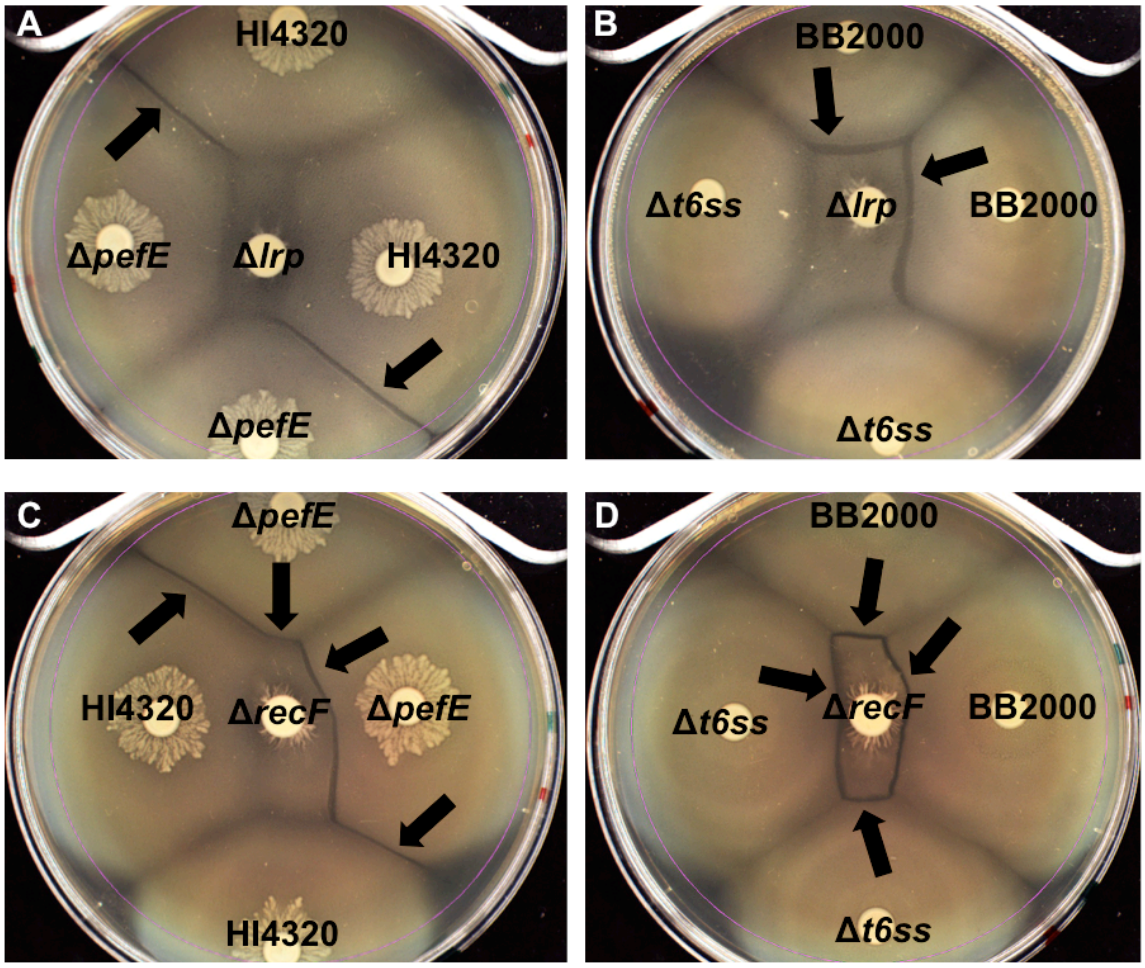

Figure 1-10. Lrp modulates the Dienes phenomenon during multicellular swarming motility in *P. mirabilis*. **(A)** A Dienes line (black arrow) is formed between wild-type *P. mirabilis* HI4320 and an HI4320 *pefE* mutant. No Dienes line is formed between the *P. mirabilis* HI4320 *lrp* mutant and the *pefE* mutant and wild-type HI4320. **(B)** A Dienes line (black arrow) is formed between *P. mirabilis* strain BB2000 and the *P. mirabilis* HI4320 *lrp* mutant. No Dienes line is formed between the *P. mirabilis* BB2000 type 6 secretion system mutant (Δ*t6ss*) and the *lrp* mutant. **(C)** A Dienes line forms between a *P. mirabilis* HI4320 *recF* mutant (similar swarming pattern to the *lrp* mutant) and the HI4320 *pefE* mutant above. No Dienes line forms between *P. mirabilis* HI4320 and the HI4320 *recF* mutant. **(D)** A Dienes line forms between the *P. mirabilis* HI4320 *recF* mutant and *P. mirabilis* BB2000 wild-type and the type 6 secretion system mutant.

reconcile this result with the apparent switching phenomenon described above (*i.e.*, allow a cell to express *lrp* even with abundant H-NS). Nevertheless, differential prominence of H-NS and Lrp, is an element of adherence and motility reciprocal regulation and the switch between multicellular and unicellular lifestyles.

tosA regulation, a hypothesis of unification between the *E. coli* nucleoid and adhesin and motility reciprocal regulation

Statement of the problem

As described above, the *tos* operon is poorly expressed when UPEC strain CFT073 is cultured under typical laboratory conditions, but well expressed during an experimental murine UTI (208, 227). Thus, this finding represents a gap in our understanding of the environment UPEC encounters during a UTI and how nucleoid states couples with environmental cues to regulate the *tos* operon. As TosA is an adhesin, studying regulation of the *tos* operon could identify and fill gaps in our knowledge of reciprocal regulation of adhesin and motility system genes and potentially how nucleoid state couples to these lifestyles.

Hypothesis one: regulators encoded by genes in the *tos* operon, in addition to nucleoid structure, contribute to *tos* operon positive and negative regulation. Environmental stimuli are also coupled to specific regulators of the *tos* operon.

The *tos* operon itself encodes three putative DNA binding proteins. These DNA binding proteins include TosR (PapB family homolog), TosE, and TosF (both LuxR family members). PapB family members are known to be both positive and negative regulators of cognate adhesin operons (179, 180, 202, 203, 404). Therefore, TosR is one logical candidate for being a dual regulator of the *tos* operon, and dual regulation of this operon may be achieved through differential TosR DNA binding in the vicinity of P_{tos}. However, it is not clear whether TosE and TosF have functions associated with *tos* operon regulation or additional regulatory functions (*e.g.*, reciprocal regulation). In addition, the nucleoid structuring proteins H-NS and Lrp both contribute to adhesin operon regulation (148, 150, 179, 181, 202, 204, 279, 326, 348, 351, 352, 354, 371-374, 376, 379, 381). Thus, regulation of the *tos* operon may be mediated through a perturbation of nucleoid structure. The H-NS and Lrp regulatory switch discussed above is also a logical candidate for underpinning regulation of the *tos* operon. H-NS may repress the *tos* operon, as is the case for other adhesins operons (148, 179, 181, 202, 351, 354, 371-374), and Lrp may mediate positive or negative regulation of the *tos* operon, as is the case for other adhesin operons (148, 179, 202, 204, 279, 348, 352, 354). Sensing environmental stimuli may be relayed to the cell in the form of differential regulation by NAPs, which may also contribute to *tos* operon regulation. Upon completion of this work, a clearer picture of *tos* operon regulation will emerge. This work will also expand our understanding of nucleoid structuring proteins and regulators associated with the *tos* operon.

Hypothesis two: regulators associated with the *tos* operon and *tos* operon regulation (including NAPs) are involved in the switch governing *E. coli* adhesin and flagellar operon reciprocal regulation.

The regulation of many adhesin operons is coupled to a switch underlying reciprocal regulation between adhesin and flagellar operons (186-188, 239, 240, 355). A great deal of this reciprocal regulation is mediated by both nucleoid structuring proteins and genes encoded at the 3' end of adhesin operons (148, 150, 179, 181, 186-188, 202, 204, 239, 240, 279, 326, 348, 351, 352, 354, 355, 371-379, 381). Thus, H-NS- and Lrp-mediated regulation of additional genes, coupled with regulation of the *tos* operon, could contribute to the phenomenon of reciprocal regulation of adherence and motility in *E. coli*. Likewise, as TosE and TosF are encoded by the terminal regulator genes of the *tos* operon, these regulators may also contribute to decreased motility. PapB family members, in addition, could cross-regulate each other (203, 204, 248-250). Therefore, exploring how TosR-mediated regulation of the *tos* operon is coupled to adhesin reciprocal regulation is a logical area of study. Upon completion of this work, both a clearer and broader picture of *E. coli* reciprocal regulation will emerge. This work will also present a clearer picture of how nucleoid structure couples *tos* operon regulation with the phenomenon of reciprocal regulation, in a unified manner.

CHAPTER 2

A CONSERVED PAPB FAMILY MEMBER, TOSR, REGULATES EXPRESSION OF THE UROPATHOGENIC *ESCHERICHIA COLI* RTX NONFIMBRIAL ADHESIN TOSA WHILE CONSERVED LUXR FAMILY MEMBERS, TOSE AND TOSF, SUPPRESS MOTILITY

Modified from: M. D. Engstrom, C. J. Alteri, and H. L. T. Mobley. 2014. Infection and Immunity 82: 3644-56.

Abstract

A heterogeneous subset of extraintestinal pathogenic *Escherichia coli* (ExPEC) strains, referred to as uropathogenic *E. coli* (UPEC), cause most uncomplicated urinary tract infections. However, no core set of virulence factors exists among UPEC strains. Instead, the focus of urovirulence has shifted to studying broad classes of virulence factors and the interactions between them. For example, the RTX nonfimbrial adhesin TosA mediates adherence to host cells derived from the upper urinary tract. The associated *tos* operon is well expressed *in vivo*, but poorly expressed *in vitro,* and encodes TosCBD, a predicted type 1 secretion system. TosR and TosEF are PapB and LuxR family transcription factors, respectively; however, no function has been assigned to these potential regulators. Thus, the focus of this study was to determine how TosR and TosEF regulate *tosA* and affect the reciprocal expression of adhesins and flagella.

32% (101/317) of sequenced UPEC strains were found to encode TosA. Nearly all strains carrying *tosA* [91% (92/101)] simultaneously carry the putative regulatory genes. Deletion of *tosR* alleviates *tosA* repression. The *tos* promoter was localized upstream of *tosR* using transcriptional fusions of putative promoter regions with *lacZ*. TosR binds to this region affecting a gel shift. A 100 bp fragment 220-319 bp upstream of *tosR* inhibits binding suggesting localization of the TosR binding site. TosEF, on the other hand, down-modulate motility when overexpressed by preventing expression of *fliC* encoding flagellin. Deletion of *tosEF* increased motility. Thus, we present an additional example of the reciprocal control of adherence and motility.

Introduction

Urinary tract infections (UTIs) are the second most common bacterial infection in humans (405). UTIs can be classified as complicated or uncomplicated infections. Uncomplicated UTIs, occurring in otherwise healthy individuals, are self-limited infections of the bladder, referred to as cystitis (12, 13, 406). However, upon bacterial ascension into the kidney, a more serious infection referred to as pyelonephritis can develop (12, 13). Pyelonephritis, in turn, can lead to the development of bacteremia and sometimes fatal urosepsis (14, 15).

UTIs normally occur when uropathogens that colonize the intestine alongside commensal organisms gain access to the periurethral area and then ascend to the urinary

bladder (407, 408). A heterogeneous subset of extraintestinal pathogenic *Escherichia coli* (ExPEC) strains, referred to as uropathogenic *E. coli* (UPEC), cause the overwhelming majority of uncomplicated UTIs (13). UPEC strains carry a battery of virulence factors including adhesins, toxins, and iron acquisition systems, which promote uropathogenesis (23, 44). However, no core set of virulence factors has been identified. Instead, any given UPEC strain appears to use various virulence factors from these three classes of virulence determinants to colonize the urinary tract (33, 35). This thesis requires that we consider established, newly discovered, and putative virulence factors, as well as the interactions among them, to better understand urovirulence.

Adhesins represent one broad class of virulence determinants. Fimbrial adhesins assembled via the chaperone-usher pathway are the most extensively studied adherence factors (143, 409). Indeed, the genes necessary to synthesize two chaperone-usher fimbriae, Type 1 and P fimbriae (pyelonephritis associated pili), were among the first cloned virulence factor genes (145, 410) and are important during experimental and human UTI, respectively (34, 164-167, 170, 172). In addition, seven other putative chaperone-usher fimbriae are encoded by prototype UPEC strain CFT073 alone (36). On the other hand, nonfimbrial adhesins have garnered less attention than chaperone-usher adhesins. These adhesins, nevertheless, can also contribute to uropathogenesis (39, 208, 411-413), underscoring the importance of continued study of this adhesin class.

In addition to adhesins, flagella-mediated motility also contributes to the development of ascending infection to the upper urinary tract (110-112). It is now recognized, however, that adherence genes and flagellar genes can be reciprocally

coordinated (187, 188, 240, 355, 414, 415). In this network, it is logical that an adherent bacterium should not be motile and a motile bacterium should not be adherent. That is, when fimbrial genes are expressed, flagellar genes should be repressed and vice versa.

With respect to nonfimbrial adhesins, we have previously described that UPEC strain CFT073 encodes within its *aspV* pathogenicity island (PAI-*aspV*), an RTX (*r*epeats-in-*to**x*in) nonfimbrial adhesin, referred to as TosA (or *t*ype *o*ne *s*ecretion protein A; originally annotated UpxA) (36, 39). RTX proteins are typically thought of as toxins that are secreted through a type 1 secretion system and diffuse away from the bacterium to mediate effects on the host. This is exemplified by the family prototype α-hemolysin (46, 48, 63, 64, 66, 68, 416). However, adhesins secreted in the same manner, but remaining associated with the bacterial cell surface are a growing group of RTX proteins composed of at least six other well characterized members (50). We presume that TosA contributes to uropathogenesis by binding to receptors on the surface of host epithelial cells derived from the upper urinary tract. Indeed, deletion of *tosA* creates a fitness and virulence defect for *E. coli* CFT073 during an experimental transurethral co-challenge of mice with the parental wild-type strain (39) or independent challenge (227). This same mutant also shows a fitness defect in the spleens and livers during bacteremia, suggesting a function for TosA during urosepsis (208). The *tosA* gene was previously found in an estimated one-fourth of UPEC strains (35, 39).

An intriguing feature of the *tos* operon is its strong *in vivo* expression, but poor *in vitro* expression (208). Indeed, TosA was discovered in an IVIAT (*in vivo* induced antigen technology) screen that identified gene products preferentially expressed *in vivo*

(227). The mechanism that explains tight regulation is not understood. Therefore, the focus of this study was to identify regulatory elements associated with *tosA* expression and the consequences of this regulation as it relates to the reciprocal regulation of motility and adherence. We found that TosR, a PapB family member, represses expression of *tosA,* while TosE and TosF, two members of the LuxR family, mediate the repression of motility. This work furthers our understanding of how adhesins are regulated and helps to describe the underlying network governing the interplay between adherence and motility.

Materials and Methods

Strain construction

E. coli CFT073 deletion mutants Δ*tosR*, Δ*tosE*Δ*tosF*, and Δ*tosR*Δ*tosE*Δ*tosF*, were generated and screened via PCR in an unmarked Δ*lacZ* background using primers described in **Table 2-1** and the lambda red recombineering method previously described (417). The original Δ*lacZ* construct was selected for on lysogeny broth (LB) agar (10 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 15 g/L agar) containing chloramphenicol (20 μg/mL), and unmarked as previously described (417). All other mutants were selected for on LB plates containing kanamycin (25 μg/mL). In the case of Δ*tosR aph*$^{+}$, this deletion mutation was unmarked as above to produce Δ*tosR*. The *tosR* mutation was also moved into a clean background of wild-type *E. coli* CFT073 by transduction using phage

Table 2-1. Primers used in this study.

Primer[a]	Sequence (5'-3')
Δ*lacZ* F	GAAATTGTGAGCGGATAACAATTTCACACAGG ATACAGCTGTGTAGGCTGGAGCTGCTTC
Δ*lacZ* R	CTTACGCGAAATACGGGCAGACATAGCCTGCCC GGTTATTAATGGGAATTAGCCATGGTCC
Δ*lacZ* Screen F	GAAAGCAGACCAAACAGCGG
Δ*lacZ* Screen R	TAACAGAACGGGAAGGCGAC
Δ*tosR* F	ATAATAAATTAAACATTGAATAATGTGTAATGG TATGGCAGTGTAGGCTGGAGCTGCTTC
Δ*tosR* R	ACTAAAAACTATTATTATAATATTCACTTAGCA ATGCGCAATGGGAATTAGCCATGGTCC
Δ*tosR* Screen F	CGACGTGCGCCATCGTGTCTG
Δ*tosR* Screen R	GATTGTGCCGAAGTTAACTCCGCCC
Δ*tosE*Δ*tosF* F	TATATACTTCTTGTAGAAGGCATAATGTATGAA TATAATGGTGTAGGCTGGAGCTGCTTC
Δ*tosE*Δ*tosF* R	CTTATCTACATAATAATAGACCTTTGTAAAATA ACTGTATATGGGAATTAGCCATGGTCC
Δ*tosE*Δ*tosF* Screen F	GGCTGACGGAGCGGGAAGTCTG
Δ*tosE*Δ*tosF* Screen R	GCCCACTCATCAGTGAGTACCC
pBAD-*tosR*-HisA F	NNNNCCATGGCTTGTAATGGTATGGCAGATCAT ATACAG
pBAD-*tosR*-HisA R	NNNNNAAGCTTCGCCCGAAAACTATTATTATAA TATTCACTTAGCAATGCGCA
pBAD-*tosEF* F	NNNNCTCGAGTAATATAATGATTGTTACGCACA ATAAATATC

pBAD-*tosEF* R	NNNNCTGCAGTTATCTACATAATAATAGACC
pBAD Screen F	TGCCATAGCATTTTTATCC
pBAD Screen R	CTGATTTAATCTGTATCAGG
P_R F	NNNNNGAATTCGTCAGTCGAAACTCAGGAGTG TGGAGG
P_R R	NNNNNGGATCCCTGTATATGATCTGCCATACCA TTACACAT
P_C F	NNNNGAATTCATTTTTATATCCACCCCCCCTTTA A
P_C R	NNNNGGATCCTTTTATGATTTTTATTTAAAATAT T
P_A F	NNNNGAATTCTTTATTATATTATTAATATCATG GC
P_A R	NNNNGGATCCATAAAATCCTTAGGCTAATTAAA AC
Promoter Screen 1 F (P_R)	NNNNGGTACCATAAACTGCCAGGAATTGGGGA TCG
Promoter Screen 2 F (P_C and P_A)	CCGCCGGGAGCGGATTTGAA
Promoter Screen R (P_R, P_C, and P_A)	GATCGGTGCGGGCCTCTTCG
P_{tosR} Shift F (P_{tosR}1 F)	AAGTTTTGGGGTGCAGTCCAC
P_{tosR} Shift R (P_{tosR}7 R)	CTGTATATGATCTGCCATACCATTACACAT
lacZ Shift F	GCGAATACCTGTTCCGTCATAGCG
lacZ Shift R	CATCGCCAATCCACATCTGTGAAAG
P_{tosR}1 R	TAGATATTATTGTTATCCATCATGT

P_{tosR}2 F	TTTAATCACTACCGCCTTGGTCGCT
P_{tosR}2 R	GCATTTTTTTGGTAAAAATCAATTTTTATA
P_{tosR}3 F	TAATATAGATATTATCTGCATATAA
P_{tosR}3 R	AAAAAGTGAAATCTCAAAACAAAAAAT
P_{tosR}4 F	CCATTTGTTTTATTTTATAAATAATTTTTTG
P_{tosR}4 R	TACTAGAGATTACATCTAAAAAATT
P_{tosR}5 F	TTAGATAAAAACCCTACAGAGAAGT
P_{tosR}5 R	CCTCAATCAAAAAACCATTAAATGAAATTTA
P_{tosR}6 F	TTATTGGTTTTATTGGTTTTAAATTTCATTT
P_{tosR}6 R	TATTGATTCACATTATAAATACATATT
P_{tosR}7 F	GCAAAAAAAATTTGATGCAAACAAATATG
tosA-tosE F	CTCAGTTAGTCAAGTTAACGGCATCGG
tosA-tosE R	GATGACAGGCTACTTATTGATTCTACTGG
tosE-tosF F	CCATGGGTGGAATGTAGCAAGTATTGC
tosE-tosF R	GCGTGGATAATATCCCTGAGAAAATC

[a] F, forward; R, reverse

ΦEB49 (418), with the following modification: phage lysate and overnight culture were incubated together at room temperature for 20 min at a ratio of 1:5. These constructs were verified with PCR using the primers listed in **Table 2-1**. The Δ*tosR* mutation was also verified with DNA sequencing.

TosR-His_6 was constructed by cloning *tosR* into the NcoI and HindIII sites of pBAD-*myc*-HisA (Invitrogen). A *tosEF* overexpression construct was generated by cloning *tosEF* into the PstI and XhoI sites of pBAD-*myc*-HisA. All constructs were verified with PCR, and the pBAD-*tosR*-His_6 construct was verified by DNA sequencing. Plasmids were maintained in LB containing ampicillin (100 μg/mL). Primers used for generating and screening plasmid constructs are described in **Table 2-1**.

lacZ transcriptional fusions of intergenic and intragenic regions within the *tos* operon were generated by cloning the 600 , 233 , and 198 bp regions upstream of *tosR*, *tosC*, and *tosA*, respectively, into the EcoRI and BamHI sites of pRS551 (187), generating pRS551-(P_R, P_C, P_A)-*lacZ*. Constructs were verified by PCR and DNA sequencing in the case of pRS551-P_C-*lacZ* and pRS551-P_A-*lacZ*. Plasmids were maintained in LB containing kanamycin (25 μg/mL). The primers used to generate and screen these transcriptional fusions are listed in **Table 2-1**.

Bioinformatics

A structural prediction of the TosA 335 amino acid tandem repeats (208) was constructed by entering this sequence into the Phyre^2 server (419) under the normal modeling mode option. Likewise, the 101, 105, and 110 amino acid sequences of TosR,

PapB, and FocB, respectively were entered in the Phyre[2] server as above. The highest scoring predicted structure was selected as the putative structure of the TosA repeat amino acid sequence. In addition, the highest scoring models for TosR, PapB and FocB were similarly selected as the putative structure of these proteins. To construct a GC sliding window plot of the *tos* operon, 12,200 bp including the entire *tos* operon and adjacent nucleotide sequences were entered into the seqinr R environment (420). We modified a sliding window plot of GC content program (http://a-little-book-of-r-for-bioinformatics.readthedocs.org/en/latest/src/chapter2.html) to construct a 200 bp sliding window GC content plot of the *tos* operon in R (version 3.0.1), which was fit to a representation of the *tos* locus. Average GC content for each gene and the total *E. coli* CFT073 chromosome was estimated using the sequence statistics feature of SeqBuilder (DNASTAR).

The prevalence of the genes encoding predicted regulators TosR, TosE, and TosF, was estimated by entering the first 100 amino acids of TosA from *E. coli* CFT073 into the BLAST query tool available on the Broad Institutes' UTI Bacteremia initiative website (https://olive.broadinstitute.org/comparisons/ecoli_uti_bacteremia.3) [*E. coli* UTI Bacteremia initiative, Broad Institute (broadinstitute.org), unpublished data]. A search was performed against the genomes present in this database, with an arbitrary E-value cutoff of $1x10^{-20}$. The results from this search represent strains carrying the *tosA* gene. The *E. coli* CFT073 amino acid sequences for TosR, TosE, and TosF were then subjected to the same BLAST search. The resulting hits from these searches were correlated with *tosA* prevalence. Predicted amino acid sequences of TosR, TosE, and TosF variants were

aligned with MegAlign (DNASTAR) using Clustal V. In addition, within the seqinr (420), Biostrings (421), and gdata (422) R environments, we analyzed the GC content of *tos* genes from the UPEC strains described above using several algorithms that we developed.

Deletion mutant and overexpression construct experimental culture conditions

The *tos* operon deletion constructs were cultured at 37 °C in LB containing kanamycin (25 μg/mL) to mid-log phase ($A_{600} \approx 0.5$). Bacteria were harvested at 6000 x *g* for 10 min. The cell pellet was again resuspended in 10mM HEPES, pH 8.3-8.9, and centrifuged again. The bacterial cell pellet was resuspended in 10 mM HEPES, pH 8.3-8.9. The cell suspension was stored at -30 °C prior to quantification with a Pierce BCA Protein Assay Kit (Thermo Scientific) and Western blot.

The pBAD-*tosR*-His_6 construct was induced in the unmarked Δ*tosR* background, with 0.05, 0.2, 0.6, 3, and 10 mM L-arabinose in LB containing ampicillin (100 μg/mL) until the culture reached mid-log phase ($A_{600} \approx 0.5$). Whole cell proteins were collected, stored, and quantified.

To assay *fliC* expression, the pBAD-*tosEF* construct was induced in *E. coli* CFT073 with no or 30 mM L-arabinose in tryptone broth (10 g/L tryptone and 5 g/L NaCl) containing ampicillin (100 μg/mL) for 2.5 h. Material was harvested as described above, with the exception that the culture was centrifuged only once at 1100 x *g* for 10 min prior to resuspension in 10 mM HEPES, pH 8.3-8.9, and stored at -30 °C.

Western blots of deletion mutants and overexpression constructs

To detect TosA, total protein from *tos* deletion mutants, *E. coli* CFT073, or *E. coli* CFT073 (pBAD overexpression constructs) was collected and Western blots were performed. Briefly, equal amounts of total proteins, as determined using a Pierce BCA Protein Assay Kit (Thermo Scientific), from specific constructs were resolved, transferred, and blotted with polyclonal Anti-TosA antibodies (208), polyclonal Anti-FliC antibodies (414), or an anti-His_6 antibody (Invitrogen).

β-galactosidase assay

Miller assays were performed as previously described (423) with the exception that bacteria harboring the *lacZ* transcriptional fusions described above were cultured to mid-log phase ($A_{600} \approx$ 0.6-0.8) in LB containing kanamycin (25 μg/mL) and, after resting on ice and centrifugation, were resuspended in Z Buffer (pH 7.0; 60 mM $Na_2HPO_4 \cdot 7H_2O$, 40 mM NaH_2PO4, 1 M KCl, 1 mM $MgSO_4$, 50mM β-mercaptoethanol).

TosR-His_6 purification

TosR-His_6 protein was isolated by incubating 50 mL cultures of CFT073 harboring pBAD-*tosR*-His_6 to mid-log phase ($A_{600} \approx$ 0.5) in LB containing ampicillin (100 μg/mL) and subsequently inducing expression with 10 mM arabinose for 2.5 hours. Bacteria from this culture were pelleted at 2700 x *g* and stored at -30 °C. TosR-His_6 was extracted using a modified QIAexpressionist protocol (Qiagen). Briefly, the cell pellet was resuspended in lysis buffer (pH 8.0; 50 mM NaH_2PO_4, 300 mM NaCl, and 40 mM

imidazole) and passed three times through a French Pressure Cell at 1200 lbs/in^2. Cellular debris was cleared by centrifugation at 10,000 x *g*. The Ni-NTA agarose (Invitrogen) was equilibrated as described in the QIAexpressionist protocol (Qiagen), with the exception that half of the volume of Ni-NTA and lysis buffer were used in this step. Cleared lysate was incubated with Ni-NTA agarose at room temperature for 30 min and subsequently ran through a column. The column bed was washed three times with washing buffer (pH 8.0; 50 mM NaH_2PO_4, 300 mM NaCl, and 60 mM imidazole) and bound proteins were eluted with elution buffer (pH 8.0; 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM imidazole). Eluted proteins were concentrated with 10 kDa Amicon Ultra Centrifugal Filters (Millipore), and quantified using a 2-D Quant Kit (GE Healthcare); this concentrate was dissolved in 10 mM HEPES, pH 8.3-8.9. Purity of the TosR-His_6 concentrate was assessed on a 12% SDS-polyacrylamide and staining with SimplyBlue SafeStain (Life Technologies). The presence of TosR-His_6 was confirmed by Western blot as described above.

Electrophoretic mobility shift assay of the *tos* operon promoter

PCR was performed to amplify P_{tosR} and *lacZ* DNA probes using the gel shift primers described in **Table 2-1** and Easy-A high fidelity enzyme (Agilent). These probes were terminally labeled with Digoxigenin-11-ddUTP (DIG-ddUTP) using a 2nd generation DIG Gel Shift Kit (Roche Applied Science). Assessing probe labeling efficiency, TosR-His_6 DNA binding reactions, and resolving and detecting shifted DNA probes were all performed as described in the same kit protocol, with the exception that

the 25-min binding reactions contained only binding buffer (pH 7.6; 100 mM HEPES, 5 mM EDTA, 50 mM $(NH_4)_2SO_4$, 5 mM DTT, 1% (vol/vol) Tween, 150 mM KCl) in addition to proteins and labeled or unlabeled DNA probes. The concentrations of gel running and transfer buffers were increased to 1X TBE (89.0 mM Tris, 89.0 mM boric acid, 2.0 mM EDTA, pH 8.0), and the Anti-DIG-AP detection antibody dilution used was decreased to 1:1000, from 1:10,000.

Unlabeled ~100bp P_{tosR} fragments were generated using primers described in **Table 2-1** as above. An electrophoretic mobility shift assay was again performed as above on DIG-ddUTP P_{tosR} using the aforementioned PCR products as unlabeled competitors. However, 5% the amount of labeled probe was used in these competition electrophoretic mobility shift competition assays, compared with the above reactions.

RNA extraction and RT-PCR

cDNAs were synthesized from equal amounts of RNAs extracted from wild type *E. coli* CFT073 and Δ*tosR* constructs in exponential phase (A_{600}≈0.4-0.5) cultured in LB as previously described (227). An exception to this extraction protocol was to stop half the amount of cellular material previously described with half the amount of stopping solution (5% phenol in ethanol). RT-PCR was performed on equal amounts of the above cDNAs using the primers directed against the *tosAE* intergenic region or the *tosEF* intergenic region are described in **Table 2-1**. Equal volumes of each PCR were loaded into sample lanes, and DNA amplicons were resolved on a 1% (wt/vol) agarose gel.

Motility assays

Overnight cultures of *E. coli* CFT073 harboring pBAD-*myc*-HisA, pBAD-*tosR*-His_6, and pBAD-*tosEF* were normalized to A_{600}=1.0 in 10 mM HEPES, pH 8.3-8.9 and stabbed into soft agar (10 g/L tryptone, 5 g/L NaCl, 2.5 g/L agar) containing ampicillin (100 μg/mL) with 33.3 mM L-arabinose. After 17 hours of incubation at 30 °C, the diameter of the zone of swimming was measured. Δ*tosR*, Δ*tosE*Δ*tosF*, Δ*tosR*Δ*tosE*Δ*tosF*, and wild-type constructs were also assayed as above, with the exception that the soft agar did not contain antibiotics or L-arabinose.

Growth curve generation

Overnight cultures of *E. coli* CFT073 harboring pBAD-*myc*-HisA, pBAD-*tosR*-His_6, and pBAD-*tosEF* were diluted 1:100 into tryptone broth (10 g/L tryptone and 5 g/L NaCl) containing ampicillin (100 μg/mL) and 30 mM L-arabinose. Constructs were cultured at 30 °C for 24 hours in a Bioscreen C Automated Growth Curve System, with A_{600} readings recorded every 15 min. This procedure was the same for Δ*tosR*, Δ*tosE*Δ*tosF*, Δ*tosR*Δ*tosE*Δ*tosF*, and wild-type constructs, with the exception that the tryptone broth did not contain antibiotics or L-arabinose.

Statistical analysis

Statistical significance of all single comparisons were determined using an unpaired Student's *t*-test. Multiple comparisons were made using an unpaired ANOVA followed by a Tukey's multiple comparisons test. With the exception of the *tos* operon

GC content plot graphed in R (version 3.0.1), all other graphs and statistical testing were performed in GraphPad Prism (version 6.0).

Results

The *tos* operon of *E. coli* CFT073 encodes the TosA adhesin, type 1 secretion system, and three putative regulators

In *E. coli* CFT073, the *tos* operon encodes the high molecular weight RTX nonfimbrial adhesin, TosA (**Figure 2-1A** and **2-1B**). In addition, the *tos* operon encodes genes for a putative type 1 secretion system, *tosCBD*. TosC is predicted to form the outer membrane pore through which TosA is released from the secretion system, TosB is the predicted ATPase/TosA recognition factor, and TosD forms the predicted periplasmic channel through which TosA passes (69, 208). Three ORFs of previously unknown function, now annotated *tosR*, *tosE*, and *tosF*, are also located within the *tos* locus. TosR is a homolog of the PapB family of adhesin regulators, and TosE and TosF align with LuxR family members (208).

To assess whether TosA may possess structural features found in other nonfimbrial adhesins (50), a bioinformatics approach was taken to predict TosA features. We previously identified that TosA contained five tandem 335 amino acid sequence repeats (208). A Phyre^2 model (**Figure 2-1A**) (419) predicts that the structure of these repeats is similar to that of the bacterial immunoglobulin-like domain group three (BIg 3)

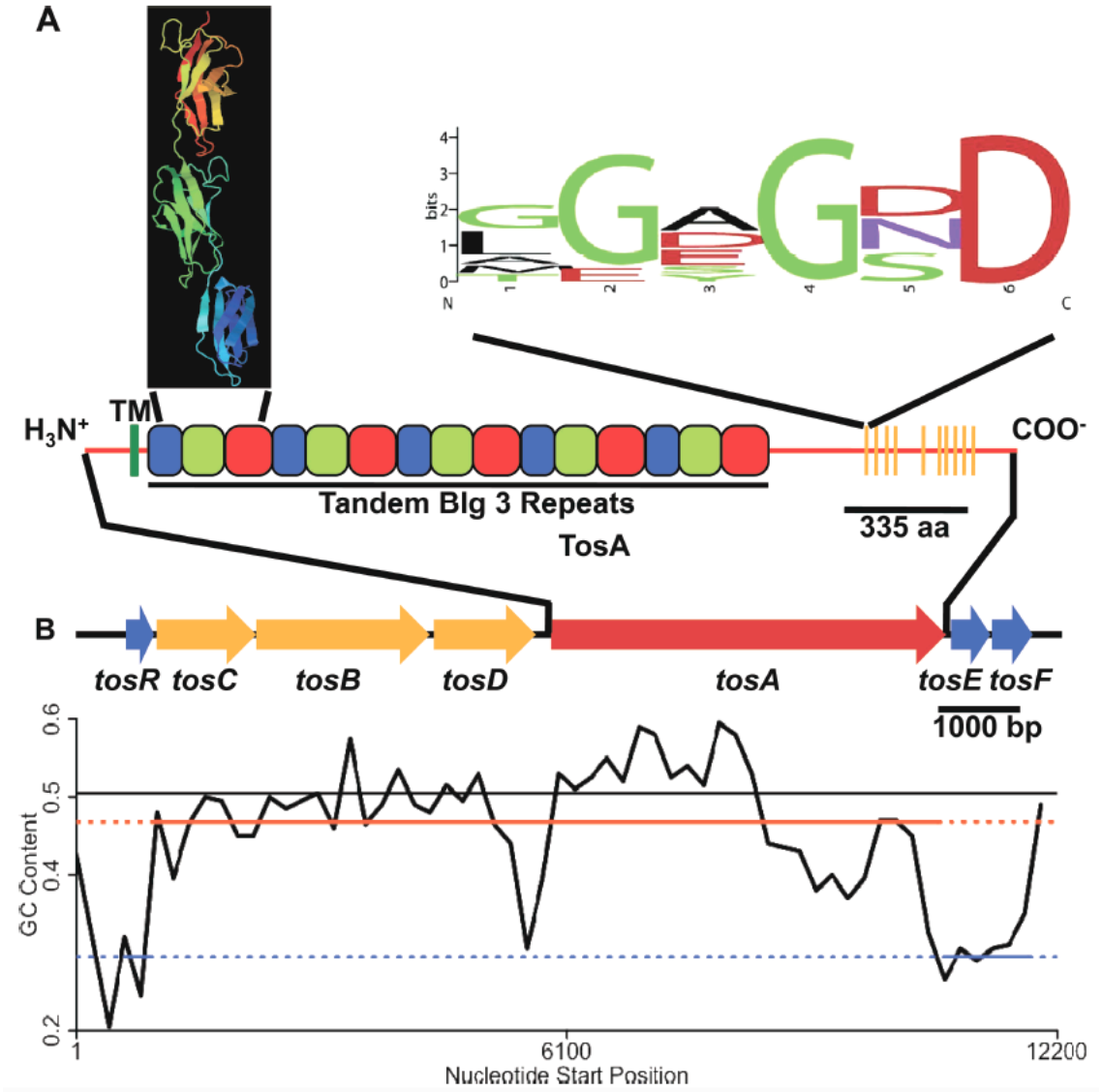

Figure 2-1. The *tos* operon encodes the genes for the RTX non-fimbrial adhesin TosA, a secretion system, and putative regulators, but does not have uniform GC content. (A) The entire predicted 2516 amino acid (aa) sequence of TosA is represented by a red horizontal line. Near the amino-terminus, a predicted transmembrane domain is designated 'TM' (vertical green line). Tandem blue, green, and red boxes represent the predicted bacterial immunoglobulin-like (BIg) family 3 folds. The corresponding predicted Ig fold structures (modeled using Phyre[2] against SiiE from *S. enterica*, with 98.5% confidence) are represented in the black box insert. Near the carboxyl-terminus, the positions of ten tandem RTX repeats are denoted with orange vertical lines. The sequence logo of the RTX repeats is noted (208). **(B)** Within the *tos* locus, blue arrows represent genes (*tosR* and *tosEF*) encoding predicted DNA binding proteins, orange

arrows represent genes (*tosCBD*) encoding a predicted type 1 secretion system, and a red arrow represents the gene (*tosA*) encoding the RTX non-fimbrial above. The entire *tos* locus is fit to a 200 bp sliding window plot of the GC content. A black line denotes the average GC content of 80 UPEC genomes, while orange-red and blue lines represent the average GC content of *tosCBDA,* and *tosR* and *tosEF* in the same 80 genomes, respectively.

repeats found in the *Salmonella enterica* nonfimbrial adhesin SiiE (216). Like other RTX nonfimbrial adhesins (50), TosA also contains 10 RTX repeats near its carboxyl-terminus (208) and a putative transmembrane domain near its amino terminus.

tos operon genes are broadly conserved among UPEC strains

As the GC content of genes within an operon tends to be similar (424), to determine whether this was also the case for the *tos* operon, we tracked the *tos* operon GC content using a 200 bp sliding window (**Figure 2-1B**). The GC content of the structural genes (*tosCBDA*), 48.5%, is similar to that of the *E. coli* CFT073 backbone (averaging 50.5%). However, the GC content of the putative regulatory genes, *tosR, tosE,* and *tosF*, is 29.1%, significantly lower than that of the chromosome, in general, and that of *tosCBDA*, in particular. Thus, these data reveal that the putative *tos* operon regulatory genes have distinct GC content.

Given the differences in GC content among *tosR* and *tosEF* in *E. coli* CFT073, to determine whether these same regulatory genes are conserved in *tos* operons of other UPEC strains, we analyzed 317 sequenced genomes found in the Broad Institute's UTI Bacteremia initiative database [*E. coli* UTI Bacteremia initiative, Broad Institute (https://olive.broadinstitute.org/comparisons/ecoli_uti_bacteremia.3)]. Among UPEC strains associated with this study, *tosA* was found in 32% (101 of 317) of isolates, a slightly higher prevalence than previously estimated by us in a PCR-based survey of our UPEC strain collection (35). Of the strains encoding *tosA*, the overwhelming majority [92 of 101 (91%)] also encoded *tosR* and *tosEF*. Among strains with single copies of

tosA and at least one gene among *tosR* and *tosEF* (96/101), our GC content analysis program was able to determine GC content for 83.3% (80/96) of these strains. The GC content disparities among *tos* structural and regulatory genes described above held among these strains encoding the *tos* genes (**Figure 2-1B**). Thus, while the GC content of *tosR* and *tosEF* (29.5%) is considerably lower than that of *tosCBDA* (46.8%), these genes are conserved and linked to each other within UPEC strains. In agreement with this, we did not find these putative regulators in a UPEC background not encoding *tosA*. However, 4.0% (4/101) and 7.9% (8/101) of strains that encode *tosA* do not encode *tosR* and *tosEF*, respectively. Therefore, the *tosR* and *tosEF* regulators are conserved and linked among the overwhelming majority of sequenced UPEC strains encoding *tosA*. This conservation among UPEC strains, despite GC content differences, strongly suggests that *tosR* and *tosEF* encode factors promoting fitness in some situations.

TosR, TosE, and TosF amino acid sequences show a clonal nature to the *tos* operon

To determine whether the TosR amino acid sequence is conserved among UPEC strains, the predicted TosR amino acid sequence was aligned against the UPEC strains encoding genes for TosR in the Broad Institute's UTI Bacteremia initiative database (**Figure 2-2A**). Based on the unique TosR amino acid sequences, we assigned each TosR sequence into one of five sequence variant groups. All variants have an overall predicted sequence identity of 59.0% among each other. Variant 1 accounts for 84.5% of TosR sequences and is the variant found in *E. coli* CFT073. Variants 2, 3, 4, and 5 are less prevalent, accounting for 3.1%, 7.2%, 3.1%, and 2.1% of TosR sequences, respectively.

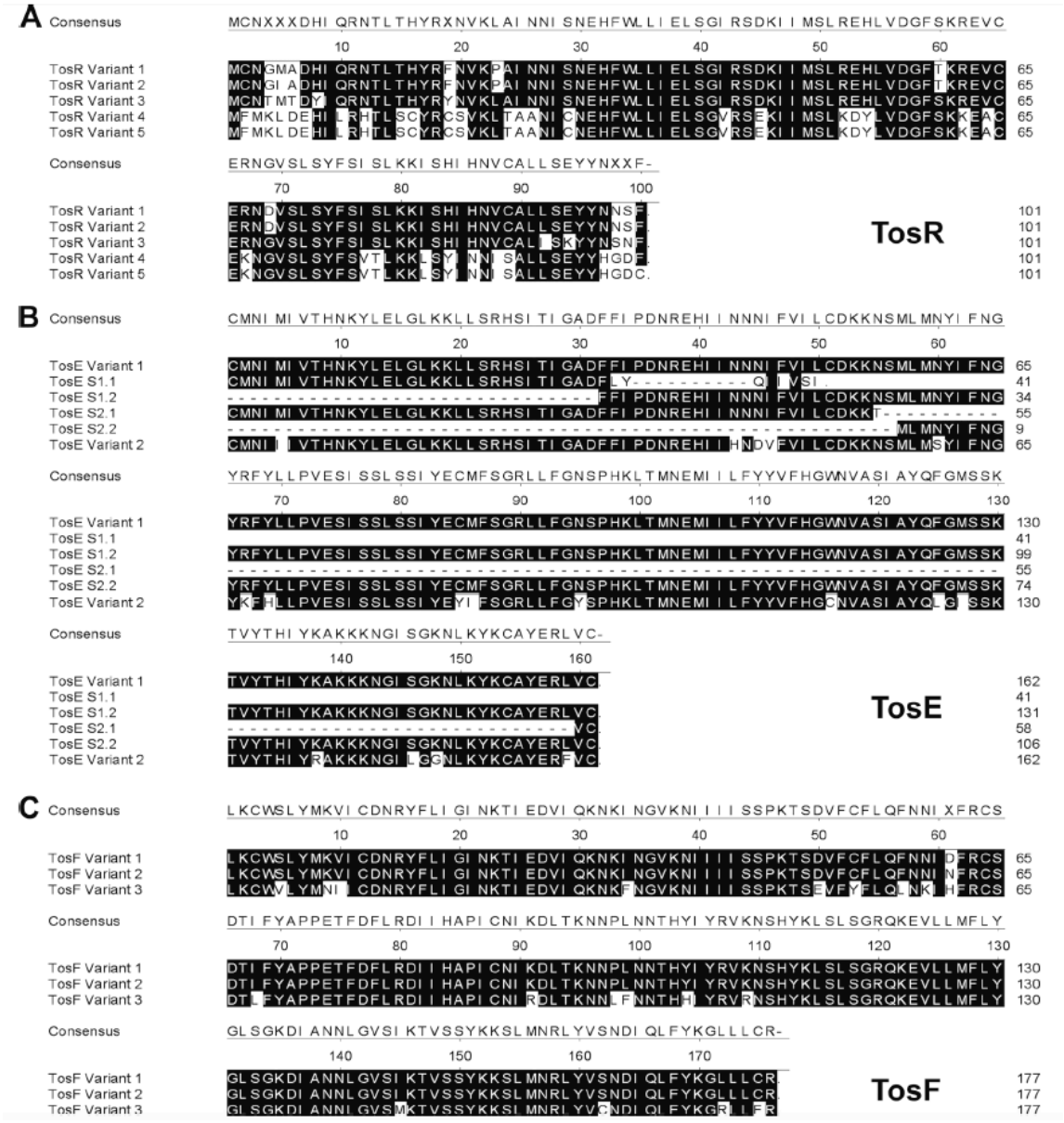

Figure 2-2. Multiple sequence alignment of TosR, TosE, and TosF amino acid sequence variants. (A) Shaded residues denote conservation with the predicted TosR consensus sequence where at least three residues at a given position are conserved. **(B)** The shaded residues are those residues conserved with the predicted TosE consensus sequence where at least four residues at a given position are conserved. The sequences denoted 'S' are predicted to result from a TosE frameshift mutation. **(C)** Residues that

match the TosF consensus, where at least two residues at a given position are conserved, are shaded.

It is interesting to note that L35, L36, L55, V56, Y74, F75, and S76 are completely conserved among all TosR sequence variants. These residues have been previously shown to be involved in PapB oligomerization (425). In addition, E53 and H54, of TosR sequence variants 1-3, and D53 and Y54, of TosR sequence variants 4 and 5, have identical or similar properties to D53 and Y54 of PapB, which are again involved in PapB oligomerization (425). Further conservation is observed at residue C65, which is important for PapB DNA binding (425). Likewise, K61 is completely conserved among all TosR sequence variants and has similar properties to R61 of PapB, which is important for DNA binding (425). Thus, from these observations, it can be suggested that strong selective pressure drives sequence conservation among TosR sequence variants. It is also intriguing to note that TosR sequence variants 4 and 5 appear to be encoded on a putative plasmid (see **Appendix A**, **Table A-1**)

TosE and TosF are also conserved among UPEC isolates. As with TosR, TosE and TosF were grouped based on their unique amino acid sequences. For TosE, there are four variants, including two frameshift mutants that are predicted to disrupt TosE function (**Figure 2-2B**). Variant 1 accounts for 79.6% of TosE sequences and is the variant found in *E. coli* CFT073. Frameshift mutants 1, 2, and variant 2 account for 7.5%, 5.4%, and 7.5% of TosE sequences, respectively. TosF has three sequence variants, albeit two are virtually identical (**Figure 2-2C**). Variant 1 accounts for 86.0% of sequences and is also the variant found in *E. coli* CFT073. TosF variants 2 and 3 account for 7.5% and 6.5% of amino acid sequences, respectively. A strain that encodes TosR variant 1 encodes TosE variant 1, frameshift mutant 1, or frameshift mutant 2 and TosF

variant 1 or 2. If a strain encodes TosR variant 2, it always encodes TosE frameshift mutant 2 and TosF variant 1. All strains that encode TosR variant 3 encode TosE variant 2 and TosF variant 3. However, little consistent homology with non-hypothetical LuxR family members makes functional characterization of conserved residues in TosEF variants difficult, compared to TosR above.

TosR is a negative regulator of *tosA* and a PapB family homolog

To test the hypothesis that one or more of the identified putative regulators associated with the *tos* operon exert a regulatory function on *tosA* expression, deletion mutations of these putative regulatory genes were constructed in *E. coli* CFT073. Deletion of *tosR* resulted in a substantial increase in TosA production as assessed by Western blots of whole cell proteins using anti-TosA serum (**Figure 2-3A**). However, no change in TosA production was observed after deletion of *tosE* and *tosF*. Additionally, overexpression of *tosE* and *tosF* did not result in altered TosA levels (see **Appendix B**, **Figure B-1**). Overexpression of TosR-His_6 from an arabinose-inducible construct partially complemented the *tosR* deletion in the unmarked mutant background, repressing *tosA* expression (**Figure 2-3B**), consistent with TosR being a negative regulator of *tosA* and that TosR-His_6 is biologically activity.

To rule out whether increased *tosA* expression was due to a secondary mutation, we transduced the *tosR* deletion mutation into a clean *E. coli* CFT073 background. This Δ*tosR* mutation transductant still overproduced TosA, compared to the wild-type *E. coli* CFT073 parental strain (**Figure 2-3C**). However, removing the *aph* cassette from the

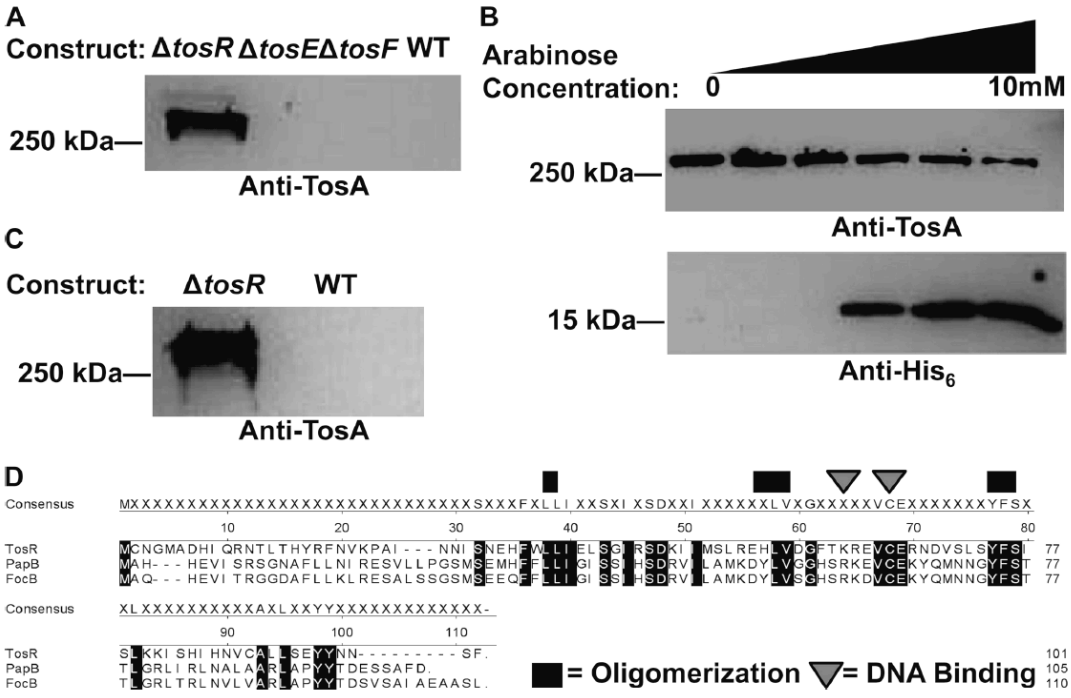

Figure 2-3. *tosA* is negatively regulated by the PapB family member, TosR. (A) A Western blot with polyclonal anti-TosA serum reveals that TosA (~250-kDa) is overproduced in a Δ*tosR* construct, but is poorly produced in both the Δ*tosE*Δ*tosF* construct and WT [wild-type (Δ*lacZ*)] *E. coli* CFT073. **(B)** A *trans*-complementation assay in a Δ*tosR* background using TosR-His_6 induced from plasmid (pBAD-*tosR*-HisA) with arabinose concentrations ranging from 0 to 10 mM shows that TosA levels (detected as above), are inversely related to TosR-His_6 (~15 kDa) levels (detected on a Western blot using a His_6 antibody). **(C)** TosA levels [detected as in **(A)** and **(B)**] remain high in a Δ*tosR aph*$^+$ phage transduced construct, as compared to the wild-type control. All lanes in a respective Western blot were loaded with equal amounts of whole cell protein as determined using a Pierce BCA Protein Assay Kit (Thermo Scientific). **(D)** Alignment of TosR (variant 1), FocB, and PapB reveals that all three share amino acid sequence identity at domains previously shown to be important for oligomerization (black boxes) and DNA binding (grey arrows). The consensus sequence represents residues conserved between the shown PapB family members.

same transductant reveals that this *tosR* mutation did not result in overproduction of TosA (see **Appendix C**, **Figure C-1**), which may suggest an additional mutation could promote TosA synthesis and be a caveat for the TosA overproduction and complementation studies above.

The predicted TosR amino acid sequence of *E. coli* strain CFT073, when subjected to BLAST, identified PapB family regulators as potential homologs. Using sequence alignment against other PapB family regulators found in *E. coli* strain CFT073 (**Figure 2-3D**), TosR shares 27.7% amino acid sequence identity with PapB and 26.7% identity with FocB. As noted above, TosR, PapB, and FocB also carry conserved residues previously shown to contribute to oligomerization and DNA binding (**Figure 2-3D**) (425). All three proteins share significant sequence similarity with each other, and the predicted TosR protein structure is nearly identical to the structure of FocB (202) (**Figure 2-4A**, **Figure 2-4B**, and **Figure 2-4C**).

TosR-His$_6$ binds to P$_{tosR}$, which contains the *tos* operon promoter

To determine whether the promoter driving *tosA* expression could be identified, we generated *lacZ* transcriptional fusions of *tos* intergenic and intragenic regions (**Figure 2-5A**), which, from their location upstream of *tosA* or its cognate secretion system, were predicted to be the most probable location of P$_{tos}$ (**Figure 2-5B**). These included the 600 bp upstream of *tosR* (with the addition of the first 30 bp of *tosR*), the 233 bp upstream of *tosC* (including the final 181 bp of *tosR*), and the 199 bp between *tosD* and *tosA*. Using Western blot, we attempted to find an optimal condition for TosA synthesis. Conditions

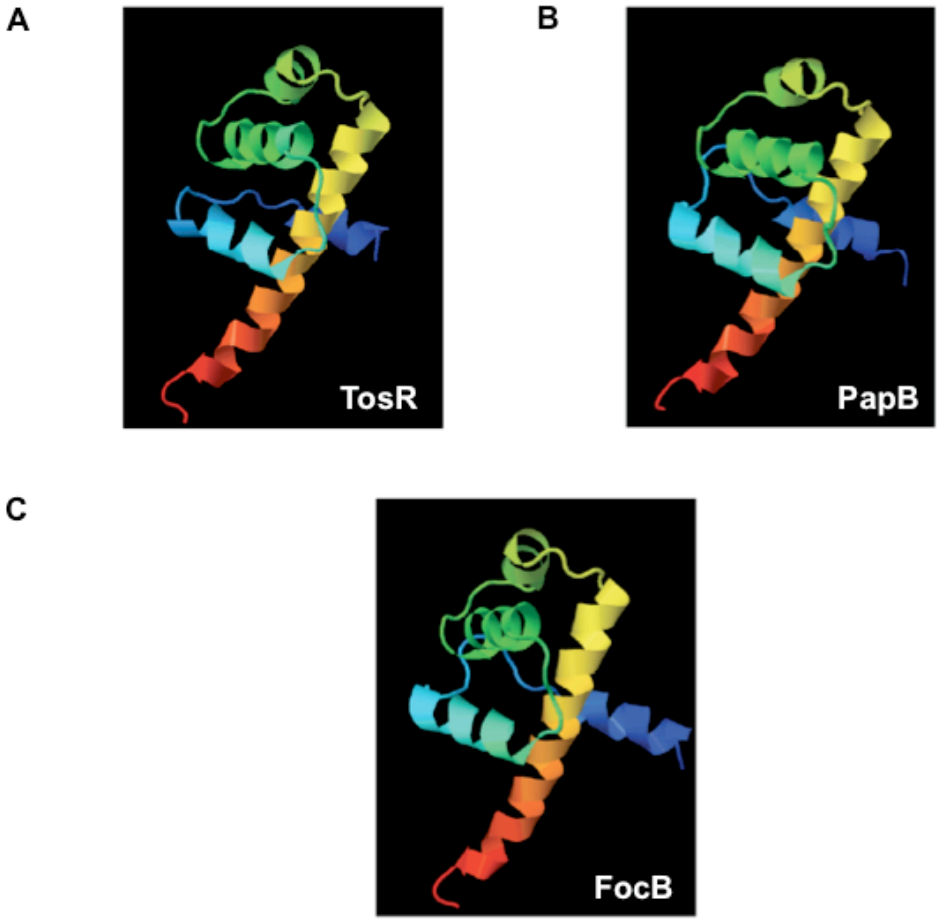

Figure 2-4. Predicted structures of PapB family homologs. Phyre[2] models based off of the amino acid sequences of TosR **(A)** and PapB **(B)** are indicated (both modeled against FocB with 100% confidence). The previously solved structure of FocB **(C)**, which was used to model TosR and PapB, above, is also depicted.

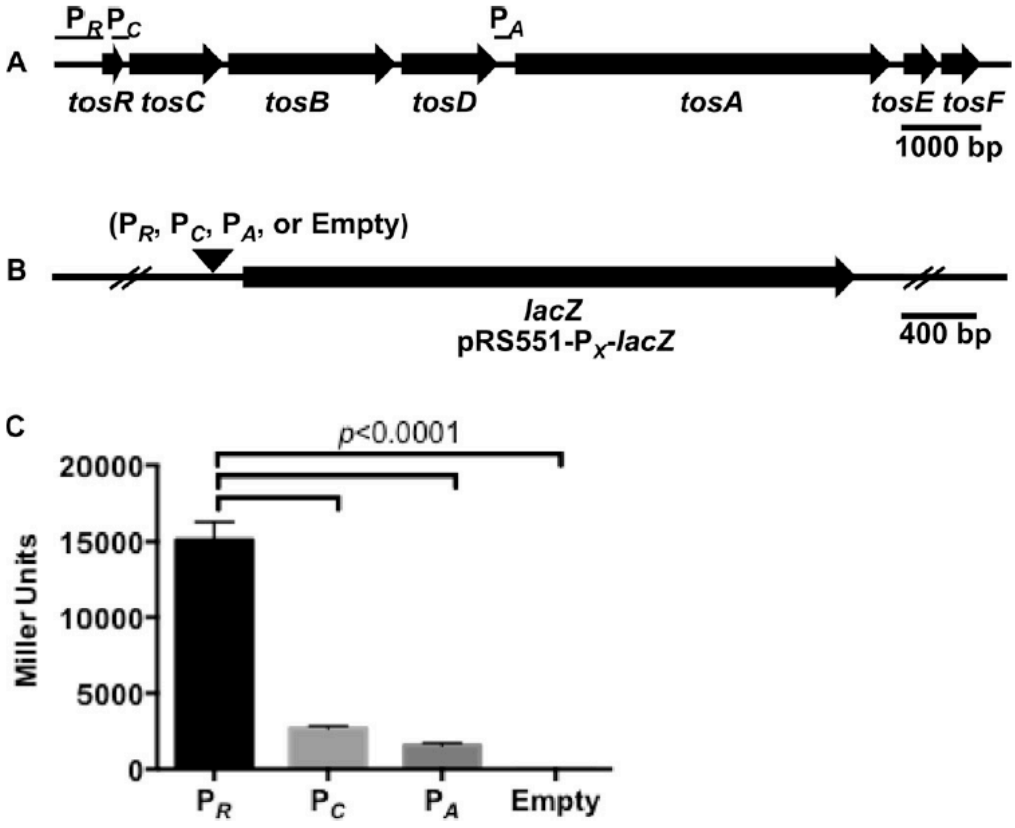

Figure 2-5. The region upstream of *tosR* exhibits transcriptional activity. (A) Within a representation of the *tos* operon, sequences harboring possible sites for *tos* promoters are denoted P_R, P_C, and P_A. **(B)** A black arrow indicates the position where the P_R, P_C, and P_A sequences are inserted into the BamHI and EcoRI sites upstream of *lacZ* in pRS551 to produce the indicated transcriptional fusions. The empty construct was native pRS551 plasmid. **(C)** The activity of each transcriptional fusion was assayed by the Miller assay. Transcriptional activity, measured indirectly through β-galactosidase activity (in Miller units), associated with the P_R construct is significantly higher than that of P_C and P_A constructs ($p<0.0001$). Black or grey bars indicate average values of Miller units for each construct ($n=6$). Significance was determined by using Tukey's multiple comparisons test following ANOVA ($p<0.0001$). Error bars indicate SD about the mean.

included culturing wild-type *E. coli* CFT073 to stationary phase, during exponential phase, cultured statically, exposed to different osmotic stresses, exposure to human urine, low iron, and different carbon sources. However, none of the aforementioned conditions resulted in reproducibly high TosA synthesis, compared with the Δ*tosR* mutation cultured to exponential phase (data not shown). Thus, the transcriptional fusions were all assayed using the Miller assay in the Δ*tosR* background cultured to exponential phase, which is an optimal condition for *tos* operon expression. All three putative constructs have elevated transcriptional activity, compared to the empty vector control (**Figure 2-5C**). However, the construct designated P_R, which contains P_{tosR}, was significantly upregulated among all constructs ($p<0.0001$). In full agreement with our previous findings regarding *tos* operon structure (208), we concluded that P_{tosR} contains P_{tos}, the *tos* operon promoter.

PapB family members usually mediate a regulatory function through binding DNA upstream of a PapB family member gene. To determine whether this is the case for TosR, we performed an electrophoretic mobility shift assay on the DNA sequence upstream of *tosR*, P_{tosR} (**Figure 2-6A**). TosR-His_6 shifted labeled P_{tosR} DNA, and competition with unlabeled P_{tosR} in the same binding reaction inhibited labeled probe shifting (**Figure 2-6B**). TosR-His_6 did not bind to an unrelated labeled *lacZ* sequence, demonstrating that TosR-His_6 binds specifically to P_{tosR} DNA.

The sequence immediately upstream of a gene encoding a PapB family member is often AT-rich (182, 426), where tandem repeated AT-rich nonomers often demarcate PapB family member binding sites (182, 247). Indeed, the region immediately upstream of *tosR* is AT-rich (**Figure 2-1B**). Therefore, to identify the putative TosR binding site,

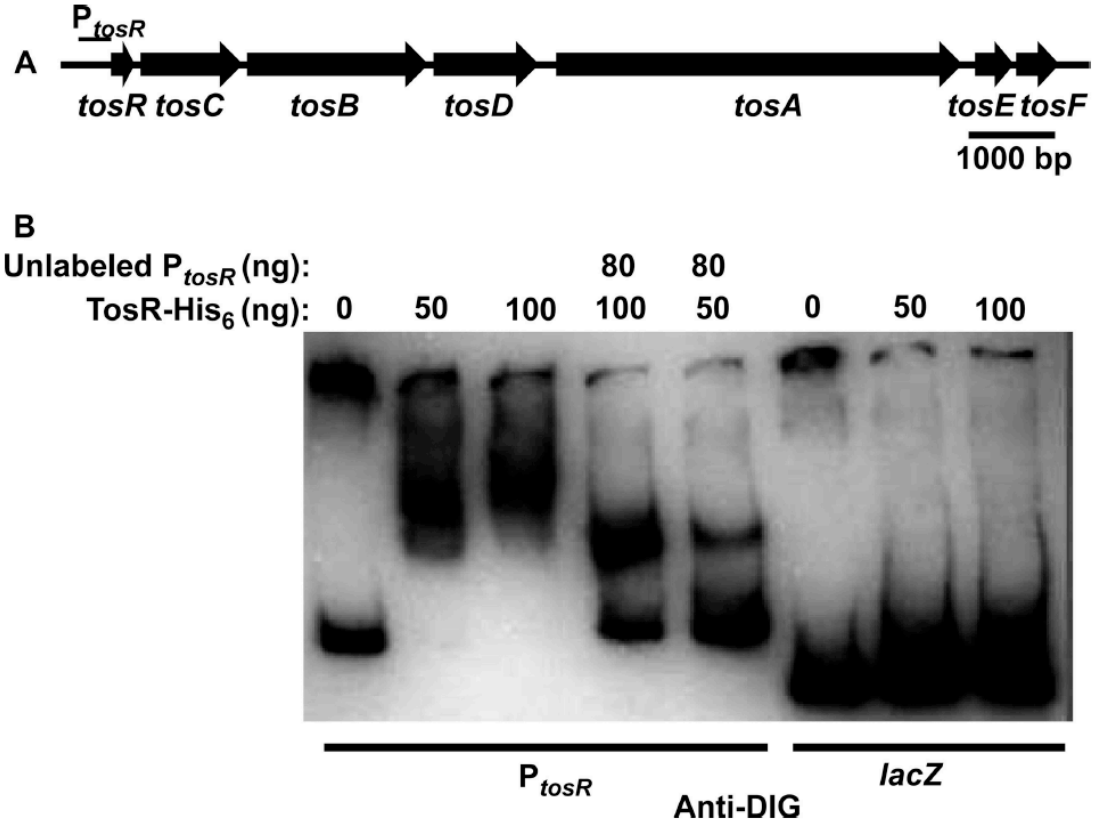

Figure 2-6. TosR-His_6 binds DNA derived from the region upstream of *tosR*. (A) Within a representation of the *tos* operon, the location of the 399 bp P_{tosR} sequence used for the DNA binding assay below is indicated. **(B)** Digoxigenin (DIG) terminally labeled P_{tosR} or a *lacZ* fragment was treated with the indicated amounts of TosR-His_6 and with or without excess unlabeled P_{tosR}, as indicated. Shifted and unshifted probes were detected with an anti-DIG antibody.

we PCR-amplified seven ~100 bp fragments of P_{tosR}, each overlapping by 50 bp. An electrophoretic mobility shift assay was performed using these P_{tosR} fragments as unlabeled competitors (**Figure 2-7A**). Unlabeled $P_{tosR}2$ (220-319 bp upstream of *tosR*) unshifted labeled P_{tosR}, while all adjacent fragments did not. Plotting the positions of the seven P_{tosR} fragments reveals that an AT-rich repetitive sequence is complete only in $P_{tosR}2$, with only partial presence in the two adjacent fragments $P_{tosR}1$ and $P_{tosR}3$ (**Figure 2-7B**). Thus, we predict that the TosR binding site is centered within this sequence.

TosE and TosF contribute to the reciprocal regulation of adherence and motility

It has been previously observed that adhesin operons regulated by PapB family members harbor genes that encode proteins that suppress motility. These genes are located at the 3' end of adhesin operons (187, 188, 355, 427). To first confirm that *tosEF* are indeed part of the *tos* transcript, we performed RT-PCR using primers directed against the junctions between *tosA* and *tosE* and *tosE* and *tosF*. We found that *tosEF* were part of the *tos* transcript, as a strain that overexpresses *tosA* (Δ*tosR*) has a corresponding higher amount of transcript with the *tosAE* (**Figure 2-8A**) and *tosEF* junctions (**Figure 2-8B**).

To determine whether TosEF affect motility, we performed motility assays in the presence and absence of *tosE* and *tosF* expression. TosEF overproduction resulted in a substantial decrease in swimming motility when this construct was stabbed into soft agar and incubated for 17 hours, as compared to an empty vector control ($p<0.0001$) (**Figures 2-9A** and **2-9B**). Likewise, compared to the Δ*tosR* and WT backgrounds, a

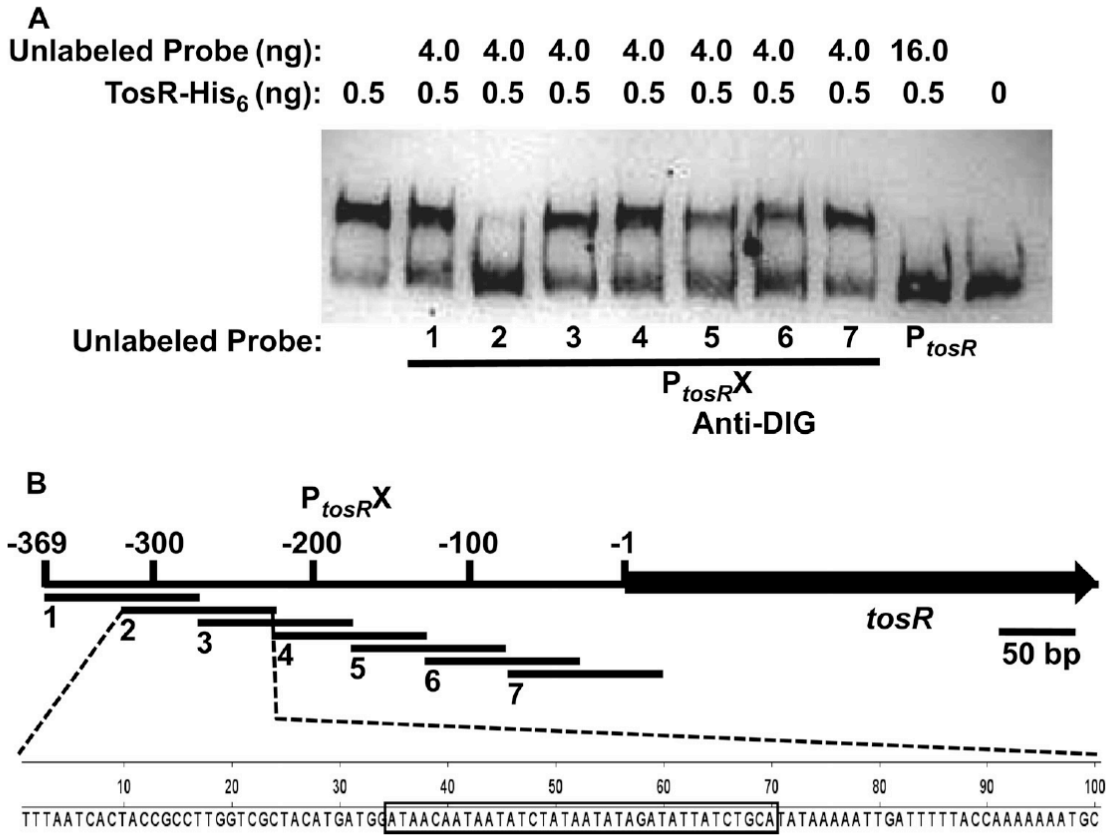

Figure 2-7. The TosR binding site is centered on $P_{tosR}2$. (A) The Digoxigenin (DIG) terminally labeled P_{tosR} probe is unshifted only by addition of excess unlabeled $P_{tosR}2$ or full length unlabeled P_{tosR}. **(B)** P_{tosR} fragments 1-7 are indicated on a P_{tosR} region schematic. The complete sequence of $P_{tosR}2$ is indicated with a boxed region indicating the predicted TosR binding site. The *tosR* open reading frame (ORF) is indicated.

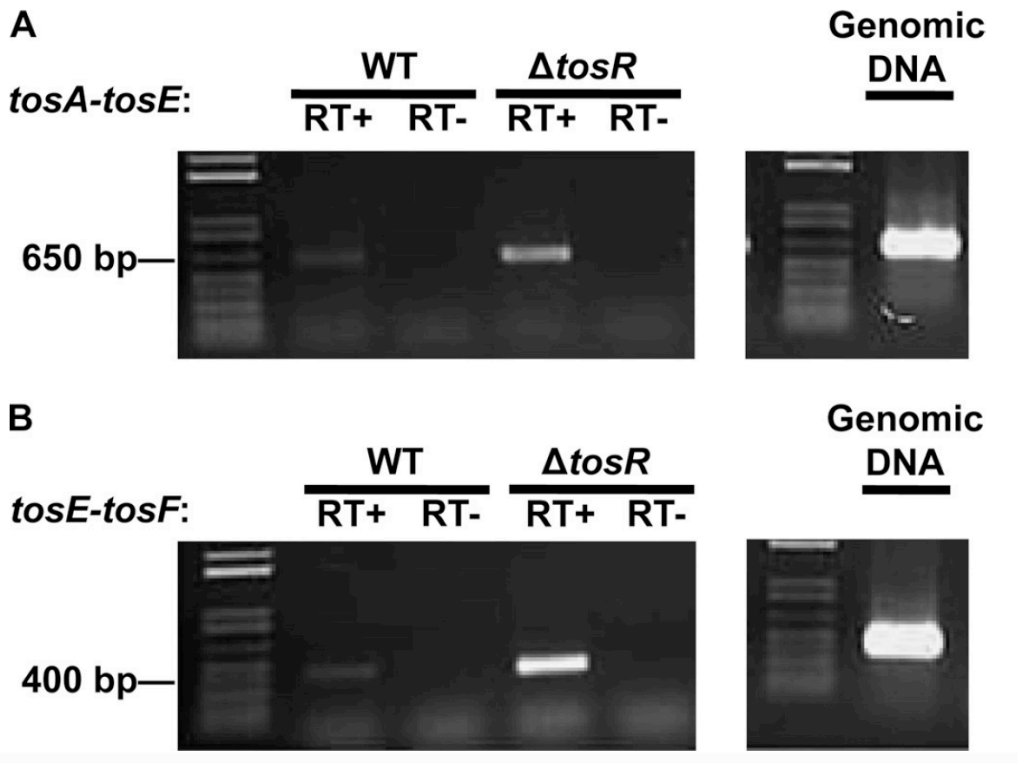

Figure 2-8. *tosEF* are part of the *tos* operon. RT-PCR assays with primers directed against the intergenic region between *tosAE* **(A)** and *tosEF* **(B)** were performed in the indicated backgrounds. The expected fragment sizes of 619 bp **(A)** and 402 bp **(B)** are only observed in reactions containing cDNAs (RT+), but not in reactions containing only input RNA (RT-). Equal amounts of input RNAs were used to synthesize all cDNAs, and equal amounts of all cDNAs were used as inputs in the PCRs shown. Likewise, equal volumes of all PCRs were loaded in each lane.

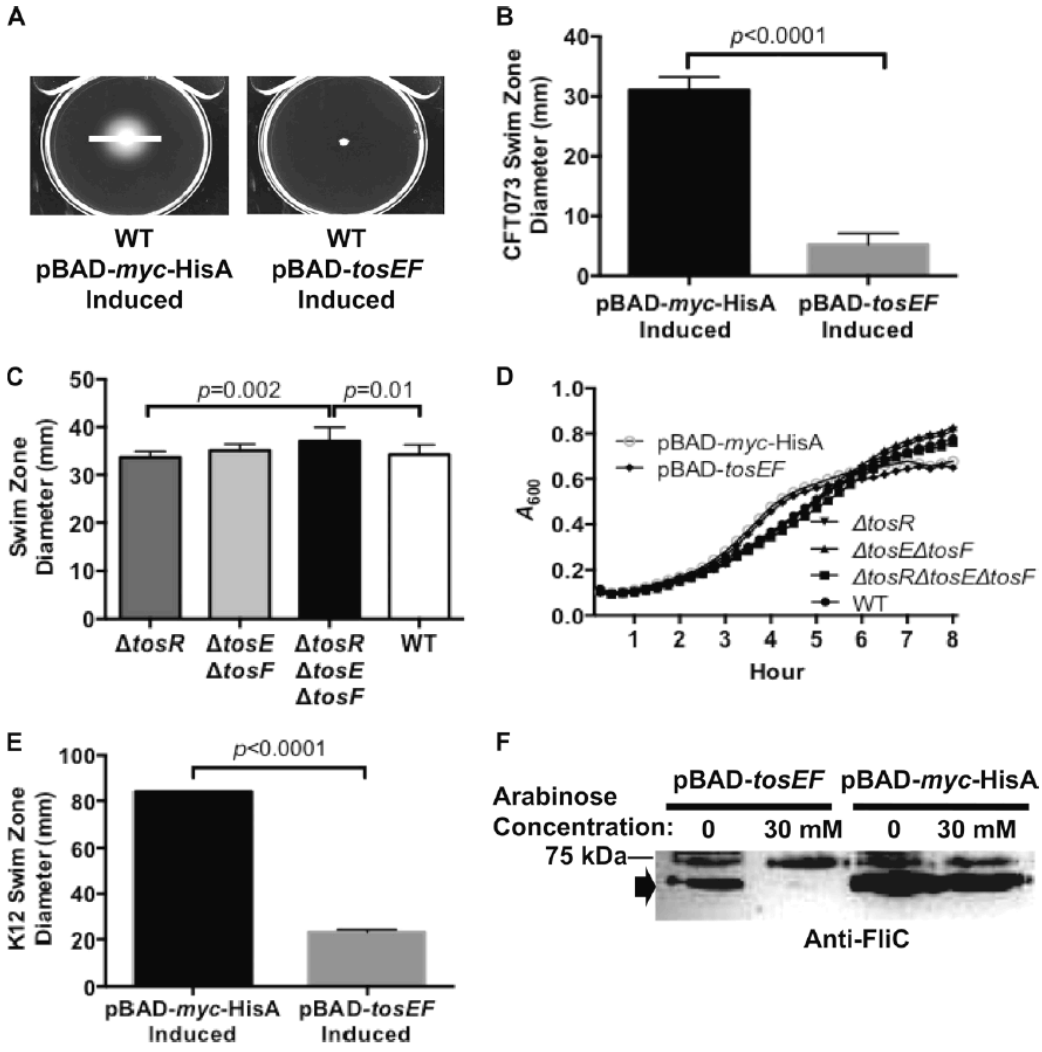

Figure 2-9. TosEF down-modulate motility. **(A)** UPEC CFT073 harboring either pBAD-*tosEF* or pBAD-*myc*-HisA was stabbed into soft agar and incubated for 17 hours with 33.3 mM L-arabinose. White lines represent the swimming zone diameter for each induced construct. **(B)** Average swimming zone diameters are represented by black or grey bars (n=9). Error bars represent SD about the mean, and significance between diameter differences was determined using Student's t-test (p-values indicated). **(C)** Average (n=9) swimming zone diameters of the indicated deletion backgrounds or wild-type UPEC CFT073, measured after 17 hours, are represented by black, grey, or white bars (n=9). Error bars are the same as above. Significance, indicated by associated p-values, was determined by using Tukey's multiple comparisons test following ANOVA (p=0.002). Error bars indicate SD about the mean **(D)** Growth curves of the indicated constructs are traced over a period of eight hours. Each point represents an average A_{600} reading at a given time point (n=12). **(E)** *E. coli* K12 MG1655 harboring either pBAD-

myc-HisA or pBAD-*tosEF* was stabbed into soft agar and incubated 17 hrs with 33.3 mM L-arabinose. Average (n=9) swimming zone diameters of each overexpression construct are represented by black or grey bars. Error bars are the same as above, and significance between diameter differences was determined using Student's *t*-test (p-values indicated). **(F)** Western blot with polyclonal anti-FliC serum reveals that FliC (~65 kDa) (black arrow) levels are reduced in wild-type CFT073 harboring pBAD-*tosEF* induced with the indicated concentrations of L-arabinose, as compared to the same background harboring pBAD-*myc*-HisA. All lanes of this Western blot were loaded with equal amounts of whole cell protein as determined by using a Pierce BCA Protein Assay Kit (Thermo Scientific).

Δ*tosR*Δ*tosE*Δ*tosF* mutant was statistically significantly, but modestly, more motile in soft agar ($p \leq 0.01$) (**Figure 2-9C**). The differences in swimming motility, in soft agar, between Δ*tosR*, Δ*tosE*Δ*tosF*, and WT CFT073 were not significant. It remains unclear why such a disparity exists between the TosEF overproduction construct and *tos* deletion constructs. However, no differences in growth rates between constructs could account for the differential motility observed above (**Figure 2-9D**). Intriguingly, we found that overexpression of *tosEF* in *E. coli* K12 MG1655 (a non-UPEC strain) also inhibits motility (**Figure 2-9E**), suggesting that TosEF suppress motility through a broadly conserved mechanism. Indeed, we found that overexpression of *tosEF* in *E. coli* CFT073 cultured in tryptone broth results in reduced production of FliC (**Figure 2-9F**). Thus, this shows that TosEF, coincident with *tosA* expression, mediate motility repression by reducing flagellin expression. Furthermore, to test whether both TosEF together suppress motility, we performed the same motility assays above with *tosE* and *tosF* individually cloned (**Figure 2-10**). Intriguingly, both TosEF are required for full motility suppression.

Discussion

UTIs are common human infections. While most uncomplicated UTIs are caused by UPEC, no core set of virulence factors has been identified. Thus, a comprehensive understanding of uropathogenesis demands an understanding of broad virulence factor classes and the underlying networks connecting them. We previously identified a novel

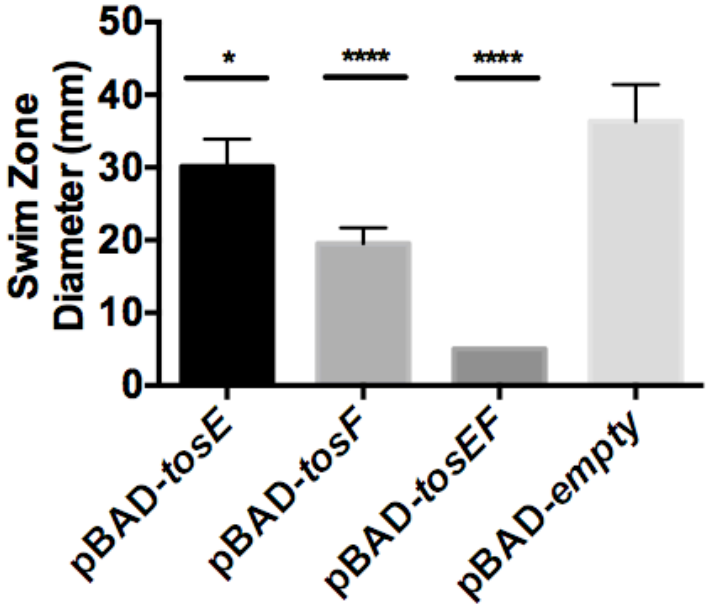

Figure 2-10. TosEF together suppress motility. *E. coli* CFT073 harboring the indicated pBAD was stabbed into soft agar and incubated 17 hours with 33.3 mM L-arabinose. Average (n=6) swimming zone diameters of the indicated constructs are shown, and error bars represent SD about the mean. Significances of the differences between swim zone diameters of the empty vector and vectors harboring *tosE* and *tosF* were determined using Tukey's multiple comparisons test following ANOVA (p<0.0001), and * and **** represent p<0.05 and p<0.0001, respectively.

E. coli adhesin, referred to as TosA, which is expressed only *in vivo* during experimental infection (208, 227). At that time, it was unclear how the *tos* operon is regulated. Here, we have shed light on *tosA* regulation and its function in the reciprocal regulation between adherence and motility. TosR negatively regulates expression of the *tos* operon, while TosEF downregulate motility when the TosA adhesin is expressed.

We previously found that TosA contains an internal repetitive region of around 1675 amino acids comprised of five repeats of 335 amino acids. Protein structure prediction revealed that these internal repeats may have a structure similar to the bacterial immunoglobulin-like domain group three (BIg 3) repeats found in another nonfimbrial adhesin of *S. enterica*, SiiE (216). These immunoglobulin folds mediate protein ligand interactions, which endow adhesive properties to SiiE (428). In addition, these immunoglobulin folds, coupled with Ca^{2+}-binding, could also promote SiiE length extension (216), bringing it into proximity with its cognate receptor on the host cell. However, whether the BIg 3 repeats found in TosA mediate adherence, extend TosA such that an element in the carboxyl-terminus can mediate adherence, or some combination of these two is still unknown.

tosR, *tosE*, and *tosF*, in addition to *tosCBDA*, are part of the *tos* locus and are well conserved among UPEC strains also encoding *tosA*. As differences in GC content often demarcate operon boundaries (424), the GC content differences between the *tos* regulatory and structural genes suggest that *tosCBDA* might be modular, where the respective regulator genes might be deleted from or inserted into the *tos* operon. However, we do not preclude the possibility that these GC content differences also reflect

an additional *tos* operon regulatory mechanism, such as differential nucleoid structuring or mRNA stability perturbations. Nonetheless, all three regulator genes are broadly conserved among UPEC strains also carrying *tosA*, and the typical operon structure is represented by *tosRCBDAEF*. This broad conservation of these regulatory genes may suggest that each serves some fitness function in different situations where UPEC may be found.

The *tos* operon is conserved among UPEC strains. The vast majority of these *tos* operons, harbored by UPEC strains, fit tightly into one of five closely related variants, based on the predicted TosR amino acid sequence encoded by the respective operon. Further support for this clonal nature of the *tos* operon comes from the fact that TosR variants are associated with specific TosE and TosF variants. We conclude that the *tos* operon present in *E. coli* CFT073 represents the archetype of the *tos* operon, as its TosR and TosEF sequence variants are the most prevalent among UPEC strains. However, the origin of the *tos* operon in its present form is still a matter of conjecture. Nevertheless, its presence on the *E. coli* CFT073 PAI-*aspV* makes acquisition by horizontal gene transfer likely. PAIs, such as PAI-*aspV*, are often acquired in such a manner.

The gene encoding TosR is located immediately upstream of *tosCBDA*. Deletion of *tosR* results in robust *tosA* expression, which leads us to conclude that TosR is a negative regulator of *tosA* expression. Phage transduction of the original Δ*tosR* mutation into a clean background resulted in high expression of *tosA*, suggesting that this phenotype is not the result of an unknown secondary mutation. In addition, a His-tagged

version of TosR complements the *tosR* deletion mutation demonstrating that TosR itself mediates this negative regulation.

Members of the PapB family regulate adhesin operons by binding in the DNA minor groove (182, 202). Minor groove binding proteins often take advantage of the inherent DNA structure of AT-rich regions to mediate target recognition and binding (277, 311, 429, 430), which is the likely case for PapB family members (182, 247, 426). Indeed, the region upstream of the *tos* operon, which contains the *tos* promoter, is an AT-rich sequence. A biologically active His-tagged TosR specifically binds an AT-rich sequence within P_{tosR}. Therefore, we propose that TosR mediates its negative regulatory effect by binding in the minor groove of an AT-rich sequence within the *tos* operon promoter.

Based on the present study, we can begin to speculate how *tos* regulation fits into other genetic networks within uropathogenic *E. coli*. While focusing on these underlying networks, such as motility and adherence reciprocal regulation, provides a more complete picture of *tosA* regulation, this broad mechanism is beyond the scope of the present study. Nevertheless, TosA is an adhesin, and the reciprocal regulation that exists between adhesins and flagellar motility (187, 188, 240, 355, 415) represents a starting point for understanding the relationship between *tosA* and the complete *E. coli* virulence network. Immediately downstream of *tosRCBDA* are the two genes encoding TosE and TosF, whose predicted amino acid sequences qualify them as members of the LuxR helix-turn-helix family of transcriptional regulators (208). Simultaneous deletion or overexpression of *tosEF* does not affect *tosA* expression. However, genes encoding factors that down-

modulate motility are often found downstream of adhesin operons (187, 188, 355). When a respective adhesin operon is induced, these motility down-modulating genes are also expressed. Indeed, consistent with this prediction, TosE and TosF overproduction results in a down-modulation of motility. Motility repression has also been observed among other LuxR family members, such as CsgD, FimZ, and RcsB (431-433).

It was previously unknown, however, whether *tosE* and *tosF* are transcriptionally coupled to *tosRCBDA* under any conditions; it was only known that their expression, like that of *tosRCBDA*, is poor *in vitro* (208). We predict that expression of *tosRCBDA* and *tosEF* are coordinated to support reciprocal regulation of adherence and motility. For example, it is possible that the loss of *tosR* results in increased expression of *tosCBDA* and *tosEF*, consistent with the *tos* transcript structure described above. Subsequent loss of *tosEF*, in the Δ*tosR* background, results in enhanced swimming motility by virtue of the loss of *tosEF*. This increased expression and possible co-transcription is also supported by the fact that a Δ*tosE*Δ*tosF* mutant alone does not replicate the phenotypes of a Δ*tosR*Δ*tosE*Δ*tosF* mutant.

The reason for the modest motility differences observed in the *tos* operon deletion mutants, compared with the *tosEF* overexpression construct remains unclear. Nevertheless, one intriguing possibility might be that *tosEF* expression in *tos* operon mutants is not as uniformly high as an overexpression vector construct may be. In this hypothesis, only certain cell sub-populations might express sufficiently high levels of *tosEF* to suppress motility, and others expressing lower levels of *tosEF* might not fully suppress motility. Thus, the composite of these two phenotypes is a swim zone of

reduced diameter, but not a severely reduced diameter. It is also possible that more modest expression of *tosEF* from the chromosome is insufficient to override the dependency for an additional, yet undefined, signal to mediate motility suppression. Therefore, the effect that TosEF has on motility in this scenario would be more modest.

We postulate that motility repression, in the case of TosEF, is an event that occurs upstream of *fliC* expression, as evidenced by the fact that TosEF overproduction results in reduced FliC synthesis. However, *fliC* is expressed as a class III gene late in the flagellar assembly gene network (119). Thus, while the *tos* operon and its expression appear to be part of the network underlying reciprocal regulation between adherence and motility, more work is required to elucidate the precise mechanism of TosEF regulation of the flagellar assembly gene network.

Acknowledgements

This work was supported in part by Public Health Service grant AI059722 from the National Institutes of Health.

CHAPTER 3

REGULATION OF THE EXPRESSION OF UROPATHOGENIC *ESCHERICHIA COLI* NONFIMBRIAL ADHESIN TOSA BY PAPB HOMOLOG TOSR, IN CONJUNCTION WITH H-NS AND LRP

Modified from: M. D. Engstrom and H. L. T. Mobley. 2016. Infection and Immunity 84: 811-21.

Abstract

Urinary tract infections are a major burden to human health. The overwhelming majority of UTIs are caused by uropathogenic *Escherichia coli* (UPEC). Unlike some pathogens, UPEC do not have a fixed core set of virulence and fitness factors, but do have a variety of adhesins and regulatory pathways. One such UPEC adhesin is the nonfimbrial adhesin TosA, which mediates adherence to the epithelium of the upper urinary tract. The *tos* operon is AT-rich, resides on pathogenicity island-*aspV*, and is not expressed under laboratory conditions. Because of this, we hypothesized that *tosA* expression is silenced by H-NS. Lrp, based on its prominent function in the regulation of other adhesins, is also hypothesized to contribute to *tos* operon regulation. Using a variety of *in vitro* techniques, we mapped both the *tos* operon promoter and TosR bindings sites. We have now identified TosR as a dual regulator of the *tos* operon, which regulates the *tos* operon in association with H-NS and Lrp. H-NS is a negative regulator

of the *tos* operon, and Lrp is a positive regulator of the *tos* operon. Leucine inhibits Lrp-mediated *tos* operon positive regulation. In addition, TosR binds to the *pap* operon, which encodes another important UPEC adhesin, P fimbria. Induction of TosR synthesis reduces production of P fimbria. At the same time, ectopic expression of *tosR* promotes synthesis of curli and/or cellulose. H-NS- and Lrp-mediated regulation also makes key contributions to reciprocal regulation of flagellar and adhesin genes. These studies advance our knowledge of regulation of adhesin expression associated with uropathogen colonization of a host.

Introduction

Urinary tract infections (UTIs), among the most common bacterial infections of humans (12), can occur in otherwise healthy individuals when bacteria colonizing the gastrointestinal tract gain access to the periurethral area. Most individuals with UTI develop an infection of the bladder, referred to as cystitis (12). However, the infecting bacterium may ascend the ureters to infect the kidneys (pyelonephritis) and, in some cases, enter the bloodstream leading to bacteremia and sometimes fatal urosepsis (12-15).

A diverse group of extraintestinal pathogenic *Escherichia coli* strains, referred to as uropathogenic *E. coli* (UPEC), cause the overwhelming majority of uncomplicated UTIs (13, 434). While numerous UPEC virulence factors have been identified, including adhesins, motility systems, toxins, and iron acquisition systems, a core set of virulence

factors has not been strictly defined (33, 35, 44). However, it is critical to understand specific virulence factors and how they are regulated.

Previous work identified and characterized the *E. coli* RTX (repeats-in-toxin) nonfimbrial adhesin TosA (for *t*ype *o*ne *s*ecretion protein A as the predicted secretion mechanism) (35, 36, 39, 208, 227, 435). In particular, *tosA* and the other *tos* operon genes have poor *in vitro* expression (208, 227, 435). TosA, a >250 kDa surface-exposed protein, mediates UPEC adherence to epithelial cells derived from the upper urinary tract (208). This is in contrast to a number of other RTX proteins, which are fully secreted into the extracellular milieu and act as toxins (49, 72, 436-438). We estimated that ~32% of UPEC strains carry genes encoding TosA and its cognate type 1 secretion system, TosCBD (435). In strain CFT073, the *tos* operon resides on PAI-*aspV* (pathogenicity island-*aspV*) (39). The *tos* operon, in addition to *tosA* and predicted cognate secretion system genes *tosCBD*, also encodes the regulatory genes *tosR*, *tosE*, and *tosF* (208, 435). TosE and TosF together suppress motility (435), a feature also found in other adhesin operon regulators (187, 188). TosR, a member of the PapB family, was previously identified as a negative regulator of the *tos* operon (435).

PapB, the prototypical member of its family, is a well-characterized positive and negative transcriptional regulator of the *pap* operon (179, 180, 384) that encodes the structural and secretion machinery necessary for P fimbria assembly (145, 439). P fimbriae are epidemiologically associated with UPEC strains (170) and have been shown to be important during experimental UTI (170, 192, 440, 441). PapB mediates transcriptional regulation by binding within the DNA minor groove (182), which suggests

that PapB might recognize structured DNA in a manner proposed for nucleoid-associated proteins (429, 442-445). In addition, the well-known nucleoid structuring protein, Lrp, also contributes to both positive and negative regulation of the *pap* operon (179, 347, 350, 354).

H-NS regulates the expression of many genes through binding structured AT-rich DNA sequences, compacting the bacterial chromosome into defined nucleoid macrodomains (41-43, 94, 309). PAIs are often identified by their AT-richness (40-43), and AT-rich genes and PAIs are often silenced by H-NS (41-43). In addition, H-NS also contributes to negative regulation of adhesin operons and dual regulation of motility operons (181, 326, 346, 354, 372, 375, 377, 378, 385-389). Indeed, PapB mediates positive regulation of the *pap* operon by anti-silencing H-NS repression (181).

Lrp and H-NS are key regulators associated with a variety of other genes, including adhesins in addition to P fimbria (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 372, 375, 377, 378, 385-389). In agreement with this, it has been predicted that Lrp and H-NS may antagonize the activity of each other, or they could act together to also potentiate gene regulation (351, 354, 444, 446). This type of regulation resembles a regulatory switch, in which one nucleoid structuring protein switches in predominance or occupancy at key regulatory elements to perturb gene regulation. This switch may be mediated by varying protein composition during different growth phases (*i.e.*, Lrp levels increase during mid-exponential phase and decrease thereafter) (94, 95, 109). However, in the case of the *pap* operon, switch regulation between H-NS and Lrp may not be mediated by direct competition for DNA binding sites (354), but instead this switch

regulation could be the result of indirect effects between H-NS and Lrp. It is unknown whether H-NS and Lrp switch regulation is responsible for *tos* operon regulation and whether H-NS and Lrp regulation of the *tos* operon is direct or indirect.

Whether TosR, like PapB, might function in the capacity of an activator in addition to repressing the *tos* operon was previously unknown. Thus, in this study, we examined the capacity of TosR to function as both a *tos* operon activator under certain conditions and as a repressor under others. Additionally, we propose that an H-NS and Lrp regulatory switch, similar to the one described above, is responsible for *tos* operon regulation. We also examined the capacity of TosR to negatively regulate P fimbria production. To our knowledge, TosA is the first nonfimbrial adhesin and RTX protein to be integrated into the reciprocal regulation network between different adhesins and between adhesins and motility systems. This cross-regulation also suggests that hierarchical regulation of adhesins and motility is much broader than previously thought.

Materials and Methods

Bacterial strains

A phage transductant of the original *tosR* deletion mutation (435) was unmarked using the FLP recombinase as previously described (435). The Δ*hns*Δ*lrp* CFT073 strain was engineered through phage-mediated transduction of a previous CFT073 Δ*lrp* mutation (355) into a previous lambda Red-engineered CFT073 Δ*hns* mutant unmarked

as described above. Transductants were selected for on lysogeny broth (LB) agar (10 g/liter tryptone, 5 g/liter yeast extract, 0.5 g/liter NaCl, 15 g/liter agar) containing kanamycin (25 μg/mL). Deletion mutations were verified by PCR.

Engineered plasmids

Untagged *tosR* and *lrp* genes were cloned into pBAD-*myc*-HisA (Invitrogen) as previously described (435). The pBAD-*tosR*-His_6 and pRS551-P_{tos}-*lacZ* were previously engineered (435). The pBAD empty vector, pBAD-*tosR*, pBAD-*tosR*-His_6, and pBAD-*lrp* constructs were maintained in LB (10 g/liter tryptone, 5 g/liter yeast extract, and 0.5 g/liter NaCl) containing 100 μg/mL ampicillin, while the pBAD-*lrp* construct was also maintained in M9 medium (12.8 g/liter $Na_2HPO_4•7H_2O$, 3 g/liter KH_2PO_4, 0.5 g/liter NaCl, 1.0 g/liter NH_4Cl, 2 mM $MgSO_4$, 0.4% glycerol, and 0.1mM $CaCl_2$) containing ampicillin (100 μg/mL) (CFT073 Δ*tosR* strain). In addition, the pRS551-P_{tos}-*lacZ* and pRS551 empty vector constructs were maintained in LB containing ampicillin (50 μg/mL), except as noted below.

5' RACE

Plasmid pRS551-P_{tos}-*lacZ* was transformed into CFT073 Δ*lacZ* and CFT073 Δ*tosR*Δ*lacZ* and maintained in LB containing ampicillin (100 μg/mL). The 5' RACE (rapid amplification of cDNA ends) procedure was performed as similar to previous methods (447). cDNA was produced using the *lacZ* cDNA primer listed in **Table 3-1** and SuperScript II reverse transcriptase as previously described (435). Input RNAs were

Table 3-1. Primers used in this study.

Primer[a]	Sequence (5'-3')
LacZ cDNA (R)	GCGGATTGACCGTAATGGGATAGGT
3' Linker	TTTAGTGAGGGTTAATAAGCGGCCGCGTCGTGA CTGGGAGCGC
Linker Forward (F)	GCCGCTTATTAACCCTCACTAAA
LacZ Nested One (R)	GACGACGACAGTATCGGCCTCAGGAAG
LacZ Nested Two (R)	CATTCAGGCTGCGCAACTGTTGGGAAGG
P_{tos13} (F)	AAGTTTTGGGGTGCAGTCCAC
P_{tos13} (R)	AAAAAGTGAAATCTCAAAACAAAAAAT
P_{tos34} (F)	TAATATAGATATTATCTGCATATAA
P_{tos34} (R)	TACTAGAGATTACATCTAAAAAATT
P_{tos57} (F)	TTAGATAAAAACCCTACAGAGAAGT
P_{tos57} (R)	CTGTATATGATCTGCCATACCATTACACAT
P_{papBA} (F)	CTCACTGTAACAAAGTTTCTTCGAATA
P_{papBA} (R)	GTTTCCCCCTTCTGTCGGGCCCCTG
lacZ (F)	GCGAATACCTGTTCCGTCATAGCG
lacZ (R)	CATCGCCAATCCACATCTGTGAAAG

[a] F, forward; R, reverse

hydrolyzed by adding NaOH (final concentration 0.16 mM) and boiled for 10 minutes. This reaction was neutralized by the addition of HCl (0.16 mM). A 3' Linker, listed in **Table 3-1**, was ligated to the above cDNA using T4 RNA ligase (New England Biolabs). After ligation, the enzyme was inactivated by incubation at 65 °C for 20 minutes. First round nested PCR was performed with the forward linker primer and *lacZ* nested primer one listed in **Table 3-1**. The second round nested PCR was performed with the forward linker primer and *lacZ* nested primer two. The resulting PCR fragment from the second round of nested PCR was sequenced.

Electrophoretic mobility shift assays (EMSAs)

tosR-His_6 was induced in wild-type CFT073 and extracted using a QIAexpressionist protocol (Qiagen) and Ni-nitrilotriacetic acid (NTA) agarose (Invitrogen) as previously described (435). Input DNAs for EMSAs were generated using P_{tos13}, P_{tos34}, P_{tos57}, P_{papBA}, and *lacZ* primers listed in **Table 3-1**. Input DNAs were terminally labeled with a Digoxigenin-11-ddUTP (DIG-ddUTP) using a 2nd generation DIG Gel Shift Kit (Roche Applied Science) as described previously (435). DNA binding reactions and detection of shifted DNA fragments were performed using a modified Roche DIG Shift Kit protocol and anti-DIG-alkaline phosphatase detection antibody (Roche Applied Science) as previously described (435), with the exception that between 400 nM and 4 μM TosR-His_6 was used in each DNA binding reaction; between 2 pg/μL (P_{papBA} and *lacZ*) and 10 pg/μL (P_{tos13}, P_{tos34}, and P_{tos57}) DIG-ddUTP labeled fragments were used in the DNA binding reactions.

Western blots

To detect TosA from induced overexpression constructs, pBAD-*tosR,* pBAD-*tosR*-His_6, and pBAD-*lrp* were transformed into a variety of CFT073 backgrounds and induced in LB containing 0, 0.06, 0.6, or 10 mM L-arabinose (pBAD-*tosR*, pBAD-*tosR*-His_6, and the pBAD empty vector) or 0, 0.6, 1.2, 2.4, or 10 mM L-arabinose (pBAD-*lrp*) for four hours. Four hours was chosen to allow *E. coli* to transit through exponential phase, ensure high titer to maximize the likelihood of observing TosA, TosR, and PapA among bacterial cells in the culture, and avoid prolonged incubation of the cultures within stationary phase. Prior to induction, overnight bacterial cultures were diluted 1:100 (CFT073 wild-type) and 1:40 (CFT Δ*lrp* and Δ*hns* strains). The pBAD-*lrp* construct transformed into CFT073 Δ*tosR* was induced for 4.5 hours with 0, 0.6, 1.2, or 10 mM L-arabinose in M9 minimal medium either containing 10 mM L-leucine or no exogenous L-leucine. Prior to induction, CFT073 Δ*tosR* harboring pBAD-*lrp* was cultured overnight in LB, pelleted at 6,000 × *g*, washed in M9 medium, and diluted 1:20. Total proteins from the inductions were collected in 10 mM HEPES (pH 8.3-8.9), quantified with a Pierce BCA Protein Assay Kit (Thermo Scientific), and assayed by Western blot with polyclonal anti-TosA antibodies or an anti-His_6 antibody (Invitrogen) as previously described (435). To detect PapA, total proteins were assayed as above in a CFT073 wild-type background harboring pBAD-*tosR*-His_6, induced in LB containing 0, 0.6, or 10 mM L-arabinose; the only exception was that polyclonal anti-PapA antibodies (Rockland) were used in place of anti-TosA antibodies.

To detect TosA from CFT073 wild-type and CFT073 Δ*hns*, Δ*lrp*, and Δ*hns*Δ*lrp*, each background construct was cultured in LB for approximately 2.5 hours to exponential phase ($A_{600} \approx$ 0.3-0.5). Prior to being cultured to exponential phase, CFT073 wild-type and Δ*lrp* overnight cultures were diluted 1:100 in LB, and Δ*hns* and Δ*hns*Δ*lrp* overnight cultures were diluted 1:40 prior to culturing in LB. Total proteins were collected in 10 mM HEPES (pH 8.3-8.9), quantified with a Pierce BCA Protein Assay Kit (Thermo Scientific), and assayed by Western blot with polyclonal anti-TosA antibodies as described above.

Promoter activity assay

Promoter activities from the pRS551-P_{tos}-*lacZ* or pRS551 empty construct, transformed into wild-type CFT073, Δ*hns*, Δ*lrp*, and Δ*hns*Δ*lrp* strains were determined using a modified Miller assay as previously described (435). The modification to the Miller assay was the use of β-methylumbelliferyl β-D-galacopyranoside (0.5 mg/mL) as substrate instead of *o*-nitrophenyl-β-galactoside.

Growth curves

Overnight cultures of *E. coli* CFT073 harboring pRS551-P_{tos}-*lacZ*, were diluted 1:100 (wild-type and Δ*lrp*) and 1:40 (Δ*hns* and Δ*hns*Δ*lrp*) into LB (10 g/liter tryptone, 5 g/liter NaCl) containing ampicillin (50 μg/ml). Constructs were cultured at 37 °C for 24 h in a Bioscreen C automated growth curve system, with A_{600} readings being recorded every 15 min.

Results

The *tos* operon promoter is located upstream of the *tos* operon regulator gene, *tosR*

Our previous work localized the *tos* operon promoter (P_{tos}) to a 630 bp sequence upstream of *tosR* (435). To determine the precise location of P_{tos} and map associated promoter elements, we conducted analysis of both RNA-Seq (not strand-specific) data (unpublished) and 5' RACE (rapid amplification of cDNA ends). From analyzing mapped normalized *tos* operon cDNA reads obtained from *E. coli* CFT073 cultured in human urine (unpublished), we hypothesized that the *tos* operon transcriptional start site is 23 bp upstream of *tosR*, based on the presence of a gap between *tosR* and the upstream ORFs *c0358* and *c0359* (**Figure 3-1A**). However, transcripts from genes encoded on the opposite DNA strand (*c0366* and *c0367*) at the 3' end of the *tos* operon make it difficult to predict transcriptional termination sites, as the RNA-Seq technique employed here is not strand-specific. For verification of the predicted transcriptional start site, 5' RACE was performed on transcripts expressed from the pRS551-P_{tos}-*lacZ* transcriptional fusion, used to ensure a high concentration of transcripts containing the *tos* operon start site expressed from plasmid-based P_{tos}. Following two rounds of nested PCR on cDNAs with a 3' linker of a known sequence ligated to this segment, we amplified a PCR product of approximately 344 bp (**Figure 3-1B**), which was consistent with the transcriptional start site obtained from the RNA-Seq analysis.

Sequencing the 5' RACE PCR product, we identified the identical distal 5' sequence (transcriptional start site) as the RNA-Seq analysis, which in turn allowed us to

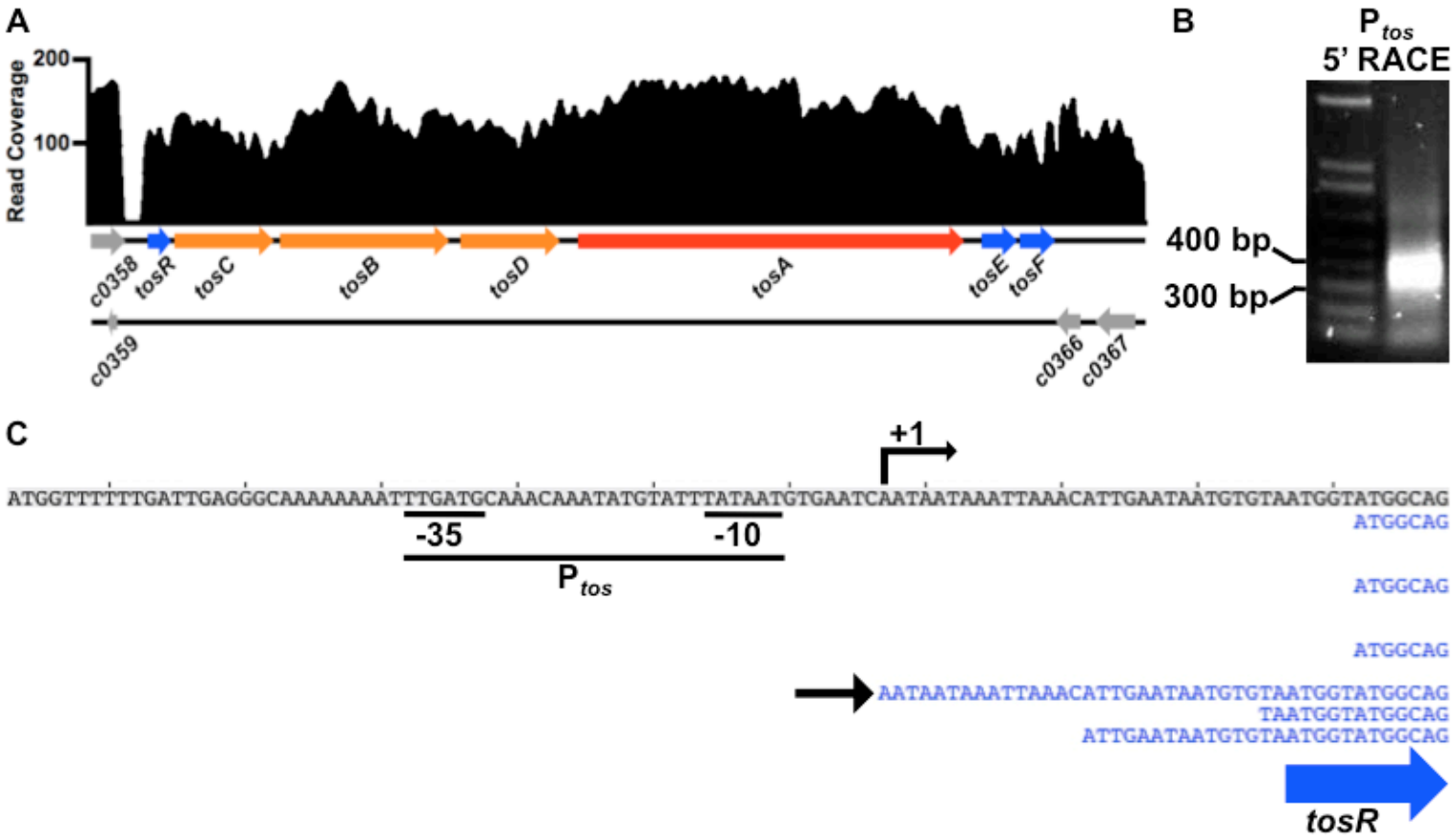
A
Read Coverage
200
100
c0358
tosR
tosC
tosB
tosD
tosA
tosE
tosF
c0359
c0366
c0367
B
P tos
5' RACE
400 bp
300 bp
C
+1
ATGGTTTTTTGATTGAGGGCAAAAAAAATTTGATGCAAACAAATATGTATTTATAATGTGAATCAATAATAAATTAAACATTGAATAATGTGTAATGGTATGGCAG
-35
-10
P tos
ATGGCAG
ATGGCAG
ATGGCAG
AATAATAAATTAAACATTGAATAATGTGTAATGGTATGGCAG
TAATGGTATGGCAG
ATTGAATAATGTGTAATGGTATGGCAG
tosR

Figure 3-1. P_{tos} is predicted to be located upstream of *tosR*. **(A)** The *tos* operon is presented along with a log-transformed cDNA read plot corresponding to cDNAs obtained from the *tos* operon of UPEC strain CFT073 cultured in filter-sterilized human urine. The scale indicates mapped reads. Two parallel lines below the read plot represent the two strands of DNA, and the directions of the arrows represent the strand on which the indicated genes are encoded. Only a partial sequence of *c0358* is depicted in the read plot. **(B)** Resolving 5' RACE products obtained from transcripts expressed from the vector pRS551-P_{tos}-*lacZ* yields a product between the indicated 300 bp and 400 bp weight markers. **(C)** Mapped cDNAs (in blue) are depicted below the top shaded DNA sequence; a blue arrow indicates the location of *tosR*. A black arrow at the left depicts the upstream most read obtained from the RNA-Seq experiment, which was also precisely the same sequence identified from sequencing the PCR product obtained in **(B)**. An angled black arrow indicates the transcriptional start site of the *tos* operon. The predicted -35 and -10 sequences of the *tos* operon promoter, P_{tos}, are depicted.

map a modified σ^{70} promoter upstream of the transcriptional start site (**Figure 3-1C**). This promoter shows 67% identity (TTGAtg) with the canonical σ^{70} -35 sequence and 100% identity with the canonical σ^{70} -10 sequence (TATAAT). Consistent with other σ^{70} promoters (448, 449), these -35 and -10 sequences are separated by 16 nucleotides. The transcriptional start site is 7 bp downstream from the end of last nucleotide -10 sequence, a spacing also consistent with σ^{70} promoters (449). In addition, the first base in the predicted transcript, adenosine, is typical of many transcriptional start sites (450, 451). However, a putative ribosome-binding site upstream of the predicted TosR translational start site could not be clearly identified, which could suggest that TosR translation is inefficient.

TosR is both a positive and negative regulator of the *tos* operon

We have previously identified a repressor function for TosR (435). As the location of the previously identified TosR-binding site (435) is not near P_{tos} (160 bp upstream of the newly identified promoter), we hypothesized that there could be additional, weaker binding sites near the *tos* operon promoter. To test this prediction, we performed an electrophoretic mobility shift assay (EMSA) on digoxigenin (DIG)-labeled DNA fragments of P_{tos} containing the strong TosR-binding site, an intergenic region between the strong TosR binding site and P_{tos}, and a region containing P_{tos} (**Figure 3-2A** and **2B**). As expected, we found that the region of P_{tos} containing the strong TosR-binding site had a reduced electrophoretic mobility (*i.e.*, was shifted) when incubated with TosR. Additionally, we found that TosR shifted the P_{tos} fragment containing the *tos*

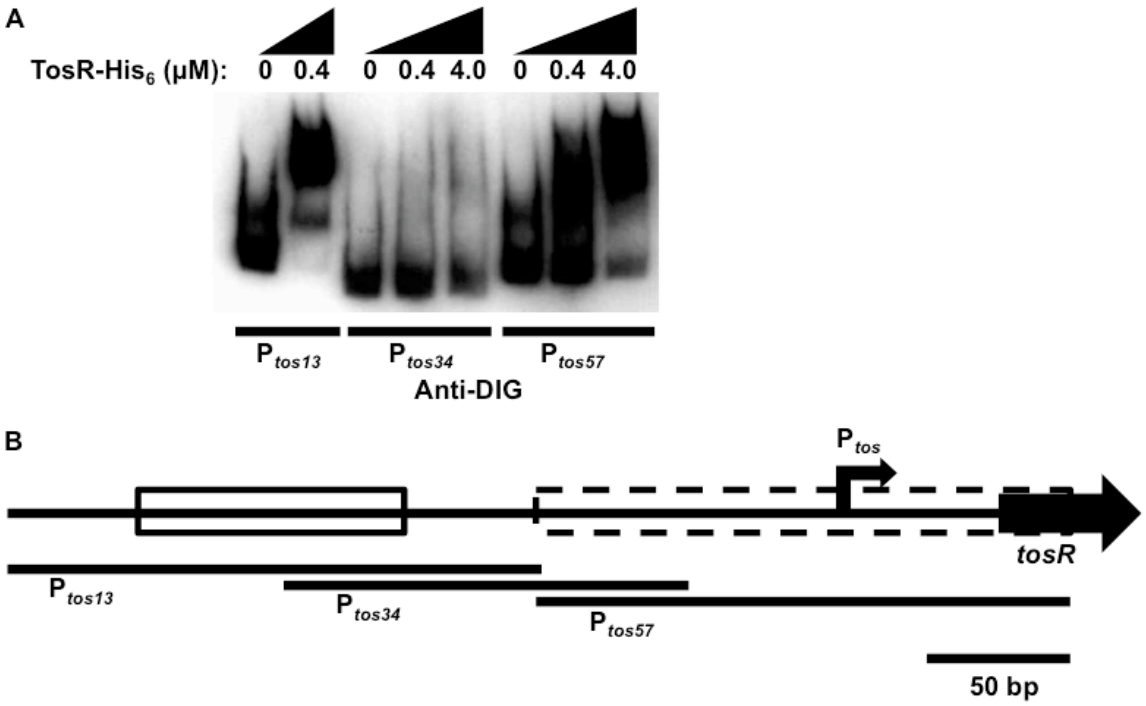

Figure 3-2. **An EMSA indicates that TosR binds P_{tos} at promoter distal and proximal positions and with varying strengths. (A)** The indicated amounts of TosR-His_6 were incubated with terminally DIG-labeled P_{tos} fragments. Shifted and unshifted DNA fragments were detected with an Anti-DIG antibody. The EMSA is representative of two independent experiments. **(B)** A schematic of P_{tos} region indicates the positions of the P_{tos} fragments used for the above EMSA, the location of the operon promoter (angled black arrow), the stronger TosR-His_6 binding site distal to the promoter (white box with solid margin), and the weaker TosR-His_6 binding site proximal to the promoter (white box with dashed margin).

operon promoter. As this latter fragment was almost fully shifted only at the highest levels of TosR (4 μM), we also reasoned that TosR weakly binds this region, compared with the strong binding site previously identified. At least 50 bp separates the strong and weak TosR binding sites in P_{tos} **(Figure 3-2B**). In addition, it is possible that TosR has some affinity for AT-rich sequences, as is also the case for other PapB family members (182). This is supported by the observation that TosR slightly shifts the intergenic region between the strong and weak binding site (*i.e.,* the intensity of the P_{tos34} unshifted fragment is weaker at the highest TosR concentration). Although, affinity for AT-rich sequences alone cannot explain all of the TosR binding activity, as AT-rich regions of P_{tos} failed to be effective competitors for TosR binding to the strong binding site in the vicinity of P_{tos} (435). However, this does not rule out the possibility that TosR recognizes a structural element, especially as another promoter, P_{papBA} (see **Appendix D**, **Figures D-1A** and **1B**), is regulated by the prototype member of the PapB family (179, 180, 384). In agreement with this, BLASTN revealed no significant sequence similarity between the weak and strong TosR binding sites. The region of DNA in the vicinity of P_{lacZ}, unrelated to both P_{tos} and P_{papBA}, does not exhibit a curved architecture (**Figure D-1C**).

Other PapB family members have been described as dual regulators of their cognate operons (179, 180, 384). Thus, based on the various degrees of TosR binding strengths for sites in the vicinity of P_{tos}, we speculated that TosR could also have an additional positive regulatory function on the *tos* operon. To test whether TosR could induce expression of the *tos* operon, we used a pBAD-*tosR*-His_6 construct and a pBAD-*tosR*

untagged construct to assay TosA synthesis at various *tosR* induction levels. Using a Western blot of proteins from whole cell preparations obtained from these pBAD overexpression constructs, we found that TosA levels are inversely related to induced *tosR* levels (**Figure 3-3A**). For induced levels of TosR below the detectable limit of our anti-His_6 antibody, we observed high levels of TosA synthesis; with high levels of TosR-His_6, detectable with anti-His_6 antibody, TosA levels were low. Likewise, these functions appear independent of the presence of the His_6 tag, as both tagged and untagged TosR proteins yielded similar results. Expression was also independent of the presence of arabinose alone, as an empty pBAD vector failed to regulate TosA synthesis. It is important to note, however, that each of these findings is based on ectopic expression of *tosR*. Therefore, it may be the case that additional regulators could supplement TosR-mediated activation and repression under native conditions. *E. coli* CFT073 also has no arabinose utilization gene mutations, and arabinose will be metabolized during these assays, which may contribute to the absence of TosR at some induction levels. Induction from the pBAD vector may also be subject to the all-or-nothing phenomenon, where cells may be binned into either uniformly high or low expression of a gene under transcriptional control of the arabinose inducible promoter (452-454). Titration from this vector, therefore, may be limited (455), especially when considering both the all-or-nothing phenomenon and arabinose utilization. Nevertheless, in terms of the newly identified TosR binding sites and concentration dependence on its regulatory functions, we predict that TosR-mediated positive regulation occurs at the strong binding site, and TosR-mediated negative regulation occurs at the weak binding site (**Figure 3-3B**).

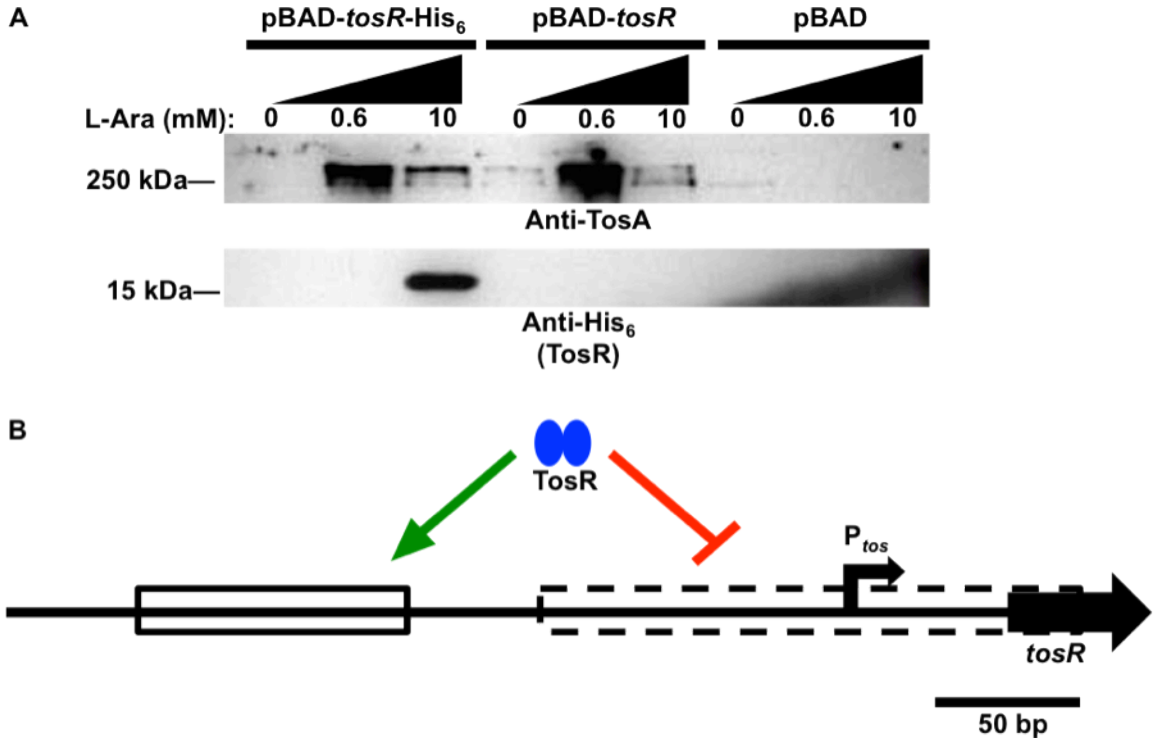

Figure 3-3. TosR is a dual positive and negative regulator of TosA. (A) Western blots using polyclonal anti-TosA antibodies or an anti-His_6 antibody were performed to detect TosA (>250 kDa) or TosR (~15 kDa). Total proteins for the Western blot were obtained from wild-type UPEC strain CFT073 harboring the indicated pBAD constructs induced with the noted concentrations of L-arabinose. Equal amounts of proteins were loaded as determined using a Pierce BCA Protein Assay Kit. The Western blot is representative of two independent experiments. **(B)** TosR (blue ovals) is predicted to mediate positive regulation at P_{tos} through binding the strong TosR-binding site (green arrow and white box with a solid margin) and negative regulation through binding the weak TosR-binding site (red bars and white box with a dashed margin)**.**

Nucleoid structuring proteins contribute to TosR regulation of the *tos* operon

The *tos* operon is localized to the PAI-*aspV* pathogenicity island in UPEC strain CFT073 (39). It is well accepted that genes on PAIs, and other AT-rich sequences, are often bound and regulated by nucleoid structuring proteins including H-NS and Lrp (41-43, 296, 359). A 400 bp region containing P_{tos} is AT-rich (74% AT). In addition, both of these nucleoid-structuring proteins regulate the expression of many genes, including adhesin and flagellar genes (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 370, 372, 375, 377, 378, 385-389). To determine whether or not a 240 bp AT-rich region near P_{tos} is similarly curved to an analogous region in P_{papBA}, suggesting Lrp and H-NS-mediated nucleoid structuring, we utilized a web-based tool (http://www.lfd.uci.edu/~gohlke/dnacurve/) to predict DNA curvature. We found that both regions have a similar predicted curved geometry (**Figure D-1A** and **1B**), which suggests Lrp and H-NS could regulate the *tos* operon. To further predict whether H-NS and Lrp bind to P_{tos}, we examined this sequence for putative H-NS and Lrp binding sites (see **Appendix E**, **Figure E-1**). There are four clusters of putative Lrp binding sites ($GN_{2-3}TTT$), based on P_{papBA} (358), downstream and partially overlapping the predicted strong TosR binding site and upstream and partially overlapping the predicted weak TosR binding site. One of the predicted Lrp binding sites also overlaps P_{tos}. In addition, there are also two putative high affinity H-NS binding sites with 80% (aCaATAAATT) and 70% (ataATAAATT) identity to a sequence with known high affinity for H-NS (258, 456) located upstream of the weak TosR binding site and downstream of P_{tos}, near the predicted transcriptional start site. Intriguingly, running BLASTN on this same P_{tos}

sequence reveals a putative CitB binding site 22 bp upstream of P_{tos} (**Figure E-1**) (457). However, the function of this sequence in relation to *tos* operon regulation remains unclear.

To determine whether H-NS and Lrp do indeed regulate the *tos* operon, we performed Western blots on proteins from whole cell preparations obtained from CFT073 Δ*hns* and Δ*lrp* backgrounds (**Figure 3-4A**). TosA levels were dramatically increased in the Δ*hns* background compared to wild-type CFT073, suggesting that H-NS could function as a negative regulator of the *tos* operon. However, it is important to note that H-NS perturbs the expression of a number of different genes (42, 179, 326, 346, 372, 375, 377, 378, 385-389); therefore, it remains unclear if additional regulators supplement H-NS-mediated negative regulation of the *tos* operon. The loss of Lrp failed to increase *tos* operon expression, as was observed for loss of H-NS.

With respect to the multitude of PapB family members, however, it is not always obvious how nucleoid structure and the cognate PapB family members integrate to govern expression of the adhesin. To determine whether H-NS and Lrp contribute to TosR regulation of the *tos* operon, we performed the same pBAD-*tosR*-His_6 overexpression experiment (with reduced ampicillin concentration) described above in the CFT073 Δ*hns* (**Figure 3-4B**) and Δ*lrp* (**Figure 3-4C**) backgrounds. As above, loss of *hns* resulted in increased TosA synthesis, but high levels of TosR did not decrease TosA levels in the CFT073 Δ*hns* background. Conversely, we found that the TosR-mediated positive regulation was dependent on Lrp; no change in TosA levels could be detected regardless of TosR level in the CFT073 Δ*lrp* background. Likewise, a shift to a lower

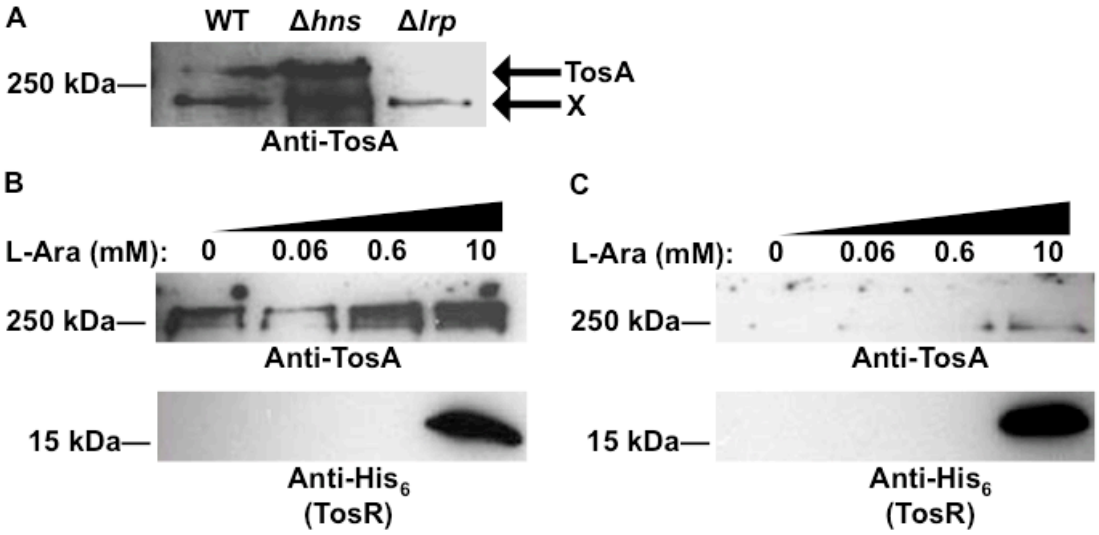

Figure 3-4. TosR-mediated negative and positive regulation is perturbed in UPEC CFT073 Δ*hns* and Δ*lrp* backgrounds. (A) A Western blot using polyclonal anti-TosA antibodies was performed on total proteins obtained from the indicated UPEC CFT073 backgrounds. Bands corresponding to TosA are indicated in the figure, and a cross-reacting band is indicated with an X. The Western blot is representative of two independent experiments. Equal amounts of proteins were loaded as determined using a Pierce BCA Protein Assay Kit. Western blots were also performed as above using polyclonal anti-TosA antibodies or an anti-His_6 antibody in the UPEC CFT073 Δ*hns* **(B)** or Δ*lrp* **(C)** backgrounds harboring pBAD-*tosR*-His_6 induced with the indicated concentrations of L-arabinose. The Western blot is representative of two independent experiments.

antibiotic concentration for the Δ*hns* and Δ*lrp* backgrounds harboring pBAD-*tosR*-His_6 does not perturb TosR regulation itself; the lower antibiotic concentration did not perturb TosR-mediated regulation in the wild-type *E. coli* CFT073 background (data not shown). As for H-NS, Lrp is also a global regulator (105, 152, 179, 183, 250, 296, 347-355). Therefore, it remains unclear if additional gene products also supplement Lrp and TosR-mediated positive regulation of the *tos* operon.

Induction of *lrp* expression is sufficient to drive TosA synthesis

Observing that Lrp is required for *tos* operon expression, we tested whether exogenous expression of Lrp alone would be sufficient to induce *tos* operon expression. To determine whether Lrp acts as a positive regulator of *tos* operon, as is the case with the *pap* operon (179, 347, 354), we performed a pBAD-*lrp* overexpression experiment in wild-type CFT073. Western blot of whole cell proteins from this overexpression construct revealed that low levels of *lrp* induction increased TosA levels (**Figure 3-5A**). In turn, high levels of *lrp* induction diminished TosA levels. However, this effect was dependent on the presence of TosR (**Figure 3-5B**). Likewise, in M9 minimal medium Lrp overexpression mediates less production of TosA in the wild-type CFT073 (**Figure 3-6**), compared with the same assay performed with the *tosR* mutant (**Figure 3-7**). Taken together, Lrp appears to be a positive regulator of the *tos* operon, but can also contribute to *tos* operon negative regulation in the presence of TosR. It is important to note, however, that the same caveats of *lrp* overexpression should also be considered as with *tosR* overexpression above.

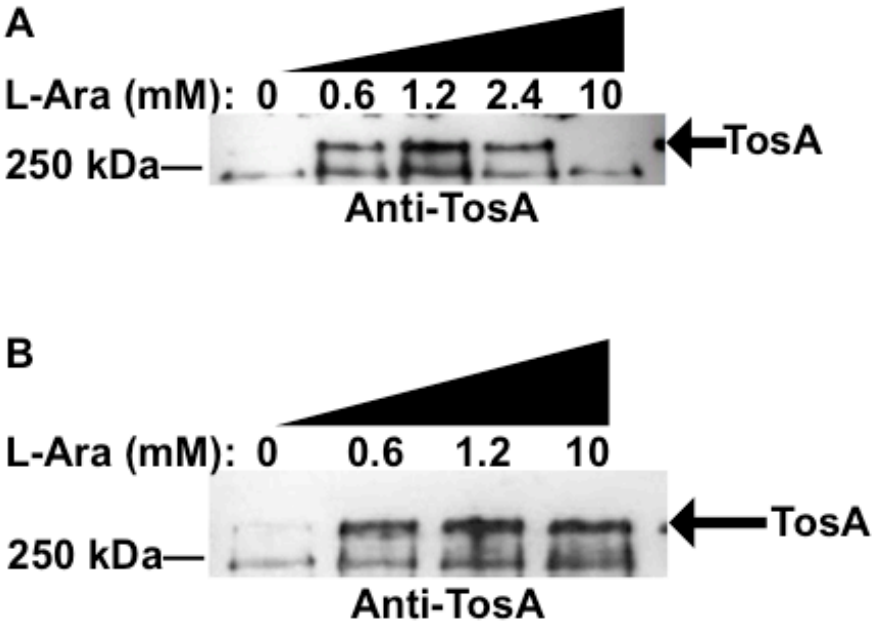

Figure 3-5. Lrp is a positive regulator of TosA. (A) A Western blot using polyclonal anti-TosA antibodies was performed on total proteins obtained from UPEC CFT073 harboring pBAD-*lrp* and induced with the indicated concentrations of L-arabinose. Bands corresponding to TosA are indicated in the figure. Equal amounts of proteins were loaded as determined using a Pierce BCA Protein Assay Kit. The Western blot is representative of two independent experiments. **(B)** A Western blot was performed as above using total proteins obtained from a Δ*tosR* UPEC CFT073 strain harboring pBAD-*lrp* induced with the indicated concentrations of L-arabinose. TosA was detected as above.

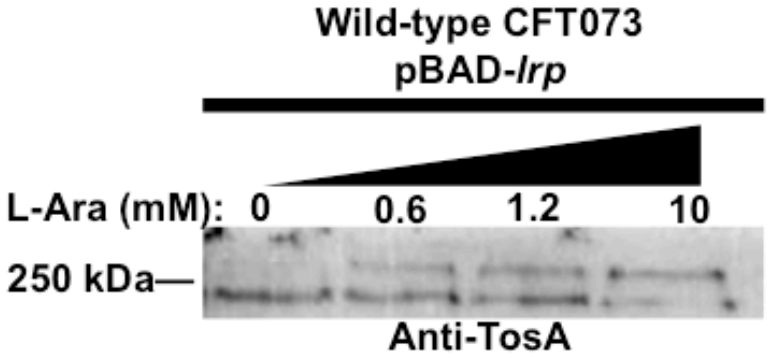

Figure 3-6. *lrp* overexpression does not support high TosA production in wild-type UPEC CFT073 cultured in M9 minimal medium. A Western blot using polyclonal anti-TosA antibodies was performed on total proteins obtained from UPEC CFT073 harboring pBAD-*lrp* and induced with the indicated concentrations of L-arabinose in M9 minimal medium. Equal amounts of proteins were loaded as determined using a Pierce BCA Protein Assay Kit.

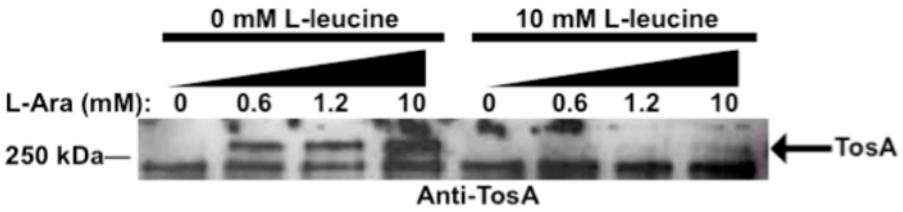

Figure 3-7. Exogenous L-leucine negatively regulates Lrp-mediated *tos* operon positive regulation in M9 minimal medium. A Western blot was performed using total proteins obtained from a UPEC CFT073 Δ*tosR* strain harboring pBAD-*lrp* induced with the indicated concentrations of L-arabinose in M9 minimal medium with and without 10 mM L-leucine. Bands corresponding to TosA, detected with polyclonal anti-TosA antibodies, are indicated in the figure.

Exogenous leucine inhibits *tos* operon regulation

Some genes regulated by Lrp are positively or negatively regulated by exogenous leucine (152, 296, 348, 352). To test whether exogenous leucine positively or negatively regulates the *tos* operon, we performed our pBAD-*lrp* overexpression assay using the CFT073 Δ*tosR* background in M9 minimal medium with and without exogenous leucine (10 mM). In the CFT073 Δ*tosR* background, induction of *lrp* expression resulted in higher TosA levels only in the absence of exogenous leucine (**Figure 3-7**). This demonstrates that Lrp-mediated positive regulation of the *tos* operon is subject to regulation by leucine, especially in M9 minimal medium. Furthermore, overexpression of pBAD-*tosR*-His_6 in LB with and without additional leucine reveals that TosR-mediated positive regulation is sensitive to exogenous leucine levels (see **Appendix F**, **Figure F-1**). To verify that low leucine, compared with LB, is an environmental condition encountered by various UPEC strains *in vivo* (32), we analyzed the differential expression of genes responsive to exogenous leucine (**Table F-1** and **Table F-2**). We found that UPEC may indeed respond to lower leucine levels *in vivo*, as many genes downregulated or inhibited by Lrp-leucine were upregulated in the human urinary tract. However, this upregulation is not uniform among all strains in the host urinary tract. It may be that this differential expression occurs at different times throughout infection or that stochastic regulation by nucleoid proteins (371, 458-462) reduces the ability of RNA-Seq to make consistent descriptions of gene regulation between UPEC strains. Nevertheless, the fact that at least some of these genes are upregulated by UPEC strains

during a urinary tract infection suggests that there is at least some resolution to describe regulation by nucleoid regulators in response to leucine.

An H-NS and Lrp regulatory switch drives *tos* operon transcriptional regulation

It has been previously proposed and noted that Lrp might act to anti-silence H-NS repression (351, 354, 444, 446). We hypothesize, therefore, that an H-NS and Lrp regulation switch (*i.e.*, predominance of either H-NS or Lrp) explains the observed regulation of the *tos* operon. In particular, if H-NS-mediated negative regulation is abolished, Lrp is no longer required for *tos* operon expression. To test this hypothesis, we performed a Western blot on a CFT073 Δ*hns*Δ*lrp* mutant (**Figure 3-8A**). TosA levels remain high in the Δ*hns*Δ*lrp*, which strengthens the premise that Lrp functions to overcome H-NS negative regulation of the *tos* operon. Coupled with the finding that TosR regulation is abolished in the Δ*lrp* background, these results suggest that an H-NS and Lrp regulation switch likely contributes to *tos* operon regulation, and TosR has a function within this regulatory switch. It is important to note, however, that both H-NS and Lrp are global regulators (152, 179, 183, 250, 296, 347-355). Therefore, it cannot be ruled out that regulation between H-NS and Lrp, at P_{tos}, is indirect.

We next tested whether the proposed H-NS and Lrp regulatory switch would function at the transcriptional level at P_{tos}. To determine whether P_{tos} is transcriptionally responsive to H-NS and Lrp, we measured the activity of our pRS551-P_{tos}-*lacZ* transcriptional fusion in both wild-type CFT073 and CFT073 nucleoid-structuring mutants using a Miller assay (**Figure 3-8B**). As previously observed (435), the P_{tos}

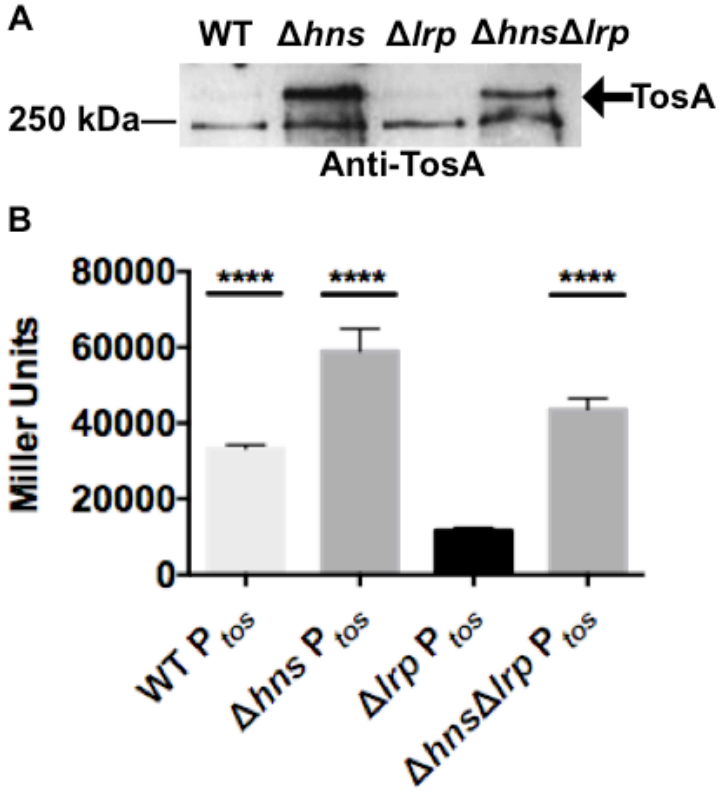

Figure 3-8. Lrp is not required for P_{tos} transcriptional activation in the Δ*hns* background. (A) A Western blot using polyclonal anti-TosA antibodies was performed on total proteins obtained from the indicated CFT073 backgrounds. Equal amounts of proteins were loaded as determined using a Pierce BCA Protein Assay Kit. **(B)** A Miller assay was performed using β-galactosidase translated from the *lacZ* gene of the pRS551-P_{tos}-*lacZ* vector harbored in the indicated backgrounds. Bars represent mean values of Miller units obtained from two biological replicates, with two technical replicates each. Error bars represent one standard deviation around the mean, and **** represents *p*-values <0.0001 obtained by comparing *lacZ* expression from the pRS551-P_{tos}-*lacZ* construct harbored in the respective mutant or wild-type UPEC CFT073 background with the Δ*lrp* background (determined using ANOVA followed by Tukey's multiple comparisons test).

promoter showed high activity in wild-type CFT073. Additionally, P_{tos} showed high activity in the CFT073 Δ*hns* mutant. P_{tos} promoter activity was greatly reduced in the CFT073 Δ*lrp* background. In the CFT073 Δ*hns*Δ*lrp* background, however, P_{tos} activity was restored to slightly higher than wild-type levels. No growth differences were observed among bacterial strains harboring pRS551 (**Figure 3-9**). These findings suggest that native Lrp levels induce P_{tos} on the pRS551 construct by overcoming H-NS-mediated negative regulation. Thus, regulation of the *tos* operon by the H-NS and Lrp regulatory switch is at the transcriptional level. From its P_{tos} DNA binding activities and association with H-NS and Lrp regulation of the *tos* operon, we further suggest that TosR also transcriptionally regulates the *tos* operon. Intriguingly, a number of genes downregulated by H-NS were generally upregulated in the human urinary tract (**Table F-3**). However, like the leucine responsive genes above, the same caveats are true for genes negatively regulated by H-NS. Nevertheless, our findings suggest that the H-NS portion of the regulatory switch above is less prominent *in vivo*, consistent with the observed increase in *tos* operon expression (208, 227, 435).

TosR contributes to *pap* operon regulation and curli and/or cellulose production

Both H-NS and Lrp are global regulators that affect the expression of a variety of genes, including adhesin and flagellar genes (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 370, 372, 375, 377, 378, 385-389). It is, therefore, not surprising that fimbrial regulators associated with H-NS and Lrp could also participate in cross-regulation between adherence and motility genes (203, 204, 248-250). Both PapB (P fimbrial

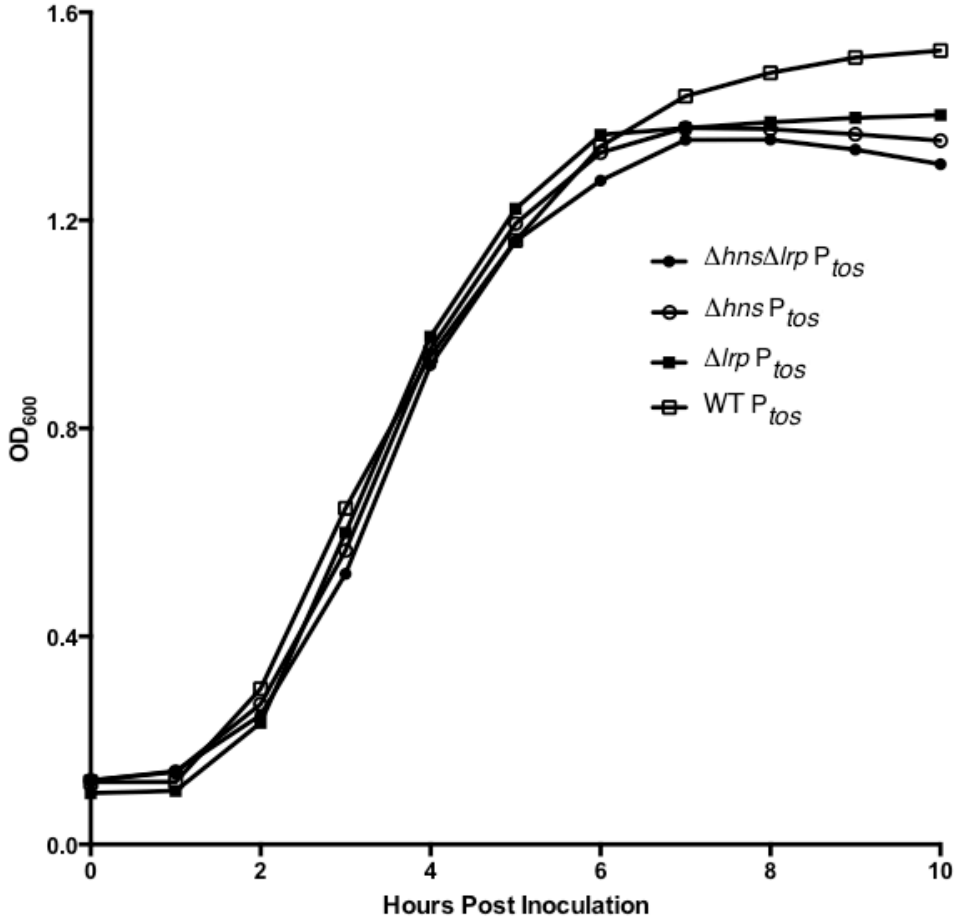

Figure 3-9. All UPEC CFT073 strains harboring pRS551-P$_{tos}$-*lacZ* have an equivalent growth phenotype. A plot of a growth curve experiment conducted over ten hours, with the indicated mutant and wild-type UPEC CFT073 strains harboring pRS551-P$_{tos}$-*lacZ*, is depicted.

operon) and FocB (F1C fimbrial operon) share approximately 80% amino acid sequence identity and regulate their respective *pap* and *foc* operons. Cross-regulation between PapB and FocB is a well-characterized phenomenon (203). In contrast, TosR has only 28% amino acid sequence identity with PapB. To determine whether TosR can also regulate the *pap* operon, we performed our pBAD-*tosR*-His_6 overexpression assay and Western blot with an anti-PapA antibody (**Figure 3-10A**). With increased TosR levels, PapA2 levels were decreased.

Previous work characterizing PapB and FocB cross-regulation explored the ability of both proteins to mediate this regulation through binding to P_{papBA} and P_{focBA} (203). To determine whether TosR might also mediate *pap* operon cross-regulation through binding P_{papBA}, we performed an EMSA on a DIG-labeled P_{papBA} fragment and, as a control, a fragment of *lacZ* (**Figure 3-10B**). TosR shifted the P_{papBA} fragment but failed to shift the control *lacZ* fragment. Thus, we conclude that despite markedly low amino acid identity between PapB and TosR, TosR mediates negative regulation of the *pap* operon through specific binding of P_{papBA}. Intriguingly, like the weak P_{tos} binding site, BLASTN reveals no substantial sequence homology with the strong P_{tos} binding site and P_{papBA}.

PapB family members have not previously been ascribed a regulatory function in the production of curli fibers and/or cellulose. Nevertheless, to determine whether ectopic *tosR* expression could contribute to the regulation of other adhesins and the switch to multicellularity, we performed a pBAD-*tosR*-His_6 expression assay on YESTA (yeast extract and tryptone agar) plates containing ampicillin (100 μg/mL) and 10 mM L-arabinose. Spots of *E. coli* strain CFT073 harboring pBAD-*tosR*-His_6, pBAD-*papB*,

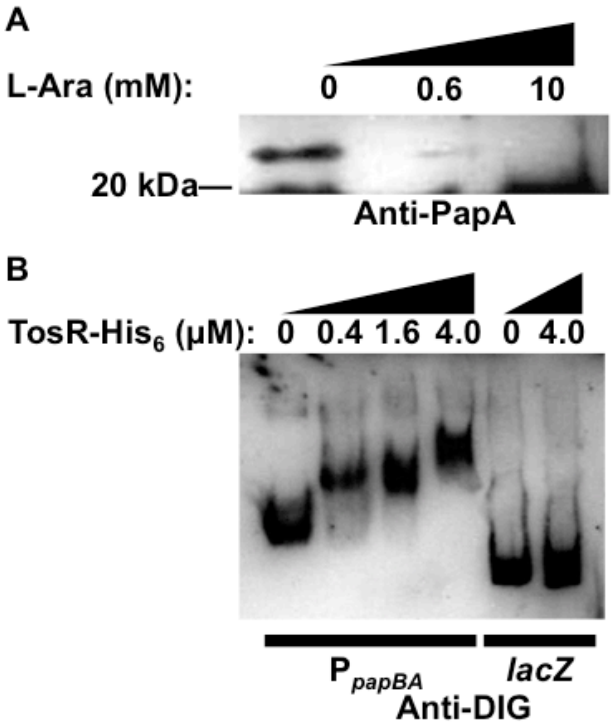

Figure 3-10. TosR negatively regulates P fimbriae synthesis. (A) A Western blot using polyclonal anti-PapA antibodies, to detect PapA2 (~23 kDa), was performed on total proteins obtained from UPEC CFT073 harboring pBAD-*tosR*-His_6 and induced with the indicated concentrations of L-arabinose. This blot is representative of two biological replicates. Equal amounts of proteins were loaded as determined using a Pierce BCA Protein Assay Kit. **(B)** The indicated amounts of TosR-His_6 were treated along with terminally DIG-labeled P_{papBA} or *lacZ* fragments. Shifted and unshifted DNA fragments were detected using an Anti-DIG antibody. The EMSA is representative of two independent experiments.

pBAD-*focB*, or the empty pBAD plasmid were incubated on the above YESTA plates for 48 hours, and after this incubation period only *E.* coli CFT073 harboring the pBAD-*tosR*-His_6 construct exhibited the rdar (red, dry, and rough) phenotype indicative of the production of curli and/or cellulose (230-233) (see **Appendix G**, **Figure G-1A**). Thus, *tos* operon expression may be coordinated with the expression of biofilm components, and TosR may exert regulatory control over these biofilm components in addition to the *tos* and *pap* operons. Intriguingly, ectopic expression of FocB may weakly promote *E. coli* CFT073 binding the Congo red dye (**Figure G-1B**), which suggests that other PapB family members and adhesins could be weakly coordinated with curli and/or cellulose production.

Lrp and H-NS serve key functions in motility and adherence reciprocal regulation

Lrp is a putative negative regulator of motility in a Type 1 fimbria locked ON *E. coli* CFT073 strain (355). To determine whether an *lrp* mutation could suppress the low production of FliC in an *hns* mutant (375, 376), we performed a Western blot using polyclonal anti-FliC antibodies on total proteins obtain from the Δ*hns*Δ*lrp* CFT073 strain (see **Appendix H**, **Figure H-1A**). We found that an *lrp* mutation is insufficient to restore FliC levels to near wild-type levels, which suggests that nucleoid structure, particularly mediated by H-NS, is vital for reciprocal regulation to occur. To determine whether reciprocal regulation between adhesins occurs in the same Δ*hns*Δ*lrp* CFT073 strain, we again performed a pBAD-*tosR*-His_6 overexpression assay and Western blot using polyclonal anti-PapA antibodies on total proteins obtained from this strain (**Figure H-1B**).

P_{papBA} is expressed in another Δ*hns*Δ*lrp E. coli* strain not harboring the *tos* operon (354). However, we observed no PapA synthesis in the UPEC CFT073 Δ*hns*Δ*lrp* strain, a stain where high TosA synthesis was observed (**Figure 3-8A**). Consistent with the results above, nucleoid structure is a key component of adhesin reciprocal regulation (*i.e.*, to prevent a few adhesin regulators, such as TosR, from simply dominating over the other adhesin operons, such as the *pap* operon).

Discussion

Here, we present a model of *tos* operon regulation involving PapB family member, TosR, and two global gene regulators, H-NS and Lrp (**Figure 3-11**). TosR is a positive and negative transcriptional regulator of the *tos* operon. We predict that TosR mediates this dual regulation through differential binding to the region of DNA containing P_{tos}, the *tos* operon promoter, by TosR. The global regulator, H-NS, transcriptionally silences expression of the *tos* operon, while another global regulator, Lrp, overcomes H-NS silencing to mediate positive regulation of the *tos* operon. When TosR levels are low, TosR promotes Lrp-mediated positive regulation of the *tos* operon. However, when TosR levels are high, TosR promotes H-NS-mediated negative regulation of the *tos* operon. We also predict that H-NS and Lrp interact either directly or indirectly to modify *tos* operon positive regulation. Additionally, TosR negatively regulates expression of the P fimbrial (*pap*) operon. Using RNA-Seq and 5' RACE, we identified the transcriptional

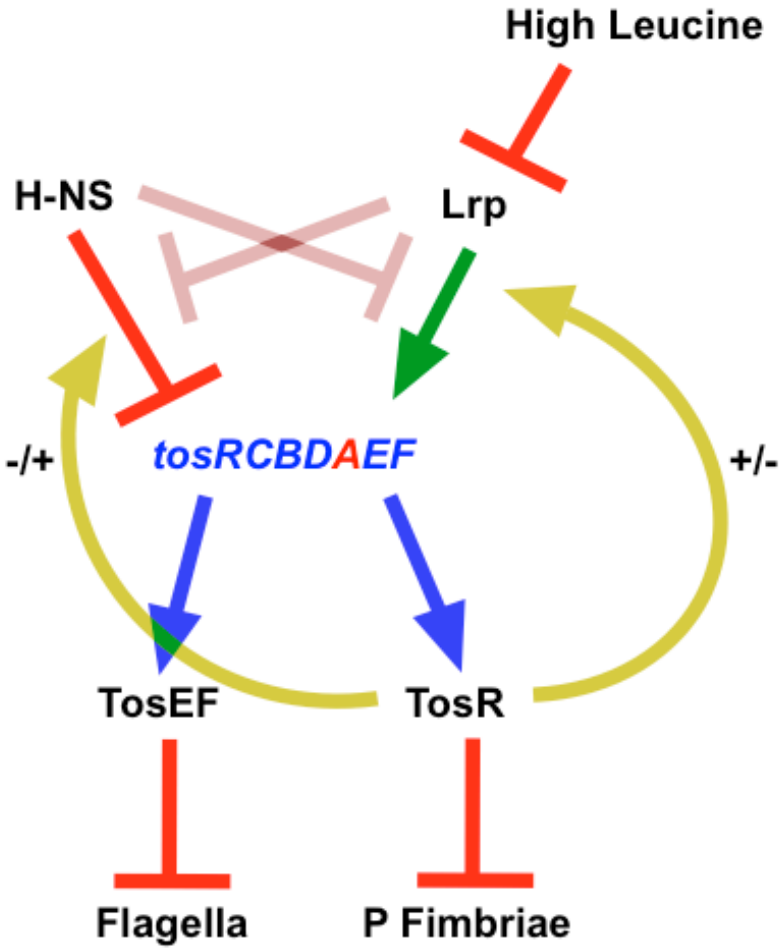

Figure 3-11. Model of *tos* operon regulation and its involvement in reciprocal regulation of adhesins and flagella. The *tos* operon is indicated with blue text, and *tosA* is represented with red text. A blue arrow indicates that TosR, TosE, and TosF are translated from genes transcriptionally linked to the *tos* operon. Under typical laboratory conditions H-NS silences expression of the *tos* operon (red bar). Under low and high concentrations, respectively, TosR positively and negatively (+/-) regulates (yellow arrow) Lrp-mediated positive regulation of the *tos* operon (green arrow). The inverse is true for H-NS-mediated regulation (yellow arrow). Low levels of TosR inhibit H-NS-mediated negative regulation and high levels promote H-NS-mediated negative of the *tos* operon (-/+). Lrp may also indirectly relieve H-NS-negative regulation of the *tos* operon, and H-NS may also indirectly inhibit Lrp-mediated positive regulation of the *tos* operon (pink bars). In turn, TosR also negatively regulates P-fimbriae synthesis (red bar), and together cognate regulators, TosE and TosF, negatively regulate flagella synthesis (red

bar). High levels of leucine, in addition, negatively regulate Lrp-mediated positive regulation of the *tos* operon (red bar).

start site of the *tos* operon 23 bp upstream of *tosR* and P_{tos} 30 bp upstream of *tosR*. The promoter sequence has only a few modifications from the canonical σ^{70}, which include two base substitutions from the canonical -35 sequence and spacing between the -35 and -10 sequences one base pair shorter than for the average σ^{70} promoter (448, 449). Additionally, the first base of the transcript, adenosine, is also typical of other σ^{70} promoters (450, 451), and the spacing between the promoter and the start of the transcript (7 bp) is also observed with other σ^{70} promoters (449). Thus, together these results suggest that P_{tos} could be a strong promoter, which is consistent with our previous observation of strong activity from the pRS551-P_{tos}-*lacZ* transcriptional fusion (435). This finding, however, is confounded by the weak expression of the *tos* operon previously observed (208, 227, 435), which points to negative transcriptional regulation at P_{tos} by other proteins as a possible mechanism of *tos* operon regulation.

Our group previously reported that the *tos* operon is repressed when cultured under laboratory conditions (LB broth, both aerated and static, at 37°C) (208, 227, 435). Some of this negative regulation is attributed to TosR (435). However, other PapB family members act as dual regulators (both activator and repressor) of their cognate operons (179, 180, 384). We found this was also the case for TosR, which shows a reciprocal relationship with TosA levels: if TosR levels are low, TosA levels are high; when TosR levels are high, TosA levels are significantly reduced. Thus, TosR is a dual regulator of the *tos* operon. We predict that at least some of this differential behavior is mediated through TosR binding to two sites within P_{tos}, one site when TosR levels are low (strong binding site) and the other when TosR levels are higher (weak binding site).

Regulation of the *tos* operon, involves not only TosR, but includes both H-NS and Lrp. We predict that TosR-positive regulation of the *tos* operon, when TosR levels are low, may be mediated through an alleviation of negative regulation by H-NS, thereby promoting Lrp-mediated positive regulation. In terms of the predicted H-NS binding site upstream of P_{tos} (**Figure E-1**), we speculate that TosR binding either displaces an H-NS filament or prevents further H-NS polymerization at this site, both known mechanisms of overcoming H-NS silencing (315). Subsequently, this activity may allow Lrp to bind to predicted binding sites in the vicinity of P_{tos} to promote positive regulation. This model is further supported by TosR-mediated positive regulation of the *tos* operon no longer being required in the Δ*hns* background. Furthermore, Lrp-mediated positive regulation is no longer necessary in the same Δ*hns* background, and TosR positive regulation is abolished in the Δ*lrp* background. Lrp alone is also sufficient to promote expression of the *tos* operon, especially in the absence of leucine, which further supports our prediction of Lrp-mediated positive regulation of the *tos* operon. Thus, in terms of the predicted Lrp binding sites (**Figure E-1**) in the vicinity of P_{tos} and previous work on Lrp by others (364-366), we speculate that Lrp binding may facilitate RNA polymerase contact with either Lrp itself or additional unknown elements near P_{tos}. Similarly, we propose that TosR-negative regulation of the *tos* operon, when TosR levels are high, may be mediated through interference of Lrp-promoted positive regulation. We also predict that TosR-mediated negative regulation is dependent on H-NS. We speculate, in terms of the second predicted TosR and H-NS binding sites (**Figure E-1**) in vicinity of P_{tos}, that TosR binding to this site occludes Lrp binding and subsequent occlusion of RNA polymerase

from P_{tos}. Further support for this conclusion comes from the finding that TosR is no longer a negative regulator of the *tos* operon in a Δ*hns* background. Also integrating the two predicted H-NS binding sites into this regulation model, it is possible that bridging at these sites could promote negative regulation (306, 314, 315, 463, 464).

From the strong activity of P_{tos} in a Δ*hns*Δ*lrp* background, we also predict that an H-NS and Lrp regulation switch is responsible for much of the *tos* operon regulation. The switch may act through alteration of the predominance of H-NS and Lrp regulation at P_{tos}, which is consistent with modulation of nucleoid levels of each during different growth phases (94, 95, 109). However, as H-NS and Lrp are pleiotropic regulators (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 370, 372, 375, 377, 378, 385-389), this switch may also be through indirect interactions. To emphasize the possibility that H-NS and Lrp indirectly interact, pink bars are depicted in our model (**Figure 3-11**). In agreement with this idea of a switch between the two nucleoid proteins, the strong decrease in P_{tos} activity observed with the loss of Lrp, is increased when H-NS is also absent. This leads us to predict that Lrp functions to overcome H-NS negative regulation of the *tos* operon, consistent with our belief that an H-NS and Lrp switch governs *tos* operon expression. It is also an intriguing possibility that this same switch is similar to a previous description of nucleoid contributions to reciprocal regulation of adherence and motility (388). However, whether this H-NS and Lrp regulation switch is mediated by direct antagonism of each component or indirect effects will need to be examined further.

We note that the estimated leucine content of pooled human urine (~0.01 mM) (401) is much lower than that of LB (~8 mM) (465). This suggests that UPEC would

adjust gene expression to accommodate the low leucine levels found in human urine. Thus, growth in an environment with relatively low levels of leucine is an environmental stress encountered by UPEC during an infection. As exogenous Lrp, specifically in M9 minimal medium, does not positively regulate the *tos* operon in the presence of high leucine levels, we also conclude that low leucine is an environmental cue that upregulates the *tos* operon. We also propose that the presence of higher levels of leucine in LB at least partially accounts for poor *tos* operon expression when cultured in this medium.

As evident from the variety of genes, including those localized to adhesin operons in addition to flagellum-mediated motility genes regulated by H-NS and Lrp (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 370, 372, 375, 377, 378, 385-389), cross-regulation is a possible feature of this regulatory switch (203, 204, 248-250). We found that TosR is a negative regulator of P fimbria production. We predict that this negative regulation is potentiated through TosR binding to P_{papBA}, the *pap* operon promoter. This is a surprising finding in that the previously well described cross-regulation between PapB family members occurred between PapB and FocB, which shared 80% amino acid sequence identity (203). TosR and PapB share only 28% amino acid sequence identity. We hypothesize that these results have important implications for studying adhesin expression. Further work should explore whether TosR, like PapB and FocB, also regulates FimA and FocA levels (203) and whether PapB and FocB contribute to *tos* operon regulation. Nevertheless, it is now our conclusion that such cross-regulation between PapB family members and different types of adhesins (*i.e.*, fimbrial and nonfimbrial adhesins) is a broader phenomenon than previously thought. Thus, a more

detailed exploration of adhesin cross-regulation, especially between unrelated or poorly related adhesins and adhesin regulators, should be undertaken to gain a more accurate picture of microbial adhesin regulation. These future explorations should include determining whether reciprocal regulation between adhesins is an important fitness trait during infection. For example, P fimbria and TosA both make contributions during experimental UTI (170, 192, 440, 441), but it is unknown whether TosR inactivation could suppress a *tosA* mutation through allowing UPEC to continue to synthesize P fimbria instead of simultaneously inhibiting P fimbria production and attempting to produce a nonexistent TosA adhesin.

Finally, previous work (435) has already established that TosEF, expressed when the *tos* operon is expressed, negatively regulate FliC levels. Together with the TosR findings above, we have also found that *tos* operon regulation participates in reciprocal regulation of adherence and motility. It is intriguing to note that a protein encoded by the terminal gene of the *pap* operon, *papX*, suppresses motility in UPEC strain CFT073 (187, 188). Future work may, thus, also explore whether overexpressing *lrp* in *tosEF* and *papX* mutant constructs decreases motility. Taken together, these results could delineate the function of the H-NS and Lrp regulatory switch in reciprocal regulation and during infection.

Acknowledgments

We wish to thank Christopher Alteri, Sargurunathan Subashchandrabose, David Friedman, Sébastien Crépin, and Chelsie Armbruster for insightful discussions about our work. This work was funded by National Institutes of Health Public Health Service grant AI059722 (to HLTM).

CHAPTER 4

CONCLUSIONS AND PERSPECTIVES

Summary of results

In these studies, I have presented the finding that TosR, TosE, and TosF are conserved among UPEC (uropathogenic *Escherichia coli*) strains. Furthermore, I have elucidated the transcriptional organization of the *tos* operon. I have also described the function of TosR as a *tos* operon dual regulator, determined that H-NS (heat-stable nucleoid structuring protein) and Lrp (leucine-responsive regulatory protein) serve important functions during regulation of the *tos* operon, and determined that exogenous leucine contributes to negative regulation of the *tos* operon. I also determined that TosE and TosF repress motility, TosR contributes to reciprocal regulation of adhesins, and H-NS and Lrp make important contributions to adherence and motility reciprocal regulation. Here, the main findings of this work are summarized:

- TosR, TosE, and TosF are broadly conserved among UPEC strains encoding structural genes *tosCBDA*. There are five variants TosR and four variants of TosE and TosF. These classifications are based on the predicted amino acid sequences of TosR and the TosEF regulators.

- The GC contents of *tosR*, *tosE*, and *tosF* are distinct from the remainder of the *tos* operon (*tosCBDA*). This suggests that regulatory genes *tosR*, *tosE*, and *tosF* were acquired by the *tos* operon after the structural genes.
- *tosRCBDAEF* form a single transcript.
- The *tos* operon promoter, P_{tos}, is 30 bp upstream of *tosR*, consistent with the predicted transcriptional architecture of the *tos* operon.
- TosR is both a positive and negative regulator of the *tos* operon, depending on the cellular concentration of TosR. As there are at least two TosR binding sites in the vicinity of P_{tos}, with different predicted binding affinities, differential DNA binding may account for TosR dual regulation.
- H-NS is a negative regulator of the *tos* operon, and Lrp is a positive regulator of the *tos* operon. TosR positive and negative regulation of the *tos* operon depends on Lrp and H-NS, respectively. However, Lrp-mediated positive regulation of the *tos* operon appears to be independent of TosR. Exogenous leucine inhibits both Lrp- and TosR-promoted positive regulation of the *tos* operon.
- TosE and TosF together suppress motility by reducing FliC levels.
- TosR also suppresses production of P fimbria, which may proceed through binding in the vicinity of P_{papBA}.
- H-NS and Lrp contribute to reciprocal regulation of adherence and motility

Final Conclusions, perspectives, and future directions

My final perspective on *tos* operon regulation, based on the work reviewed below and proposed hypotheses (with tests), follows. *In vitro* [LB medium], the *tos* operon is both expressed and repressed, with repression overwhelmingly predominating among the UPEC cell population. The high cellular copy number of H-NS (95) normally represses the *tos* operon. Lrp, at lower concentrations than H-NS (95), is unable to induce the chromosomal copy of the *tos* operon. Reduced positive regulation by Lrp is further compounded by the fact that leucine levels are relatively high in LB medium (465), which partially inhibits Lrp-positive regulation of the *tos* operon. However, some cells in the UPEC population still express the *tos* operon even under conditions not normally permissive to *tos* operon induction. That TosR overexpression could still induce the *tos* operon (*i.e.,* Lrp can overcome leucine levels in LB), and that Lrp still positively regulates the P_{tos}-*lacZ* construct, supports the conclusion that some cells in the population, even under conditions that are not permissive for *tos* operon expression, could still express the operon. I predict that TosR relieves H-NS repression and promotes Lrp binding in the vicinity of P_{tos}. I predict that TosR overexpression also has a normalizing effect on *tos* operon expression (*i.e.,* increases the fraction of cells expressing the *tos* operon from a vanishingly small number to nearly uniform expression). Without *tosR* overexpression, I hypothesize that Lrp may promote *tos* operon expression, with TosR functioning in a feedforward loop to further favor *tos* operon expression. A high level of Lrp, itself promoting *tos* operon expression, favors this particular explanation. This, and P_{tos}-*lacZ* expression being mediated by Lrp, as discussed above, is also evidence for the

H-NS and Lrp switch processes discussed in this work (*i.e.,* regulation of the *tos* operon in the UPEC cell is either predominated by H-NS or Lrp and switches between the two). At the same time, TosR binds in the vicinity of P_{papBA} to reduce P fimbria levels, and TosEF levels accumulate to suppress flagellar levels in the UPEC cell. Thus, as TosA levels increase, motility and other adhesin levels decrease.

As TosR levels accumulate in the UPEC cell with overexpression of *tosR* in LB, I propose that TosR binds to a weaker affinity site in the vicinity of P_{tos} and inhibits expression of the *tos* operon, likely through occluding Lrp and favoring H-NS binding. H-NS levels, as noted above, are also high in the UPEC cell and could regain prominence in *tos* operon repression without TosR (*i.e.*, the unmarked *tosR* mutation does not overexpress the *tos* operon). In addition, exogenous leucine inhibits TosR-mediated positive regulation of the *tos* operon, most likely through abrogating Lrp binding in the vicinity of P_{tos}.

tos operon regulation in M9 minimal medium is different, compared to culturing in LB medium above. In a minimal medium, the cellular Lrp levels are higher (104) and exogenous leucine is essentially absent, and as such the fraction of cells in the UPEC population expressing the *tos* operon is predicted to be slightly higher than in LB. However, as overproduction of Lrp in some situations (*e.g.,* in the UPEC CFT073 Δ*tosR* background) can still induce the *tos* operon, nucleoid regulation is not abolished in M9 medium. Thus, *tos* operon expression could still be lower than necessary for detection in these assays. As Lrp levels increase in wild-type UPEC CFT073, such as from pBAD-*lrp*, I predict that *tos* operon expression also increases resulting in rapid accumulation of

TosR and subsequent inhibition of Lrp-mediated positive regulation. Without the ability to increase TosR levels (as would be the case in the Δ*tosR* background), the *tos* operon could be expressed in a manner similar to LB. Nevertheless, all of these findings suggest that the H-NS and Lrp regulation switch is active in M9 minimal medium.

When UPEC enters the urinary tract, some cells either previously exist in (*i.e.,* there are small fractions of cells expressing the *tos* operon as in LB and M9 minimal medium) or change to an H-NS and Lrp regulatory switch configuration conducive to *tos* operon expression. If the switch is previously on, these cells expressing the *tos* operon may be selected for in the urinary tract, and Lrp promotes *tos* operon expression. TosR binding in the vicinity of P_{tos}, in addition, continues to promote expression of the *tos* operon in these cells. If the switch is off in some UPEC cells, an environmental cue such as reduced leucine levels may switch the *tos* operon on. In this case, either TosR overcomes H-NS repression to promote Lrp binding, which further induces *tosR* in a feedforward loop, or Lrp overcomes H-NS repression to induce *tosR* expression in the same feedforward loop. At the same time, motility and other adhesin genes may be suppressed in these UEPC cells, ensuring variability in the UPEC population (*i.e.,* cells not producing TosA may produce other adhesins or flagella, instead of all cells making all adhesins simultaneously with flagella). As TosR levels accumulate, the *tos* operon may switch off in some other UPEC cells in the population. This may be compounded by H-NS also regaining repressive prominence of the *tos* operon. Motility, at this point, may switch on and an invasive infection could be promoted. The *tos* operon may switch back

on at a later time, or, as the UPEC population synchronized, other cells could switch the *tos* operon on.

I also hypothesize that NAPs (nucleoid-associated proteins), such as H-NS and Lrp, serve the key function in generating non-genetic variability among cells to exploit a variety of host niches or evade the host immune system (*i.e.,* allow reciprocal regulation between adhesin- and motility-related genes to occur). Indeed, the *tos* operon is optimally expressed when H-NS regulation does not predominate at P_{tos}. This condition, in conjunction with low leucine levels, Lrp predominating at P_{tos}, and modest TosR levels (all thought to inhibit H-NS predominance at P_{tos}), likely occurs in the human urinary tract.

In this work, I first set out to determine whether regulators encoded by the *E. coli* strain CFT073 *tos* operon were similarly encoded in other UPEC genomes. In a survey of the genomic sequences of 317 UPEC bacteremia isolates from the Broad Institute (https://olive.broadinstitute.org/comparisons/ecoli_uti_bacteremia.3), it was discovered that nearly one-third of these strains harbor the *tos* operon. The overwhelming majority (91%) of the *tos* operon containing UPEC strains simultaneously encode regulatory genes *tosR*, *tosE*, and *tosF*. There are five variants of TosR (based on predicted amino acid sequence) followed by four variants of TosE and three variants of TosF. TosR is a member of the PapB family, and TosE and TosF are both LuxR family members. As discussed in this work, specific variants are associated with each other. This overwhelmingly supports the conclusion that *tosR*, *tosE*, and *tosF* are genetically linked. It is also intriguing to note that *tos* operons encoding TosR variants four and five appear

to be localized to a self-transmissible plasmid, which confirms the hypothesis that the *tos* operon could be acquired by horizontal gene transfer. However, even though the evidence in support of this hypothesis is overwhelming, to establish whether the plasmid transfer genes are functional, mating between different *E. coli* strains with and without these plasmids could be performed.

The *tos* operon also appears to be dynamic, meaning that some genes seem to have been acquired or lost from the operon at different times than other genes in the operon. The GC contents of the *tos* regulatory genes (*tosREF*) are substantially lower (29.1%), than the *tos* structural genes (*tosCBDA*) (48.5%). This is unusual for genes in an operon, which typically maintain similar GC content (such as for the case for *tosCBDA*) (424). Nevertheless, as *tosR* and *tosEF* both have similar AT-richness, and all three genes are nearly always present or absent together, it seems most likely that the *tos* operon structural genes, at some point in their evolution, were inserted into a locus containing *tosREF*. Both the former observation (some operons lack *tos* regulators) and latter hypothesis (*tosCBDA* were inserted into *tosREF*) predict that multiple pathways could regulate the *tos* operon, but TosR- and Lrp-mediated regulation would occur if P_{tos} from *E. coli* strain CFT073 was inserted into these operons lacking *tosREF*. This regulation hypothesis could be tested as follows. P_{tos} from *E. coli* CFT073 could be inserted into a *tos* operon not containing this promoter at a similar site to *E. coli* CFT073. Expression from this engineered *tos* operon could then be assessed with ectopic *tosR* and *lrp* expression.

Both RNA-Seq and 5' RACE were valuable techniques to map the *tos* operon promoter. My work showed that *tosRCBDAEF* form a single transcript, and the region upstream of *tosR* can promote expression of a promoterless *lacZ* gene when engineered into a vector upstream of this *lacZ* gene. RNA-Seq and 5' RACE both identified the same transcriptional start site 23 bp upstream of *tosR*. In turn, 7 bp upstream of the identified transcriptional start site is a canonical σ^{70} promoter. As the leader sequence from this transcriptional start site is shorter than some leader sequences exhibiting post-transcriptional regulation (466), it does not appear likely that such a mechanism accounts for *tos* operon regulation at the beginning of the *tos* transcript. However, such a mechanism may still exist at other regions in the *tos* operon transcript, especially given the large gaps between *tos* operon genes (*e.g.*, 199 bp between *tosD* and *tosA*). Given that the *tos* operon is repressed when *E. coli* CFT073 is cultured under laboratory conditions (208, 227), and the *tos* promoter itself has the consensus sequence of a canonical σ^{70} promoter, these data also strongly suggest that repression and activation (overcoming repression) are likely mechanisms of *tos* operon regulation.

Using overexpression assays, I found that TosR both induces and represses expression of the *tos* operon. This dual regulation is most likely dependent on cellular concentrations of TosR; when TosR levels are low, the *tos* operon is likely to be expressed, and when TosR levels are high, the *tos* operon is less likely to be expressed. This dual regulation of cognate operons is also seen among other PapB family members (179, 180, 202, 203, 404). I have identified at least two TosR binding sites in the vicinity of P_{tos}, and I have predicted that TosR has different affinities for these binding sites,

which may couples TosR concentrations to TosR-mediated dual regulation. The strong binding site is at least 160 bp upstream of P_{tos} and may function by obstructing the function of a repressor protein. The weak binding site is near P_{tos} and may obstruct RNA polymerase or the function of another activator (*e.g.,* Lrp). An assay to verify whether TosR binds with different affinities to these sites in the vicinity of P_{tos} could include performing ITC (isothermal calorimetry) on purified TosR and DNA fragments containing TosR binding sites. TosR binding affinities for these DNA fragments could subsequently be determined, as is already the case for other DNA binding proteins (467). To determine whether or not these binding affinities are consistent with TosR dual regulation in living cells, single cell sorting assays, for the purpose of comparing TosR and TosA levels, could be performed and compared with the ITC values above.

As PapB family member binding sites do have some degree of AT-richness (182), which also appears to be the case for TosR, it seems logical to explore whether TosR, PapB, and FocB exert global regulation on *E. coli* gene expression. This regulatory property has not previously been ascribed to the PapB family, but could be logically explored, given that other global regulators also bind AT-rich sequences (41-43, 94, 286, 295-297, 309, 359). To explore the consequences of expressing each PapB family member individually and in combination, on global gene expression, microarray, ChIP-Seq, and RNA-Seq are all assays that could be utilized.

Using single cell sorting, it could also be possible to numerically calculate the conditional entropy of *tos* operon expression at various cellular levels of TosR. In so doing, a consistent physical language could be developed to describe *tos* operon (and

other operons) regulation. For the sake of brevity, entropy describes the average amount of information associated with a probability distribution of a random variable; conditional entropy is the average amount of information associated with a probability distribution of a random variable with respect to another random variable (468). I predict that the conditional entropy, itself, could be seen as describing widths of gene expression histograms of sorted cells (under an invariant scales assumption). Narrow histograms (**Figure 4-1A** and **1B**) have low gene expression variability and low entropy. Wide histograms have high entropy (**Figure 4-1C**) and high variability in gene expression across the cell population. In terms of the *tos* operon, low expression entropy may be observed with both high and low levels of TosR induction (*i.e.,* most cells will not express *tosA*). Expression entropy will be expected to increase with levels of TosR conducive to *tosA* expression.

The global regulators H-NS and Lrp are also involved in regulation of the *tos* operon. H-NS repression of the *tos* operon, in fact, may be where nearly all *tos* operon regulation is originally derived. For instance, loss of H-NS results in a substantial increase in TosA synthesis, and neither TosR- nor Lrp-mediated regulation is required without H-NS. This also makes sense in that H-NS represses genes acquired by horizontal transfer (41-43), the presumed method of *tos* operon acquisition by *E. coli*. Lrp is the presumptive positive regulator of the *tos* operon, through which any TosR-mediated positive regulation is derived. In agreement with this, loss of Lrp abolishes TosR-mediated regulation of the *tos* operon, but loss of TosR has no effect on Lrp-mediated positive regulation of the *tos* operon. Lrp is unable to overcome TosR-

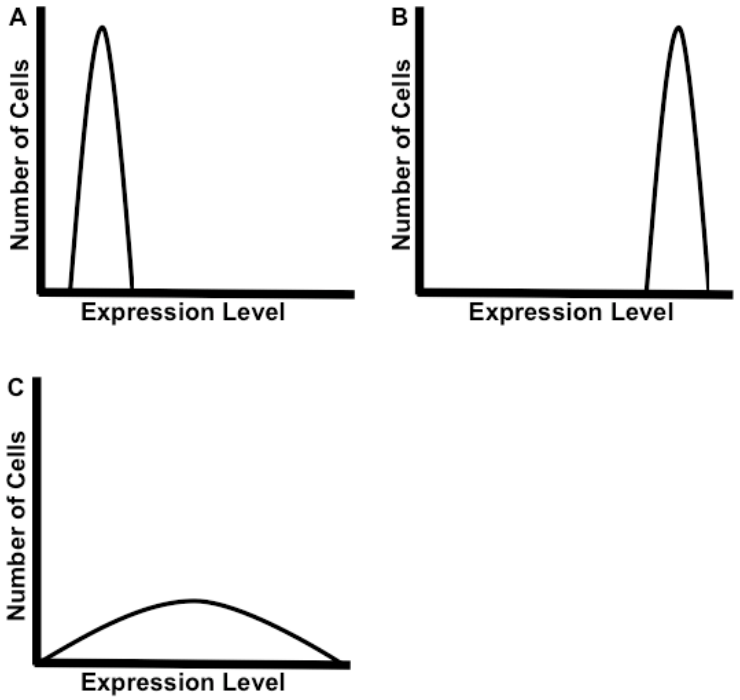

Figure 4-1. Sample curves with predicted low and high gene expression entropy. The number of cells with a strongly repressed **(A)** or strongly induce **(B)** gene will be mostly binned into low and high expression levels, respectively (low entropy). **(C)** Cells that are found in a situation where a gene could be either induced or repressed, with little certainty of either event, exhibit greater expression variability (high entropy). For simplicity, the scales and number of cells sorted are assumed to be invariant in all figures.

mediated repression at high levels of *lrp* induction in both LB and M9 media. To my knowledge, this may be the first description of a PapB family member being directly involved in Lrp-mediated regulation of its cognate operon, as opposed to regulating the expression of a gene that encodes a protein modifying Lrp regulation of the cognate PapB family operon (*e.g.*, PapI) (179-181). Thus, a new regulatory activity may have been ascribed to the PapB family.

Predictions using known H-NS and Lrp binding sites (258, 358) reveal that high affinity H-NS binding sites may be found in the vicinity of the strong TosR binding site and overlapping the *tos* operon transcriptional initiation site. Lrp binding sites may be located downstream of the strong TosR binding site and overlapping the weak TosR binding site. However, future gel shift assays must be performed with purified H-NS, Lrp, and DNA fragments containing these sequences to verify whether these NAPs (nucleoid-associated proteins) bind these sequences. In terms of the *tos* operon regulation model, it seems likely that in single cells, TosR, like other regulators (181, 307, 315), promotes positive regulation by obstructing H-NS-mediated repression of the *tos* operon when bound to the strong TosR binding site. Simultaneously, TosR binding to the strong TosR binding site may promote Lrp binding in the vicinity of P_{tos}, by possibly blocking H-NS obstruction at these sites. At high cellular levels of TosR, TosR presumably obstructs Lrp positive regulation through binding the weak TosR-binding sites, which in turn may allow H-NS binding in the vicinity of P_{tos}. H-NS, however, is an abundant NAP (95); therefore, any displacement of H-NS may only be transient. To test whether TosR competes with H-NS and Lrp for binding sites in the vicinity of P_{tos}, gel

supershift assays could be performed with differentially tagged TosR, H-NS, and Lrp proteins, mixed in combination and at various concentrations with an antibody directed against the regulator of interest in each binding reaction (469).

Expression of the *tos* operon is sensitive to exogenous levels of leucine, suggesting that this is an environmental cue involved in *tos* operon regulation. High leucine levels inhibit both TosR- and Lrp-mediated regulation of the *tos* operon. Previous work has already established that UPEC will likely encounter lower levels of leucine in the urinary tract compared with LB medium (401, 465). To further determine whether Lrp binding to P_{tos} is perturbed by leucine, gel shift assays could be performed, as above, in the presence of leucine. As TosR-mediated regulation of the *tos* operon occurs through Lrp, perturbed Lrp DNA binding would also be the presumptive mechanism of TosR sensitivity to leucine. It is important to note, however, that RNA-Seq assays on UPEC strains obtained from patients at a clinic (32) revealed that UPEC sensing leucine during an infection may be complex. For instance, although the population of UPEC cells was generally more leucine-starved in the human urinary tract, compared with LB, some strains do not upregulate genes that would otherwise be inhibited by Lrp-leucine. Thus, while lower leucine levels certainly promote *tos* operon expression, it may also be the case that at least some of the time a low exogenous leucine concentration may promote *tos* operon expression (*i.e.,* TosR could repress the *tos* operon if levels of *tosR* expression were too high). This may also agree with the finding that TosR and Lrp could induce the *tos* operon in LB, but the same regulators fail to induce the *tos* operon with a chromosomal copy of *tosR* in M9. To further test this hypothesis,

lrp and *tosR* induction could be performed in the wild-type *E. coli* CFT073 background with a concentration gradient of leucine supplemented to M9 medium and assayed for TosA levels. However, it may also be the case that gene regulation by nucleoid regulators, such as H-NS and Lrp, could be stochastic (371, 458-462) and RNA-Seq not able to consistently resolve differential expression of genes regulated by nucleoid proteins. Therefore, to ascertain whether individual cells may be responding to low leucine in the urinary tract, high-resolution techniques, such as FISSEQ (fluorescent *in situ* sequencing) and SeqFISH (sequential *in situ* hybridization) (470, 471) in conjunction with conditional entropy studies could be utilized.

An H-NS and Lrp switch (*i.e.,* a switch between H-NS- and Lrp-mediated regulation at P_{tos}) also underpins regulation of the *tos* operon (see **Figure 1-6**). For instance, when P_{tos}-*lacZ* is present on a multiple copy number plasmid, Lrp is able to promote expression of *lacZ* from this promoter. I concluded this from the fact that this overexpression phenotype is not observed in the absence of Lrp. At the same time, an *hns* mutation can suppress this lack of overexpression, from the P_{tos}-*lacZ* construct, when Lrp is also absent. The overexpression phenotype is mirrored in a Western blot to detect TosA synthesis, which further shows that a switch in H-NS- and Lrp-mediated regulation is key to *tos* operon regulation. A switch to Lrp prominence in regulation of the *tos* operon mediates positive regulation, and a switch to H-NS prominence in regulation of the *tos* operon mediates negative regulation. In addition, I predict that an H-NS predominating switch is unlikely to be observed *in vivo*, as genes negatively regulated by H-NS were either not differentially regulated or upregulated during a human UTI (32).

However, genes negatively regulated by H-NS are not uniformly upregulated by all UPEC strains during human UTI, which suggests that RNA-Seq may have lower resolution than what is necessary to consistently resolve differential gene expression as described above. To ascertain whether the H-NS occupancy part of the switch does not predominate during experimental UTI, FISSEQ and SeqFISH could again be utilized in conjunction with conditional entropy studies.

At least some of this regulation may be due to changes in the concentrations of other NAP regulators. Indeed, Lrp overexpression itself is sufficient to drive TosA synthesis, further supporting the idea that an H-NS and Lrp switch modulates *tos* operon regulation. In this instance, Lrp binding to putative Lrp-binding sites in the vicinity of P_{tos} may be sufficient to override H-NS-mediated repression of the *tos* operon. Furthermore, as outlined above, TosR-mediated regulation is abrogated in *hns* and *lrp E. coli* CFT073 mutants, which shows that this switch itself underpins *tos* operon regulation. TosR, however, may also manipulate the switch to mediate positive and negative regulation of the *tos* operon as outlined above (*i.e.,* bias the switch to favor one NAP). It is important to note, however, that H-NS and Lrp are both global regulators (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 370, 372, 375, 377, 378, 385-389). Thus, both the direct and indirect effects of this switch may be important for regulation of the *tos* operon. The importance of direct effects could be described with Lrp and H-NS competition gel supershifts as outlined above. Indirect effects could be determined through transforming a transposon mutant library (generated in wild-type *E. coli* CFT073, *hns*, *lrp* and *hns* and *lrp* mutants) with the P_{tos}-*lacZ* vector to identify secondary mutations with perturbed *tos*

operon regulation. To determine which additional environmental cues could inhibit H-NS-mediated repression, the P_{tos}-*lacZ* vector may also be used in the *lrp* mutant background for a Biolog® screen (472).

An intriguing possibility for *tos* operon regulation is that it may also be stochastically regulated. For instance, differential methylation may be stochastic and regulates the *pap* operon (458). Furthermore, H-NS and Lrp may also manipulate methylation of P_{papBA} (371, 458-460), and some level of *pap* operon expression is observed even in the presence of H-NS (*i.e.,* the negative regulator of the *pap* operon) (354). These NAPs, in addition to IHF, also contribute to phase variation of the *fimS* element of the *fim* operon, another stochastic process (166, 169, 473). H-NS and IHF may also propagate cell population noise in gene expression (*i.e.,* promote higher entropy in the expression of genes regulated by each) (461, 462). Thus, as H-NS and Lrp are regulators of the *tos* operon, it is logical to expect that some regulation of this operon will be stochastic. To test this possibility, single cell sorting could be performed on UPEC strain CFT073 cells obtained from an LB culture, M9 minimal medium culture, and the murine urinary tract for TosA synthesis. This will also be another instance where exploring the entropy of gene expression could be a useful metric for describing gene regulation both in an LB culture and the murine urinary tract. An increase in entropy suggests a high degree of variability in the cellular population (*i.e.,* cells are neither fully induced nor fully repressed), and a decrease in entropy suggests low population variability (*i.e.,* most cells are inducing or repressing expression). I predict that higher entropy suggests a more random and stochastic regulation profile. In addition, I predict

that cells obtained from the urinary tract would likely have the highest *tos* expression entropy, followed by M9 minimal medium cells, and concluding with UPEC cells cultured in LB. It is also expected that entropy will decrease for *lrp* mutations, as Lrp is the positive regulator of the *tos* operon (*i.e.,* even fewer cells will be expected to express the *tos* operon). Entropy for an *hns* mutation may, however, be less intuitive. For example it is unclear whether all cells will be binned into uniform high expression or binned into various levels of intermediate expression. Nevertheless, these assays will delineate in which direction each NAP biases the switch (*i.e.,* less variable for an *lrp* mutation and likely more variable for an *hns* mutation).

To determine whether it may be possible to enrich for cells with the H-NS and Lrp switch locked into an Lrp predominated phase, providing additional support that there is some degree of stochasticity to *tos* operon expression, a transcriptional fusion of the *tos* operon with an antibiotic resistance cassette could be engineered, the *E. coli* culture treated with the respective antibiotic, and the cells resistant to antibiotic screened for TosA production. With some additional knowledge about the *E. coli* input titer, an estimate of the number of cell generations, and the final output titer, it would be possible to estimate how many cells originally express the *tos* operon. This may even allow for some estimation of the probability that a cell switches into an Lrp predominating phase of *tos* operon regulation. The antibiotic may even be added at later time points to determine whether more or fewer cells (and the corresponding changes in the above probability) are switching into the Lrp predominating phase of *tos* operon regulation. Similar assays could also be performed *in vivo*, where antibiotics could be supplemented into the

drinking water of mice treated with this construct. Further entropy studies, examining motility and adhesin reciprocal regulation, as outlined below, could also be used to further demonstrate whether the state of this nucleoid switch could be enriched *in vivo*.

Overexpression of *tosEF* leads to a decrease in flagella-mediated motility. This finding is consistent with coordination between adhesin and flagellar operons previously discussed. For example, when the *tos* operon, encoding the TosA adhesin, is expressed motility is suppressed. It is intriguing to note that motility suppression is also attributed to the terminal gene of the *pap* operon, *papX* (187, 188), which is also regulated to some degree by PapB (188). I found that this *tos*-related motility suppression is most likely mediated through inhibiting FliC synthesis (the main structural component of flagella). Thus, the *tos* operon is the first known *E. coli* RTX nonfimbrial adhesin operon to be integrated into the *E. coli* reciprocal regulation network (*e.g.,* between adhesins and flagellar-mediated motility). It is important to note, however, that I found that full motility suppression is an emergent phenomenon (*i.e.,* it is superadditive). This motility suppression is superadditive owing to the observation that both TosE and TosF together are required for full motility suppression, and motility suppression from each regulator individually summed together does not equal motility suppression when both regulators are present at the same time. I note, however, that the precise level of TosEF control over motility is unknown (*i.e.,* there are multiple levels of flagellar gene regulation) (116-119). Thus, gel shift assays must still be performed using purified TosE, TosF, and DNA fragments derived from flagellar gene promoters. To determine whether secondary interactions are also important for TosEF-mediated motility suppression, *tosEF*

overexpression and motility assays could be performed in a library of transposon mutants. To corroborate these findings, it may also be of interest to perform RNA-Seq experiments on this overexpression construct to determine whether suppressor mutants are similarly differentially regulated by TosEF and whether motility genes are also differentially regulated in these mutants. ChIP-seq or SELEX (systematic evolution of ligands by exponential enrichment) may also be performed to map additional TosEF binding sites.

It remains unclear, however, why *tosEF* expressed from the chromosome show only a modest decrease in motility compared with *tosEF* expressed from a multiple copy number plasmid and why there is only a modest increase in motility when *tosR*, *tosE*, and *tosF* are simultaneously deleted in CFT073. One intriguing possibility is that additional environmental stimuli are necessary to mediate motility suppression (*i.e.*, this requirement could be overcome when TosE and TosF are present at high concentrations). To test whether environmental cues, especially those supporting *tos* operon expression, supplement TosEF repression of motility, soft agar motility assays may be performed on unmarked *tosR* single mutants, unmarked *tosR*, *tosE*, and *tosF* triple mutants, and wild-type *E. coli* CFT073 in the presence of putative stimuli. In addition, it is also a possibility that motility suppression may not be uniform across the CFT073 cellular population. To determine whether this may also be an explanation for TosEF-mediated motility repression, a conditional entropy experiment may again be performed. Cell sorting, with gates on TosA and FliC levels, could be performed on the *tos* operon mutants indicated above and wild-type CFT073. The conditional entropy for FliC expression would be expected to decrease at high TosA levels (*i.e.,* less variability in this

subpopulation with cells mostly suppressing motility). These results may then be corroborated with cell sorting performed on wild-type cells obtained *in vivo*. Consistent with the hypothesis that TosEF-mediated motility suppression is not uniform across the CFT073 population, cell sorting of unmarked *tosR* mutants will also be expected to reveal a conditional entropy higher at some TosA levels (*i.e.,* motility suppression may be incomplete). TosR overexpression assays, again explored with cell sorting and conditional entropy for FliC levels at specific TosR levels in both wild-type and *tosEF* mutant strains, could be performed to further connect *tos* operon regulation with *E. coli* reciprocal regulation.

To my surprise, TosR also negatively regulates P fimbria synthesis and positively regulates *E. coli* biofilm components (curli and/or cellulose). Such cross-regulation occurs between other homologous PapB family members (203, 204, 249, 250). TosR, on the other hand, does not share significant homology with PapB; the two regulators only share 27% predicted amino acid sequence identity. Nevertheless, this shows that even nonfimbrial adhesin operons could be involved in the reciprocal regulation between adhesin genes. Intriguingly, ectopic expression of *tosR* results in an increased production of *E. coli* biofilm components (curli and/or cellulose). To my knowledge, this is also the first instance of a cognate regulator of an RTX nonfimbrial adhesin operon participating in cross-regulation between adhesin genes and the first instance of a PapB family member regulating curli and/or cellulose synthesis. Thus, adhesin operon reciprocal regulation is generalizable to other adhesin operons.

To further demonstrate whether TosR is a repressor of the *pap* operon, a qRT-PCR assay may be performed on transcripts derived from cells overexpressing TosR. To map the TosR binding sites in the vicinity of P_{papBA} and determine whether these binding sites overlap other regulatory sequences, gel shift assays could again be performed with shorter segments of this promoter region. It will also be informative to explore whether TosR could regulate other adhesin operons (*e.g., fim* and *foc* operons). To determine whether this additional regulation is possible, the same assays above could be used. As *E. coli* does not simultaneously express all of its adhesin operons (251, 252), it is intuitive that adhesin regulators may also promote this phenotype and variability in the cell population (*i.e.,* not all cells will express exactly the same adhesins at exactly the same time). Although unintuitive, to test this hypothesis a cell sorting conditional entropy experiment could again be conducted to demonstrate whether low entropy in expression is observed among other adhesins (*e.g., pap, fim*, and *foc* operons) when TosA levels are high. To demonstrate the importance of this phenotype during an infection, the same assay could also be conducted on UPEC bacteria collected during a murine experimental infection.

To determine whether TosR directly binds in the vicinity of curli and/or cellulose synthesis operon promoters, gel shift assays could be performed using TosR-His_6, P_{csgD}, P_{csgBA}, and P_{bcsA}. Additionally, to demonstrate whether TosR transcriptionally activates expression of the *csg* and *bcs* operons, qRT-PCR assays could be utilized in conjunction with ectopic expression of *tosR*. In addition, to determine whether additional proteins contribute to TosR-mediated curli and/or cellulose production, Congo red binding assays

on YESTA (yeast extract and tryptone agar) plates could be performed with ectopic expression of *tosR* in a library of UPEC CFT073 transposon mutants. Likewise, to analyze the conditional entropy of *tosR* expression (expressed from a plasmid) in addition to the *tos*, curli, cellulose, and flagellar operons, transcriptional fusions of genes encoding fluorescent proteins could be engineered in association with *tosR* expression or the expression of each operon above. High-resolution microscopy could be utilized in a manner similar to another rugose biofilm study (234), and a frequentist approach to probability could be used to determine the conditional entropy of *tos*, curli, and cellulose operon expression. This assay will also allow for a determination of the coincidence of expression of *tosR* and the other operons above, within a rugose biofilm (*i.e.,* a situation with low conditional entropy). To further our understanding of these structures in terms of gene expression and the behaviors of component cells, the conditional entropy of the expression of *tosR* with the other operons above could be mapped within the geography of the rugose biofilm. FISSEQ and SeqFISH (470, 471), in conjunction with conditional entropy, could be utilized to estimate the state of the nucleoid (*i.e.,* switched into an H-NS or Lrp predominating state) of cells associated with a rugose biofilm. As ectopic expression of *focB* also promotes UPEC CFT073 weak Congo red binding, additional ectopic expression assays using other PapB family members and YESTA plates containing Congo red could be performed. To determine whether FocB and other PapB family members promoting UPEC CFT073 binding to Congo red do function similarly to TosR, all of the above assays in this section could be repeated with the aforementioned PapB family members replacing TosR.

Finally, it was demonstrated in this work and elsewhere that NAPs H-NS and Lrp, making up the switch that regulates the *tos* operon, also make contributions to the switch between adherent and motile lifestyles (148, 179, 181, 202, 204, 279, 348, 351, 352, 354, 355, 371-378). In particular, H-NS mediates dual regulation of motility-related genes and negative regulation of adherence-related genes (148, 179, 181, 202, 326, 351, 354, 371-379, 381). Lrp mediates dual regulation of adherence-related genes and negative regulation of motility-related genes (148, 179, 202, 204, 279, 348, 352, 354, 355). Thus, while both NAPs are global regulators (42, 105, 152, 179, 183, 250, 296, 326, 346-355, 370, 372, 375, 377, 378, 385-389), symmetry may be maintained between *E. coli* cells in the population (*i.e.,* H-NS mutants almost always have reduced motility, and Lrp mutants have lower expression of some adhesins). In this same pursuit, from this work it appears that typical *E. coli* reciprocal regulation between adhesin genes and between adhesin and motility genes requires NAPs. Indeed, the reduced motility of the UPEC strain CFT073 *hns* mutant cannot be suppressed by an additional *lrp* mutation. Surprisingly, PapA levels are dramatically decreased in the UPEC strain CFT073 *hns* and *lrp* double mutant, while TosA levels remain high. Expression from P_{papBA} was previously observed in another *hns* and *lrp* double mutant *E. coli* strain not harboring the *tos* operon (354, 371). Therefore, some elements of adhesin reciprocal regulation are abrogated in nucleoid structuring mutants, as I predict that the TosA adhesin simply dominates over P fimbria in the cell population (*i.e.,* little opportunity may exist to switch between adhesins). To further determine whether the switch between nucleoid state, manifesting as a switch between *E. coli* lifestyles (*e.g.,* adherent and motile), is similarly important *in vivo*, the

same *tos* operon antibiotic resistance cassette transcriptional fusion discussed above could be used to determine the conditional entropy of FliC, PapA, and other adhesins during experimental murine UTI. This would not only provide an additional control for the same transcriptional fusion experiment described above, but to my knowledge be the first assay to directly demonstrate a nucleoid switch phenomenon underpinning *E. coli* reciprocal regulation during an experimental infection. To my knowledge, this would also be the first experiment to provide evidence in strong support of reciprocal regulation at the single cell level *in vivo*.

APPENDICES

APPENDIX A:

TOSR VARIANTS 4 AND 5 ARE HARBORED ON A PUTATIVE SELF-TRANSMISSIBLE PLASMID

To verify that our developed project R search algorithm for GC content described in **Chapter 2** identified correct *tos* operon sequences, a determination of the distance between the start and end of the *tos* operon was included (*i.e.,* start nucleotide position of the operon subtracted from the last nucleotide position). For some *tos* operons encoding TosR variants 4 and 5 this distance was negative (**Table A-1)**, which suggests that the *tos* operon sequences ended before it began (*i.e.,* the contig harboring the *tos* operon was a circular molecule). Subsequent analysis revealed that *tra* genes and plasmid stability genes are also harbored on these contigs, which overwhelmingly supports the conclusion that these *tos* operons are on a self-transmissible plasmid.

Table A-1. TosR variants predicted to be encoded on a self-transmissible plasmid.

UPEC Strain	TosR Variant	Distance between *tosC* and *tosA*[a]
HVH 21	TosR Variant 4	-124,599 bp
HVH 138	TosR Variant 4	-109,149 bp
KOEGE 131	TosR Variant 5	-93,424 bp

[a] Start nucleotide position of *tosC* gene subtracted from the last nucleotide position of *tosA*

APPENDIX B:

tosEF OVEREXPRESSION DOES NOT RESULT IN TOSA PRODUCTION

To determine whether *tosEF* overexpression could result in TosA production, suggesting that one or both of these regulators are involved in positive regulation of the *tos* operon, we performed pBAD-*tosEF* ectopic expression assays in wild-type UPEC strain CFT073 cultured in LB containing ampicillin (100 μg/mL) and a variety of L-arabinose concentrations for 4 hours. Western blotting using polyclonal anti-TosA antibodies blots on total proteins derived from the above constructs revealed that TosEF do not mediate TosA production, despite ectopic expression of these regulators (**Figure B-1**). Thus, we conclude that TosEF are not involved in simple positive regulation of the *tos* operon.

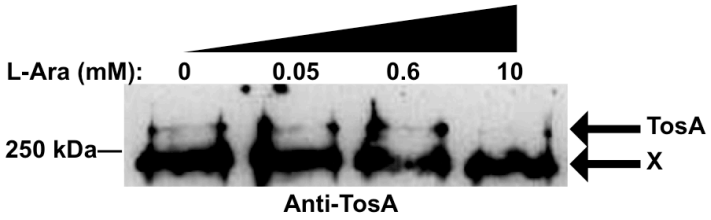

Figure B-1. TosEF are not positive regulators of the *tos* operon. A Western blot was performed using total proteins obtained from wild-type UPEC CFT073 harboring pBAD-*tosEF* induced with the indicated concentrations of L-arabinose. Bands corresponding to TosA were detected with polyclonal anti-TosA antibodies. An X indicates a cross-reacting band.

APPENDIX C:

THE PHAGE TRANSDUCED, UNMARKED *tosR* MUTATION IN UPEC STRAIN CFT073 DOES NOT OVERPRODUCE TOSA

To determine whether phage transduction of the *tosR* mutation along with unmarking the construct could reveal contributions of secondary mutations to *tos* operon overexpression observed in **Chapter 2**, phage transduction and Western blotting with polyclonal anti-TosA antibodies were performed as outlined in **Chapter 2**. It was revealed that this engineered construct did not overproduce TosA (**Figure C-1**), in contrast to the first unmarked *tosR* mutation (**Figure 2-3C**). This suggests that a secondary mutation could contribute to the *tos* operon overexpression phenotype observed in the previously unmarked *tosR* mutation. Indeed, subsequent 5' RACE (rapid analysis of cDNA ends) performed on the first unmarked *tosR* mutation revealed that this unmarked mutation was also an inversion mutation with a previously unmarked *lacZ* mutation. This may confound some conclusions from the original unmarked *tosR* mutation discussed in **Chapter 2**.

Figure C-1. The transduced, unmarked *tosR* mutation does not lead to overproduction of TosA. A Western blot was performed using total proteins obtained from a Δ*tosR*Δ*aph* UPEC CFT073 strain (derived from a *tosR* mutation transductant). TosA was detected with a polyclonal anti-TosA antibody, and bands corresponding to TosA are indicated in the figure. Equal amounts of proteins were loaded as determined using a Pierce BCA Protein Assay kit.

APPENDIX D:

THE AREA IN THE VICINITY OF P_{tos} IS PREDICTED TO BE SIMILARLY BENT TO THE AREA IN THE VICINITY OF P_{papBA}

Modified from: M. D. Engstrom and H. L. T. Mobley. 2016. Infection and Immunity 84: 811-21.

To determine whether regulation of the *tos* operon has elements similar to regulation of the *pap* operon, we utilized a web-based tool (http://www.lfd.uci.edu/~gohlke/dnacurve/) to predict DNA structure in the vicinity of P_{tos} and P_{papBA} (**Figures D-1A** and **1B**). Indeed, both P_{tos} (**Figure D-1A**) and P_{papBA} (**Figure D-1B**) regions are similarly bent. Thus, it seems likely that a PapB family member, H-NS, and Lrp could potentiate *tos* operon regulation. In addition, a region of DNA in the vicinity of P_{lacZ}, unrelated to both P_{tos} and P_{papBA}, does not exhibit a curved architecture (**Figure D-1C**).

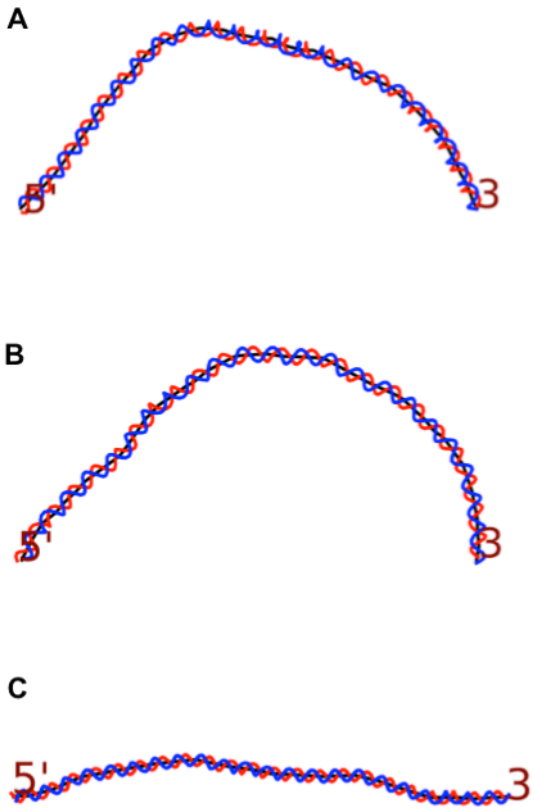

Figure D-1. An AT-rich region near P_{tos} is predicted to be bent similar to a corresponding region near P_{papBA}. **(A)** A predicted DNA curvature plot of a region near P_{tos}, obtained from http://www.lfd.uci.edu/~gohlke/dnacurve/ using an AA Wedge Model and oriented in the 5' to 3' direction, is indicated. **(B)** The same web-based tool as above was used to generate a predicted DNA curvature plot of a region near P_{papBA}. **(C)** The predicted DNA curvature plot, modeled using the same methods as above, of the region in the vicinity of P_{lacZ} is shown.

APPENDIX E:

NUMEROUS PUTATIVE REGULATORY SEQUENCES ARE FOUND IN THE VICINITY OF P_{tos}

Modified from: M. D. Engstrom and H. L. T. Mobley. 2016. Infection and Immunity 84: 811-21.

To determine whether predicted H-NS- (258, 456) and Lrp-sites (358) could be found in the vicinity of P_{tos}, this region was searched for these predicted high affinity sites. A BLASTN search was also conducted to map putative additional binding sites. We found, in addition to predicted TosR binding sites, putative high affinity H-NS binding sites and clusters of Lrp binding sites (**Figure E-1**). In addition, a putative CitB binding site (457) was also found near P_{tos}. Thus, multiple regulatory systems may be involved in *tos* operon regulation.

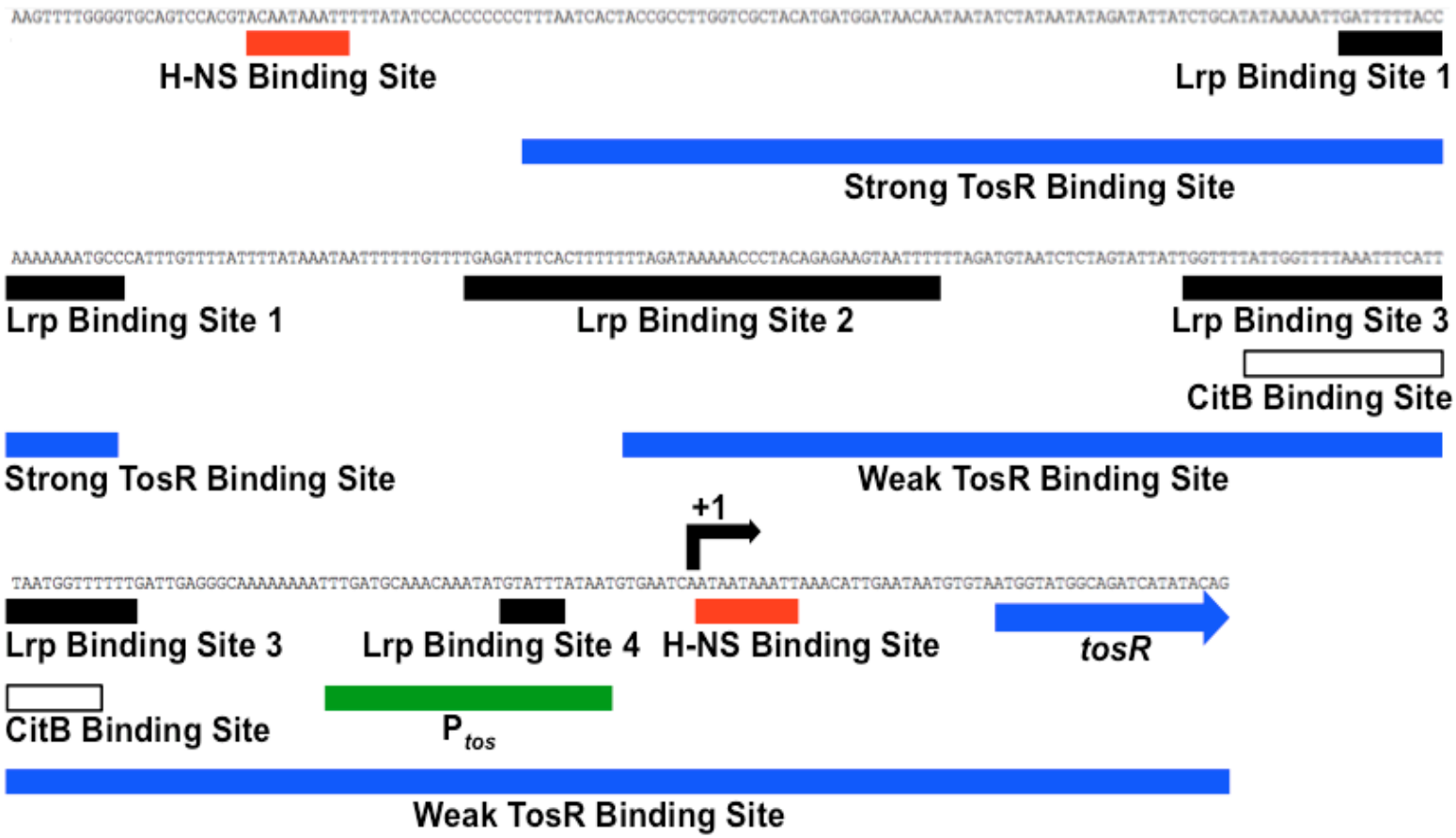
AAGTTTTGGGGTGCAGTCCACGTACAATAAATTTTTATATCCACCCCCCCTTTAATCACTACCGCCTTGGTCGCTACATGATGGATAACAATAATATCTATAATATAGATATTATCTGCATATAAAAATTGATTTTTACC
H-NS Binding Site
Lrp Binding Site 1
Strong TosR Binding Site
AAAAAAATGCCCATTTGTTTTATTTTATAAATAATTTTTTGTTTTGAGATTTCACTTTTTTTAGATAAAAACCCTACAGAGAAGTAATTTTTTAGATGTAATCTCTAGTATTATTGGTTTTATTGGTTTTAAATTTCATT
Lrp Binding Site 1
Lrp Binding Site 2
Lrp Binding Site 3
CitB Binding Site
Strong TosR Binding Site
Weak TosR Binding Site
+1
TAATGGTTTTTTGATTGAGGGCAAAAAAAATTTGATGCAAACAAATATGTATTTATAATGTGAATCAATAATAAATTAAACATTGAATAATGTGTAATGGTATGGCAGATCATATACAG
Lrp Binding Site 3
Lrp Binding Site 4
H-NS Binding Site
tosR
CitB Binding Site
P tos
Weak TosR Binding Site

Figure E-1. H-NS and Lrp binding sites are predicted in the vicinity of P_{tos}. Predicted H-NS binding sites are marked with a red bar; predicted Lrp binding sites are marked with black bars; the predicted strong and weak TosR binding sites are denoted with blue bars; a predicted CitB binding site is denoted with a white bar; P_{tos} is indicated with a green bar, and the predicted transcriptional start site is denoted with a "+1" and hooked arrow.

APPENDIX F:

LEUCINE AND NUCLEOID STRUCTURE PERTURBATIONS POTENTIALLY CONTRIBUTE TO *tos* OPERON REGULATION DURING HUMAN UTI

To determine whether TosR-mediated regulation of the *tos* operon could be modulated by the addition of exogenous leucine to LB medium, we performed pBAD-*tosR* ectopic expression assays in wild-type UPEC CFT073 cultured in LB containing ampicillin (100 μg/mL) and 0.6 mM L-arabinose for 4 hours with and without additional exogenous leucine (**Figure F-1**). From this work, TosR-mediated expression of the *tos* operon is inhibited by additional exogenous leucine. In agreement with exogenous leucine being important for *tos* operon negative regulation, several genes expected to be downregulated by Lrp-leucine were generally upregulated in clinical UTI RNA-Seq experiments (**Table F-1**), and several genes expected to be upregulated by Lrp, but inhibited by leucine, were also generally upregulated in the human urinary tract (**Table F-2**) (32). This suggests that leucine levels are low *in vivo*. Likewise, in the same RNA-Seq experiment, it was revealed that a number of genes downregulated by H-NS, like the *tos* operon, were similarly generally upregulated *in vivo* (**Table F-3**). This suggests that H-NS-mediated negative regulation is less prominent in the human urinary tract, which is a condition that supports *tos* operon expression.

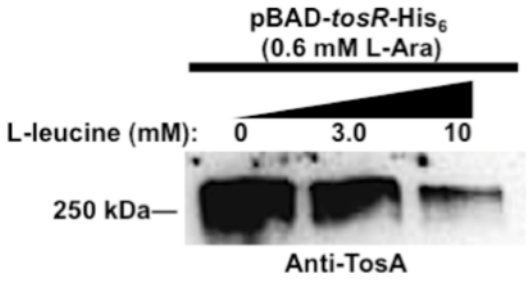

Figure F-1. Exogenous L-leucine negatively regulates TosR-mediated *tos* operon positive regulation. A Western blot was performed using total proteins obtained from wild-type UPEC CFT073 harboring pBAD-*tosR*-His_6 induced with 0.6 mM L-arabinose in the presence of the indicated concentrations of additional L-leucine in LB medium. Bands corresponding to TosA were detected with polyclonal anti-TosA antibodies.

Table F-1. Clinical UTI genes downregulated by Lrp in the presence of leucine.

UPEC Strain	Gene	Log_2(UR/LB)[a*]	Log_2(UTI/LB)[b*]	Log_2(UTI/UR)[c*]
HM26	*livK*	7.56	10.02	2.47
	livJ	4.54	7.59	3.06
	brnQ	2.61	5.67	3.06
HM27	*livK*	7.54	8.48	0.94
	livJ	5.02	5.60	0.58
	brnQ	1.94	3.95	2.01
HM46	*livK*	8.84	7.63	-1.21
	livJ	5.08	5.53	0.45
	brnQ	1.93	2.94	1.01
HM65	*livK*	6.24	2.15	-4.09
	livJ	1.70	2.79	1.09
	brnQ	-0.06	2.30	2.36
HM69	*livK*	6.79	3.66	-3.13
	livJ	5.17	1.38	-3.79
	brnQ	2.26	3.10	0.84

[a] Log_2 transformed fold increase of gene expression in urine compared to LB

[b] Log_2 transformed fold increase of gene expression in human UTI compared to LB

[c] Log_2 transformed fold increase of gene expression in human UTI compared to urine

*A log_2 transformed fold change <-2 or 2< is considered differentially expressed

Table F-2. Clinical UTI genes upregulated by Lrp in the absence of leucine.

UPEC Strain	Gene	Log_2(UR/LB)[a*]	Log_2(UTI/LB)[b*]	Log_2(UTI/UR)[c*]
HM26	*htrE*	1.27	6.38	5.11
	stpA	2.17	4.88	2.71
	serB	-0.32	4.24	4.56
	lolC	0.95	4.91	3.96
HM27	*htrE*	0.13	3.20	3.07
	stpA	1.49	4.19	2.70
	serB	-0.21	3.10	3.32
	lolC	-0.08	3.73	3.81
HM46	*htrE*	0.46	5.21	4.75
	stpA	2.01	3.85	1.84
	serB	0.18	1.69	1.51
	lolC	-0.26	1.23	1.48
HM65	*htrE*	1.75	1.11	-0.65
	stpA	1.99	2.45	0.46
	serB	-2.43	-0.68	1.75
	lolC	-2.63	-2.42	0.21
HM69	*htrE*	-0.63	0.27	0.90
	stpA	2.12	4.32	2.20
	serB	-0.35	1.29	1.64
	lolC	-0.91	1.64	2.55

[a] Log_2 transformed fold increase of gene expression in urine compared to LB

[b] Log_2 transformed fold increase of gene expression in human UTI compared to LB

[c] Log_2 transformed fold increase of gene expression in human UTI compared to urine

*A log_2 transformed fold change <-2 or 2< is considered differentially expressed

Table F-3. Clinical UTI genes downregulated by H-NS.

UPEC Strain	Gene	Log_2(UR/LB)[a*]	Log_2(UTI/LB)[b*]	Log_2(UTI/UR)[c*]
HM26	*fis*	-0.38	2.19	2.58
	hdfR	-0.23	4.26	4.50
	rcsD	0.48	2.50	2.02
	papA	—	—	—
	fimA	0.73	4.21	3.48
	csgA	1.79	7.31	5.52
HM27	*fis*	-1.43	2.26	3.68
	hdfR	0.51	4.13	3.63
	rcsD	0.49	2.54	2.05
	papA	2.00	-0.48	-2.48
	fimA	0.71	7.44	6.73
	csgA	1.66	5.58	3.92
HM46	*fis*	-0.69	3.34	4.04
	hdfR	-0.58	2.08	2.66
	rcsD	-0.09	0.79	0.87
	papA	—	—	—
	fimA	0.31	0.39	0.08
	csgA	1.52	3.83	2.31
HM65	*fis*	-5.85	-4.83	1.03
	hdfR	0.49	1.59	1.10
	rcsD	1.81	2.14	0.32
	papA	—	—	—
	fimA	-3.56	-0.21	3.35
	csgA	-2.31	-0.13	2.18
HM69	*fis*	0.06	3.15	3.09
	hdfR	-0.46	1.67	2.13
	rcsD	-0.23	1.01	1.24

papA	-0.20	7.70	7.90
fimA	0.30	-2.89	-3.19
csgA	1.05	-0.35	-1.40

[a] Log_2 transformed fold increase of gene expression in urine compared to LB

[b] Log_2 transformed fold increase of gene expression in human UTI compared to LB

[c] Log_2 transformed fold increase of gene expression in human UTI compared to urine

*A log_2 transformed fold change <-2 or 2< is considered differentially expressed

— Gene is not encoded by this UPEC strain

APPENDIX G:

TOSR POSITIVELY REGULATES CURLI AND/OR CELLULOSE PRODUCTION IN UPEC STRAIN CFT073

To determine whether ectopic expression of *tosR* mediates curli and/or cellulose production, showing that other adherence-mediating components may be under TosR regulation, UPEC strain CFT073 harboring pBAD-*tosR*-His_6, pBAD-*papB*, pBAD-*focB*, or the empty pBAD plasmid were spotted onto YESTA (yeast extract and tryptone agar) Congo red plates (10 g/L tryptone, 1 g/L yeast extract, 15 g/L agar, 50 μg/mL Congo red, and 1 μg/mL Coomassie Brilliant Blue G-250) containing ampicillin (100 μg/mL) and 10 mM L-arabinose. After incubation for 48 hours at 37 °C, strain CFT073 harboring pBAD-*tosR*-His_6 strongly exhibits the rdar (red, dry, and rough) phenotype (**Figure G-1A**). This suggests curli and/or cellulose were produced by a UPEC CFT073 strain that expressed *tosR* (230-233). Thus, production of *E. coli* biofilm components may be coordinated with TosA synthesis and further underscores how *tos* operon regulation may be coordinated with other adherence-mediating components. In addition, ectopic expression of *focB* also weakly promotes UPEC CFT073 binding of the Congo red dye (**Figure G-1B**), which suggests that FocB, and possibly other PapB family members, may coordinate synthesis of curli and/or cellulose with other adhesins.

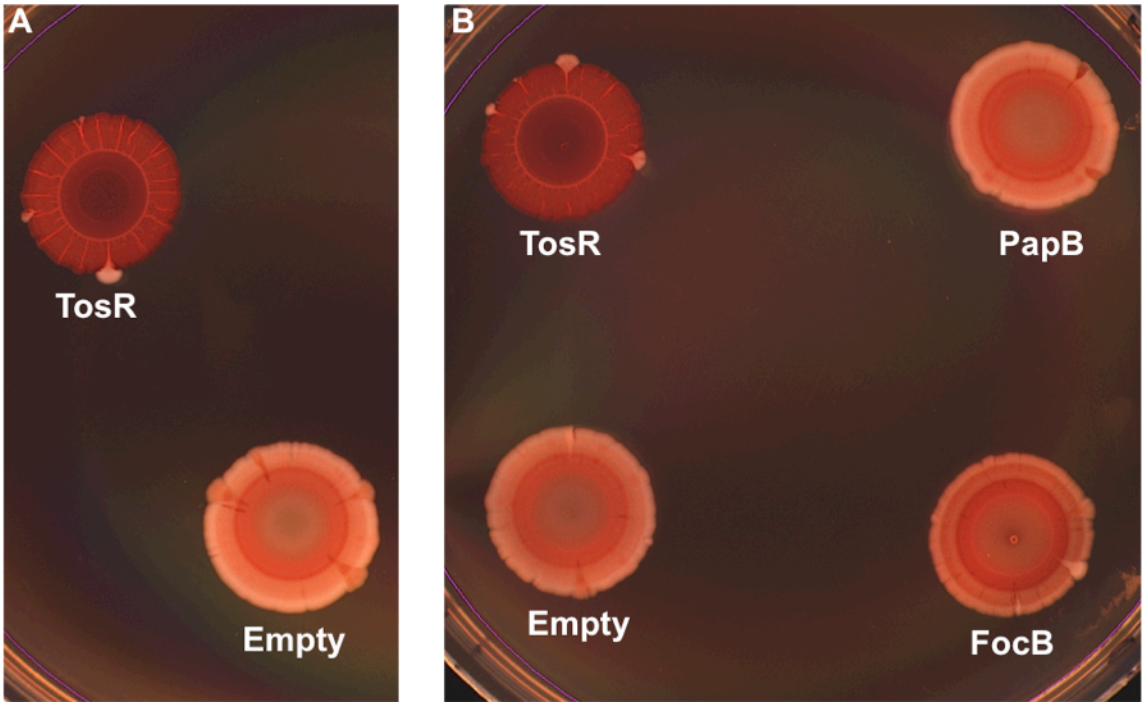

Figure G-1. Ectopic expression of TosR and FocB promote curli and/or cellulose synthesis. (A) After 48 hours of incubation at 37 °C on a YESTA (yeast extract and tryptone agar) plate containing 40 μg/mL Congo red, 1 μg/mL Coomassie Brilliant Blue R-250, ampicillin (100 μg/mL), and 10 mM L-arabinose, UPEC CFT073 harboring pBAD-*tosR*-His_6 exhibits the rdar (red, dry, and rough) phenotype indicative of curli and/or cellulose production. The CFT073 strain harboring the empty pBAD plasmid does not bind the Congo red dye. **(B)** After 48 hours of incubation under the same conditions as above, UPEC CFT073 harboring pBAD-*tosR*-His_6 exhibits the rdar phenotype. UPEC CFT073 harboring pBAD-*focB* weakly binds the Congo red dye, which suggests low production of curli and/or cellulose. The CFT073 stains harboring pBAD-*papB* and the empty pBAD plasmid do not bind the Congo red dye.

APPENDIX H:

RECIPROCAL REGULATION OF ADHESIN- AND MOTILITY-RELATED GENES IS ABROGATED IN UPEC STRAIN CFT073 *hns* AND *lrp* MUTANTS

To determine whether *fliC* expression, reduced in an H-NS mutant (375, 376), could be restored with the additional mutation of *lrp* [a mutation previously observed to enhance motility (355)], we performed Western blots on whole cell proteins obtained from wild-type CFT073, Δ*hns*, Δ*lrp*, and Δ*hns*Δ*lrp* using anti-FliC antibodies (**Figure H-1A)**. FliC levels cannot be restored in the Δ*hns*Δ*lrp* background. Likewise, to determine whether reciprocal regulation between adhesins is possible in the CFT073 Δ*hns*Δ*lrp* (a background suggested to favor *pap* operon expression), we performed a pBAD-*tosR*-His_6 overexpression assay (**Figure H-1B**). Unexpectedly, PapA levels were low regardless of ectopic expression of *tosR*, which suggests that adhesin reciprocal regulation is abolished (*i.e.,* I hypothesize that the mechanism to ensure some expression of alternative adhesin operons is lost without H-NS and Lrp).

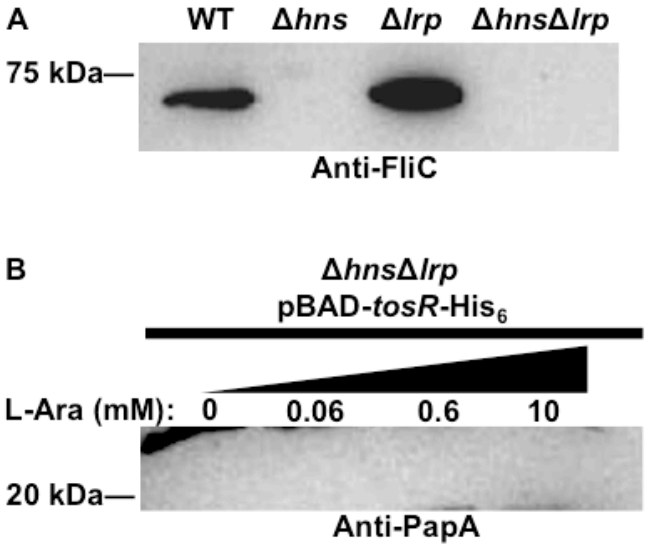

Figure H-1. H-NS and Lrp serve a key function mediating reciprocal regulation between adhesin- and motility-related genes. (A) A Western blot was performed using total proteins obtained from the indicated CFT073 strains cultured in tryptone medium. Bands corresponding to FliC were detected using polyclonal anti-FliC antibodies. **(B)** A Western blot was performed using total proteins obtained from the CFT073 Δ*hns*Δ*lrp* mutant harboring pBAD-*tosR*-His_6 induced with the indicated concentrations of L-arabinose in LB medium. Polyclonal anti-PapA antibodies were used to detect PapA.

REFERENCES

1. **Foxman B, Brown P.** 2003. Epidemiology of urinary tract infections: transmission and risk factors, incidence, and costs. Infect Dis Clin North Am **17:**227-241.
2. **Foxman B, Gillespie B, Koopman J, Zhang L, Palin K, Tallman P, Marsh JV, Spear S, Sobel JD, Marty MJ, Marrs CF.** 2000. Risk factors for second urinary tract infection among college women. Am J Epidemiol **151:**1194-1205.
3. **Litwin MS, Saigal CS, Yano EM, Avila C, Geschwind SA, Hanley JM, Joyce GF, Madison R, Pace J, Polich SM, Wang M, Urologic Diseases in America P.** 2005. Urologic diseases in America Project: analytical methods and principal findings. J Urol **173:**933-937.
4. **Keating KN, Perfetto EM, Subedi P.** 2005. Economic burden of uncomplicated urinary tract infections: direct, indirect and intangible costs. Expert Rev Pharmacoecon Outcomes Res **5:**457-466.
5. **Ciani O, Grassi D, Tarricone R.** 2013. An economic perspective on urinary tract infection: the "costs of resignation". Clin Drug Investig **33:**255-261.
6. **Abrahamian FM, Krishnadasan A, Mower WR, Moran GJ, Coker JR, Talan DA.** 2011. The association of antimicrobial resistance with cure and quality of life among women with acute uncomplicated cystitis. Infection **39:**507-514.
7. **Edlin RS, Shapiro DJ, Hersh AL, Copp HL.** 2013. Antibiotic resistance patterns of outpatient pediatric urinary tract infections. J Urol **190:**222-227.
8. **Karlowsky JA, Thornsberry C, Jones ME, Sahm DF.** 2003. Susceptibility of antimicrobial-resistant urinary Escherichia coli isolates to fluoroquinolones and nitrofurantoin. Clin Infect Dis **36:**183-187.
9. **Talan DA, Krishnadasan A, Abrahamian FM, Stamm WE, Moran GJ, Group EMINS.** 2008. Prevalence and risk factor analysis of trimethoprim-sulfamethoxazole- and fluoroquinolone-resistant Escherichia coli infection among emergency department patients with pyelonephritis. Clin Infect Dis **47:**1150-1158.
10. **Katouli M.** 2010. Population structure of gut Escherichia coli and its role in development of extra-intestinal infections. Iran J Microbiol **2:**59-72.
11. **Czaja CA, Stamm WE, Stapleton AE, Roberts PL, Hawn TR, Scholes D, Samadpour M, Hultgren SJ, Hooton TM.** 2009. Prospective cohort study of microbial and inflammatory events immediately preceding Escherichia coli recurrent urinary tract infection in women. J Infect Dis **200:**528-536.
12. **Foxman B.** 2002. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am J Med **113 Suppl 1A:**5S-13S.
13. **Foxman B.** 2010. The epidemiology of urinary tract infection. Nat Rev Urol **7:**653-660.
14. **Ikaheimo R, Siitonen A, Karkkainen U, Mustonen J, Heiskanen T, Makela PH.** 1994. Community-acquired pyelonephritis in adults: characteristics of E. coli isolates in bacteremic and non-bacteremic patients. Scand J Infect Dis **26:**289-296.
15. **Kalra OP, Raizada A.** 2009. Approach to a patient with urosepsis. J Glob Infect Dis **1:**57-63.
16. **Bent S, Nallamothu BK, Simel DL, Fihn SD, Saint S.** 2002. Does this woman have an acute uncomplicated urinary tract infection? JAMA **287:**2701-2710.

17. **Fairley KF, Carson NE, Gutch RC, Leighton P, Grounds AD, Laird EC, McCallum PH, Sleeman RL, O'Keefe CM.** 1971. Site of infection in acute urinary-tract infection in general practice. Lancet **2:**615-618.
18. **Nahar A, Akom M, Hanes D, Briglia A, Drachenberg CB, Weinman EJ.** 2004. Pyelonephritis and acute renal failure. Am J Med Sci **328:**121-123.
19. **Kleeman CR, Hewitt W, Guze LB.** 1960. Pyelonephritis. J Am Med Assoc **173:**257-259.
20. **Najar MS, Saldanha CL, Banday KA.** 2009. Approach to urinary tract infections. Indian J Nephrol **19:**129-139.
21. **Mazzulli T.** 2012. Diagnosis and management of simple and complicated urinary tract infections (UTIs). Can J Urol **19 Suppl 1:**42-48.
22. **Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ.** 2015. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. Nat Rev Microbiol **13:**269-284.
23. **Reiss D, Engstrom M, Mobley H.** 2013. Urinary Tract Infections, p 323-351. *In* Rosenberg E, DeLong E, Lory S, Stackebrandt E, Thompson F (ed), The Prokaryotes, 4 ed, vol 5: Human Microbiology. Springer-Verlag Berlin Heidelberg, Heidelberg, New York, Dordrecht, London.
24. **Ronald A.** 2003. The etiology of urinary tract infection: traditional and emerging pathogens. Dis Mon **49:**71-82.
25. **Zhang L, Foxman B.** 2003. Molecular epidemiology of Escherichia coli mediated urinary tract infections. Front Biosci **8:**e235-244.
26. **Moriel DG, Rosini R, Seib KL, Serino L, Pizza M, Rappuoli R.** 2012. Escherichia coli: great diversity around a common core. MBio **3**.
27. **Kaper JB, Nataro JP, Mobley HL.** 2004. Pathogenic Escherichia coli. Nat Rev Microbiol **2:**123-140.
28. **Toval F, Kohler CD, Vogel U, Wagenlehner F, Mellmann A, Fruth A, Schmidt MA, Karch H, Bielaszewska M, Dobrindt U.** 2014. Characterization of Escherichia coli isolates from hospital inpatients or outpatients with urinary tract infection. J Clin Microbiol **52:**407-418.
29. **Johnson JR, Russo TA.** 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) Escherichia coli. Int J Med Microbiol **295:**383-404.
30. **Bielecki P, Muthukumarasamy U, Eckweiler D, Bielecka A, Pohl S, Schanz A, Niemeyer U, Oumeraci T, von Neuhoff N, Ghigo JM, Haussler S.** 2014. In vivo mRNA profiling of uropathogenic Escherichia coli from diverse phylogroups reveals common and group-specific gene expression profiles. MBio **5:**e01075-01014.
31. **Subashchandrabose S, Mobley HL.** 2015. Virulence and Fitness Determinants of Uropathogenic Escherichia coli. Microbiol Spectr **3**.
32. **Subashchandrabose S, Hazen TH, Brumbaugh AR, Himpsl SD, Smith SN, Ernst RD, Rasko DA, Mobley HL.** 2014. Host-specific induction of Escherichia coli fitness genes during human urinary tract infection. Proc Natl Acad Sci U S A **111:**18327-18332.

33. **Spurbeck RR, Dinh PC, Jr., Walk ST, Stapleton AE, Hooton TM, Nolan LK, Kim KS, Johnson JR, Mobley HL.** 2012. Escherichia coli isolates that carry vat, fyuA, chuA, and yfcV efficiently colonize the urinary tract. Infect Immun **80:**4115-4122.
34. **Spurbeck RR, Stapleton AE, Johnson JR, Walk ST, Hooton TM, Mobley HL.** 2011. Fimbrial profiles predict virulence of uropathogenic Escherichia coli strains: contribution of Ygi and Yad Fimbriae. Infection and Immunity **79:**4753-4763.
35. **Vigil PD, Stapleton AE, Johnson JR, Hooton TM, Hodges AP, He Y, Mobley HL.** 2011. Presence of putative repeat-in-toxin gene tosA in Escherichia coli predicts successful colonization of the urinary tract. MBio **2:**e00066-00011.
36. **Welch RA, Burland V, Plunkett G, 3rd, Redford P, Roesch P, Rasko D, Buckles EL, Liou SR, Boutin A, Hackett J, Stroud D, Mayhew GF, Rose DJ, Zhou S, Schwartz DC, Perna NT, Mobley HL, Donnenberg MS, Blattner FR.** 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. Proc Natl Acad Sci U S A **99:**17020-17024.
37. **Schubert S, Dufke S, Sorsa J, Heesemann J.** 2004. A novel integrative and conjugative element (ICE) of Escherichia coli: the putative progenitor of the Yersinia high-pathogenicity island. Mol Microbiol **51:**837-848.
38. **Davison J.** 1999. Genetic exchange between bacteria in the environment. Plasmid **42:**73-91.
39. **Lloyd AL, Henderson TA, Vigil PD, Mobley HL.** 2009. Genomic islands of uropathogenic Escherichia coli contribute to virulence. Journal of Bacteriology **191:**3469-3481.
40. **Lloyd AL, Rasko DA, Mobley HL.** 2007. Defining genomic islands and uropathogen-specific genes in uropathogenic Escherichia coli. J Bacteriol **189:**3532-3546.
41. **Dorman CJ.** 2007. H-NS, the genome sentinel. Nat Rev Microbiol **5:**157-161.
42. **Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC.** 2006. Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. Science **313:**236-238.
43. **Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M, Hinton JC.** 2006. H-NS mediates the silencing of laterally acquired genes in bacteria. PLoS Pathog **2:**e81.
44. **Nielubowicz GR, Mobley HL.** 2010. Host-pathogen interactions in urinary tract infection. Nat Rev Urol **7:**430-441.
45. **Brumbaugh AR, Mobley HL.** 2012. Preventing urinary tract infection: progress toward an effective Escherichia coli vaccine. Expert Rev Vaccines **11:**663-676.
46. **Welch RA, Dellinger EP, Minshew B, Falkow S.** 1981. Haemolysin contributes to virulence of extra-intestinal E. coli infections. Nature **294:**665-667.
47. **Welch RA, Hull R, Falkow S.** 1983. Molecular cloning and physical characterization of a chromosomal hemolysin from Escherichia coli. Infect Immun **42:**178-186.

48. **Welch RA.** 1991. Pore-forming cytolysins of gram-negative bacteria. Molecular Microbiology **5:**521-528.
49. **Linhartova I, Bumba L, Masin J, Basler M, Osicka R, Kamanova J, Prochazkova K, Adkins I, Hejnova-Holubova J, Sadilkova L, Morova J, Sebo P.** 2010. RTX proteins: a highly diverse family secreted by a common mechanism. FEMS Microbiol Rev **34:**1076-1112.
50. **Satchell KJ.** 2011. Structure and function of MARTX toxins and other large repetitive RTX proteins. Annu Rev Microbiol **65:**71-90.
51. **Felmlee T, Welch RA.** 1988. Alterations of amino acid repeats in the Escherichia coli hemolysin affect cytolytic activity and secretion. Proc Natl Acad Sci U S A **85:**5269-5273.
52. **Bhakdi S, Mackman N, Nicaud JM, Holland IB.** 1986. Escherichia coli hemolysin may damage target cell membranes by generating transmembrane pores. Infection and Immunity **52:**63-69.
53. **Boehm DF, Welch RA, Snyder IS.** 1990. Calcium is required for binding of Escherichia coli hemolysin (HlyA) to erythrocyte membranes. Infection and Immunity **58:**1951-1958.
54. **Ludwig A, Jarchau T, Benz R, Goebel W.** 1988. The repeat domain of Escherichia coli haemolysin (HlyA) is responsible for its Ca2+-dependent binding to erythrocytes. Mol Gen Genet **214:**553-561.
55. **Felmlee T, Pellett S, Welch RA.** 1985. Nucleotide sequence of an Escherichia coli chromosomal hemolysin. Journal of Bacteriology **163:**94-105.
56. **Hardie KR, Issartel JP, Koronakis E, Hughes C, Koronakis V.** 1991. In vitro activation of Escherichia coli prohaemolysin to the mature membrane-targeted toxin requires HlyC and a low molecular-weight cytosolic polypeptide. Molecular Microbiology **5:**1669-1679.
57. **Hughes C, Issartel JP, Hardie K, Stanley P, Koronakis E, Koronakis V.** 1992. Activation of Escherichia coli prohemolysin to the membrane-targetted toxin by HlyC-directed ACP-dependent fatty acylation. FEMS Microbiol Immunol **5:**37-43.
58. **Issartel JP, Koronakis V, Hughes C.** 1991. Activation of Escherichia coli prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. Nature **351:**759-761.
59. **Ludwig A, Garcia F, Bauer S, Jarchau T, Benz R, Hoppe J, Goebel W.** 1996. Analysis of the in vivo activation of hemolysin (HlyA) from Escherichia coli. Journal of Bacteriology **178:**5422-5430.
60. **Stanley P, Koronakis V, Hughes C.** 1998. Acylation of Escherichia coli hemolysin: a unique protein lipidation mechanism underlying toxin function. Microbiol Mol Biol Rev **62:**309-333.
61. **Benz R, Dobereiner A, Ludwig A, Goebel W.** 1992. Haemolysin of Escherichia coli: comparison of pore-forming properties between chromosome and plasmid-encoded haemolysins. FEMS Microbiol Immunol **5:**55-62.

62. **Herlax V, Mate S, Rimoldi O, Bakas L.** 2009. Relevance of fatty acid covalently bound to Escherichia coli alpha-hemolysin and membrane microdomains in the oligomerization process. J Biol Chem **284:**25199-25210.
63. **Gadeberg OV, Orskov I.** 1984. In vitro cytotoxic effect of alpha-hemolytic Escherichia coli on human blood granulocytes. Infection and Immunity **45:**255-260.
64. **Gadeberg OV, Orskov I, Rhodes JM.** 1983. Cytotoxic effect of an alpha-hemolytic Escherichia coli strain on human blood monocytes and granulocytes in vitro. Infection and Immunity **41:**358-364.
65. **Mobley HL, Green DM, Trifillis AL, Johnson DE, Chippendale GR, Lockatell CV, Jones BD, Warren JW.** 1990. Pyelonephritogenic Escherichia coli and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. Infection and Immunity **58:**1281-1289.
66. **Smith YC, Rasmussen SB, Grande KK, Conran RM, O'Brien AD.** 2008. Hemolysin of uropathogenic Escherichia coli evokes extensive shedding of the uroepithelium and hemorrhage in bladder tissue within the first 24 hours after intraurethral inoculation of mice. Infection and Immunity **76:**2978-2990.
67. **Wiles TJ, Dhakal BK, Eto DS, Mulvey MA.** 2008. Inactivation of host Akt/protein kinase B signaling by bacterial pore-forming toxins. Mol Biol Cell **19:**1427-1438.
68. **Dhakal BK, Mulvey MA.** 2012. The UPEC pore-forming toxin alpha-hemolysin triggers proteolysis of host proteins to disrupt cell adhesion, inflammatory, and survival pathways. Cell Host Microbe **11:**58-69.
69. **Holland IB, Schmitt L, Young J.** 2005. Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review). Mol Membr Biol **22:**29-39.
70. **Koronakis V, Stanley P, Koronakis E, Hughes C.** 1992. The HlyB/HlyD-dependent secretion of toxins by gram-negative bacteria. FEMS Microbiol Immunol **5:**45-53.
71. **Wagner W, Vogel M, Goebel W.** 1983. Transport of hemolysin across the outer membrane of Escherichia coli requires two functions. Journal of Bacteriology **154:**200-210.
72. **Wandersman C, Delepelaire P.** 1990. TolC, an Escherichia coli outer membrane protein required for hemolysin secretion. Proc Natl Acad Sci U S A **87:**4776-4780.
73. **Benabdelhak H, Kiontke S, Horn C, Ernst R, Blight MA, Holland IB, Schmitt L.** 2003. A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A. J Mol Biol **327:**1169-1179.
74. **Koronakis V, Hughes C, Koronakis E.** 1993. ATPase activity and ATP/ADP-induced conformational change in the soluble domain of the bacterial protein translocator HlyB. Molecular Microbiology **8:**1163-1175.
75. **Thanabalu T, Koronakis E, Hughes C, Koronakis V.** 1998. Substrate-induced assembly of a contiguous channel for protein export from E.coli: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. EMBO J **17:**6487-6496.

76. **Gray L, Mackman N, Nicaud JM, Holland IB.** 1986. The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from Escherichia coli. Mol Gen Genet **205:**127-133.
77. **Koronakis V, Koronakis E, Hughes C.** 1989. Isolation and analysis of the C-terminal signal directing export of Escherichia coli hemolysin protein across both bacterial membranes. EMBO J **8:**595-605.
78. **Pimenta AL, Racher K, Jamieson L, Blight MA, Holland IB.** 2005. Mutations in HlyD, part of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin. Journal of Bacteriology **187:**7471-7480.
79. **Mills M, Meysick KC, O'Brien AD.** 2000. Cytotoxic necrotizing factor type 1 of uropathogenic Escherichia coli kills cultured human uroepithelial 5637 cells by an apoptotic mechanism. Infection and Immunity **68:**5869-5880.
80. **Davis JM, Carvalho HM, Rasmussen SB, O'Brien AD.** 2006. Cytotoxic necrotizing factor type 1 delivered by outer membrane vesicles of uropathogenic Escherichia coli attenuates polymorphonuclear leukocyte antimicrobial activity and chemotaxis. Infection and Immunity **74:**4401-4408.
81. **Hofman P, Le Negrate G, Mograbi B, Hofman V, Brest P, Alliana-Schmid A, Flatau G, Boquet P, Rossi B.** 2000. Escherichia coli cytotoxic necrotizing factor-1 (CNF-1) increases the adherence to epithelia and the oxidative burst of human polymorphonuclear leukocytes but decreases bacteria phagocytosis. J Leukoc Biol **68:**522-528.
82. **Guyer DM, Henderson IR, Nataro JP, Mobley HL.** 2000. Identification of sat, an autotransporter toxin produced by uropathogenic Escherichia coli. Molecular Microbiology **38:**53-66.
83. **Guyer DM, Radulovic S, Jones FE, Mobley HL.** 2002. Sat, the secreted autotransporter toxin of uropathogenic Escherichia coli, is a vacuolating cytotoxin for bladder and kidney epithelial cells. Infection and Immunity **70:**4539-4546.
84. **Heimer SR, Rasko DA, Lockatell CV, Johnson DE, Mobley HL.** 2004. Autotransporter genes pic and tsh are associated with Escherichia coli strains that cause acute pyelonephritis and are expressed during urinary tract infection. Infection and Immunity **72:**593-597.
85. **Parreira VR, Gyles CL.** 2003. A novel pathogenicity island integrated adjacent to the thrW tRNA gene of avian pathogenic Escherichia coli encodes a vacuolating autotransporter toxin. Infection and Immunity **71:**5087-5096.
86. **Dutta PR, Cappello R, Navarro-Garcia F, Nataro JP.** 2002. Functional comparison of serine protease autotransporters of enterobacteriaceae. Infection and Immunity **70:**7105-7113.
87. **Kostakioti M, Stathopoulos C.** 2004. Functional analysis of the Tsh autotransporter from an avian pathogenic Escherichia coli strain. Infection and Immunity **72:**5548-5554.
88. **Provence DL, Curtiss R, 3rd.** 1994. Isolation and characterization of a gene involved in hemagglutination by an avian pathogenic Escherichia coli strain. Infection and Immunity **62:**1369-1380.

89. **Stathopoulos C, Provence DL, Curtiss R, 3rd.** 1999. Characterization of the avian pathogenic Escherichia coli hemagglutinin Tsh, a member of the immunoglobulin A protease-type family of autotransporters. Infection and Immunity **67:**772-781.
90. **Alteri CJ, Mobley HL.** 2012. Escherichia coli physiology and metabolism dictates adaptation to diverse host microenvironments. Curr Opin Microbiol **15:**3-9.
91. **Alteri CJ, Smith SN, Mobley HL.** 2009. Fitness of Escherichia coli during urinary tract infection requires gluconeogenesis and the TCA cycle. PLoS Pathog **5:**e1000448.
92. **Stuger R, Woldringh CL, van der Weijden CC, Vischer NO, Bakker BM, van Spanning RJ, Snoep JL, Westerhoff HV.** 2002. DNA supercoiling by gyrase is linked to nucleoid compaction. Mol Biol Rep **29:**79-82.
93. **Dorman CJ.** 2014. Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. J Mol Microbiol Biotechnol **24:**316-331.
94. **Dorman CJ.** 2013. Genome architecture and global gene regulation in bacteria: making progress towards a unified model? Nat Rev Microbiol **11:**349-355.
95. **Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A.** 1999. Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. J Bacteriol **181:**6361-6370.
96. **Navarro Llorens JM, Tormo A, Martinez-Garcia E.** 2010. Stationary phase in gram-negative bacteria. FEMS Microbiol Rev **34:**476-495.
97. **Dorman CJ.** 1996. Flexible response: DNA supercoiling, transcription and bacterial adaptation to environmental stress. Trends Microbiol **4:**214-216.
98. **Kusano S, Ding Q, Fujita N, Ishihama A.** 1996. Promoter selectivity of Escherichia coli RNA polymerase E sigma 70 and E sigma 38 holoenzymes. Effect of DNA supercoiling. J Biol Chem **271:**1998-2004.
99. **Liu LF, Wang JC.** 1987. Supercoiling of the DNA template during transcription. Proc Natl Acad Sci U S A **84:**7024-7027.
100. **Drlica K.** 1992. Control of bacterial DNA supercoiling. Mol Microbiol **6:**425-433.
101. **Mizushima T, Natori S, Sekimizu K.** 1993. Relaxation of supercoiled DNA associated with induction of heat shock proteins in Escherichia coli. Mol Gen Genet **238:**1-5.
102. **Fulcrand G, Dages S, Zhi X, Chapagain P, Gerstman BS, Dunlap D, Leng F.** 2016. DNA supercoiling, a critical signal regulating the basal expression of the lac operon in Escherichia coli. Sci Rep **6:**19243.
103. **Dorman CJ, Corcoran CP.** 2009. Bacterial DNA topology and infectious disease. Nucleic Acids Res **37:**672-678.
104. **Landgraf JR, Wu J, Calvo JM.** 1996. Effects of nutrition and growth rate on Lrp levels in Escherichia coli. J Bacteriol **178:**6930-6936.
105. **Tani TH, Khodursky A, Blumenthal RM, Brown PO, Matthews RG.** 2002. Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. Proc Natl Acad Sci U S A **99:**13471-13476.

106. **Mangan MW, Lucchini S, Danino V, Croinin TO, Hinton JC, Dorman CJ.** 2006. The integration host factor (IHF) integrates stationary-phase and virulence gene expression in Salmonella enterica serovar Typhimurium. Mol Microbiol **59:**1831-1847.
107. **Prieto AI, Kahramanoglou C, Ali RM, Fraser GM, Seshasayee AS, Luscombe NM.** 2012. Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in Escherichia coli K12. Nucleic Acids Res **40:**3524-3537.
108. **Claret L, Rouviere-Yaniv J.** 1997. Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival. J Mol Biol **273:**93-104.
109. **Ishihama A, Kori A, Koshio E, Yamada K, Maeda H, Shimada T, Makinoshima H, Iwata A, Fujita N.** 2014. Intracellular concentrations of 65 species of transcription factors with known regulatory functions in Escherichia coli. J Bacteriol **196:**2718-2727.
110. **Lane MC, Lockatell V, Monterosso G, Lamphier D, Weinert J, Hebel JR, Johnson DE, Mobley HL.** 2005. Role of motility in the colonization of uropathogenic Escherichia coli in the urinary tract. Infection and Immunity **73:**7644-7656.
111. **Wright KJ, Seed PC, Hultgren SJ.** 2005. Uropathogenic Escherichia coli flagella aid in efficient urinary tract colonization. Infection and Immunity **73:**7657-7668.
112. **Lane MC, Alteri CJ, Smith SN, Mobley HL.** 2007. Expression of flagella is coincident with uropathogenic Escherichia coli ascension to the upper urinary tract. Proc Natl Acad Sci U S A **104:**16669-16674.
113. **Snyder JA, Haugen BJ, Buckles EL, Lockatell CV, Johnson DE, Donnenberg MS, Welch RA, Mobley HL.** 2004. Transcriptome of uropathogenic Escherichia coli during urinary tract infection. Infection and Immunity **72:**6373-6381.
114. **Hagan EC, Lloyd AL, Rasko DA, Faerber GJ, Mobley HL.** 2010. Escherichia coli global gene expression in urine from women with urinary tract infection. PLoS Pathog **6:**e1001187.
115. **Walters MS, Lane MC, Vigil PD, Smith SN, Walk ST, Mobley HL.** 2012. Kinetics of uropathogenic Escherichia coli metapopulation movement during urinary tract infection. MBio **3**.
116. **Chilcott GS, Hughes KT.** 2000. Coupling of flagellar gene expression to flagellar assembly in Salmonella enterica serovar typhimurium and Escherichia coli. Microbiol Mol Biol Rev **64:**694-708.
117. **Frye J, Karlinsey JE, Felise HR, Marzolf B, Dowidar N, McClelland M, Hughes KT.** 2006. Identification of new flagellar genes of Salmonella enterica serovar Typhimurium. J Bacteriol **188:**2233-2243.
118. **Chevance FF, Hughes KT.** 2008. Coordinating assembly of a bacterial macromolecular machine. Nat Rev Microbiol **6:**455-465.

119. **Kalir S, McClure J, Pabbaraju K, Southward C, Ronen M, Leibler S, Surette MG, Alon U.** 2001. Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. Science **292:**2080-2083.
120. **Baker MD, Wolanin PM, Stock JB.** 2006. Signal transduction in bacterial chemotaxis. Bioessays **28:**9-22.
121. **Parkinson JS.** 2003. Bacterial chemotaxis: a new player in response regulator dephosphorylation. J Bacteriol **185:**1492-1494.
122. **Grebe TW, Stock J.** 1998. Bacterial chemotaxis: the five sensors of a bacterium. Curr Biol **8:**R154-157.
123. **Hedblom ML, Adler J.** 1980. Genetic and biochemical properties of Escherichia coli mutants with defects in serine chemotaxis. Journal of Bacteriology **144:**1048-1060.
124. **Kondoh H, Ball CB, Adler J.** 1979. Identification of a methyl-accepting chemotaxis protein for the ribose and galactose chemoreceptors of Escherichia coli. Proc Natl Acad Sci U S A **76:**260-264.
125. **Manson MD, Blank V, Brade G, Higgins CF.** 1986. Peptide chemotaxis in E. coli involves the Tap signal transducer and the dipeptide permease. Nature **321:**253-256.
126. **Silverman M, Simon M.** 1977. Chemotaxis in Escherichia coli: methylation of che gene products. Proc Natl Acad Sci U S A **74:**3317-3321.
127. **Lane MC, Lloyd AL, Markyvech TA, Hagan EC, Mobley HL.** 2006. Uropathogenic Escherichia coli strains generally lack functional Trg and Tap chemoreceptors found in the majority of E. coli strains strictly residing in the gut. Journal of Bacteriology **188:**5618-5625.
128. **Donnelly MA, Steiner TS.** 2002. Two nonadjacent regions in enteroaggregative Escherichia coli flagellin are required for activation of toll-like receptor 5. J Biol Chem **277:**40456-40461.
129. **Stewart MK, Cookson BT.** 2012. Non-genetic diversity shapes infectious capacity and host resistance. Trends Microbiol **20:**461-466.
130. **Schirmer T, Jenal U.** 2009. Structural and mechanistic determinants of c-di-GMP signalling. Nat Rev Microbiol **7:**724-735.
131. **Hengge R.** 2009. Principles of c-di-GMP signalling in bacteria. Nat Rev Microbiol **7:**263-273.
132. **Romling U, Galperin MY, Gomelsky M.** 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev **77:**1-52.
133. **Paul K, Nieto V, Carlquist WC, Blair DF, Harshey RM.** 2010. The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. Mol Cell **38:**128-139.
134. **Spurbeck RR, Alteri CJ, Himpsl SD, Mobley HL.** 2013. The multifunctional protein YdiV represses P fimbria-mediated adherence in uropathogenic Escherichia coli. J Bacteriol **195:**3156-3164.
135. **Ryan RP, Fouhy Y, Lucey JF, Dow JM.** 2006. Cyclic di-GMP signaling in bacteria: recent advances and new puzzles. J Bacteriol **188:**8327-8334.

136. **Hengge R, Galperin MY, Ghigo JM, Gomelsky M, Green J, Hughes KT, Jenal U, Landini P.** 2015. Systematic Nomenclature for GGDEF and EAL Domain-Containing Cyclic Di-GMP Turnover Proteins of Escherichia coli. J Bacteriol **198:**7-11.
137. **Galperin MY.** 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. J Bacteriol **188:**4169-4182.
138. **Spangler C, Bohm A, Jenal U, Seifert R, Kaever V.** 2010. A liquid chromatography-coupled tandem mass spectrometry method for quantitation of cyclic di-guanosine monophosphate. J Microbiol Methods **81:**226-231.
139. **Amsler CD, Cho M, Matsumura P.** 1993. Multiple factors underlying the maximum motility of Escherichia coli as cultures enter post-exponential growth. J Bacteriol **175:**6238-6244.
140. **Beloin C, Roux A, Ghigo JM.** 2008. Escherichia coli biofilms. Curr Top Microbiol Immunol **322:**249-289.
141. **Costerton JW, Stewart PS, Greenberg EP.** 1999. Bacterial biofilms: a common cause of persistent infections. Science **284:**1318-1322.
142. **Martinez LC, Vadyvaloo V.** 2014. Mechanisms of post-transcriptional gene regulation in bacterial biofilms. Front Cell Infect Microbiol **4:**38.
143. **Waksman G, Hultgren SJ.** 2009. Structural biology of the chaperone-usher pathway of pilus biogenesis. Nat Rev Microbiol **7:**765-774.
144. **Orndorff PE, Bloch CA.** 1990. The role of type 1 pili in the pathogenesis of Escherichia coli infections: a short review and some new ideas. Microb Pathog **9:**75-79.
145. **Hull RA, Gill RE, Hsu P, Minshew BH, Falkow S.** 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection Escherichia coli isolate. Infection and Immunity **33:**933-938.
146. **Hultgren SJ, Normark S, Abraham SN.** 1991. Chaperone-assisted assembly and molecular architecture of adhesive pili. Annu Rev Microbiol **45:**383-415.
147. **Schilling JD, Mulvey MA, Hultgren SJ.** 2001. Structure and function of Escherichia coli type 1 pili: new insight into the pathogenesis of urinary tract infections. J Infect Dis **183 Suppl 1:**S36-40.
148. **Schwan WR.** 2011. Regulation of fim genes in uropathogenic Escherichia coli. World J Clin Infect Dis **1:**17-25.
149. **McClain MS, Blomfield IC, Eisenstein BI.** 1991. Roles of fimB and fimE in site-specific DNA inversion associated with phase variation of type 1 fimbriae in Escherichia coli. J Bacteriol **173:**5308-5314.
150. **Olsen PB, Klemm P.** 1994. Localization of promoters in the fim gene cluster and the effect of H-NS on the transcription of fimB and fimE. FEMS Microbiol Lett **116:**95-100.
151. **Abraham JM, Freitag CS, Clements JR, Eisenstein BI.** 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of Escherichia coli. Proc Natl Acad Sci U S A **82:**5724-5727.

152. **Calvo JM, Matthews RG.** 1994. The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. Microbiol Rev **58:**466-490.
153. **Brinton CC, Jr.** 1965. The structure, function, synthesis and genetic control of bacterial pili and a molecular model for DNA and RNA transport in gram negative bacteria. Trans N Y Acad Sci **27:**1003-1054.
154. **Choudhury D, Thompson A, Stojanoff V, Langermann S, Pinkner J, Hultgren SJ, Knight SD.** 1999. X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. Science **285:**1061-1066.
155. **Johnson JR.** 1991. Virulence factors in Escherichia coli urinary tract infection. Clin Microbiol Rev **4:**80-128.
156. **Jones CH, Pinkner JS, Nicholes AV, Slonim LN, Abraham SN, Hultgren SJ.** 1993. FimC is a periplasmic PapD-like chaperone that directs assembly of type 1 pili in bacteria. Proc Natl Acad Sci U S A **90:**8397-8401.
157. **Jones CH, Pinkner JS, Roth R, Heuser J, Nicholes AV, Abraham SN, Hultgren SJ.** 1995. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc Natl Acad Sci U S A **92:**2081-2085.
158. **Abraham SN, Beachey EH.** 1987. Assembly of a chemically synthesized peptide of Escherichia coli type 1 fimbriae into fimbria-like antigenic structures. Journal of Bacteriology **169:**2460-2465.
159. **Abraham SN, Goguen JD, Beachey EH.** 1988. Hyperadhesive mutant of type 1-fimbriated Escherichia coli associated with formation of FimH organelles (fimbriosomes). Infection and Immunity **56:**1023-1029.
160. **Hung CS, Bouckaert J, Hung D, Pinkner J, Widberg C, DeFusco A, Auguste CG, Strouse R, Langermann S, Waksman G, Hultgren SJ.** 2002. Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection. Molecular Microbiology **44:**903-915.
161. **Krogfelt KA, Bergmans H, Klemm P.** 1990. Direct evidence that the FimH protein is the mannose-specific adhesin of Escherichia coli type 1 fimbriae. Infection and Immunity **58:**1995-1998.
162. **Wu XR, Sun TT, Medina JJ.** 1996. In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: relation to urinary tract infections. Proc Natl Acad Sci U S A **93:**9630-9635.
163. **Sokurenko EV, Chesnokova V, Dykhuizen DE, Ofek I, Wu XR, Krogfelt KA, Struve C, Schembri MA, Hasty DL.** 1998. Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin. Proc Natl Acad Sci U S A **95:**8922-8926.
164. **Bahrani-Mougeot FK, Buckles EL, Lockatell CV, Hebel JR, Johnson DE, Tang CM, Donnenberg MS.** 2002. Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic Escherichia coli virulence determinants in the murine urinary tract. Molecular Microbiology **45:**1079-1093.
165. **Connell I, Agace W, Klemm P, Schembri M, Marild S, Svanborg C.** 1996. Type 1 fimbrial expression enhances Escherichia coli virulence for the urinary tract. Proc Natl Acad Sci U S A **93:**9827-9832.

166. **Gunther NWt, Snyder JA, Lockatell V, Blomfield I, Johnson DE, Mobley HL.** 2002. Assessment of virulence of uropathogenic Escherichia coli type 1 fimbrial mutants in which the invertible element is phase-locked on or off. Infection and Immunity **70:**3344-3354.
167. **Snyder JA, Lloyd AL, Lockatell CV, Johnson DE, Mobley HL.** 2006. Role of phase variation of type 1 fimbriae in a uropathogenic Escherichia coli cystitis isolate during urinary tract infection. Infection and Immunity **74:**1387-1393.
168. **Struve C, Krogfelt KA.** 1999. In vivo detection of Escherichia coli type 1 fimbrial expression and phase variation during experimental urinary tract infection. Microbiology **145 (Pt 10):**2683-2690.
169. **Gunther NWt, Lockatell V, Johnson DE, Mobley HL.** 2001. In vivo dynamics of type 1 fimbria regulation in uropathogenic Escherichia coli during experimental urinary tract infection. Infect Immun **69:**2838-2846.
170. **Kallenius G, Mollby R, Svenson SB, Helin I, Hultberg H, Cedergren B, Winberg J.** 1981. Occurrence of P-fimbriated Escherichia coli in urinary tract infections. Lancet **2:**1369-1372.
171. **Dowling KJ, Roberts JA, Kaack MB.** 1987. P-fimbriated Escherichia coli urinary tract infection: a clinical correlation. South Med J **80:**1533-1536.
172. **Hagberg L, Jodal U, Korhonen TK, Lidin-Janson G, Lindberg U, Svanborg Eden C.** 1981. Adhesion, hemagglutination, and virulence of Escherichia coli causing urinary tract infections. Infection and Immunity **31:**564-570.
173. **Jacobsen SH, Lins LE, Svenson SB, Kallenius G.** 1985. P fimbriated Escherichia coli in adults with acute pyelonephritis. J Infect Dis **152:**426-427.
174. **Johnson JR, Roberts PL, Stamm WE.** 1987. P fimbriae and other virulence factors in Escherichia coli urosepsis: association with patients' characteristics. J Infect Dis **156:**225-229.
175. **Latham RH, Stamm WE.** 1984. Role of fimbriated Escherichia coli in urinary tract infections in adult women: correlation with localization studies. J Infect Dis **149:**835-840.
176. **O'Hanley P, Low D, Romero I, Lark D, Vosti K, Falkow S, Schoolnik G.** 1985. Gal-Gal binding and hemolysin phenotypes and genotypes associated with uropathogenic Escherichia coli. N Engl J Med **313:**414-420.
177. **Vaisanen V, Elo J, Tallgren LG, Siitonen A, Makela PH, Svanborg-Eden C, Kallenius G, Svenson SB, Hultberg H, Korhonen T.** 1981. Mannose-resistant haemagglutination and P antigen recognition are characteristic of Escherichia coli causing primary pyelonephritis. Lancet **2:**1366-1369.
178. **Hung CS, Dodson KW, Hultgren SJ.** 2009. A murine model of urinary tract infection. Nat Protoc **4:**1230-1243.
179. **Hernday A, Krabbe M, Braaten B, Low D.** 2002. Self-perpetuating epigenetic pili switches in bacteria. Proc Natl Acad Sci U S A **99 Suppl 4:**16470-16476.
180. **Forsman K, Goransson M, Uhlin BE.** 1989. Autoregulation and multiple DNA interactions by a transcriptional regulatory protein in E. coli pili biogenesis. EMBO J **8:**1271-1277.

181. **Forsman K, Sonden B, Goransson M, Uhlin BE.** 1992. Antirepression function in Escherichia coli for the cAMP-cAMP receptor protein transcriptional activator. Proc Natl Acad Sci U S A **89:**9880-9884.
182. **Xia Y, Forsman K, Jass J, Uhlin BE.** 1998. Oligomeric interaction of the PapB transcriptional regulator with the upstream activating region of pili adhesin gene promoters in Escherichia coli. Mol Microbiol **30:**513-523.
183. **Braaten BA, Platko JV, van der Woude MW, Simons BH, de Graaf FK, Calvo JM, Low DA.** 1992. Leucine-responsive regulatory protein controls the expression of both the pap and fan pili operons in Escherichia coli. Proc Natl Acad Sci U S A **89:**4250-4254.
184. **van der Woude M, Hale WB, Low DA.** 1998. Formation of DNA methylation patterns: nonmethylated GATC sequences in gut and pap operons. J Bacteriol **180:**5913-5920.
185. **Kuehn MJ, Heuser J, Normark S, Hultgren SJ.** 1992. P pili in uropathogenic E. coli are composite fibres with distinct fibrillar adhesive tips. Nature **356:**252-255.
186. **Li X, Rasko DA, Lockatell CV, Johnson DE, Mobley HL.** 2001. Repression of bacterial motility by a novel fimbrial gene product. EMBO J **20:**4854-4862.
187. **Reiss DJ, Mobley HL.** 2011. Determination of the target sequence bound by PapX, a repressor of bacterial motility, in the flhD promoter using SELEX and high-throughput sequencing. J Biol Chem doi:10.1074/jbc.M111.290684.
188. **Simms AN, Mobley HL.** 2008. PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic Escherichia coli. Infect Immun **76:**4833-4841.
189. **Korhonen TK, Vaisanen V, Saxen H, Hultberg H, Svenson SB.** 1982. P-antigen-recognizing fimbriae from human uropathogenic Escherichia coli strains. Infection and Immunity **37:**286-291.
190. **Korhonen TK, Virkola R, Holthofer H.** 1986. Localization of binding sites for purified Escherichia coli P fimbriae in the human kidney. Infection and Immunity **54:**328-332.
191. **Mobley HL, Jarvis KG, Elwood JP, Whittle DI, Lockatell CV, Russell RG, Johnson DE, Donnenberg MS, Warren JW.** 1993. Isogenic P-fimbrial deletion mutants of pyelonephritogenic Escherichia coli: the role of alpha Gal(1-4) beta Gal binding in virulence of a wild-type strain. Molecular Microbiology **10:**143-155.
192. **Buckles EL, Luterbach CL, Wang X, Lockatell CV, Johnson DE, Mobley HL, Donnenberg MS.** 2015. Signature-tagged mutagenesis and co-infection studies demonstrate the importance of P fimbriae in a murine model of urinary tract infection. Pathog Dis **73**.
193. **Melican K, Sandoval RM, Kader A, Josefsson L, Tanner GA, Molitoris BA, Richter-Dahlfors A.** 2011. Uropathogenic Escherichia coli P and Type 1 fimbriae act in synergy in a living host to facilitate renal colonization leading to nephron obstruction. PLoS Pathog **7:**e1001298.
194. **Kisielius PV, Schwan WR, Amundsen SK, Duncan JL, Schaeffer AJ.** 1989. In vivo expression and variation of Escherichia coli type 1 and P pili in the urine of adults with acute urinary tract infections. Infection and Immunity **57:**1656-1662.

195. **de Ree JM, van den Bosch JF.** 1987. Serological response to the P fimbriae of uropathogenic Escherichia coli in pyelonephritis. Infection and Immunity **55:**2204-2207.
196. **Donnenberg MS, Welch RA.** 1996. Virulence determinants of uropathogenic *Escherichia coli*. *In* Mobley HL, Warren JW (ed), Urinary Tract Infections: Molecular Pathogenesis and Clinical Management. ASM Press, Washington DC.
197. **Pere A, Leinonen M, Vaisanen-Rhen V, Rhen M, Korhonen TK.** 1985. Occurrence of type-1C fimbriae on Escherichia coli strains isolated from human extraintestinal infections. J Gen Microbiol **131:**1705-1711.
198. **Zingler G, Blum G, Falkenhagen U, Orskov I, Orskov F, Hacker J, Ott M.** 1993. Clonal differentiation of uropathogenic Escherichia coli isolates of serotype O6:K5 by fimbrial antigen typing and DNA long-range mapping techniques. Med Microbiol Immunol **182:**13-24.
199. **Zingler G, Ott M, Blum G, Falkenhagen U, Naumann G, Sokolowska-Kohler W, Hacker J.** 1992. Clonal analysis of Escherichia coli serotype O6 strains from urinary tract infections. Microb Pathog **12:**299-310.
200. **Riegman N, Kusters R, Van Veggel H, Bergmans H, Van Bergen en Henegouwen P, Hacker J, Van Die I.** 1990. F1C fimbriae of a uropathogenic Escherichia coli strain: genetic and functional organization of the foc gene cluster and identification of minor subunits. J Bacteriol **172:**1114-1120.
201. **Sjostrom AE, Sonden B, Muller C, Rydstrom A, Dobrindt U, Wai SN, Uhlin BE.** 2009. Analysis of the sfaX(II) locus in the Escherichia coli meningitis isolate IHE3034 reveals two novel regulatory genes within the promoter-distal region of the main S fimbrial operon. Microb Pathog **46:**150-158.
202. **Hultdin UW, Lindberg S, Grundstrom C, Huang S, Uhlin BE, Sauer-Eriksson AE.** 2010. Structure of FocB--a member of a family of transcription factors regulating fimbrial adhesin expression in uropathogenic Escherichia coli. FEBS J **277:**3368-3381.
203. **Lindberg S, Xia Y, Sonden B, Goransson M, Hacker J, Uhlin BE.** 2008. Regulatory Interactions among adhesin gene systems of uropathogenic Escherichia coli. Infect Immun **76:**771-780.
204. **van der Woude MW, Low DA.** 1994. Leucine-responsive regulatory protein and deoxyadenosine methylase control the phase variation and expression of the sfa and daa pili operons in Escherichia coli. Mol Microbiol **11:**605-618.
205. **Ott M, Hoschutzky H, Jann K, Van Die I, Hacker J.** 1988. Gene clusters for S fimbrial adhesin (sfa) and F1C fimbriae (foc) of Escherichia coli: comparative aspects of structure and function. J Bacteriol **170:**3983-3990.
206. **Khan AS, Kniep B, Oelschlaeger TA, Van Die I, Korhonen T, Hacker J.** 2000. Receptor structure for F1C fimbriae of uropathogenic Escherichia coli. Infect Immun **68:**3541-3547.
207. **Korhonen TK, Virkola R, Westurlund B, Holthofer H, Parkkinen J.** 1990. Tissue tropism of Escherichia coli adhesins in human extraintestinal infections. Curr Top Microbiol Immunol **151:**115-127.

208. **Vigil PD, Wiles TJ, Engstrom MD, Prasov L, Mulvey MA, Mobley HL.** 2012. The repeat-in-toxin family member TosA mediates adherence of uropathogenic Escherichia coli and survival during bacteremia. Infect Immun **80:**493-505.
209. **Guo S, Garnham CP, Karunan Partha S, Campbell RL, Allingham JS, Davies PL.** 2013. Role of Ca(2)(+) in folding the tandem beta-sandwich extender domains of a bacterial ice-binding adhesin. FEBS J **280:**5919-5932.
210. **Guo S, Garnham CP, Whitney JC, Graham LA, Davies PL.** 2012. Re-evaluation of a bacterial antifreeze protein as an adhesin with ice-binding activity. PLoS One **7:**e48805.
211. **Hinsa SM, Espinosa-Urgel M, Ramos JL, O'Toole GA.** 2003. Transition from reversible to irreversible attachment during biofilm formation by Pseudomonas fluorescens WCS365 requires an ABC transporter and a large secreted protein. Mol Microbiol **49:**905-918.
212. **Fuqua C.** 2010. Passing the baton between laps: adhesion and cohesion in Pseudomonas putida biofilms. Mol Microbiol **77:**533-536.
213. **Martinez-Gil M, Yousef-Coronado F, Espinosa-Urgel M.** 2010. LapF, the second largest Pseudomonas putida protein, contributes to plant root colonization and determines biofilm architecture. Mol Microbiol **77:**549-561.
214. **Moor H, Teppo A, Lahesaare A, Kivisaar M, Teras R.** 2014. Fis overexpression enhances Pseudomonas putida biofilm formation by regulating the ratio of LapA and LapF. Microbiology **160:**2681-2693.
215. **Latasa C, Roux A, Toledo-Arana A, Ghigo JM, Gamazo C, Penades JR, Lasa I.** 2005. BapA, a large secreted protein required for biofilm formation and host colonization of Salmonella enterica serovar Enteritidis. Mol Microbiol **58:**1322-1339.
216. **Griessl MH, Schmid B, Kassler K, Braunsmann C, Ritter R, Barlag B, Stierhof YD, Sturm KU, Danzer C, Wagner C, Schaffer TE, Sticht H, Hensel M, Muller YA.** 2013. Structural insight into the giant Ca(2)(+)-binding adhesin SiiE: implications for the adhesion of Salmonella enterica to polarized epithelial cells. Structure **21:**741-752.
217. **Gerlach RG, Jackel D, Stecher B, Wagner C, Lupas A, Hardt WD, Hensel M.** 2007. Salmonella Pathogenicity Island 4 encodes a giant non-fimbrial adhesin and the cognate type 1 secretion system. Cell Microbiol **9:**1834-1850.
218. **Morgan E, Bowen AJ, Carnell SC, Wallis TS, Stevens MP.** 2007. SiiE is secreted by the Salmonella enterica serovar Typhimurium pathogenicity island 4-encoded secretion system and contributes to intestinal colonization in cattle. Infect Immun **75:**1524-1533.
219. **Chatterjee R, Nag S, Chaudhuri K.** 2008. Identification of a new RTX-like gene cluster in Vibrio cholerae. FEMS Microbiol Lett **284:**165-171.
220. **Harris TJ, Tepass U.** 2010. Adherens junctions: from molecules to morphogenesis. Nat Rev Mol Cell Biol **11:**502-514.
221. **Syed KA, Beyhan S, Correa N, Queen J, Liu J, Peng F, Satchell KJ, Yildiz F, Klose KE.** 2009. The Vibrio cholerae flagellar regulatory hierarchy controls expression of virulence factors. J Bacteriol **191:**6555-6570.

222. **Zhou G, Yuan J, Gao H.** 2015. Regulation of biofilm formation by BpfA, BpfD, and BpfG in Shewanella oneidensis. Front Microbiol **6:**790.
223. **D'Auria G, Jimenez N, Peris-Bondia F, Pelaz C, Latorre A, Moya A.** 2008. Virulence factor rtx in Legionella pneumophila, evidence suggesting it is a modular multifunctional protein. BMC Genomics **9:**14.
224. **Cirillo SL, Bermudez LE, El-Etr SH, Duhamel GE, Cirillo JD.** 2001. Legionella pneumophila entry gene rtxA is involved in virulence. Infect Immun **69:**508-517.
225. **Cirillo SL, Lum J, Cirillo JD.** 2000. Identification of novel loci involved in entry by Legionella pneumophila. Microbiology **146 (Pt 6):**1345-1359.
226. **Cirillo SL, Yan L, Littman M, Samrakandi MM, Cirillo JD.** 2002. Role of the Legionella pneumophila rtxA gene in amoebae. Microbiology **148:**1667-1677.
227. **Vigil PD, Alteri CJ, Mobley HL.** 2011. Identification of in vivo-induced antigens including an RTX family exoprotein required for uropathogenic Escherichia coli virulence. Infect Immun **79:**2335-2344.
228. **Rollins SM, Peppercorn A, Hang L, Hillman JD, Calderwood SB, Handfield M, Ryan ET.** 2005. In vivo induced antigen technology (IVIAT). Cell Microbiol **7:**1-9.
229. **Hobley L, Harkins C, MacPhee CE, Stanley-Wall NR.** 2015. Giving structure to the biofilm matrix: an overview of individual strategies and emerging common themes. FEMS Microbiol Rev **39:**649-669.
230. **Romling U.** 2005. Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae. Cell Mol Life Sci **62:**1234-1246.
231. **Hufnagel DA, DePas WH, Chapman MR.** 2014. The disulfide bonding system suppresses CsgD-independent cellulose production in Escherichia coli. J Bacteriol **196:**3690-3699.
232. **Zogaj X, Nimtz M, Rohde M, Bokranz W, Romling U.** 2001. The multicellular morphotypes of Salmonella typhimurium and Escherichia coli produce cellulose as the second component of the extracellular matrix. Mol Microbiol **39:**1452-1463.
233. **Romling U, Sierralta WD, Eriksson K, Normark S.** 1998. Multicellular and aggregative behaviour of Salmonella typhimurium strains is controlled by mutations in the agfD promoter. Mol Microbiol **28:**249-264.
234. **DePas WH, Hufnagel DA, Lee JS, Blanco LP, Bernstein HC, Fisher ST, James GA, Stewart PS, Chapman MR.** 2013. Iron induces bimodal population development by Escherichia coli. Proc Natl Acad Sci U S A **110:**2629-2634.
235. **Barnhart MM, Chapman MR.** 2006. Curli biogenesis and function. Annu Rev Microbiol **60:**131-147.
236. **Chen AY, Deng Z, Billings AN, Seker UO, Lu MY, Citorik RJ, Zakeri B, Lu TK.** 2014. Synthesis and patterning of tunable multiscale materials with engineered cells. Nat Mater **13:**515-523.
237. **Whitchurch CB, Tolker-Nielsen T, Ragas PC, Mattick JS.** 2002. Extracellular DNA required for bacterial biofilm formation. Science **295:**1487.

238. **Spurbeck RR, Tarrien RJ, Mobley HL.** 2012. Enzymatically active and inactive phosphodiesterases and diguanylate cyclases are involved in regulation of Motility or sessility in Escherichia coli CFT073. MBio **3**.
239. **Bode NJ, Debnath I, Kuan L, Schulfer A, Ty M, Pearson MM.** 2015. Transcriptional analysis of the MrpJ network: modulation of diverse virulence-associated genes and direct regulation of mrp fimbrial and flhDC flagellar operons in Proteus mirabilis. Infect Immun **83:**2542-2556.
240. **Pearson MM, Mobley HL.** 2008. Repression of motility during fimbrial expression: identification of 14 mrpJ gene paralogues in Proteus mirabilis. Molecular Microbiology **69:**548-558.
241. **Pesavento C, Becker G, Sommerfeldt N, Possling A, Tschowri N, Mehlis A, Hengge R.** 2008. Inverse regulatory coordination of motility and curli-mediated adhesion in Escherichia coli. Genes Dev **22:**2434-2446.
242. **Lopez HM, Gachelin J, Douarche C, Auradou H, Clement E.** 2015. Turning Bacteria Suspensions into Superfluids. Phys Rev Lett **115:**028301.
243. **Nickerson CA, Ott CM, Wilson JW, Ramamurthy R, Pierson DL.** 2004. Microbial responses to microgravity and other low-shear environments. Microbiol Mol Biol Rev **68:**345-361.
244. **Alsharif G, Ahmad S, Islam MS, Shah R, Busby SJ, Krachler AM.** 2015. Host attachment and fluid shear are integrated into a mechanical signal regulating virulence in Escherichia coli O157:H7. Proc Natl Acad Sci U S A **112:**5503-5508.
245. **Thomas WE, Trintchina E, Forero M, Vogel V, Sokurenko EV.** 2002. Bacterial adhesion to target cells enhanced by shear force. Cell **109:**913-923.
246. **Hospenthal MK, Redzej A, Dodson K, Ukleja M, Frenz B, Rodrigues C, Hultgren SJ, DiMaio F, Egelman EH, Waksman G.** 2016. Structure of a Chaperone-Usher Pilus Reveals the Molecular Basis of Rod Uncoiling. Cell **164:**269-278.
247. **Xia Y, Gally D, Forsman-Semb K, Uhlin BE.** 2000. Regulatory cross-talk between adhesin operons in Escherichia coli: inhibition of type 1 fimbriae expression by the PapB protein. EMBO J **19:**1450-1457.
248. **Holden NJ, Uhlin BE, Gally DL.** 2001. PapB paralogues and their effect on the phase variation of type 1 fimbriae in Escherichia coli. Mol Microbiol **42:**319-330.
249. **Morschhauser J, Vetter V, Emody L, Hacker J.** 1994. Adhesin regulatory genes within large, unstable DNA regions of pathogenic Escherichia coli: cross-talk between different adhesin gene clusters. Mol Microbiol **11:**555-566.
250. **Casadesus J, Low D.** 2006. Epigenetic gene regulation in the bacterial world. Microbiol Mol Biol Rev **70:**830-856.
251. **Totsika M, Beatson SA, Holden N, Gally DL.** 2008. Regulatory interplay between pap operons in uropathogenic Escherichia coli. Mol Microbiol **67:**996-1011.
252. **Holden NJ, Gally DL.** 2004. Switches, cross-talk and memory in Escherichia coli adherence. J Med Microbiol **53:**585-593.

253. **Volkmer B, Heinemann M.** 2011. Condition-dependent cell volume and concentration of Escherichia coli to facilitate data conversion for systems biology modeling. PLoS One **6:**e23126.
254. **Lu M, Campbell JL, Boye E, Kleckner N.** 1994. SeqA: a negative modulator of replication initiation in E. coli. Cell **77:**413-426.
255. **Skarstad K, Boye E, Steen HB.** 1986. Timing of initiation of chromosome replication in individual Escherichia coli cells. EMBO J **5:**1711-1717.
256. **Cooper S, Helmstetter CE.** 1968. Chromosome replication and the division cycle of Escherichia coli B/r. J Mol Biol **31:**519-540.
257. **Pul Ü, Wagner R.** 2010. Nucleoid-Associated Proteins: Structural Properties, p 149-173. *In* Dame R, Dorman C (ed), Bacterial Chromatin doi:10.1007/978-90-481-3473-1_8. Springer Netherlands.
258. **Lang B, Blot N, Bouffartigues E, Buckle M, Geertz M, Gualerzi CO, Mavathur R, Muskhelishvili G, Pon CL, Rimsky S, Stella S, Babu MM, Travers A.** 2007. High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. Nucleic Acids Res **35:**6330-6337.
259. **Postow L, Hardy CD, Arsuaga J, Cozzarelli NR.** 2004. Topological domain structure of the Escherichia coli chromosome. Genes Dev **18:**1766-1779.
260. **Wang X, Montero Llopis P, Rudner DZ.** 2013. Organization and segregation of bacterial chromosomes. Nat Rev Genet **14:**191-203.
261. **Delius H, Worcel A.** 1974. Letter: Electron microscopic visualization of the folded chromosome of Escherichia coli. J Mol Biol **82:**107-109.
262. **Kavenoff R, Ryder OA.** 1976. Electron microscopy of membrane-associated folded chromosomes of Escherichia coli. Chromosoma **55:**13-25.
263. **Woldringh CL, Nanninga N.** 2006. Structural and physical aspects of bacterial chromosome segregation. J Struct Biol **156:**273-283.
264. **Woldringh C.** 2010. Nucleoid Structure and Segregation, p 71-96. *In* Dame R, Dorman C (ed), Bacterial Chromatin doi:10.1007/978-90-481-3473-1_5. Springer Netherlands.
265. **Umbarger MA, Toro E, Wright MA, Porreca GJ, Bau D, Hong SH, Fero MJ, Zhu LJ, Marti-Renom MA, McAdams HH, Shapiro L, Dekker J, Church GM.** 2011. The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. Mol Cell **44:**252-264.
266. **Hadizadeh Yazdi N, Guet CC, Johnson RC, Marko JF.** 2012. Variation of the folding and dynamics of the Escherichia coli chromosome with growth conditions. Mol Microbiol **86:**1318-1333.
267. **Berlatzky IA, Rouvinski A, Ben-Yehuda S.** 2008. Spatial organization of a replicating bacterial chromosome. Proc Natl Acad Sci U S A **105:**14136-14140.
268. **Boccard F, Esnault E, Valens M.** 2005. Spatial arrangement and macrodomain organization of bacterial chromosomes. Mol Microbiol **57:**9-16.
269. **Valens M, Penaud S, Rossignol M, Cornet F, Boccard F.** 2004. Macrodomain organization of the Escherichia coli chromosome. EMBO J **23:**4330-4341.

270. **Espeli O, Mercier R, Boccard F.** 2008. DNA dynamics vary according to macrodomain topography in the E. coli chromosome. Mol Microbiol **68:**1418-1427.
271. **Dame RT, Kalmykowa OJ, Grainger DC.** 2011. Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in gram negative bacteria. PLoS Genet **7:**e1002123.
272. **Worcel A, Burgi E.** 1972. On the structure of the folded chromosome of Escherichia coli. J Mol Biol **71:**127-147.
273. **Baker TA, Sekimizu K, Funnell BE, Kornberg A.** 1986. Extensive unwinding of the plasmid template during staged enzymatic initiation of DNA replication from the origin of the Escherichia coli chromosome. Cell **45:**53-64.
274. **Peter BJ, Ullsperger C, Hiasa H, Marians KJ, Cozzarelli NR.** 1998. The structure of supercoiled intermediates in DNA replication. Cell **94:**819-827.
275. **Wigley DB, Davies GJ, Dodson EJ, Maxwell A, Dodson G.** 1991. Crystal structure of an N-terminal fragment of the DNA gyrase B protein. Nature **351:**624-629.
276. **Vivero A, Banos RC, Mariscotti JF, Oliveros JC, Garcia-del Portillo F, Juarez A, Madrid C.** 2008. Modulation of horizontally acquired genes by the Hha-YdgT proteins in Salmonella enterica serovar Typhimurium. J Bacteriol **190:**1152-1156.
277. **Rice PA, Yang S, Mizuuchi K, Nash HA.** 1996. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell **87:**1295-1306.
278. **Khrapunov S, Brenowitz M, Rice PA, Catalano CE.** 2006. Binding then bending: a mechanism for wrapping DNA. Proc Natl Acad Sci U S A **103:**19217-19218.
279. **Blomfield IC, Kulasekara DH, Eisenstein BI.** 1997. Integration host factor stimulates both FimB- and FimE-mediated site-specific DNA inversion that controls phase variation of type 1 fimbriae expression in Escherichia coli. Mol Microbiol **23:**705-717.
280. **Rice PA.** 1997. Making DNA do a U-turn: IHF and related proteins. Curr Opin Struct Biol **7:**86-93.
281. **Hoover TR, Santero E, Porter S, Kustu S.** 1990. The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. Cell **63:**11-22.
282. **Dorman CJ, Higgins CF.** 1987. Fimbrial phase variation in Escherichia coli: dependence on integration host factor and homologies with other site-specific recombinases. J Bacteriol **169:**3840-3843.
283. **Eisenstein BI, Sweet DS, Vaughn V, Friedman DI.** 1987. Integration host factor is required for the DNA inversion that controls phase variation in Escherichia coli. Proc Natl Acad Sci U S A **84:**6506-6510.
284. **Dworkin J, Jovanovic G, Model P.** 1997. Role of upstream activation sequences and integration host factor in transcriptional activation by the constitutively active prokaryotic enhancer-binding protein PspF. J Mol Biol **273:**377-388.

285. **Wigneshweraraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K, Buck M.** 2008. Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. Mol Microbiol **68:**538-546.
286. **Goodrich JA, Schwartz ML, McClure WR.** 1990. Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for Escherichia coli integration host factor (IHF). Nucleic Acids Res **18:**4993-5000.
287. **Nystrom T.** 1995. Glucose starvation stimulon of Escherichia coli: role of integration host factor in starvation survival and growth phase-dependent protein synthesis. J Bacteriol **177:**5707-5710.
288. **Swinger KK, Lemberg KM, Zhang Y, Rice PA.** 2003. Flexible DNA bending in HU-DNA cocrystal structures. EMBO J **22:**3749-3760.
289. **van Noort J, Verbrugge S, Goosen N, Dekker C, Dame RT.** 2004. Dual architectural roles of HU: formation of flexible hinges and rigid filaments. Proc Natl Acad Sci U S A **101:**6969-6974.
290. **Perez-Martin J, de Lorenzo V.** 1995. The sigma 54-dependent promoter Ps of the TOL plasmid of Pseudomonas putida requires HU for transcriptional activation in vivo by XylR. J Bacteriol **177:**3758-3763.
291. **Carmona M, Magasanik B.** 1996. Activation of transcription at sigma 54-dependent promoters on linear templates requires intrinsic or induced bending of the DNA. J Mol Biol **261:**348-356.
292. **Pontiggia A, Negri A, Beltrame M, Bianchi ME.** 1993. Protein HU binds specifically to kinked DNA. Mol Microbiol **7:**343-350.
293. **Tanaka H, Goshima N, Kohno K, Kano Y, Imamoto F.** 1993. Properties of DNA-binding of HU heterotypic and homotypic dimers from Escherichia coli. J Biochem **113:**568-572.
294. **Castaing B, Zelwer C, Laval J, Boiteux S.** 1995. HU protein of Escherichia coli binds specifically to DNA that contains single-strand breaks or gaps. J Biol Chem **270:**10291-10296.
295. **Hengen PN, Bartram SL, Stewart LE, Schneider TD.** 1997. Information analysis of Fis binding sites. Nucleic Acids Res **25:**4994-5002.
296. **Cho BK, Knight EM, Barrett CL, Palsson BO.** 2008. Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. Genome Res **18:**900-910.
297. **Pan CQ, Finkel SE, Cramton SE, Feng JA, Sigman DS, Johnson RC.** 1996. Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. J Mol Biol **264:**675-695.
298. **Bradley MD, Beach MB, de Koning AP, Pratt TS, Osuna R.** 2007. Effects of Fis on Escherichia coli gene expression during different growth stages. Microbiology **153:**2922-2940.
299. **Nilsson L, Emilsson V.** 1994. Factor for inversion stimulation-dependent growth rate regulation of individual tRNA species in Escherichia coli. J Biol Chem **269:**9460-9465.

300. **Ross W, Thompson JF, Newlands JT, Gourse RL.** 1990. E.coli Fis protein activates ribosomal RNA transcription in vitro and in vivo. EMBO J **9:**3733-3742.
301. **Weinstein-Fischer D, Altuvia S.** 2007. Differential regulation of Escherichia coli topoisomerase I by Fis. Mol Microbiol **63:**1131-1144.
302. **Schneider R, Travers A, Muskhelishvili G.** 2000. The expression of the Escherichia coli fis gene is strongly dependent on the superhelical density of DNA. Mol Microbiol **38:**167-175.
303. **Schneider R, Lurz R, Luder G, Tolksdorf C, Travers A, Muskhelishvili G.** 2001. An architectural role of the Escherichia coli chromatin protein FIS in organising DNA. Nucleic Acids Res **29:**5107-5114.
304. **Falconi M, Brandi A, La Teana A, Gualerzi CO, Pon CL.** 1996. Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of hns expression. Mol Microbiol **19:**965-975.
305. **Hardy CD, Cozzarelli NR.** 2005. A genetic selection for supercoiling mutants of Escherichia coli reveals proteins implicated in chromosome structure. Mol Microbiol **57:**1636-1652.
306. **Navarre WW, McClelland M, Libby SJ, Fang FC.** 2007. Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. Genes Dev **21:**1456-1471.
307. **Navarre W.** 2010. H-NS as a Defence System, p 251-322. *In* Dame R, Dorman C (ed), Bacterial Chromatin doi:10.1007/978-90-481-3473-1_13. Springer Netherlands.
308. **Tupper AE, Owen-Hughes TA, Ussery DW, Santos DS, Ferguson DJ, Sidebotham JM, Hinton JC, Higgins CF.** 1994. The chromatin-associated protein H-NS alters DNA topology in vitro. EMBO J **13:**258-268.
309. **Rimsky S, Zuber F, Buckle M, Buc H.** 2001. A molecular mechanism for the repression of transcription by the H-NS protein. Mol Microbiol **42:**1311-1323.
310. **Lim CJ, Lee SY, Kenney LJ, Yan J.** 2012. Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing. Sci Rep **2:**509.
311. **Owen-Hughes TA, Pavitt GD, Santos DS, Sidebotham JM, Hulton CS, Hinton JC, Higgins CF.** 1992. The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. Cell **71:**255-265.
312. **Prosseda G, Fradiani PA, Di Lorenzo M, Falconi M, Micheli G, Casalino M, Nicoletti M, Colonna B.** 1998. A role for H-NS in the regulation of the virF gene of Shigella and enteroinvasive Escherichia coli. Res Microbiol **149:**15-25.
313. **Carrera P, Azorin F.** 1994. Structural characterization of intrinsically curved AT-rich DNA sequences. Nucleic Acids Res **22:**3671-3680.
314. **Arold ST, Leonard PG, Parkinson GN, Ladbury JE.** 2010. H-NS forms a superhelical protein scaffold for DNA condensation. Proc Natl Acad Sci U S A **107:**15728-15732.
315. **Stoebel DM, Free A, Dorman CJ.** 2008. Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. Microbiology **154:**2533-2545.

316. **Lim CJ, Kenney LJ, Yan J.** 2014. Single-molecule studies on the mechanical interplay between DNA supercoiling and H-NS DNA architectural properties. Nucleic Acids Res **42:**8369-8378.
317. **Grainger DC, Goldberg MD, Lee DJ, Busby SJ.** 2008. Selective repression by Fis and H-NS at the Escherichia coli dps promoter. Mol Microbiol **68:**1366-1377.
318. **Jordi BJ, Higgins CF.** 2000. The downstream regulatory element of the proU operon of Salmonella typhimurium inhibits open complex formation by RNA polymerase at a distance. J Biol Chem **275:**12123-12128.
319. **Nagarajavel V, Madhusudan S, Dole S, Rahmouni AR, Schnetz K.** 2007. Repression by binding of H-NS within the transcription unit. J Biol Chem **282:**23622-23630.
320. **Schroder O, Wagner R.** 2000. The bacterial DNA-binding protein H-NS represses ribosomal RNA transcription by trapping RNA polymerase in the initiation complex. J Mol Biol **298:**737-748.
321. **Dame RT, Wyman C, Wurm R, Wagner R, Goosen N.** 2002. Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the rrnB P1. J Biol Chem **277:**2146-2150.
322. **Landick R, Wade JT, Grainger DC.** 2015. H-NS and RNA polymerase: a love-hate relationship? Curr Opin Microbiol **24:**53-59.
323. **Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM.** 1998. Force and velocity measured for single molecules of RNA polymerase. Science **282:**902-907.
324. **Dame RT, Noom MC, Wuite GJ.** 2006. Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. Nature **444:**387-390.
325. **Shin M, Lagda AC, Lee JW, Bhat A, Rhee JH, Kim JS, Takeyasu K, Choy HE.** 2012. Gene silencing by H-NS from distal DNA site. Mol Microbiol **86:**707-719.
326. **Olsen A, Arnqvist A, Hammar M, Sukupolvi S, Normark S.** 1993. The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in Escherichia coli. Mol Microbiol **7:**523-536.
327. **Shin M, Song M, Rhee JH, Hong Y, Kim YJ, Seok YJ, Ha KS, Jung SH, Choy HE.** 2005. DNA looping-mediated repression by histone-like protein H-NS: specific requirement of Esigma70 as a cofactor for looping. Genes Dev **19:**2388-2398.
328. **Higgins CF, Hinton JC, Hulton CS, Owen-Hughes T, Pavitt GD, Seirafi A.** 1990. Protein H1: a role for chromatin structure in the regulation of bacterial gene expression and virulence? Mol Microbiol **4:**2007-2012.
329. **Mojica FJ, Higgins CF.** 1997. In vivo supercoiling of plasmid and chromosomal DNA in an Escherichia coli hns mutant. J Bacteriol **179:**3528-3533.
330. **Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, May G, Bremer E.** 1988. A physiological role for DNA supercoiling in the osmotic regulation of gene expression in S. typhimurium and E. coli. Cell **52:**569-584.

331. **Kotlajich MV, Hron DR, Boudreau BA, Sun Z, Lyubchenko YL, Landick R.** 2015. Bridged filaments of histone-like nucleoid structuring protein pause RNA polymerase and aid termination in bacteria. Elife **4**.
332. **Dole S, Nagarajavel V, Schnetz K.** 2004. The histone-like nucleoid structuring protein H-NS represses the Escherichia coli bgl operon downstream of the promoter. Mol Microbiol **52:**589-600.
333. **Atlung T, Ingmer H.** 1997. H-NS: a modulator of environmentally regulated gene expression. Mol Microbiol **24:**7-17.
334. **Dame RT, Goosen N.** 2002. HU: promoting or counteracting DNA compaction? FEBS Lett **529:**151-156.
335. **Porter ME, Dorman CJ.** 1997. Positive regulation of Shigella flexneri virulence genes by integration host factor. J Bacteriol **179:**6537-6550.
336. **Turner EC, Dorman CJ.** 2007. H-NS antagonism in Shigella flexneri by VirB, a virulence gene transcription regulator that is closely related to plasmid partition factors. J Bacteriol **189:**3403-3413.
337. **Wyborn NR, Stapleton MR, Norte VA, Roberts RE, Grafton J, Green J.** 2004. Regulation of Escherichia coli hemolysin E expression by H-NS and Salmonella SlyA. J Bacteriol **186:**1620-1628.
338. **Chen CC, Wu HY.** 2005. LeuO protein delimits the transcriptionally active and repressive domains on the bacterial chromosome. J Biol Chem **280:**15111-15121.
339. **Walthers D, Carroll RK, Navarre WW, Libby SJ, Fang FC, Kenney LJ.** 2007. The response regulator SsrB activates expression of diverse Salmonella pathogenicity island 2 promoters and counters silencing by the nucleoid-associated protein H-NS. Mol Microbiol **65:**477-493.
340. **Diekmann S.** 1987. Temperature and salt dependence of the gel migration anomaly of curved DNA fragments. Nucleic Acids Res **15:**247-265.
341. **Driessen RP, Sitters G, Laurens N, Moolenaar GF, Wuite GJ, Goosen N, Dame RT.** 2014. Effect of temperature on the intrinsic flexibility of DNA and its interaction with architectural proteins. Biochemistry **53:**6430-6438.
342. **Wang X, Lim HJ, Son A.** 2014. Characterization of denaturation and renaturation of DNA for DNA hybridization. Environ Health Toxicol **29:**e2014007.
343. **Ono S, Goldberg MD, Olsson T, Esposito D, Hinton JC, Ladbury JE.** 2005. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochem J **391:**203-213.
344. **Lease RA, Cusick ME, Belfort M.** 1998. Riboregulation in Escherichia coli: DsrA RNA acts by RNA:RNA interactions at multiple loci. Proc Natl Acad Sci U S A **95:**12456-12461.
345. **Singh SS, Singh N, Bonocora RP, Fitzgerald DM, Wade JT, Grainger DC.** 2014. Widespread suppression of intragenic transcription initiation by H-NS. Genes Dev **28:**214-219.
346. **Muller CM, Dobrindt U, Nagy G, Emody L, Uhlin BE, Hacker J.** 2006. Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic Escherichia coli. J Bacteriol **188:**5428-5438.

347. **Braaten BA, Nou X, Kaltenbach LS, Low DA.** 1994. Methylation patterns in pap regulatory DNA control pyelonephritis-associated pili phase variation in E. coli. Cell **76:**577-588.
348. **Gally DL, Rucker TJ, Blomfield IC.** 1994. The leucine-responsive regulatory protein binds to the fim switch to control phase variation of type 1 fimbrial expression in Escherichia coli K-12. J Bacteriol **176:**5665-5672.
349. **Huisman TT, Bakker D, Klaasen P, de Graaf FK.** 1994. Leucine-responsive regulatory protein, IS1 insertions, and the negative regulator FaeA control the expression of the fae (K88) operon in Escherichia coli. Mol Microbiol **11:**525-536.
350. **Khandige S, Kronborg T, Uhlin BE, Moller-Jensen J.** 2015. sRNA-Mediated Regulation of P-Fimbriae Phase Variation in Uropathogenic Escherichia coli. PLoS Pathog **11:**e1005109.
351. **McFarland KA, Lucchini S, Hinton JC, Dorman CJ.** 2008. The leucine-responsive regulatory protein, Lrp, activates transcription of the fim operon in Salmonella enterica serovar typhimurium via the fimZ regulatory gene. J Bacteriol **190:**602-612.
352. **Roesch PL, Blomfield IC.** 1998. Leucine alters the interaction of the leucine-responsive regulatory protein (Lrp) with the fim switch to stimulate site-specific recombination in Escherichia coli. Mol Microbiol **27:**751-761.
353. **Shimada T, Saito N, Maeda M, Tanaka K, Ishihama A.** 2015. Expanded roles of leucine-responsive regulatory protein in transcription regulation of the Escherichia coli genome: Genomic SELEX screening of the regulation targets. M Gen **1**.
354. **van der Woude MW, Kaltenbach LS, Low DA.** 1995. Leucine-responsive regulatory protein plays dual roles as both an activator and a repressor of the Escherichia coli pap fimbrial operon. Mol Microbiol **17:**303-312.
355. **Simms AN, Mobley HL.** 2008. Multiple genes repress motility in uropathogenic Escherichia coli constitutively expressing type 1 fimbriae. Journal of Bacteriology **190:**3747-3756.
356. **de los Rios S, Perona JJ.** 2007. Structure of the Escherichia coli leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. J Mol Biol **366:**1589-1602.
357. **Beloin C, Jeusset J, Revet B, Mirambeau G, Le Hegarat F, Le Cam E.** 2003. Contribution of DNA conformation and topology in right-handed DNA wrapping by the Bacillus subtilis LrpC protein. J Biol Chem **278:**5333-5342.
358. **Nou X, Braaten B, Kaltenbach L, Low DA.** 1995. Differential binding of Lrp to two sets of pap DNA binding sites mediated by Pap I regulates Pap phase variation in Escherichia coli. EMBO J **14:**5785-5797.
359. **Cui Y, Wang Q, Stormo GD, Calvo JM.** 1995. A consensus sequence for binding of Lrp to DNA. J Bacteriol **177:**4872-4880.
360. **Chen S, Calvo JM.** 2002. Leucine-induced dissociation of Escherichia coli Lrp hexadecamers to octamers. J Mol Biol **318:**1031-1042.

361. **Chen S, Rosner MH, Calvo JM.** 2001. Leucine-regulated self-association of leucine-responsive regulatory protein (Lrp) from Escherichia coli. J Mol Biol **312:**625-635.
362. **Cho BK, Barrett CL, Knight EM, Park YS, Palsson BO.** 2008. Genome-scale reconstruction of the Lrp regulatory network in Escherichia coli. Proc Natl Acad Sci U S A **105:**19462-19467.
363. **Hart BR, Blumenthal RM.** 2011. Unexpected coregulator range for the global regulator Lrp of Escherichia coli and Proteus mirabilis. J Bacteriol **193:**1054-1064.
364. **Paul L, Blumenthal RM, Matthews RG.** 2001. Activation from a distance: roles of Lrp and integration host factor in transcriptional activation of gltBDF. J Bacteriol **183:**3910-3918.
365. **Paul L, Mishra PK, Blumenthal RM, Matthews RG.** 2007. Integration of regulatory signals through involvement of multiple global regulators: control of the Escherichia coli gltBDF operon by Lrp, IHF, Crp, and ArgR. BMC Microbiol **7:**2.
366. **Weyand NJ, Braaten BA, van der Woude M, Tucker J, Low DA.** 2001. The essential role of the promoter-proximal subunit of CAP in pap phase variation: Lrp- and helical phase-dependent activation of papBA transcription by CAP from -215. Mol Microbiol **39:**1504-1522.
367. **Yokoyama K, Ishijima SA, Clowney L, Koike H, Aramaki H, Tanaka C, Makino K, Suzuki M.** 2006. Feast/famine regulatory proteins (FFRPs): Escherichia coli Lrp, AsnC and related archaeal transcription factors. FEMS Microbiol Rev **30:**89-108.
368. **Potrykus K, Cashel M.** 2008. (p)ppGpp: still magical? Annu Rev Microbiol **62:**35-51.
369. **Newman EB, D'Ari R, Lin RT.** 1992. The leucine-Lrp regulon in E. coli: a global response in search of a raison d'etre. Cell **68:**617-619.
370. **Zinser ER, Kolter R.** 2000. Prolonged stationary-phase incubation selects for lrp mutations in Escherichia coli K-12. J Bacteriol **182:**4361-4365.
371. **White-Ziegler CA, Angus Hill ML, Braaten BA, van der Woude MW, Low DA.** 1998. Thermoregulation of Escherichia coli pap transcription: H-NS is a temperature-dependent DNA methylation blocking factor. Mol Microbiol **28:**1121-1137.
372. **White-Ziegler CA, Villapakkam A, Ronaszeki K, Young S.** 2000. H-NS controls pap and daa fimbrial transcription in Escherichia coli in response to multiple environmental cues. J Bacteriol **182:**6391-6400.
373. **May G, Dersch P, Haardt M, Middendorf A, Bremer E.** 1990. The osmZ (bglY) gene encodes the DNA-binding protein H-NS (H1a), a component of the Escherichia coli K12 nucleoid. Mol Gen Genet **224:**81-90.
374. **O'Gara JP, Dorman CJ.** 2000. Effects of local transcription and H-NS on inversion of the fim switch of Escherichia coli. Mol Microbiol **36:**457-466.

375. **Bertin P, Terao E, Lee EH, Lejeune P, Colson C, Danchin A, Collatz E.** 1994. The H-NS protein is involved in the biogenesis of flagella in Escherichia coli. J Bacteriol **176:**5537-5540.
376. **Kim EA, Blair DF.** 2015. Function of the Histone-Like Protein H-NS in Motility of Escherichia coli: Multiple Regulatory Roles Rather than Direct Action at the Flagellar Motor. J Bacteriol **197:**3110-3120.
377. **Krin E, Danchin A, Soutourina O.** 2010. RcsB plays a central role in H-NS-dependent regulation of motility and acid stress resistance in Escherichia coli. Res Microbiol **161:**363-371.
378. **Soutourina O, Kolb A, Krin E, Laurent-Winter C, Rimsky S, Danchin A, Bertin P.** 1999. Multiple control of flagellum biosynthesis in Escherichia coli: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the flhDC master operon. J Bacteriol **181:**7500-7508.
379. **Arnqvist A, Olsen A, Normark S.** 1994. Sigma S-dependent growth-phase induction of the csgBA promoter in Escherichia coli can be achieved in vivo by sigma 70 in the absence of the nucleoid-associated protein H-NS. Mol Microbiol **13:**1021-1032.
380. **Hammar M, Arnqvist A, Bian Z, Olsen A, Normark S.** 1995. Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in Escherichia coli K-12. Mol Microbiol **18:**661-670.
381. **Ogasawara H, Yamada K, Kori A, Yamamoto K, Ishihama A.** 2010. Regulation of the Escherichia coli csgD promoter: interplay between five transcription factors. Microbiology **156:**2470-2483.
382. **Brunet YR, Khodr A, Logger L, Aussel L, Mignot T, Rimsky S, Cascales E.** 2015. H-NS Silencing of the Salmonella Pathogenicity Island 6-Encoded Type VI Secretion System Limits Salmonella enterica Serovar Typhimurium Interbacterial Killing. Infect Immun **83:**2738-2750.
383. **Lahooti M, Roesch PL, Blomfield IC.** 2005. Modulation of the sensitivity of FimB recombination to branched-chain amino acids and alanine in Escherichia coli K-12. J Bacteriol **187:**6273-6280.
384. **Baga M, Goransson M, Normark S, Uhlin BE.** 1985. Transcriptional activation of a pap pilus virulence operon from uropathogenic Escherichia coli. EMBO J **4:**3887-3893.
385. **Donato GM, Lelivelt MJ, Kawula TH.** 1997. Promoter-specific repression of fimB expression by the Escherichia coli nucleoid-associated protein H-NS. J Bacteriol **179:**6618-6625.
386. **Jordi BJ, Dagberg B, de Haan LA, Hamers AM, van der Zeijst BA, Gaastra W, Uhlin BE.** 1992. The positive regulator CfaD overcomes the repression mediated by histone-like protein H-NS (H1) in the CFA/I fimbrial operon of Escherichia coli. EMBO J **11:**2627-2632.
387. **Torres AG, Slater TM, Patel SD, Popov VL, Arenas-Hernandez MM.** 2008. Contribution of the Ler- and H-NS-regulated long polar fimbriae of Escherichia coli O157:H7 during binding to tissue-cultured cells. Infect Immun **76:**5062-5071.

388. **Martinez-Antonio A, Janga SC, Thieffry D.** 2008. Functional organisation of Escherichia coli transcriptional regulatory network. J Mol Biol **381:**238-247.
389. **Girgis HS, Liu Y, Ryu WS, Tavazoie S.** 2007. A comprehensive genetic characterization of bacterial motility. PLoS Genet **3:**1644-1660.
390. **Blomfield IC, Calie PJ, Eberhardt KJ, McClain MS, Eisenstein BI.** 1993. Lrp stimulates phase variation of type 1 fimbriation in Escherichia coli K-12. J Bacteriol **175:**27-36.
391. **Dudin O, Geiselmann J, Ogasawara H, Ishihama A, Lacour S.** 2014. Repression of flagellar genes in exponential phase by CsgD and CpxR, two crucial modulators of Escherichia coli biofilm formation. J Bacteriol **196:**707-715.
392. **Morgan JL, McNamara JT, Zimmer J.** 2014. Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. Nat Struct Mol Biol **21:**489-496.
393. **Simm R, Morr M, Kader A, Nimtz M, Romling U.** 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Mol Microbiol **53:**1123-1134.
394. **Romling U, Rohde M, Olsen A, Normark S, Reinkoster J.** 2000. AgfD, the checkpoint of multicellular and aggregative behaviour in Salmonella typhimurium regulates at least two independent pathways. Mol Microbiol **36:**10-23.
395. **Blumer C, Kleefeld A, Lehnen D, Heintz M, Dobrindt U, Nagy G, Michaelis K, Emody L, Polen T, Rachel R, Wendisch VF, Unden G.** 2005. Regulation of type 1 fimbriae synthesis and biofilm formation by the transcriptional regulator LrhA of Escherichia coli. Microbiology **151:**3287-3298.
396. **Pratt LA, Kolter R.** 1998. Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili. Mol Microbiol **30:**285-293.
397. **Alteri CJ, Himpsl SD, Pickens SR, Lindner JR, Zora JS, Miller JE, Arno PD, Straight SW, Mobley HL.** 2013. Multicellular bacteria deploy the type VI secretion system to preemptively strike neighboring cells. PLoS Pathog **9:**e1003608.
398. **Hay NA, Tipper DJ, Gygi D, Hughes C.** 1997. A nonswarming mutant of Proteus mirabilis lacks the Lrp global transcriptional regulator. J Bacteriol **179:**4741-4746.
399. **Alteri CJ, Himpsl SD, Engstrom MD, Mobley HL.** 2012. Anaerobic respiration using a complete oxidative TCA cycle drives multicellular swarming in Proteus mirabilis. MBio **3**.
400. **Pearson MM, Rasko DA, Smith SN, Mobley HL.** 2010. Transcriptome of swarming Proteus mirabilis. Infect Immun **78:**2834-2845.
401. **Armbruster CE, Hodges SA, Mobley HL.** 2013. Initiation of swarming motility by Proteus mirabilis occurs in response to specific cues present in urine and requires excess L-glutamine. J Bacteriol **195:**1305-1319.
402. **Jones BV, Young R, Mahenthiralingam E, Stickler DJ.** 2004. Ultrastructure of Proteus mirabilis swarmer cell rafts and role of swarming in catheter-associated urinary tract infection. Infect Immun **72:**3941-3950.

403. **Oshima T, Ito K, Kabayama H, Nakamura Y.** 1995. Regulation of lrp gene expression by H-NS and Lrp proteins in Escherichia coli: dominant negative mutations in lrp. Mol Gen Genet **247:**521-528.
404. **Goransson M, Forsman P, Nilsson P, Uhlin BE.** 1989. Upstream activating sequences that are shared by two divergently transcribed operons mediate cAMP-CRP regulation of pilus-adhesin in Escherichia coli. Mol Microbiol **3:**1557-1565.
405. **Schappert SM.** Ambulatory medical care utilization estimates for 2006.
406. **Gupta K, Hooton TM, Naber KG, Wullt B, Colgan R, Miller LG, Moran GJ, Nicolle LE, Raz R, Schaeffer AJ, Soper DE.** 2011. International clinical practice guidelines for the treatment of acute uncomplicated cystitis and pyelonephritis in women: A 2010 update by the Infectious Diseases Society of America and the European Society for Microbiology and Infectious Diseases. Clin Infect Dis **52:**e103-120.
407. **Hooton TM.** 2001. Recurrent urinary tract infection in women. Int J Antimicrob Agents **17:**259-268.
408. **Russo TA, Stapleton A, Wenderoth S, Hooton TM, Stamm WE.** 1995. Chromosomal restriction fragment length polymorphism analysis of Escherichia coli strains causing recurrent urinary tract infections in young women. J Infect Dis **172:**440-445.
409. **Busch A, Waksman G.** 2012. Chaperone-usher pathways: diversity and pilus assembly mechanism. Philos Trans R Soc Lond B Biol Sci **367:**1112-1122.
410. **Orndorff PE, Falkow S.** 1984. Organization and expression of genes responsible for type 1 piliation in Escherichia coli. J Bacteriol **159:**736-744.
411. **Nesta B, Spraggon G, Alteri C, Moriel DG, Rosini R, Veggi D, Smith S, Bertoldi I, Pastorello I, Ferlenghi I, Fontana MR, Frankel G, Mobley HL, Rappuoli R, Pizza M, Serino L, Soriani M.** 2012. FdeC, a novel broadly conserved Escherichia coli adhesin eliciting protection against urinary tract infections. MBio **3**.
412. **Allsopp LP, Beloin C, Moriel DG, Totsika M, Ghigo JM, Schembri MA.** 2012. Functional heterogeneity of the UpaH autotransporter protein from uropathogenic Escherichia coli. J Bacteriol **194:**5769-5782.
413. **Allsopp LP, Totsika M, Tree JJ, Ulett GC, Mabbett AN, Wells TJ, Kobe B, Beatson SA, Schembri MA.** 2010. UpaH is a newly identified autotransporter protein that contributes to biofilm formation and bladder colonization by uropathogenic Escherichia coli CFT073. Infect Immun **78:**1659-1669.
414. **Lane MC, Simms AN, Mobley HL.** 2007. complex interplay between type 1 fimbrial expression and flagellum-mediated motility of uropathogenic Escherichia coli. J Bacteriol **189:**5523-5533.
415. **Cooper LA, Simmons LA, Mobley HL.** 2012. Involvement of mismatch repair in the reciprocal control of motility and adherence of uropathogenic Escherichia coli. Infect Immun **80:**1969-1979.
416. **Gur C, Coppenhagen-Glazer S, Rosenberg S, Yamin R, Enk J, Glasner A, Bar-On Y, Fleissig O, Naor R, Abed J, Mevorach D, Granot Z, Bachrach G, Mandelboim O.** 2013. Natural killer cell-mediated host defense against

uropathogenic E. coli is counteracted by bacterial hemolysinA-dependent killing of NK cells. Cell Host Microbe **14:**664-674.

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

418. **Battaglioli EJ, Baisa GA, Weeks AE, Schroll RA, Hryckowian AJ, Welch RA.** 2011. Isolation of generalized transducing bacteriophages for uropathogenic strains of Escherichia coli. Appl Environ Microbiol **77:**6630-6635.

419. **Kelley LA, Sternberg MJ.** 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc **4:**363-371.

420. **Charif D, Lobry JR.** 2007. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis, p 207-232. *In* Bastolla U, Porto M, Roman HE, Vendruscolo M (ed), Structural approaches to sequence evolution: Molecules, networks, populations. Springer Verlag, New York.

421. **Pages H, Aboyoun P, Gentleman R, DebRoy S.** Biostrings: String objects representing biological sequences, and matching algorithms, vR package version 2.30.1.

422. **Warnes GR, Bolker B, Gorjanc G, Grothendieck G, Korosec A, Lumley T, MacQueen D, Magnusson A, Rogers J, et al.** 2014. gdata: Various R programming tools for data manipulation, vR package version 2.13.3. http://CRAN.R-project.org/package=gdata.

423. **Miller J.** 1992. A Short Course in Bacterial Genetics. Cold Springer Harbor Laboratory Press, Plainview.

424. **Zheng Y, Szustakowski JD, Fortnow L, Roberts RJ, Kasif S.** 2002. Computational identification of operons in microbial genomes. Genome Res **12:**1221-1230.

425. **Xia Y, Uhlin BE.** 1999. Mutational analysis of the PapB transcriptional regulator in Escherichia coli. Regions important for DNA binding and oligomerization. J Biol Chem **274:**19723-19730.

426. **Lugering A, Benz I, Knochenhauer S, Ruffing M, Schmidt MA.** 2003. The Pix pilus adhesin of the uropathogenic Escherichia coli strain X2194 (O2 : K(-): H6) is related to Pap pili but exhibits a truncated regulatory region. Microbiology **149:**1387-1397.

427. **Marklund BI, Tennent JM, Garcia E, Hamers A, Baga M, Lindberg F, Gaastra W, Normark S.** 1992. Horizontal gene transfer of the Escherichia coli pap and prs pili operons as a mechanism for the development of tissue-specific adhesive properties. Mol Microbiol **6:**2225-2242.

428. **Wagner C, Barlag B, Gerlach RG, Deiwick J, Hensel M.** 2013. The Salmonella enterica giant adhesin SiiE binds to polarized epithelial cells in a lectin-like manner. Cell Microbiol doi:10.1111/cmi.12253.

429. **Stella S, Cascio D, Johnson RC.** 2010. The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. Genes Dev **24:**814-826.

430. **Zuber F, Kotlarz D, Rimsky S, Buc H.** 1994. Modulated expression of promoters containing upstream curved DNA sequences by the Escherichia coli nucleoid protein H-NS. Mol Microbiol **12:**231-240.
431. **Ogasawara H, Yamamoto K, Ishihama A.** 2011. Role of the biofilm master regulator CsgD in cross-regulation between biofilm formation and flagellar synthesis. J Bacteriol **193:**2587-2597.
432. **Clegg S, Hughes KT.** 2002. FimZ is a molecular link between sticking and swimming in Salmonella enterica serovar Typhimurium. J Bacteriol **184:**1209-1213.
433. **Francez-Charlot A, Laugel B, Van Gemert A, Dubarry N, Wiorowski F, Castanie-Cornet MP, Gutierrez C, Cam K.** 2003. RcsCDB His-Asp phosphorelay system negatively regulates the flhDC operon in Escherichia coli. Mol Microbiol **49:**823-832.
434. **Stamm WE, Hooton TM.** 1993. Management of urinary tract infections in adults. N Engl J Med **329:**1328-1334.
435. **Engstrom MD, Alteri CJ, Mobley HL.** 2014. A conserved PapB family member, TosR, regulates expression of the uropathogenic Escherichia coli RTX nonfimbrial adhesin TosA while conserved LuxR family members TosE and TosF suppress motility. Infect Immun **82:**3644-3656.
436. **Glaser P, Sakamoto H, Bellalou J, Ullmann A, Danchin A.** 1988. Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of Bordetella pertussis. EMBO J **7:**3997-4004.
437. **Boardman BK, Satchell KJ.** 2004. Vibrio cholerae strains with mutations in an atypical type I secretion system accumulate RTX toxin intracellularly. J Bacteriol **186:**8137-8143.
438. **Fullner KJ, Mekalanos JJ.** 2000. In vivo covalent cross-linking of cellular actin by the Vibrio cholerae RTX toxin. EMBO J **19:**5315-5323.
439. **Sauer FG, Remaut H, Hultgren SJ, Waksman G.** 2004. Fiber assembly by the chaperone-usher pathway. Biochim Biophys Acta **1694:**259-267.
440. **Hagberg L, Hull R, Hull S, Falkow S, Freter R, Svanborg Eden C.** 1983. Contribution of adhesion to bacterial persistence in the mouse urinary tract. Infect Immun **40:**265-272.
441. **Lane MC, Mobley HL.** 2007. Role of P-fimbrial-mediated adherence in pyelonephritis and persistence of uropathogenic Escherichia coli (UPEC) in the mammalian kidney. Kidney Int **72:**19-25.
442. **Sette M, Spurio R, Trotta E, Brandizi C, Brandi A, Pon CL, Barbato G, Boelens R, Gualerzi CO.** 2009. Sequence-specific recognition of DNA by the C-terminal domain of nucleoid-associated protein H-NS. J Biol Chem **284:**30453-30462.
443. **Yang CC, Nash HA.** 1989. The interaction of E. coli IHF protein with its specific binding sites. Cell **57:**869-880.
444. **Pul U, Wurm R, Lux B, Meltzer M, Menzel A, Wagner R.** 2005. LRP and H-NS--cooperative partners for transcription regulation at Escherichia coli rRNA promoters. Mol Microbiol **58:**864-876.

445. **Wang Q, Calvo JM.** 1993. Lrp, a major regulatory protein in Escherichia coli, bends DNA and can organize the assembly of a higher-order nucleoprotein structure. EMBO J **12:**2495-2501.
446. **Will WR, Lu J, Frost LS.** 2004. The role of H-NS in silencing F transfer gene expression during entry into stationary phase. Mol Microbiol **54:**769-782.
447. **Troutt AB, McHeyzer-Williams MG, Pulendran B, Nossal GJ.** 1992. Ligation-anchored PCR: a simple amplification technique with single-sided specificity. Proc Natl Acad Sci U S A **89:**9823-9825.
448. **Hawley DK, McClure WR.** 1983. Compilation and analysis of Escherichia coli promoter DNA sequences. Nucleic Acids Res **11:**2237-2255.
449. **Harley CB, Reynolds RP.** 1987. Analysis of E. coli promoter sequences. Nucleic Acids Res **15:**2343-2361.
450. **Wu CW, Goldthwait DA.** 1969. Studies of nucleotide binding to the ribonucleic acid polymerase by a fluoresence technique. Biochemistry **8:**4450-4458.
451. **Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, Jimenez-Jacinto V, Salgado H, Juarez K, Contreras-Moreira B, Huerta AM, Collado-Vides J, Morett E.** 2009. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. PLoS One **4:**e7526.
452. **Smolke CD, Khlebnikov A, Keasling JD.** 2001. Effects of transcription induction homogeneity and transcript stability on expression of two genes in a constructed operon. Appl Microbiol Biotechnol **57:**689-696.
453. **Khlebnikov A, Risa O, Skaug T, Carrier TA, Keasling JD.** 2000. Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. J Bacteriol **182:**7029-7034.
454. **Siegele DA, Hu JC.** 1997. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. Proc Natl Acad Sci U S A **94:**8168-8172.
455. **Morra R, Shankar J, Robinson CJ, Halliwell S, Butler L, Upton M, Hay S, Micklefield J, Dixon N.** 2015. Dual transcriptional-translational cascade permits cellular level tuneable expression control. Nucleic Acids Res doi:10.1093/nar/gkv912.
456. **Bouffartigues E, Buckle M, Badaut C, Travers A, Rimsky S.** 2007. H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nat Struct Mol Biol **14:**441-448.
457. **Scheu PD, Witan J, Rauschmeier M, Graf S, Liao YF, Ebert-Jung A, Basche T, Erker W, Unden G.** 2012. CitA/CitB two-component system regulating citrate fermentation in Escherichia coli and its relation to the DcuS/DcuR system in vivo. J Bacteriol **194:**636-645.
458. **Zhou B, Beckwith D, Jarboe LR, Liao JC.** 2005. Markov Chain modeling of pyelonephritis-associated pili expression in uropathogenic Escherichia coli. Biophys J **88:**2541-2553.
459. **Hernday AD, Braaten BA, Broitman-Maduro G, Engelberts P, Low DA.** 2004. Regulation of the pap epigenetic switch by CpxAR: phosphorylated CpxR

inhibits transition to the phase ON state by competition with Lrp. Mol Cell **16:**537-547.
460. **Blyn LB, Braaten BA, Low DA.** 1990. Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states. EMBO J **9:**4045-4054.
461. **Wolf L, Silander OK, van Nimwegen E.** 2015. Expression noise facilitates the evolution of gene regulation. Elife **4**.
462. **Silander OK, Nikolic N, Zaslaver A, Bren A, Kikoin I, Alon U, Ackermann M.** 2012. A genome-wide analysis of promoter-mediated phenotypic noise in Escherichia coli. PLoS Genet **8:**e1002443.
463. **Dame RT, Wyman C, Goosen N.** 2000. H-NS mediated compaction of DNA visualised by atomic force microscopy. Nucleic Acids Res **28:**3504-3510.
464. **Dame RT.** 2005. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Mol Microbiol **56:**858-870.
465. **Sezonov G, Joseleau-Petit D, D'Ari R.** 2007. Escherichia coli physiology in Luria-Bertani broth. J Bacteriol **189:**8746-8749.
466. **Raghavan R, Groisman EA, Ochman H.** 2011. Genome-wide detection of novel regulatory RNAs in E. coli. Genome Res **21:**1487-1497.
467. **Xu C, Bian C, Lam R, Dong A, Min J.** 2011. The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. Nat Commun **2:**227.
468. **Shannon CE.** 2001. A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review **5:**3-55.
469. **Holden NS, Tacon CE.** 2011. Principles and problems of the electrophoretic mobility shift assay. J Pharmacol Toxicol Methods **63:**7-14.
470. **Lubeck E, Cai L.** 2012. Single-cell systems biology by super-resolution imaging and combinatorial labeling. Nat Methods **9:**743-748.
471. **Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Ferrante TC, Terry R, Turczyk BM, Yang JL, Lee HS, Aach J, Zhang K, Church GM.** 2015. Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. Nat Protoc **10:**442-458.
472. **Bochner BR.** 2009. Global phenotypic characterization of bacteria. FEMS Microbiol Rev **33:**191-205.
473. **Adiciptaningrum AM, Blomfield IC, Tans SJ.** 2009. Direct observation of type 1 fimbrial switching. EMBO Rep **10:**527-532.