**The repeat-in-toxin (RTX) family member TosA mediates adherence of Uropathogenic *Escherichia coli* to urinary tract epithelial cells and enhances survival and toxicity during bacteremia and sepsis**

**by**

**Patrick Vigil**

**A dissertation submitted in partial fulfillment**
**of the requirements for the degree of**
**Doctor of Philosophy**
**(Microbiology and Immunology)**
**in The University of Michigan**
**2011**

**Doctoral Committee:**

**Professor Harry L.T. Mobley, chair**
**Professor Victor DiRita**
**Professor Michele S. Swanson**
**Associate Professor Philip C. Hanna**
**Assistant Professor David Miller**

**TABLE OF CONTENTS**

**LIST OF FIGURES**

**LIST OF TABLES**

**ABSTRACT**

Uropathogenic *Escherichia coli* (UPEC) are responsible for the majority of uncomplicated urinary tract infections (UTI) and represent the most common bacterial infection of the adult population. UPEC utilize a wide range of virulence factors to colonize the host environment, including the secreted protein toxin α-hemolysin. While a great deal is known about this protein, which serves as the prototype of a large family of secreted bacterial proteins termed the repeats-in-toxin (RTX) family, the protein itself has only a minor contribution to UPEC pathogenesis. The work contained in this dissertation describes a novel RTX family member, TosA, which was discovered in a screen of proteins synthesized by a human UPEC strain that were recognized as antigenic during chronic UTI in an animal model of infection. TosA was determined to be specifically expressed in the host urinary tract and contributed significantly to the virulence and survival of UPEC. TosA was only found in the B2 phylogenetic subgroup of *E. coli* and served as a marker for a group of isolates that also contained a large number of well characterized UPEC virulence factors. The presence of the gene *tosA* was determined to predict the success of fecal isolates in the murine model of ascending UTI regardless of the source of the isolate. Finally, a detailed analysis of the function of *tosA* revealed that this gene is linked to genes encoding a conserved type 1 secretion system similar to other RTX family members and may be part of large regulatory network that

includes other UPEC adherence factors. The TosA protein was found to mediate *i*) adherence to host cells derived from the upper urinary tract, *ii*) survival in disseminated infections, and *iii*) enhanced lethality during sepsis (as discovered in two different animal models of infection). An experimental vaccine using purified protein protected vaccinated animals against urosepsis. From this work, it was concluded that TosA belongs to a novel group of RTX proteins that mediate adherence and host damage during UTI and urosepsis, and may be a novel target for the development of new therapeutics to treat ascending urinary tract infections.

# CHAPTER 1

# INTRODUCTION

## *Escherichia coli* pathogenesis in the urinary tract

The species *Escherichia coli* comprises a wide range isolates that can be found associated with the gastrointestinal tract of numerous species of mammals (91). Most isolates can exist in a commensal, non-pathogenic state in the host GI tract, but a subset can cause disease in several different body sites. Many different subtypes of GI pathogenic *E. coli* exist, and these groups are distinct from extra-intestinal pathogenic *E. coli (*ExPEC) (98). ExPEC include all isolates that can cause disease in healthy hosts outside of the GI tract, with the majority of human isolates originating from urinary tract infections (UTI) (65). The differences observed among *E. coli* of different phylogenetic background (131) and containing different accessory genomic content (126) confers the unique behaviors that characterize each group.

Uncomplicated UTI caused by *E. coli* represents the most common bacterial infection in healthy adults (31). The economic impact of these infections is immense, with direct and indirect medical costs in the United States exceeding $1.6 billion annually (32). While these infections could be well controlled with antibiotics in the past, a global rise in antibiotic resistance among human UPEC isolates (1, 12, 16, 25) threatens to make these therapeutic options ineffective in the near future. Clearly

there is a great need for an enhanced understanding of the mechanisms that UPEC utilize during infection in order to generate the next generation of UTI therapeutics.

*E. coli*-mediated UTI proceeds in an ascending manner, starting most commonly with contamination of the perineal area with fecal material containing *E. coli* that ascend the urethra to gain access to the bladder. This route of infection was established by multiple studies that identified the same strain of *E. coli* in the urine and fecal material of infected individuals, as well as close contacts, such as sexual partners (33, 137). Remarkably, this same pattern is observed among dogs with UTI (64) and strains that infect human urinary tracts have been recovered from fecal samples of household pets (62, 87), arguing for an extensive reservoir of UPEC that could potentially include the GI tract of many mammalian species.

Once in the bladder, bacteria replicate and initiate a lower urinary tract infection. In a subset of patients this colonization may not lead to overt symptoms, a condition termed asymptomatic bacteriuria (57). However, in most cases this colonization results in the clinical signs and symptoms of cystitis. Bacteria may then ascend the ureters and colonize the kidneys, resulting in a more severe infection termed pyelonephritis (7).

Bacteremia may result as a sequela of severe urinary tract infections from spread of bacteria from the kidneys into the bloodstream. In one survey covering 18 years of *E. coli* bacteremia monitoring at one hospital, urinary tract infections ranked as the most common medical event preceding bacteremia (40). While organisms other than UPEC are more common sources of bacteremia in nosocomial infections, *E. coli* originating

specifically in the urinary tract have been identified as the most common source of infection in community-acquired bacteremia cases that later result in hospitalization (2, 18). Hence, when considering the mechanisms of UTI pathogenesis and UPEC-mediated disease, the contribution of individual virulence factors to disease should not focus solely on the urinary tract, but should also include consideration of the impact on bacteremia and sepsis, the final steps in the ascending route of infection.

Many of the same mechanisms that aid survival in the urinary tract can enhance survival during disseminated infections. For example, a recently developed murine model of bacteremia has discovered that many virulence factors characterized in the setting of UTI, including adherence factors and nutrient acquisition processes, also enhance fitness during bacteremia (109). Excellent reviews on the current understanding of UPEC pathogenesis are available (65, 85, 89, 136) and only the information directly relating to the studies presented in this work are outlined below.

**α-hemolysin**

One of the earliest UPEC virulence factors characterized in both the urinary tract and during disseminated infections is the secreted protein α-hemolysin. The secreted protein toxin responsible for hemolysis was determined to be encoded on plasmids in many intestinal isolates (39), but the gene was later found to exist on the chromosome in ExPEC, particularly those isolated from the urinary tract (29). In both cases, the gene content and order was conserved and consisted of the operon structure *hlyCABD* (Figure 1-1A). *hlyA* encodes the secreted toxin, and *hlyB* and *hlyD* encode the inner membrane

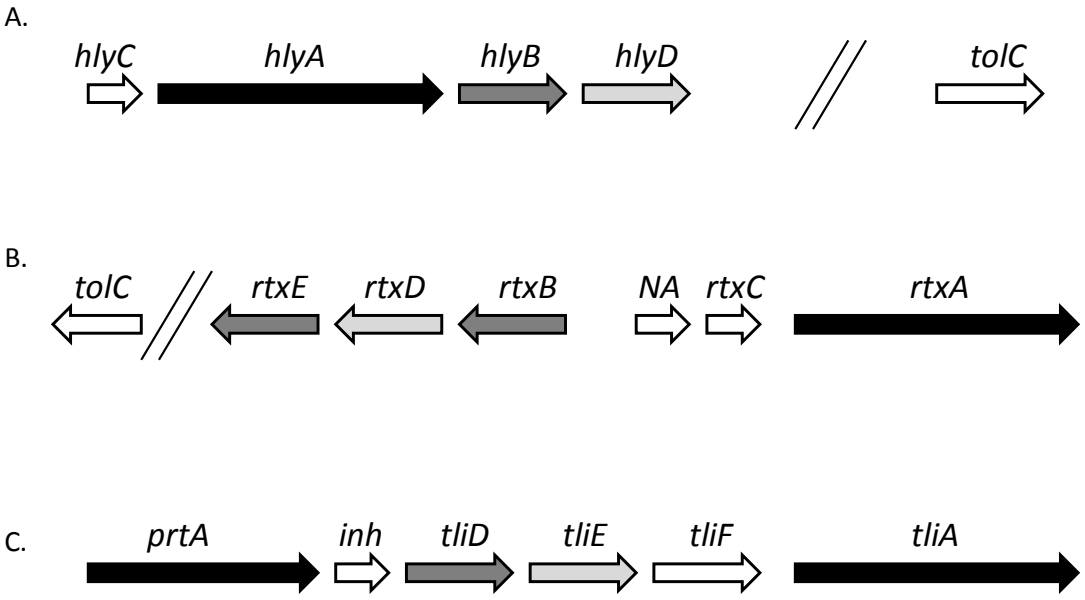



**Figure 1-1. Operon structure of representative RTX members.** A. Operon structure of *E. coli* α-hemolysin. B. Operon structure of $MARTX_{VC}$ of *V. cholerae.* C. Operon structure of *P. fluorescens* protease PrtA and lipase TliA operon. Secreted RTX member genes are coded black, ABC transporters coded dark grey, and membrane fusion proteins coded light grey. TolC encoding gene is located outside of operon in A. and B. and encoded by *tliF* in C. NA – non-annotated, no function yet ascribed to this gene.

components of a type 1 secretion system (T1SS). These proteins work with TolC, an outer membrane porin encoded elsewhere on the chromosome, to form a dedicated secretion apparatus (6). The presence of *hlyC* is necessary for toxicity (39), and the gene encodes an enzyme that adds acyl side groups to HlyA prior to secretion (60). This modification is necessary for the activity of the toxin; however, the exact purpose of the post-translational modification remains unclear.

While originally characterized as a bacterial hemolysin, the function of activated, secreted HlyA is not limited to the lysis of red blood cells. The protein can insert into host cell membranes and form pores in a wide variety of cell types, including host immune cells such as granulocytes, monocytes, and lymphocytes (36), and epithelial cells of the kidney tubules (68, 83). An important role for α-hemolysin in *E. coli*-mediated infections has been established in several animal models of bacteremia and sepsis. Strains of *E. coli* demonstrate enhanced lethality in animal models of peritonitis when the genes encoding the hemolysin operon are present (127, 128) and enhanced lethality during sepsis is observed when multiple copies of the *hly* operon are present in the same strain (88). Despite this, the role of α-hemolysin in the urinary tract environment has been more difficult to establish. While reduced lethality is observed in Δ*hly* UPEC strains in an infant mouse model of UTI (88), few defects are observed in adult mice. An increase in inflammation and more severe histopathology is observed in bladders of adult mice when infected with strains that produce the toxin, but equal numbers of wild type and α-hemolysin deficient bacteria were observed in the urinary tract of infected animals (110).

Work conducted at the same time on the potent secreted toxin LktA found in *Pasteurella haemolytica* determined that the nucleotide sequence of *hlyA* and *lktA* were homologous (114) and hinted at a similar operon structure for the two systems (79). After comparing these two toxins with those found in other human pathogens, including *Bordetella pertussis* and *Proteus vulgaris*, it was realized that these secreted toxins belonged to much larger family of proteins found throughout gram negative bacterial pathogens (125). Due to the fact that most of these proteins exhibited hemolytic or cytotoxic effects on host cells and the presence of a characteristic glycine- and aspartate-rich repeats in the proteins, the family was termed repeats in toxins (RTX) (125). This name proved to be misleading as subsequent work has established that the proteins are even more widely distributed than previously appreciated and can also be found in non-pathogenic species (76). As discussed below, many of the features of the toxin members of the RTX family are shared by all family members and help to define the RTX family of secreted proteins.

**RTX family of secreted bacterial proteins**

Regardless of the source, RTX family members share two defining features, a characteristic repeat structure in the C terminus of the protein and a type 1 secretion system (T1SS) that exports the proteins out of the cell. The T1SS shared among RTX members is highly conserved and in most RTX systems studied to date consist of three components. The inner membrane components consist of an ATP-binding cassette (ABC)-containing protein that powers the secretion apparatus, designated the B gene of

toxin members but often marked as the D gene of protease and lipase RTX operons (Figure 1-1). A membrane fusion protein that links this inner membrane transporter with the outer membrane porin is designated as the D gene of toxin containing operons, but often given an E designation in other systems (76). The outer membrane porin of the TolC family is often encoded outside of the RTX operon (138). However in a few examples of toxin members, such as the cytotoxic RTX system in *B. pertussis* (38), and the majority of the protease and lipase RTX operons (76), the TolC homolog is encoded by a gene contained within the same operon. The MARTX family, observed in a variety of *Vibrio spp.,* is the exception and consists of a fourth component, a second ABC - containing protein that is necessary for secretion of MARTX members (11) (Figure 1-1).

The secretion of RTX proteins is thought to proceed in several distinct steps; however, the precise nature of these mechanisms remains unresolved. What is known is that an uncleaved signal sequence most likely exists in the last 50 amino acids of the C terminus, as the deletion of these residues ablates secretion (112), but addition of the residues to normally non-secreted proteins allows for secretion (37). As the primary amino acid sequence of this portion of RTX family members is not well conserved, it is thought that the signal recognized by the ABC transporter protein is not a specific string of amino acid residues, but rather a secondary or tertiary fold that can be adopted by the variety of sequences observed in this region (125). The membrane fusion protein associates with the ABC transporter and facilitates the association of the outer membrane porin with the complex, forming a channel that allows direct passage of the

RTX protein from the cytoplasm of the cell into the extracellular environment, bypassing the periplasmic space entirely (76).

The characteristic RTX repeat is found in the C terminus of the secreted protein, usually close to secretion signal recognized by the T1SS. The consensus sequence of the repeat is rich in glycine and aspartate residues, the most basic sequence consisting of repeating units of G-G-X-G-X-D separated by variable spacer regions. The number of repeats can vary, with the prototype α-hemolysin containing more than a dozen and some family members predicted to contain more than 40 (76). The solved structure of the *Pseudomonas areuginosa* RTX protease AprA identified the repeats in a unique protein fold only observed so far in RTX family members. This fold, termed a β-roll, coordinates $Ca^{2+}$ ions between a turn separated by very short β-strands (10). Each repeat forms a turn and β-strand segment, allowing multiple RTX repeats to build up a block of $Ca^{2+}$-binding sites (Figure 1-2B and C).

Proper folding of RTX proteins is thought to depend on the RTX motif binding $Ca^{2+}$ ions. The proposed purpose of the motif is that it acts as a switch, keeping the RTX unfolded while inside the bacterium and allowing the T1SS a less rigidly structured protein for secretion (138). Once outside of the bacterium, where $Ca^{2+}$ concentrations are often more than an order of magnitude higher, the RTX motif assumes the proper β-roll structure. This is theorized to trigger a conformational change in the protein that leads to the assumption of the fully folded and functional protein (Figure 1-2A). All RTX family members are thought to utilize the conserved T1SS and $Ca^{2+}$-binding mechanism to aid in proper folding once outside of the cell, however, the functions of individual RTX

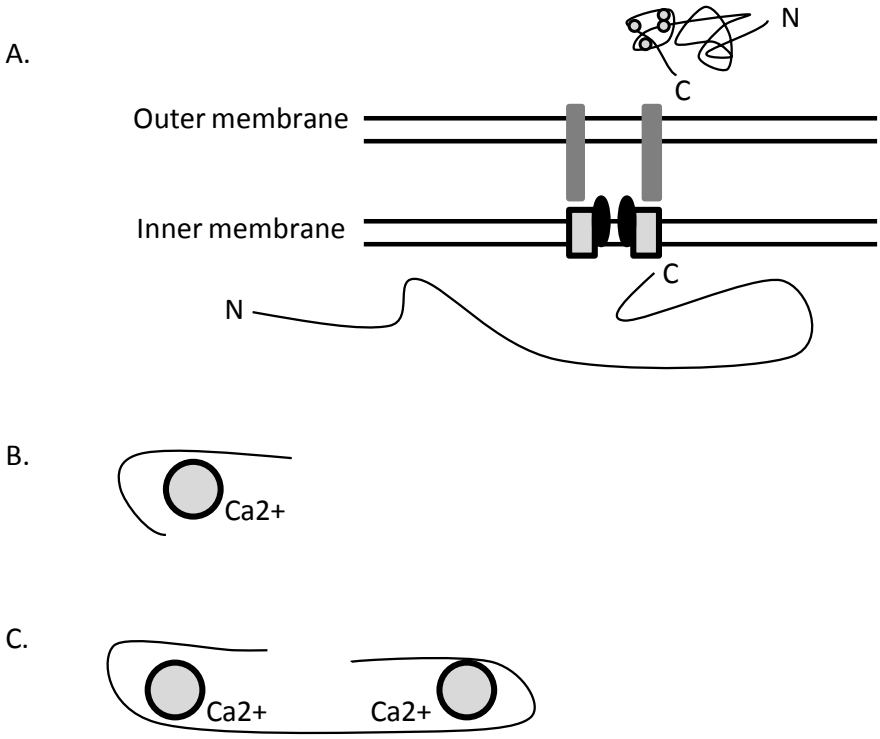



**Figure 1-2. T1SS and RTX motif.** A. The type 1 secretion system of typical RTX members is drawn with the ABC transporter embedded in inner membrane as grey boxes, the membrane fusion protein as black elipses, and the TolC homolog as long gray rectangles embedded in outer membrane. Partially unfolded RTX protein in cytoplasm of cell is drawn with N and C terminal sections noted. Fully secreted RTX member outside of cell is shown binding to $Ca^{2+}$ ions, depicted as small circles, in the folded conformation. B. A single RTX motif coordinates one $Ca^{2+}$ ion in a fold and short β-strand. C. Two RTX motifs separated by a short spacer are depicted in β-fold structure coordinating two $Ca^{2+}$ ions. Elements are not drawn to scale.

secreted proteins are extremely diverse.

The first set of RTX family members described exhibited strong toxic effects on host cells. Generally these proteins bind to host cell membranes and form pores in the membrane, disrupting cellular signaling by allowing $Ca^{2+}$ ions to leak in through the pores (66) and, with high concentrations of toxins, may result in cell lysis. These toxins may have activity against a broad spectrum of cell types and species, such as α-hemolysin, or may exhibit much narrower range of targets, suggesting that a specific cell receptor may be recognized (125). In both groups of toxins, a homolog of *hlyC* and the resulting acyl group modification appears to be critical for toxin function.

A large group of secreted RTX members exhibiting protease activity has been described in several bacterial species (76, 125). These operons often contain a protease inhibitor that is transported into the periplasm , perhaps as a countermeasure to any RTX protease that leaks back into the cells. These proteases are usually zinc-metalloproteases that act on specific targets, such as cleavage of fibrin and fibrinogen by *P. areuginosa* AprA (104) and IgG and IgA cleavage by *P. mirabilis* ZapA (81, 124). Less is known about the function of RTX lipases or the specific roles they play in bacterial adaptation to a given environment, but they are often found in conjunction with RTX proteases as part of the same genetic loci (Figure 1-1).

Recent work has uncovered a novel group of RTX proteins that exhibit multiple functions. The multifunctional autoprocessing RTX (MARTX) members of numerous *Vibrio spp.,* including the human pathogens *Vibrio cholerae* and *Vibrio vulnificus*, appear to be built from variable domains, each with unique functions and characteristic repeat

structures, but secreted through the modified RTX T1SS (100) described above. These proteins exhibit translocation of several domains into the host cell cytosol where proteolytic cleavage releases these domains and allows the toxin to function (103). While the prototype of the family, $MARTX_{Vc}$ can polymerize host actin (35), numerous other functions have been ascribed to other members (76). The full range of functions encoded by MARTX proteins could be quite diverse.

One group of RTX members, the S-layer RTX proteins, are found associated with the outer membrane of certain bacterial species and form a dense coat of protein around the cell (108). This layer of protein may serve a role protecting pathogens from the host immune system (92). But not all proteins of this family are associated with pathogens, and a subset of S-layer RTX members has been identified in *Cyanobacteria* and are involved in swimming motility (13). Additional work in *Rhizobiaceae spp.* of nitrogen-fixing bacteria has determined that RTX proteins secreted early, after association with the root structure of legumes, helps to trigger nodulation (28) and may play a key role in establishing the symbiotic relationship between the plant and the bacteria.

From this overview it is clear that RTX proteins serve a staggering variety of roles in pathogenic species and even in establishing symbiotic relationships with non-pathogens. These proteins appear to be widely distributed through gram-negative bacterial genera. Indeed, a general survey of bacterial genome sequences has uncovered more than 1000 known and putative RTX members (76). Even more surprisingly, these were observed in 251 species that span the most closely studied

gram-negative lineages and even include a handful of RTX genes present in gram-positive organisms. While it is unknown how many of these genetic loci encode functional RTX systems, the true number of RTX family members is likely to be much greater than indicated in this initial survey. Of particular interest to the work described in this document, many putative RTX members have been detected in the genomes of uropathogens (76, 94, 126). Given the importance of α-hemolysin to UPEC pathogenesis, it is likely that many of these uncharacterized RTX family members also contribute to the success of this widespread human pathogen.

**Use of RTX proteins as experimental vaccines**

Considerable interest has existed in targeting RTX family members of bacterial pathogens to develop new vaccines. An early attempt to construct a UPEC vaccine targeting α-hemolysin found reduced pathology in kidneys of vaccinated animals following UPEC challenge, but little effect on colonization levels. This model demonstrated protection when the vaccine was combined with one that targeted P fimbriae (90), but despite these results an α-hemolysin based vaccine has yet to be cleared for use in human patients. Greater success has been achieved in targeting the multiple RTX family members present in the porcine pathogen *Actinobacillus pleuropneumoniae,* and multiple studies have now demonstrated protection against subsequent challenge using purified RTX proteins (45, 101, 102, 118). When compared to inactivated whole cell vaccines, currently the only commercially available vaccines for this pathogen, the targeted vaccine demonstrated enhanced protection (102). It

remains to be seen if vaccines that specifically target RTX members will replace more traditional vaccination approaches, but these encouraging results suggest that additional development of RTX based vaccines is warranted.

**Statement of the Problem**

The importance of UPEC mediated disease in humans is highlighted by the increasing antibiotic resistance patterns and lack of suitable alternative therapeutic options. The recent development of an experimental vaccine that targets UPEC iron acquisition systems (3) and the establishment of criteria for ideal vaccine candidates for a multivalent UPEC vaccine (105) motivated the search for new vaccine candidates. An early result was the identification of a novel RTX gene, *tosA*, which had no known homology to well characterized RTX family members. The distribution and function of this gene was studied to determine whether *tosA* would be a suitable target for the development of a novel therapeutic. Finally, given the recent success of experimental vaccines to target the hemolytic RTX proteins produced by *A. pleuropneumoniae* (45, 101, 102), we investigated the potential of an experimental vaccine that targeted TosA to protect against an ascending UTI. The individual studies conducted as part of this research are outlined below.

<u>Aim 1: Identify novel antigenic UPEC virulence factors that are specifically produced during the course of a UTI</u>

While the understanding of UPEC pathogenesis that has been amassed to date is impressive, it represents an incomplete list of the molecular mechanisms utilized by this

important human pathogen. We hypothesized that additional, previously uncharacterized, virulence factors exist in UPEC that contribute significantly to the process of pathogenesis in the urinary tract. Chapter 2 describes a large screen of the majority of peptide antigens produced by the human pyelonephritis/urosepsis isolate CFT073. By treating serum taken from a murine model of chronic UTI treated to remove antibodies that recognize antigens made during *in vitro* growth, we were able to identify 93 genes that are antigenic and preferentially made in the host urinary tract. Included in this list was the putative RTX family member *tosA*, which was determined to play a significant role in the murine model of an ascending UTI.

Aim 2: Determine the distribution of *tosA* among natural isolates of *E. coli* and characterize the genetic factors shared among human UPEC isolates

Previous research has identified that RTX family members, including α-hemolysin, are not universally present in a bacterial species but are often confined to specific subgroups of isolates. We hypothesized that the presence of *tosA* would also be limited to a specific subgroup of *E. coli* and that this group would contain numerous UPEC isolates. Chapter 3 details the results of a large survey of natural *E. coli* isolates for the presence of this gene. In addition, these isolates were also characterized according to the presence of 14 additional genes known or suspected to play an important role in UPEC pathogenesis that are not found among commensal *E. coli* isolates. The results indicated that *tosA*-positive isolates comprised a distinct subset of UPEC that contained many more of these virulence associated traits than fecal isolates.

Furthermore, the presence of *tosA* in the rare fecal isolate that contained the gene was predictive of increased success in the murine model of UTI when compared to fecal isolates that lacked the gene, suggesting that *tosA* could serve as a marker for this group of UPEC. This result held true regardless of the collection source or host species from which the strain was isolated.

Aim 3: Characterize *tosA* expression in the host environment, determine potential functions of TosA, and determine the suitability of TosA as a target for development of future therapeutics

Finally, we investigated the function of TosA during an ascending UTI in a murine model and discovered that TosA is expressed during all stages of an ascending UTI, including the bacteremic state resulting from spread from the urinary tract. We hypothesized that given the lack of homology to known RTX members, TosA might confer unique functions. TosA was shown to play a direct role in mediating adherence to the epithelial cells that line the upper urinary tract, but not the lower urinary tract using *in vitro* assays and confirmed this result with a newly developed *in vivo* bladder epithelial adherence model. A possible secondary role enhancing survival during bacteremia in a murine model of non-lethal disseminated infection and enhanced lethality in a zebrafish model of ExPEC sepsis raises the possibility that TosA may have multiple functions in the different stages of ascending UTI.

## CHAPTER 2

## IDENTIFICATION OF *IN VIVO* INDUCED ANTIGENS INCLUDING A RTX FAMILY EXOPROTEIN REQUIRED FOR UROPATHOGENIC *ESCHERICHIA COLI* VIRULENCE

### INTRODUCTION

The gram-negative bacterium *Escherichia coli* is responsible for 80% of community acquired urinary tract infections (UTI) (121). Uncomplicated UTIs create a significant burden on human health with more than 50% of adult women suffering a UTI during their lifetime (41). An equally large burden is placed on our healthcare system with estimates of direct and indirect costs of $1.6 billion per year in the United States alone (31). While antibiotics have proven to be an effective treatment in the past, a worldwide increase in antibiotic resistance in uropathogenic *E. coli* (UPEC) isolates (1, 12, 16, 25) has created an urgent need for alternative treatment and prevention strategies to combat this important and widespread human pathogen.

In addition to the core genetic backbone of this species, UPEC contain a unique repertoire of genetic material distinct from what is present in commensal strains or other *E. coli* pathotypes (126). This set of accessory genes provides the mechanisms for colonizing the host urinary tract and eliciting the signs and symptoms of an uncomplicated UTI (89). The host urinary tract environment challenges bacteria with an immune response that the bacteria must overcome to persist in the host (106).

Recently, our group correlated antibody production with protection in a mucosal vaccination model using several vaccine candidates that targeted iron acquisition system proteins expressed *in vivo* by human UPEC strains (3). This suggested that eliciting a humoral immune response to UPEC surface proteins, some encoded by specific accessory genetic sequences, may be beneficial in creating novel UPEC vaccines. Clearly, new candidate vaccine targets will be required to deal with the diversity of mechanisms of pathogenesis observed among different UPEC isolates (15).

New insight into the mechanisms of bacterial pathogenesis mediated by proteins synthesized only during the course of an infection can be achieved through the use of *In vivo* Induced Antigen Technology (IVIAT) (49). IVIAT utilizes serum from infected patients or experimentally infected animals that has been adsorbed against bacteria cultured *in vitro*. This process removes antibodies that react to common antigens made during *in vitro* growth and enriches the serum for those antibodies that may recognize antigens only induced *in vivo*. This serum is then used to probe libraries that express genes from the pathogen of interest to identify which genes encode antigenic proteins that are preferentially synthesized during the course of an infection (95). This process has been successfully applied to study a diverse collection of important human pathogens including enterohemorrhagic *E. coli* (EHEC) (61), *Streptococcus pyogenes* (99), and *Mycobacterium tuberculosis* (24, 71). This approach could be particularly effective for addressing urinary tract infections as the proteins identified by IVIAT elicit a host immune response and are expressed only during the course of an infection, both properties expected to be critical for the formulation of an effective UPEC vaccine (105).

In this study, we generated a genomic expression library using DNA from the human pyelonephritis isolate *E. coli* CFT073. This library was screened against antiserum, collected from mice chronically infected with this UPEC strain, that had been adsorbed against bacteria cultured *In vitro*. In total, more than 40,000 clones were screened, resulting in 93 IVIAT-identified genes. *In vivo* expression of all of 13 representative genes was confirmed by qPCR. Five isogenic mutants of CFT073, deficient in five representative systems identified by IVIAT, were constructed and one of these, an osmoprotection system, conferred a fitness advantage when competed against wild type CFT073 in an animal UTI model. Finally, we determined that the gene encoding a repeat-in toxin (RTX) family member, *tosA* (93) [originally annotated as *upxA* (126)], identified here by IVIAT, contributes significantly to the virulence of UPEC. These results not only expand our understanding of the processes that occur during an infection, but also expand the list of candidates for the development of novel therapeutics to combat UTI.

## MATERIALS AND METHODS

**Bacteria and culture conditions.** *E. coli* CFT073, a prototypic UPEC strain, was isolated from the blood and urine of a patient with acute pyelonephritis (84); its genome has been sequenced and fully annotated (126). *E. coli* TOP10 (Invitrogen) and BL21 (DE3) pLysS were used for cloning and as the host expression strain, respectively. Isolated colonies were inoculated into LB medium and cultured aerobically overnight. Culture in human urine was conducted as described previously (4). For library screening, dilutions of the expression strain were plated onto LB agar containing 1 mM IPTG, kanamycin (25 μg/ml), and chloramphenicol (34 μg/ml) and incubated at 30ºC.

**Library construction.** Genomic DNA was purified from uropathogenic *E. coli* strain CFT073 using DNeasy columns (Qiagen) and partially digested using the enzyme *Sau*3AI. Digested DNA was size-fractionated on a 0.9% agarose gel and 1.0-1.5 kb fragments were isolated and purified using a Gel extraction kit (Qiagen). The isolated fragments were ligated into a pool of *Bam*HI-digested pET30a, pET30b, and pET30c plasmids and transformed into competent TOP10 cells. 60,000 colonies from 25 plates were pooled into LB medium and used for plasmid midi-preps (Promega). Purified plasmids were electroporated into BL21 (DE3) pLysS and all transformants, approximately 100,000 colonies from 35 plates, were pooled into 10 ml LB medium containing 35% glycerol and stored in 0.5 ml aliquots at -80ºC.

**Preparation of antisera.** Sera from 20 chronically infected mice (47) were pooled (final volume of 2.75 ml) and adsorbed with CFT073 bacterial cells cultured in LB medium to

an $OD_{600}$ of 0.5 as follows. Bacterial cells were collected from the LB medium by centrifugation, washed in PBS, suspended in PBS, and divided into aliquots containing $10^{11}$ CFU of CFT073. Aliquots were used for adsorption directly (whole cells) or lysed in a French pressure cell (whole cell lysates) as described previously (5). Pooled sera were adsorbed with aliquots of whole cells on a rocking platform at 4ºC for 1h. Unadsorbed antibodies were collected from the supernatant following centrifugation and adsorption was repeated two additional times with CFT073 cells and three times with BL21 cells. Following adsorption to whole cells, antibodies were incubated with nitrocellulose strips saturated with 1 mg whole cell lysate from CFT073 on a rocking platform at 4ºC for 1h. Soluble antibodies were recovered and applied to additional whole cell lysate three times for CFT073 followed by two adsorptions using whole cell lysate from BL21. Additional whole cell lysate was electrophoresed on a 12% SDS-polyacrylamide gel and transferred to PVDF to prepare denatured whole cell extracts for absorption. Adsorption was continued using denatured whole cell extracts as described for a total of five adsorptions to denatured extracts from CFT073 and three adsorptions to denatured extracts from BL21. Aliquots of the final adsorbed sera were stored at -20ºC.

***In vivo* induced antigen screening.** IVIAT was performed as described (49) with the following modifications. Aliquots of BL21 (DE3) pLysS containing the expression library were thawed, diluted, and plated onto LB medium containing kanamycin to obtain 300-400 colonies per plate. Master plates were incubated at 30ºC for 16h before transfer onto nitrocellulose filters. Filters were placed colony side up on LB agar containing 1mM IPTG and kanamycin (25 µg/ml) and incubated at 37ºC for 3-4h. Following

induction, nitrocellulose filters were removed and colonies were lysed in chloroform vapor. Colony-immunoblotting was performed using the pooled adsorbed sera as primary antibodies. Reactive colonies were detected using goat anti-mouse HRP secondary antibodies and visualized using chemiluminescent substrate (Pierce). Clones identified during primary screening were picked from the master plate and subjected to secondary screening by patching, in an alternating pattern, with BL21 vector control. Secondary screens were carried out as described for the primary screen. Clones that maintained reactivity on two of three plates were considered positive identifications and subjected to DNA sequencing using T7 primers that flank the multiple cloning site within the vectors.

**RNA extraction and quantitative real-time PCR.** For preparation of RNA, bacteria were collected by centrifugation from overnight cultures, washed with sterile PBS, and $10^6$ CFU were used to inoculate pre-warmed LB medium or human urine. Cultures were incubated statically at 37ºC until exponential phase ($OD_{600}$=1.0 for LB medium or $OD_{600}$=0.35 for human urine. Following incubation, 1 ml of bacterial culture was added to 0.125 ml of ice cold phenol-ethanol stop solution (5% phenol in ethanol), and bacteria were collected by centrifugation. Cells were lysed and RNA was extracted using the RNeasy kit (Qiagen) following the manufacturer's protocol. Following elution, nucleic acid concentrations were determined by spectrophotometry (NanoDrop) and residual DNA contamination was removed by incubating the samples with 4U of TURBO DNase (Ambion). After DNase inactivation, RNA was recovered, quantified, and used as a template for PCR to confirm inactivation of contaminating DNA. The PCR-negative RNA

was used for first strand cDNA synthesis using SuperScript II reverse transcriptase (Invitrogen). For cDNA synthesis, the manufacturer's protocol was followed, starting with 1.35 µg total RNA template and 50 ng random hexamers. Following synthesis, the cDNA was purified on a QIAquick column (Qiagen), quantified using the NanoDrop, and diluted to 5 ng/µl with nuclease-free water. RNA transcripts were quantified on an MX3000P real-time PCR machine (Stratagene) using Brilliant SYBR green QPCR mix (Stratagene) in 25 µl volumes containing 25 ng cDNA. Optimal primer concentrations were determined empirically. For comparative, quantitative analysis, transcript levels were normalized to the level of *gapA* (glyceraldehyde 3-phosphate dehydrogenase A) and changes were determined using an experiment-specific calibrator (LB) using MXPro v 3.00 software package (Stratagene). For *in vivo* quantitative PCR, RNA was purified as described from pooled urine from groups of mice experimentally infected with CFT073 to determine transcript levels during UTI.

**Construction of mutants in UPEC strain CFT073.** Deletion mutants were generated using the lambda red recombinase system (23). Primers homologous to sequences within the 5' and 3' ends of the target genes were designed and used to replace target genes with a nonpolar kanamycin resistance cassette derived from the template plasmid pKD4 (23). Kanamycin (25 µg/ml) was used for selection of all mutant strains. Gene deletions begin with the start codon and end with the stop codon for each gene. To determine whether the kanamycin resistance cassette recombined within the target gene site, primers that flank the target gene sequence were designed and used for PCR. After amplification, each PCR product was compared to wild-type PCR product and in

cases where size-differences are negligible; PCR products were digested with the restriction enzyme *Eag*I (New England Biolabs). Both the PCR products and restriction digests were visualized on a 0.8% agarose gel stained with ethidium bromide.

**Experimental UTI.** The murine model of ascending UTI was used to assess virulence and fitness (48). Six-to eight-week-old female CBA/J mice (20 to 22 g; Jackson Laboratories) were transurethrally inoculated with a 50 µl bacterial suspension per mouse using a sterile polyethylene catheter (I.D. 0.28 mm x O.D. 0.61 mm) connected to an infusion pump (Harvard Apparatus). For independent challenge, overnight LB cultures of CFT073 and an isogenic mutant were collected by centrifugation and resuspended in sterile PBS and adjusted to deliver 2 x $10^8$ CFU per mouse. To measure relative fitness of mutant and wild-type, overnight LB cultures for CFT073 and individual mutant strains were collected by centrifugation and resuspended in sterile PBS, mixed 1:1 and adjusted to deliver 4 x $10^8$ CFU per mouse. Dilutions of each inoculum were spiral plated onto LB with and without kanamycin using an Autoplate 4000 (Spiral Biotech) to determine the input CFU/mL. After 48 hr post-inoculation, mice from independent challenge or co-challenge were sacrificed by overdose with isoflurane and the bladder and kidneys were aseptically removed, weighed, and homogenized in sterile culture tubes containing 3 ml of PBS using an OMNI mechanical homogenizer (OMNI International). Appropriate dilutions of the homogenized tissue were then spiral plated onto duplicate LB plates with and without kanamycin to determine the output CFU/g of tissue and difference in colonization were determined using the Mann Whitney significance test. For co-challenge, plate counts obtained on kanamycin were subtracted from those on plates

lacking antibiotic to determine the number of wild-type bacteria. Statistically significant differences in colonization (*P*-value <0.05) were determined using a two-tailed Wilcoxon matched pairs test. All animal protocols were approved by the University Committee on Use and Care of Animals at the University of Michigan Medical School.

## RESULTS

**Pooled and adsorbed serum from chronically infected mice does not react with uropathogenic *E. coli* cultured *in vitro*.** Serum, collected from 20 female CBA/J mice transurethrally inoculated with UPEC isolate CFT073 in a chronic UTI infection model, was collected as previously described (47). Serum was adsorbed against whole cells of strain CFT073 cultured *in vitro* and then adsorbed against native and denatured whole cell lysates of strain CFT073 immobilized on a membrane in order to expose the serum to both surface-exposed and non-exposed proteins synthesized during *in vitro* growth. This adsorption process was repeated with host expression strain *E. coli* BL21 (DE3) pLysS that carried the genomic expression library used for the IVIAT screening. Multiple rounds of adsorption were conducted until colony immunoblots of both bacterial strains registered negligible background levels of reactivity with the treated serum when compared to unabsorbed serum (Figure 2- 1 A-C).

**Screening of a CFT073 genomic expression library with the adsorbed serum identified candidate *in vivo*-expressed genes.** Fragments of DNA (1.0 to 1.5 kb) were size-fractionated from genomic DNA isolated from the UPEC isolate CFT073, ligated into an IPTG-inducible vector that permits cloning in all three possible open reading frames and transformed into *E. coli* BL21 (DE3) pLysS. The CFT073 genome contains 5,079 predicted protein coding sequences; to achieve adequate representation, it is desirable to obtain 50,000 clones with insert sizes of 1.0 kb. After multiple transformations and selection, we obtained 60,000 clones, with 24 of 24 representative clones having inserts greater

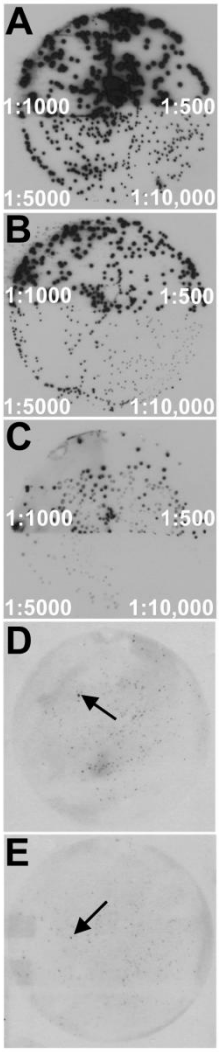



**Figure 2-1. Identification of antigens induced by UPEC *in vivo*.** Colony western blots of *E. coli* CFT073 colonies using (A) pooled sera from chronically infected mice, (B) adsorbed sera, or (C) BL21 pET30a probed with adsorbed sera. Representative primary screening results showing reactive BL21 clones expressing CFT073 *in vivo* induced antigens are indicated with arrows in (D) and (E). Dilutions of primary sera are indicated in panels (A-C).

than 1.0 kb. The resulting genomic expression library is predicted to cover most of the repertoire of protein antigens contained in the CFT073 genome.

Primary screening of the adsorbed serum was conducted using colony immunoblots of the genomic expression library. 461 clones that exhibited reactivity significantly above background (Figure 2-1 D and E) were isolated and subjected to a second round of screening. Three replicate colony immunoblots were conducted during secondary screening to minimize false positives. A total of 83 clones remained reactive to the adsorbed serum in at least two of the three replicate screens. Plasmids from these clones were isolated and sequenced to identify the gene fragments identified by the IVIAT screening (Table 1). Ten inserts contained DNA that spanned two genes, bringing the total number of candidate *in vivo*-expressed genes to 93.

A subset of the IVIAT-identified genes has known roles during UPEC colonization of the urinary tract and was not further investigated in this study. *iucC* and *fhuC*, both involved in hydroxymate sideophore transport, were identified in our screen, which is consistent with iron acquisition being a key virulence trait for UPEC during UTI. NanA, involved in the transport and catabolism of sialic acid was previously shown to be up-regulated in human urine, however a mutant had no apparent fitness defect during UTI (5). Interestingly, this same mutant strain demonstrated a significant fitness defect *in vivo*, in a bacteremia model (109). Thus, finding these antigens using IVIAT suggests other genes identified in the screen may also have important roles for UPEC pathogenesis.

**Validation of *in vivo* expression pattern by qPCR of bacterial RNA from pooled urine from the murine model.** To verify that the adsorbed serum was identifying genes that were expressed *in vivo* but not *in vitro*, 13 of the genes identified in this study were selected for validation of *in vivo* expression by qPCR (Table 1). The genes: *yddQ*, *proWX*, *kpsS*, *lolA*, *lolD*, *tosA* (*upxA*), *adhP*, c2432, *katG*, *narJI*, and *ydhX*, were selected as representative genes from six categories including those involved in transport, regulation, secreted proteins, metabolism, cellular processes, and hypothetical function. The predicted dipeptide transport component *yddQ* was selected because known peptide transport genes *dppA* and *oppA* are required for UPEC fitness *in vivo* (5). Similarly, the role of osmo-protection within the urinary tract (*proWX*) and capsule (*kpsS*) have been previously shown to be important for UPEC pathogenesis *in vivo* (9, 22). The remaining genes selected for qPCR validation represent putative or known metabolism or transport genes that have unknown contributions for UPEC virulence or fitness during UTI.

Pooled urine from three replicate groups, each containing five female CBA/J mice that had been transurethrally inoculated with strain CFT073, was collected three times a day for three days post inoculation. RNA was isolated from bacteria from within each pooled urine sample for qPCR analysis and data were compared to results from RNA extracted from strain CFT073 cultured *in vitro* under the same conditions used initially to adsorb the serum. The results demonstrated that each of the 13 genes were more highly expressed *in vivo* than in bacteria cultured under standard laboratory conditions (Figure 2-2A). The greatest increase in expression *in vivo* over *in vitro* was an

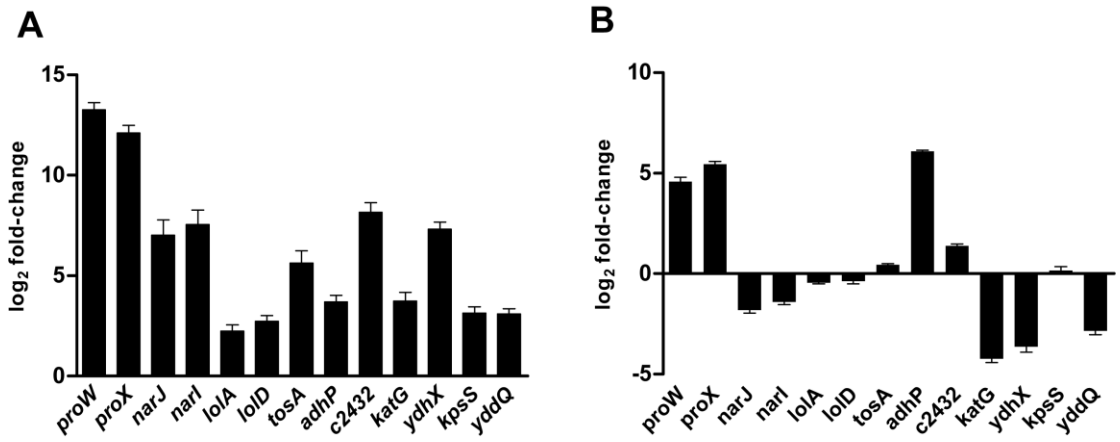



**Figure 2-2. Quantitative real-time qPCR for transcripts of UPEC *in vivo* induced antigen genes during experimental UTI or during growth in human urine.** (A) qPCR results for genes expressed by CFT073 during ascending UTI or (B) during growth in human urine. Data were normalized to *gapA* (glyceraldehyde 3-phosphate dehydrogenase) expression levels and changes were determined using expression levels during growth *in vitro* in LB medium as the calibrator to generate fold-change values. The bars represent $\log_2$ fold-changes in (A).

increase of >2000-fold for the *proW* and *proX* genes. With the exception of *lolA* and *lolD*, the increase in expression was always ≥4-fold for all 11 other genes examined by qPCR. This demonstrates the ability of the IVIAT screening approach to identify genes that are expressed many fold higher *in vivo* over expression observed under standard *in vitro* culture conditions.

Bacteria incubated in pooled filter-sterilized human urine, a condition that may mimic *in vivo* conditions, increased expression of 4 of the 13 genes examined by qPCR by more than two-fold over growth in LB (*proW, proX, adhP,* and c2432) (Figure 2-2B). The remaining 9 genes were expressed at comparable or lower levels than in bacteria cultured aerobically to exponential phase ($OD_{600}$=1.0) in LB. *katG*, *ydhX*, and *yddQ* were significantly down-regulated in human urine compared to LB. While this specialized growth medium more closely resembled expression patterns *in vivo*, the IVIAT approach was clearly more successful in identifying genes that are specific for growth *in vivo* during the course of an infection and identified many genes that would have been overlooked by methods relying strictly on *in vitro* growth to mimic *in vivo* conditions.

**IVIAT identified the putative virulence gene *tosA*.** Five isogenic mutant strains of CFT073 were constructed using lambda red recombinase methodology (*tosA, proWX, katG, adhP, narJI*) (23) and tested for colonization in a murine model of an ascending UTI. Of the five mutants under study, *narJI* and *tosA*, had the greatest *in vivo* expression levels by qPCR despite little or no up-regulation in human urine, *proWX* demonstrated the highest *in vivo* expression, *katG* was upregulated *in vivo* and down-regulated in human urine, and *adhP* displayed substantial up-regulation both *in vivo* and in human

urine (Figure 2-2). CFT073 mutants in each of these genes were competed against wild-type CFT073 in co-challenge infections or tested in independent challenge in the murine model of ascending UTI. CFU levels in bladder and kidney were quantified for each strain at 48 hours post inoculation (Figure 2-3 and 2-4).

In a previous study, our group reported on the contribution of the 100-kb pathogenicity associated island $PAI_{CFT073}$-*aspV* containing the RTX family member *tosA* to fitness in the murine urinary tract (77). In the current study, IVIAT independently identified TosA as an antigen that is preferentially expressed *in vivo* (Table 2 and Figure 2-2A). *tosA* expression was determined to be specific to the *in vivo* environment and expression above background levels was not observed when bacteria were cultured *in vitro* in pooled filter-sterilized human urine (Figure 2-2B). A Δ*tosA* mutant of CFT073 was tested in the murine model of ascending UTI and colonization levels in the bladder and kidneys were compared to colonization levels observed when the parental wild type strain was independently tested (Figure 2-3). The Δ*tosA* mutant was severely attenuated in the animal model of infection; significantly lower numbers of bacteria were recovered from both bladder ($P$=0.0016) and kidney ($P$=0.0120) tissue compared to the wild-type strain. Thus, the IVIAT approach employed in this study identified a new virulence factor of UPEC that is specifically induced *in vivo* and contributes significantly to the pathogenesis of this strain of *E. coli* in an animal model of urinary tract infection.

***proWX*, identified by IVIAT contributes to fitness in a murine model of UTI**. The *proWX* genes are involved in osmoregulation by encoding the proline betaine transport system. An isogenic mutant that contained a single deletion covering the *proW a*nd *proX* genes,

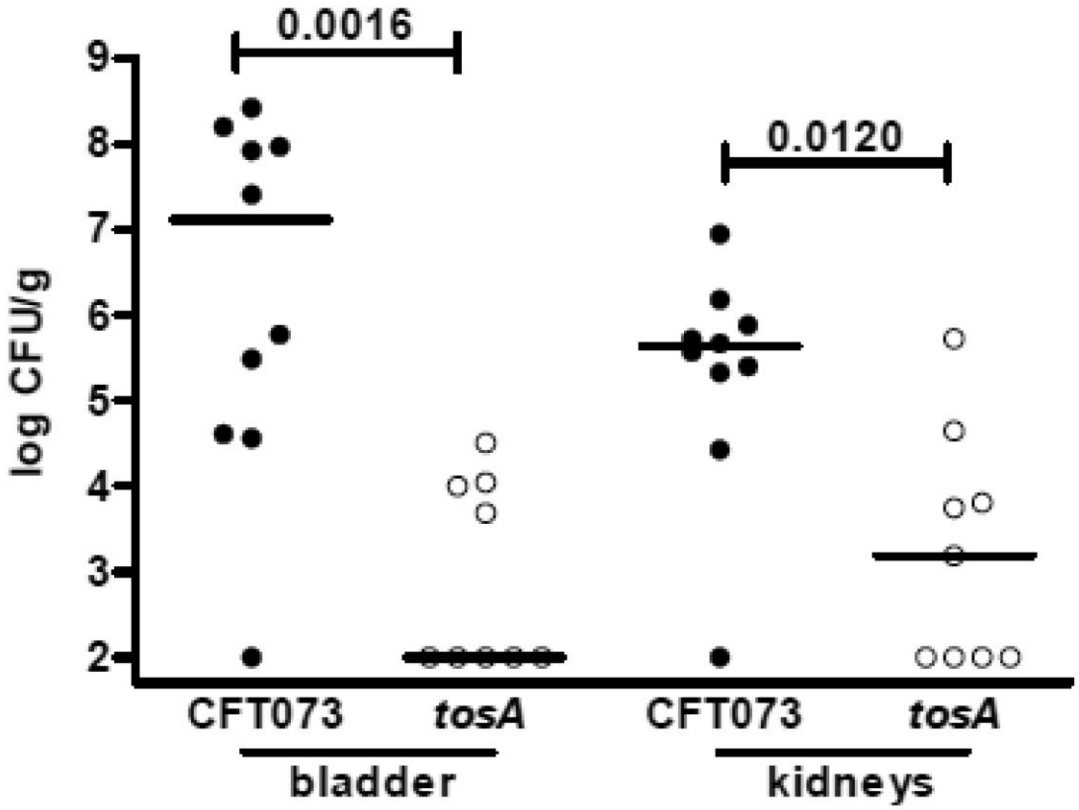



**Figure 2-3. A putative RTX family exoprotein gene, *tosA*, is required for virulence in UPEC.** Bladder and kidney colonization levels at 48 h post-transurethral inoculation following independent challenge with wild-type uropathogenic *E. coli* CFT073 and its Δ*tosA* mutant. Circles represent the $\log_{10}$ CFU/g from individual mice and horizontal bars represent the median CFU/g. *P*-values were determined using the non-parametric Mann Whitney significance test.

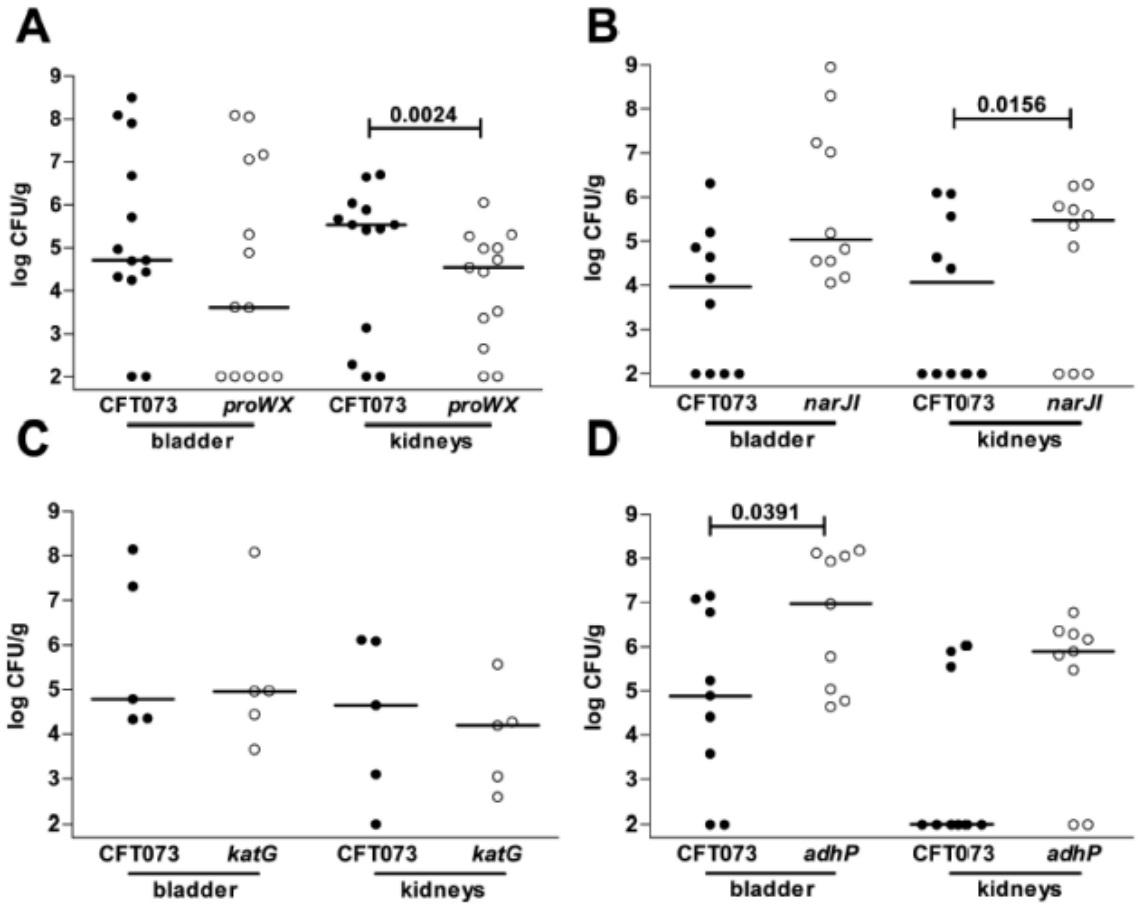



**Figure 2-4. Assessment of fitness for select *in vivo* induced antigen mutants using co-challenge infections.** Individual female mice were transurethrally inoculated with 2 x $10^8$ CFU of a 1:1 mixture of wild-type and mutant bacteria. *In vivo* fitness at 48 hr post inoculation for UPEC mutants lacking: (A) *proWX*, (B) *narJI*, (C) *katG*, or (D) *adhP*. At 48 hpi, bladders and kidneys were aseptically removed, homogenized, and plated on LB or LB containing kanamycin to determine viable counts of wild-type and mutant strains, respectively. In all cochallenges depicted, 37 of 37 mice had detectable CFU of wild-type and/or mutant in the bladder; all but 5 of 37 mice had detectable CFU of wild-type and/or mutant in the kidneys ((2 in *adhP*, 1 in *proWX*, *2 in narJI*). Each filled circle represents the $\log_{10}$ CFU/g from an individual animal. Bars represent the median CFU/g and the limit of detection is 100 CFU. Significant differences in colonization levels ($P$<0.05) are indicated and were determined using a two-tailed Wilcoxon matched pairs test.

corresponding to the region of the genome contained in a single expression clone identified by IVIAT, was outcompeted by the parental strain in the kidney (*P*=0.0024) (Figure 2-4A). The data from this co-challenge illustrate that the IVIAT screen may identify genes and systems that confer an advantage during UPEC colonization of the urinary tract.

The mutant lacking catalase encoded by *katG* was recovered in numbers equal to wild-type (Figure 2-4B) and thus did not confer a fitness advantage and could not be considered a virulence factor. Surprisingly, the Δ*adhP* mutant lacking an alcohol peroxidase (Figure 2-4C), which was very highly expressed in human urine (Figure 2-2B), and the Δ*narJI* mutant lacking an anaerobic nitrate reductase I (Figure 2-4D) were both recovered in significantly higher numbers than wild-type in the bladder (*P*=0.0391) and kidneys (*P*=0.0156), respectively. While each of these genes was properly identified by the IVIAT approach as being expressed at higher levels *in vivo* (Figure 2-2A) and eliciting a host immune response, these results indicate that not all of the identified genes confer a fitness advantage in the host environment. Taken together with the other animal model results, the remaining IVIAT-identified genes would need to be studied independently in greater detail to understand the importance of each gene to the pathogenesis of a UTI.

**DISCUSSION**

Two of the genes identified by the IVIAT screening approach employed in this study, a putative secreted RTX family exoprotein (*tosA*) and an osmoprotection system that transports glycine betaine (*proWX*), contribute to fitness of UPEC in the host urinary tract (Figure 2-3 and 2- 4). In addition, more than 90 additional genes were identified by IVIAT (Table 1) and many of these may also contribute to pathogenesis of UTI. Of particular note is the observation that out of the 93 identified genes, 23 genes have not been previously studied and were assigned a function based only on sequence homology to known proteins. Seventeen additional genes did not demonstrate sufficient homology to be assigned function. The putative RTX family member *tosA* (originally annotated as *upxA* (126)), which was determined in this study to contribute significantly to the ability of the human pyelonephritis isolate CFT073 to colonize a murine UTI model (Figure 2-3), is just one example from this group of genes of unknown function. Because the IVIAT method relies on antibodies made by the host during an infection (95), this list also represents UPEC antigens that the host immune system recognized and against which an immune response was mounted. While additional studies will be needed to clarify the role of those genes not pursued further in this study, this list may contain useful targets for the development of a UPEC vaccine.

Several limitations of the project must be addressed to clarify the results of this study. Due to the limited quantity of murine antiserum available, only 40,000 clones of the genomic expression library were screened. Because our group calculated that ~50,000 clones would be needed to obtain 95% genome coverage, Table 2 represents

only a partial but nevertheless substantial list of genes preferentially expressed *in vivo* and their proteins that are recognized as antigenic by the host. In addition, for a gene product to be recognized by IVIAT, the antibody titer must be sufficiently high to obtain a consistently strong signal on immunoblot in both primary and secondary screening. Next, any proteins produced during *in vitro* growth, even those with known contributions to fitness in the urinary tract, would be available to remove specific antibodies during the serum adsorption process and are not predicted to be recognized by IVIAT. Finally, there was difficulty constructing a stable plasmid containing the 4.8 kb *tosA* gene, and thus complementation was not confirmed in the independent challenge (Figure 2-3). These limitations may explain why many well characterized UPEC virulence factors, such as P fimbriae, which are known to be expressed during *in vitro* growth on LB agar (21), were not identified by this screen.

Despite these caveats, the results of the IVIAT screen clarify and reinforce previous observations about general metabolic processes that UPEC employ to colonize the host urinary tract. Our findings also revealed some overlap with a similar screen in enterohemorrhagic *E. coli* (EHEC) carried out using convalescent sera from patients following recovery from hemolytic uremic syndrome (61). Both the current results and the EHEC study identified *adhP*, osmolarity homeostasis genes, components of the maltose regulon, antigens representing enzymes involved in acetyl-CoA metabolism (*accC*), and hydrogenase and additional respiratory components involved in the reduction of nitrate (61). It is not surprising that these core, or backbone components of the *E. coli* chromosome were identified in both screens. For UPEC, finding respiratory

components as *in vivo* induced antigens is consistent with the fitness requirement for the TCA cycle during UTI (5) and utilization of alternate electron acceptors during "over-flow" metabolism that occurs from rapid growth *in vivo* (46).

While many of the processes identified by IVIAT may be active during the course of a UTI, not all of the systems studied in the murine model demonstrated a clear fitness advantage. For example, isogenic mutants of CFT073, deficient in the genes *adhP*, *katG*, and *narJ* and *narI,* were all successful in competing against wild type CFT073 during co-challenge experiments (Figure 2-4). Each of these genes were determined to be upregulated *in vivo* compared to growth in LB medium, and expression of the genes *narJ*, *narI*, and *katG* were specific to the *in vivo* environment and were not induced during growth in pooled filtered human urine. While the results of the Δ*katG* strain were unexpected, precedence exists in descriptions of full virulence of catalase-defective strains of *Neisseria gonorrhoeae* and *Listeria monocytogenes* in animal models of infection (72, 111). It is not clear why these mutants were able to compete equally with wild type bacteria, but this does illustrate the importance of following up the IVIAT screen with specific experiments to determine the relative importance of the genes identified in a large genome wide screen such as IVIAT to the fitness of an organism during an infection.

The genes *proW* and *proX*, part of the *proU* osmoprotection system, increase fitness of the human pyelonephritis isolate CFT073 in the kidneys of infected mice (Figure 2-4). A trend toward lower CFU of the Δ*proWX* mutant was also observed in the bladder on co-challenge with wild-type CFT073. Previous research had identified the

major substrates of the *proU* system in *E. coli* as glycine betaine and proline betaine (43). Both compounds are present in human urine and the presence of either compound increases the tolerance of *E. coli* to higher concentrations of osmotic agents than observed without the compounds present (17), suggesting an important role for the transport of these compounds as an adaptation by UPEC to survive the environment of the host urinary tract. Additional studies on the role of the *proU* system in the human pyelonephritis strains CFT073 and HU734 found Δ*proU* mutants of each strain colonized the murine model of an ascending UTI at the same levels as wild type in both bladder and kidneys of infected mice (22). Additional osmoprotection systems have been identified in *E. coli* (22, 132) and this functional redundancy may help explain why a strain of UPEC that lacks the *proU* system can still efficiently colonize the urinary tract. In contrast to the current study that mixed wild type CFT073 at an equal ratio to mutant and co-inoculated the two strains in the same animal, previous experiments were conducted as independent challenges of each strain's ability to colonize infected mice in separate infections. Indeed, we consider co-challenge a more sensitive index of fitness than independent challenge. When considered together, the results of this study help to clarify the role of the *proU* osmoprotection system during an experimental UTI. While this system is not necessary for colonization, the expression of the system is strongly induced in both an *in vivo* infection model and growth *in vitro* in human urine, and the presence of the system contributes enhanced fitness of UPEC in the kidneys of infected animals.

The IVIAT-identified gene *tosA* (*upxA*) was originally described in the genome annotation of CFT073 as a UPEC-specific gene not present in the genomes of commensal *E. coli* or other *E. coli* pathotypes (126). The 4.8-kb gene appears as part of a 9.5-kb locus on the $PAI_{CFT073}$-*aspV* that appears to be linked to a type I secretion system (93). While the components of the type one secretion system are well conserved with other RTX family members, the predicted secreted target of the system, TosA, is poorly conserved with other known RTX family members (93). Further differences from the *hlyA* locus are noted in the organization of the locus both in gene order and gene content in CFT073 (126), suggesting that *tosA* may function differently from *hlyA*. In a previous study, our group had identified $PAI_{CFT073}$-*aspV* containing the gene *tosA* as contributing enhanced fitness to CFT073 in an animal model of infection (77). In the current study, the IVIAT screening approach independently identified *tosA* as an antigenic protein synthesized in the host urinary tract.

Expression of the RTX family member gene *tosA* was determined to be specific to the *in vivo* environment of the host urinary tract. Unlike the genes of the *proU* system, *tosA* expression was only measurable in bacteria that had been growing *in vivo* (*i.e.*, from the urine of infected mice). Expression of *tosA* was not induced under any *in vitro* condition studied, including growth in human urine (Figure 2B). This limited expression pattern stands in contrast to *hlyA*, the gene that encodes the prototype RTX toxin α-hemolysin, which is expressed during *in vitro* growth in LB medium (73), the same conditions utilized for the RT-qPCR study presented in Figure 2-2B. While the UPEC strain CFT073 contains the two RTX family members *tosA* and *hlyA*, the regulatory

mechanisms governing expression of each system appear to be different. This raises the possibility that a specific host factor, not present in human urine, may trigger the expression of *tosA* observed in this study.

A Δ*tosA* strain of CFT073 was attenuated in both the bladder and kidneys of infected mice following independent challenge (Figure 2-3). While our previous study identified a role for *tosA* in enhancing fitness in the murine model of infection during co-challenge with wild type bacteria (77), the results of the current study indicate the gene serves a more critical function during infection. In independent challenge, approximately half of mice infected with the Δ*tosA* strain had undetectable levels of bacteria in bladder and kidney tissue ($\leq 10^2$ CFU/g tissue) as opposed to only 1 of 10 mice infected with the wild-type strain ($P$=0.043). In addition to the gene regulation differences, this result also indicates a role for *tosA* that is different from *hlyA*. While Δ*hlyA* mutants of UPEC are attenuated in animal models of infection outside of the urinary tract, notably in several bacteremia models (15, 127), no defect in colonization levels of bladder or kidney tissue were observed in a similar murine UTI model (110). Both in the urinary tract and during *in vitro* assays, α-hemolysin mediates damage to host cells (110, 117) but is not necessary for successful colonization of the host urinary tract. Additional studies will be necessary to elucidate the function of *tosA*, but the results of this study implicate a different and important role for this novel RTX family member during a UTI mediated by *E. coli*.

The results of this study demonstrate the utility of the IVIAT approach to clarifying the role of genes with known functions (*e.g.*, *proU* system) as well as

identifying genes of unknown function (*e.g.*, *tosA*) that play an important role in the success of bacterial pathogens while in the host environment. The results of the IVIAT screen hold particular importance for the development of new experimental vaccines as the genes identified are known to be preferentially expressed *in vivo* and elicit a host immune response in the formation of the antibodies utilized by the IVIAT screen. While many mechanisms of UPEC pathogenesis have been well characterized, the results of the IVIAT screen presented here indicate that much remains to be learned about the specific mechanisms utilized by *E. coli* during a UTI.

**Table 2-1.** Genes identified by IVIAT as induced in uropathogenic *E. coli* during experimental urinary tract infection

| | Gene | ORF | Function |
|---|---|---|---|
| **Transport** | | | |
| | *yddQ** | c5079 | Hypothetical dipeptide ABC transporter, permease subunit |
| | *proW**-*proX** | c3231-3232 | glycine betaine transporter |
| | *yjcQ* | c5086 | multidrug efflux system protein MdtO |
| | *kpsS** | c3691 | capsule polysaccharide export |
| | *cirA* | c2690 | ferric iron-catecholate outer membrane receptor |
| | *fliY* | c2335 | cystine transporter subunit |
| | *lolD** | c1392 | lipoprotein transporter ATP-binding subunit |
| | *lolA** | c1028 | outer-membrane lipoprotein carrier protein |
| | *ybeX* | c0743 | Magnesium and cobalt efflux protein corC |
| | *sapB* | c1770 | Peptide transport system permease protein sapB |
| | *btuD-btuE* | c2105-2106 | vitamin B12-transporter |

| | | |
|---|---|---|
| *malM-lamB* | c5007-5006 | maltose regulon periplasmic protein; maltoporin |
| *iucC* | c3625 | aerobactin siderophore synthesis |
| *fhuC* | c0186 | iron-hydroxamate transporter ATP-binding subunit |
| | c4487 | Putative fructose-specific phosphotransferase system |
| *ydjK* | c2179 | Hypothetical metabolite transport protein |
| *yphD* | c3068 | Hypothetical ABC transporter permease |
| *ydiM* | c2085 | Hypothetical transport protein |
| *yieO* | c4682 | Hypothetical transport protein |
| *ydhP* | c2051 | Hypothetical transport protein |
| *ybhR* | c0875 | hypothetical ABC transporter |
| *yijE* | c4901 | Hypothetical transport protein |
| *yajQ-yajR* | c0537 | hypothetical transporter |
| *yehZ* | c2661 | hyptherical transporter |
| **Regulatory** | | |
| *fhlA* | c3291 | Formate hydrogenlyase transcriptional activator |

| | | | |
|---|---|---|---|
| | *cytR* | c4887 | DNA-binding transcriptional regulator |
| | *hupB* | c0556 | transcriptional regulator HU subunit beta |
| | *yaeG* | c0199 | carbohydrate diacid transcriptional activator CdaR |
| | *nlp* | c3946 | DNA-binding transcriptional activator of maltose metabolism |
| | *yijO* | c4913 | Hypothetical transcriptional regulator |
| | *ybiH* | c0879 | putative DNA-binding transcriptional regulator |
| **Secreted** | | | |
| | *tosA**[a] | c0363 | Putative RTX family exoprotein A gene |
| **Metabolism** | | | |
| | *accA-ldcC* | c0223-0224 | acetyl-CoA carboxylase; constitutive Lysine decarboxylase |
| | *adhP** | c1911 | alcohol dehydrogenase |
| | *nanA-nanT* | c3979-3978 | N-acetylneuraminate lyase; sialic acid transporter |
| | *phnJ* | c5104 | phosphate metabolism |
| | | c2432* | Putative thioesterase |

| | | |
|---|---|---|
| *aroG* | c0830 | phospho-2-dehydro-3-heoxyheptonate aldolase |
| *hemA* | c1668 | glutamyl-tRNA reductase |
| *citC* | c0709 | [Citrate [pro-3S]-lyase] ligase |
| *cdd* | c2675 | cytidine deaminase |
| *celA* | c1955 | 6-phospho-beta-glucosidase |
| *astD* | c2146 | arginine succinyltransferase |
| | c2147 | succinylglutamic semialdehyde dehydrogenase |
| *fdhF* | c5625 | Formate dehydrogenase H |
| *adhE* | c1705 | Acetaldehyde dehydrogenase |
| *metH* | c4976 | B12-dependent methionine synthase |
| *asnA* | c4672 | asparagine synthetase AsnA |
| *yphC* | c3067 | Hypothetical zinc-type alcohol dehydrogenase |
| *ygjK* | c3838 | predicted glycosyl hydrolase |
| *rhaB* | c4853 | rhamnulokinase |
| *phnM* | c5101 | carbon-phosphate lyase |
| *accC* | c4012 | acetyl-CoA carboxylase biotin carboxylase subunit |

| | | |
|---|---|---|
| *prsA* | c1665 | ribose-phosphate pyrophosphokinase |
| *nudF* | c3780 | ADP-ribose pyrophosphatase NudF |
| *katG** | c4900 | Peroxidase/catalase |
| *ispB* | c3945 | octaprenyl diphosphate synthase |
| *ivbL-ivbB* | c5498;4596 | ilvB operon leader peptide; acetolactate synthase catalytic subunit |
| *hyaF* | c1118 | Hydrogenase-1 operon protein |
| *hypE* | c3290 | Hydrogenase isoenzyme formation protein hypE |
| *narJ*-narI** | c1687-1688 | Respiratory nitrate reductase 1 |
| *yhjA* | c4329 | Probable cytochrome C peroxidase |
| *pnp* | c3920 | polynucleotide phosphorylase/polyadenylase |
| *wecC-wecB* | c4707-4706 | UDP-N-acetyl-D-mannosamine dehydrogenase; UDP-N-acetylglucosamine 2-epimerase |
| *prfC* | c5456 | peptide chain release factor 3 |
| *rpsS* | c4083 | 30S ribosomal protein S19 |
| *ycfH* | c1372 | predicted metallodependent hydrolase |

| | | |
|---|---|---|
| *holB* | c1371 | DNA polymerase III subunit delta' |
| *yebU* | c2244 | rRNA (cytosine-C(5)-)-methyltransferase RsmF |
| *fusA* | c4112 | elongation factor G |
| *fhiA* | c0377 | FhiA protein (FhlA homolog) |
| *yegD* | c2596 | putative chaperone |
| *ygbJ* | c3297 | Hypothetical oxidoreductase |
| *yajO* | c0530 | Hypothetical oxidoreductase |
| *ydhX** | c2065 | Putative ferredoxin-like protein |
| *yjgB* | c5370 | Hypothetical zinc-type alcohol dehydrogenase-like protein |

**Hypothetical**

| | | |
|---|---|---|
| *yahJ* | c0445 | hypothetical deaminase |
| | c1590/3154 | phage related |
| *yedE* | c2344 | predicted inner membrane protein |
| | c5347-5348 | hypothetical protien; hypothetical protein |
| | c3664 | Hypothetical protein |

| | | |
|---|---|---|
| *yfeK* | c2954 | Hypothetical protein |
| | c2770 | Hypothetical protein |
| | c4743 | Hypothetical protein |
| *yqiB* | c3779 | Hypothetical protein |
| | c3293-91 | Hypothetical protein |
| | c0446 | Hypothetical protein |
| *yjjJ* | c5470 | Hypothetical protein |
| *aroM* | c0498 | Hypothetical protein |
| *yjiN* | c5419 | Hypothetical protein |
| *yedJ* | c2381 | Hypothetical protein |
| *yicH* | c4480 | Hypothetical protein yicH |
| *ycfC* | c1511 | Hypothetical protein |

*****Genes selected for qPCR validation of expression levels *in vivo*.

[a]also known as *upxA* (126).

## CHAPTER 3

## PUTATIVE REPEAT-IN-TOXIN (RTX) GENE TOSA OF *ESCHERICHIA COLI* PREDICTS SUCCESSFUL COLONIZATION OF THE URINARY TRACT

### INTRODUCTION

The majority of urinary tract infections (UTI) in otherwise healthy individuals are caused by uropathogenic *Escherichia coli* (UPEC) (121). This unique group of *E. coli* can reside in the lower gastrointestinal tract of healthy adults (33, 137), but upon entry into the urinary tract can ascend to and colonize the bladder causing cystitis. The infection may be confined to the bladder, or bacteria may ascend the ureters to infect the kidneys causing pyelonephritis. In severe cases, bacteria can further disseminate across the proximal tubular epithelium and capillary endothelium to the bloodstream causing bacteremia (85). A significant proportion of UTIs occur in patients with no known abnormalities of the urinary tract, so-called uncomplicated UTIs. However, certain host characteristics, such as the presence of an indwelling Foley catheter or congenital defect in urinary tract anatomy, are considered complicating factors for UTI, and increase susceptibility to this infection, as well as affecting diagnosis and management (31). Finally, colonization of the bladder in high numbers (>$10^5$ CFU/ml of voided urine) may occur without eliciting symptoms from the host, a condition known as asymptomatic bacteriuria (ABU) (57).

As a species, *E. coli* demonstrates significant genome plasticity, possessing a well conserved core genome, as well as highly variable and heterogeneous accessory genetic material apparently acquired through multiple horizontal gene transfer events (126). Among UPEC, individual mechanisms of pathogenesis have been well characterized (89); however, no single fixed set of virulence factors has been observed in the majority of clinical isolates. Unlike the intestinal pathogenic *E. coli*, distinct pathotypes of UPEC have yet to be defined (82). Without a distinct set of markers to differentiate UPEC from other types of *E. coli*, most of our knowledge about this important human pathogen comes from studying a few prototype strains. Determining a more accurate and reliable pathogenic signature for UPEC among the other commensal and pathogenic *E. coli* in the gastrointestinal tract, the natural reservoir of UPEC (33, 62, 137), will be invaluable for investigations of the epidemiology of UPEC, or of public health concerns, such as the impact of antimicrobial pressure on the natural diversity of UPEC.

To develop means to reliably differentiate UPEC from non-UPEC, we tested *E. coli* isolated from fecal samples and specific UTI syndromes for the presence of specific genes previously characterized as UPEC virulence factors or suspected to play a role in fitness in the urinary tract, querying whether unique combinations of these genes were enriched in UPEC isolates from distinct clinical syndromes. A single gene, *tosA*, which encodes an *in vivo*-induced repeat-in-toxin (RTX) family member, was identified as a candidate UPEC marker. We tested the hypothesis that the presence of *tosA* correctly predicts the ability of a strain to successfully colonize the urinary tract. Finally, the possibility that *tosA*-positive strains represent a phylogenetically distinct subset within

the overall *E. coli* population was examined by sequence analysis of representative isolates identified in the study.

## Materials and Methods

**Strain collection and culture conditions.** A total of 314 strains of *E. coli* was collected from a variety of clinical settings (Table 1) (46, 57, 70, 80, 83, 91, 113, 122, 123). The collection covered a range of natural isolates, from nonpathogenic strains that reside in the lower gastrointestinal tract to highly pathogenic strains that were isolated from the blood of human patients with clinical cases of pyelonephritis complicated by bacteremia. This collection included a broad range of human fecal isolates and isolates from non-human primates and other mammalian hosts of *E. coli* (83, 91). In addition to isolates recovered from human patients with lower urinary tract infections (cystitis) (46, 91, 113) and upper urinary tract infections (pyelonephritis) (91, 122), this collection included samples taken from two additional patient populations. The compromised host anatomy group included strains from patients with abnormal urinary tract anatomy, due either to a congenital defect leading to clinical problems like vesicoureteral reflux(80), or to extrinsic factors such as the presence of an indwelling Foley catheter (123). This patient group is at an increased risk for UTIs from a variety of opportunistic pathogens (31). Finally, the collection also included samples taken from patients with >105 CFU/ml of urine, but without symptoms of cystitis (57), a condition known as asymptomatic bacteriuria. Frozen glycerol stocks were streaked onto LB agar plates and cultured aerobically at 37°C overnight for PCR testing.

**Hemolysin phenotype.** Blood agar plates were prepared as described (129). For determination of the presence or absence of hemolytic activity, a colony of each strain

was streaked onto a blood agar plate for isolation of single colonies and incubated overnight at 37°C. Plates were examined the next day for a visible clearing surrounding isolated single colonies. If no clearing was observed, part of a colony was pushed aside with a sterile toothpick to look for clearing underneath the colony. Any strain that showed clearing around or underneath isolated colonies was scored as positive and all others were scored as negative. Strains were tested using the same batch of blood agar plates to minimize variation between samples.

**Multiplex PCR.** PCR primers were designed based on DNA sequences for each of the 15 genes under study. Each group of sequences was aligned and conserved regions of homology for each gene were chosen to create primers that would generate PCR products of a unique length with ≥ 100bp difference in size from the other PCR products in the multiplex assay. Two sets of multiplex primers were created to maximize the amount of difference in expected product size (Figure 3-1A). Multiplex primer sets were validated by amplifying purified genomic DNA from five sequenced strains of E. coli (MG1655, HS, UTI89, 536, and CFT073) using DNeasy columns (Qiagen). Multiplex PCR reactions used 1 µl of DNA template, 12.5 µl multiplex PCR master mix (Qiagen), 1.25 µl PCR primer mix (concentration of stock is 2pmol/µl/primer), and 10.25 µl nuclease-free water, for a final volume of 25 µl. The thermal cycling protocol was: 1) 95°C for 15min, 2) 94°C for 0.5min, 3) 62°C for 1.5min, 4) 72°C for 1.5min, 5) repeat steps 2-4 29 times, 6) 72°C for 10 min. PCR products were separated by electrophoresis on 1.5% UltraPure agarose (Invitrogen) TAE gels. Gels were stained with ethidium bromide and photographed under UV transillumination using a Gel Doc XR system (Biorad).

For multiplex screening, a single colony from each strain was scraped from a LB agar plate with a sterile toothpick and resuspended in 50 µl sterile water. Cell suspensions were heated to 100°C for 10 minutes, cooled, and frozen until use. PCR reactions were carried out in 25 µl volumes as described above substituting 1µl of cell boil prep for the DNA template.

**Phylogenetic typing.** Major *E. coli* phylogenetic group (A, B1, B2, D) was determined by an established triplex PCR method (19), using PCR reagents and cycling conditions as described above for virulence genotyping.

**MLST.** Multilocus sequence typing data were retrieved from the database maintained at the University College of Cork website (http://mlst.ucc.ie) for the 72 strains of the ECOR collection. The remaining strains tested in the murine model of ascending UTI were typed according to standard protocols (131). Briefly, genomic DNA was extracted from overnight cultures of each strain and used in separate PCR reactions with primers for each of the seven genes used in the database (adk, fumC, gyrB, icd, mdh, purA, and recA). PCR products were purified and submitted for DNA sequencing using the same primers. Data for each gene were aligned with reference sequences in the database and submitted to the online analysis program at the UCC website for determination of sequence type for each strain.

Representative DNA sequences for the ECOR collection and the strains for which MLST data was generated in this study were aligned for each of the seven gene fragments in Lasergene 8 software (DNASTAR). Alignments were imported to MEGA4

(116) and genes were aligned in frame using in silico translation of codons. Using concatenated sequences of the seven gene fragments a neighbor-joining dendrogram was constructed in MEGA4 from using the Maximum Composite Likelihood method. An E. coli clade I strain was used to root the dendrogram (45).

**Murine model of ascending urinary tract infection.** The murine model of ascending UTI was followed as previously described (107). Briefly, female C57Bl/6 mice, 4-6 weeks of age were transurethrally inoculated with $10^8$ CFU of the strain of *E. coli* being tested. Following this, mice were allowed to recover and returned to the animal care facility. 48 hours post inoculation animals were euthanized and bladder and kidneys were aseptically removed, weighed and placed into separate test tubes filled with 3 ml of sterile PBS. Organs were homogenized and dilutions of the homogenate were plated onto LB agar plates using a Spiral plater (Spiral Biotech) and incubated overnight at 37°C. The following day, colony counts for each plate were enumerated and the CFU/g of tissue for each mouse was determined. If no colonies were recovered, the sample was assigned a value of 100 CFU/g of tissue, the limit of detection for this assay.

**Statistical and data analysis.** Images of agarose gels containing multiplex PCR products were scored for the presence or absence of each gene under study for each strain. One-way ANOVA and student's *t*-tests were carried out using Prism software (Graphpad). Multivariate ANOVA testing using the ANOSIM protocol was conducted using PRIMER version 6 (Primer-E).

**Bayesian network analysis.** The Bayesian network (BN) analysis was performed using a web-based MARIMBA system (http://marimba.hegroup.org/) (53). Investigators (YH and APH) were blinded to the putative roles of the 15 genes in virulence. Binary data (0 or 1) was generated for each gene, representing the presence or absence of this gene for individual strains. A new variable called 'group' was created to represent five different clinical settings for each strain: 1) fecal, 2) Compromised Host, 3) ABU, 4) cystitis, and 5) pyelonephritis. Simulated annealing was used to search for candidate network structures where each network has a unique set of connections between variables. In total, 750 million BNs were searched. A consensus network was calculated by considering the best 11 BN models sharing the top log posterior probability.

## Results

**Prevalence of virulence and fitness genes increases in strains from clinical scenarios of increasing severity.** A collection of 314 strains of *E. coli*, representing isolates from fecal samples, ABU, uncomplicated cystitis and pyelonephritis, and complicated UTIs, (Table 1), were assessed by multiplex PCR for the presence of 15 virulence or fitness genes (Figure 3-1A). The 15 genes, chosen to broadly represent documented virulence and fitness factors and general mechanisms of pathogenesis, have been well characterized, with the exception of *tosA* (Table 2). These genes are categorized as encoding proteins that (a) mediate adherence to host cells (*fimA* and *papA*); (b) mediate acquisition of the essential micronutrient iron (*chuA*, *hma*, *iutA*, *iroN*, *fyuA*, *iha*, and *ireA*); or (c) secreted proteins that elicit toxic effects on host cells (*hlyA*, *cnf1*, *tosA*, *sat*, *picU*, and *tsh*). Eleven of the chosen genes reside on large stretches of DNA within the genome of *E. coli* CFT073 known as pathogenicity associated islands (PAI) (78), predicted to be acquired through horizontal gene transfer, while three of the genes are widely dispersed around the genome of the prototypical pyelonephritis strain CFT073 (Figure 3-1B). *cnf1* is not present in CFT073, but is prevalent in other strains.

Screening of the 314 clinical and fecal *E. coli* isolates for UPEC-associated virulence or fitness genes identified a wide range of prevalence values for the 15 targeted virulence genes, in certain instances with considerable variation according to source group. The average percentage of strains that scored positive for a given gene per isolate group is shown in Table 2.

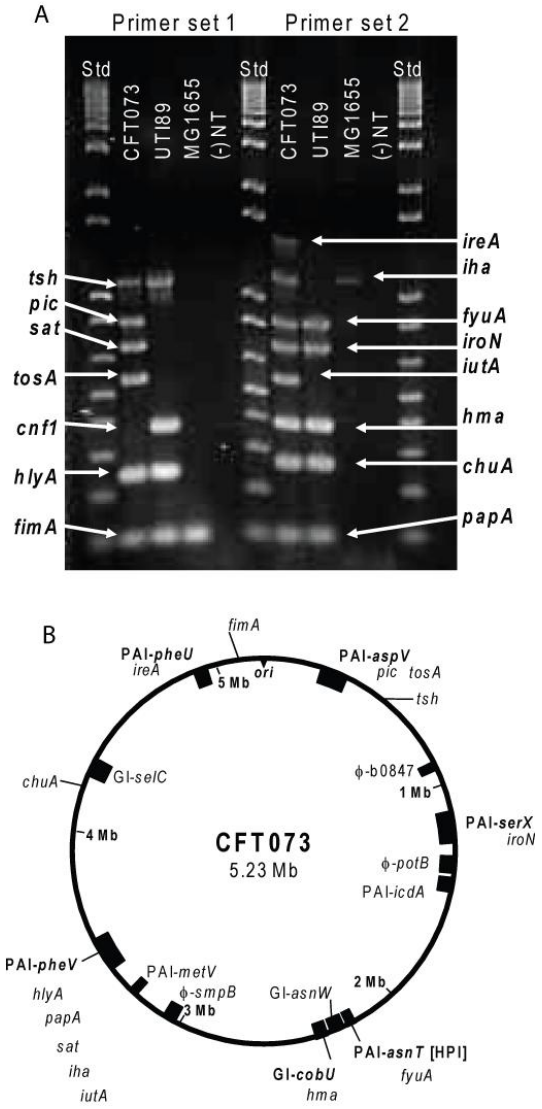



**Figure 3-1. Virulence and fitness genes assessed in this study.** A) Multiplex PCR results using genomic DNA purified from CFT073, UTI89, and MG1655. Primers specific for each virulence or fitness gene were used for PCR amplification of gene fragments from a boiled lysate of the indicated *E. coli* strain. No band observed at the expected size was counted as negative for the presence of a gene. (-)NT indicates no template control samples. Std, DNA ladder standards. B) Schematic of CFT073 genome with the location of each gene under study indicated. The origin of replication is indicated by *ori*. Previously identified pathogenicity associated islands (PAI) or genomic islands (GI) containing one or more genes under study are indicated according to nomenclature of Lloyd et al. (78). *cnf1*, which is not present in strain CFT073, is not shown. Phage insertion sites (ϕ). Three genes, *tsh, chu*, and *fimA* are not found in a PAI or GI.

***tosA* is a marker for the presence of other virulence or fitness genes**. In a previous study, our group identified *tosA* as the main contributing factor of $PAI_{CFT073}$-*aspV* to enhanced fitness of UPEC strain CFT073 in a murine model of ascending UTI (77). Here, we explored associations among *tosA* and other virulence genes by comparing the prevalence of other virulence genes in isolates positive *versus* negative for a given virulence gene (Table 3). In this analysis, *tosA*-positive strains overall showed a combination of carrying the highest number of other virulence genes, the greatest differential in overall virulence gene content, the greatest differential in percent virulence gene prevalence, and the ability of strains of fecal source marked by that gene to colonize both the murine bladder and kidney (Table 3) and thus, the following analysis focused on this gene.

*tosA*-positive isolates collectively exhibited a higher prevalence of almost every virulence gene than was observed in any of the clinical isolate groups (Table 2 and 3). For example, *hlyA* was present in 48% of pyelonephritis isolates (the highest-prevalence clinical group), but in fully 70% of *tosA*-positive isolates. Likewise, the heme receptor *chuA* was present in 89.6% of pyelonephritis isolates (again, the highest-prevalence clinical group), but in 98.4% of *tosA*-positive isolates. *tosA*-positive isolates contained, on average, 11.2 of the 15 studied virulence or fitness genes, whereas *tosA*-negative isolates averaged only 5.4 virulence or fitness genes (Table 3). This trend was observed regardless from which strain collection the *tosA*-positive isolate originated. When data were analyzed independently in this way for every virulence gene, strains grouped by *tosA*, or secondly by *picU,* were associated with the highest number of other virulence

genes assayed for in this study (Table 3). Both *tosA* and *picU* genes reside within the same pathogenicity-associated island, $PAI_{CFT073}$-*aspV* (Figure 3-1B), and thus may be genetically linked in strains within the collection.

**Other toxins genes.** The category of secreted toxins showed a similar pattern, with the prevalence of a given gene generally increasing in relation to increasing clinical severity of the source group. The ABU isolates were a notable exception. Although by definition, ABU strains elicit no symptoms from the host, the prevalence of the gene sequences in this category were comparable among ABU strains to those observed among cystitis and pyelonephritis isolates. In the case of *cnf1* (cytotoxic necrotizing factor 1), for example, ABU isolates were the highest prevalence group. However, presence of a specific gene sequence does not necessarily indicate the presence of an intact gene and functional gene product. Accordingly, additional experiments were conducted to determine if multiplex PCR results for another toxin gene, *hlyA*, which encoded the secreted toxin α-hemolysin, were predictive of its distinctive phenotype.

The presence or absence of the secreted protein toxin gene *hlyA* among strains in this study is highly correlated with the *in vitro* hemolysis phenotype. When the 314 strains under study were cultured on 5% sheep blood agar plates, a positive PCR result predicted a hemolytic pattern on blood agar plates with a sensitivity of 97.9%. A negative PCR result for this gene was determined to be 98.1% specific for a lack of hemolysis seen on blood agar plates. Previous research has established that the presence of hemolysin is predictive of the level of pathology observed in tissue culture systems (83) and correlated to the severity of pathology observed in the murine model of an ascending UTI (110). The high sensitivity

and specificity of the multiplex PCR assay for predicting the hemolysin phenotype indicated that the gene presence or absence data provided insight into the pathogenic potential of each strain. This appealing concept was further investigated using several statistical tools in an analysis of the entire dataset to gain a more detailed understanding into how the molecular mechanisms of pathogenesis are distributed among natural isolates of *E. coli* (see section on testing of three models).

**Adhesin and iron acquisition genes**. *fimA* encoding the type 1 fimbria major structural subunit was almost universally present in strains from each clinical group regardless of source with >98% of isolates testing positive. In contrast, *papA (*structural subunit of pyelonephritis-associated pili), although less prevalent overall (56.5%), was significantly more prevalent among all urine isolates combined (59.2%) than among fecal isolates (26.4%) ($P$<0.0001). Both of these findings are in general agreement with results from previous studies (85).

The genes encoding outer membrane iron receptors (*chuA*, *hma*, *iutA*, *iroN*, *fyuA*, *iha*, and *ireA*) were less common among fecal isolates (average prevalence, 30.3%; range, 14.3-47.3%) and more common among strains isolated from a symptomatic UTI (average prevalence was 59.3%; range was 10.8-93.5%), with the prevalence rates of certain genes, such as *fyuA* in pyelonephritis isolates, exceeding 90%. This result is also in agreement with previously published findings (47).

***tosA*-positive strains colonize the murine urinary tract regardless of clinical source.** If one hypothesized that the presence of *tosA* predicts virulence regardless of clinical source, then the presence of *tosA* should predict successful colonization of a *tosA*-

positive fecal *E. coli* isolate in the murine model of an ascending urinary tract infection. For example, EFC5, a fecal isolate from a healthy human volunteer (83) determined in this study to be *tosA*-positive, was transurethrally inoculated into the bladder of female C57Bl/6 mice. 48 hours post-inoculation, mice were euthanized, bladder and kidney tissue was removed, and the CFU/g of tissue was determined for each mouse to assess whether or not this fecal isolate could colonize and survive in the host urinary tract. This was repeated with strain CFT073, a human pyelonephritis isolate that is also *tosA*-positive, (Figure 3-2A). There were no differences observed among colonization levels for both bladder and kidney tissue between these two strains, indicating that the fecal isolate EFC5 was a successful colonizer in this model system.

The murine model is recognized as capable of differentiating fecal and uropathogenic strains of *E. coli* (48, 83). However, fecal *E. coli* isolates that carry *tosA* colonize the murine model in higher numbers in both bladder and kidney tissue than *tosA*-negative fecal isolates. Seven *tosA*-positive fecal strains, including EFC5, were tested in the murine model and compared to ten *tosA*-negative fecal isolates for CFU/g of bladder (Figure 3-2B) or kidney tissue (Figure 3-2C) at 48 hours post infection. Statistically significant differences were observed using a two-tailed student's *t*-test with higher average CFU levels in the *tosA*-positive group for both bladder ($p<0.001$) and kidney ($p<0.01$) tissue.

The presence of *tosA* also predicted a greater CFU load in bladder and kidney tissue of infected mice regardless of the host species from which a strain was isolated.

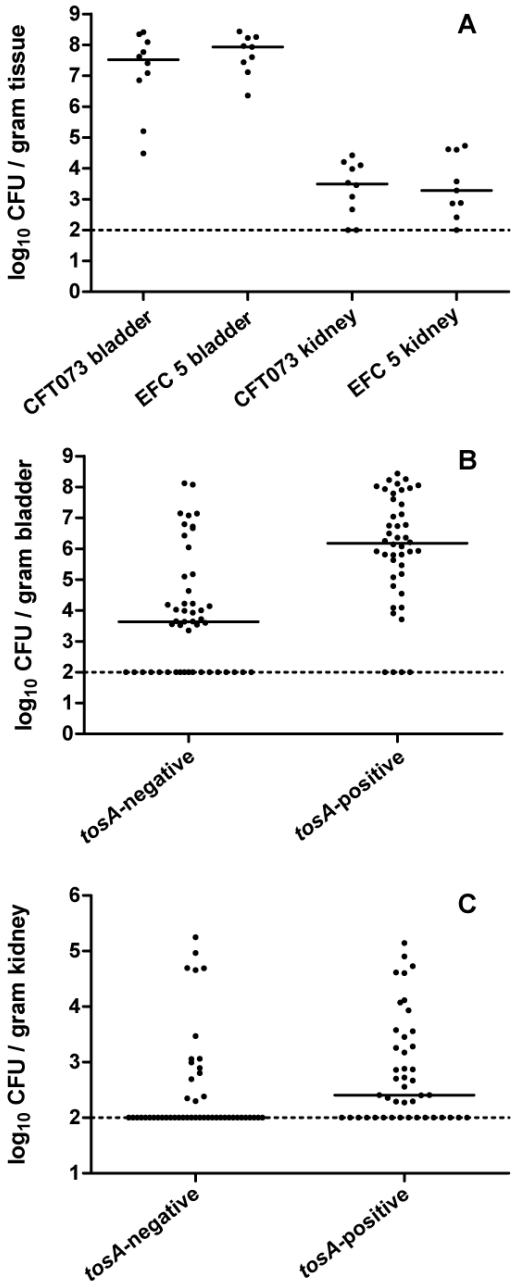



**Figure 3-2. *tosA* presence predicts successful colonization of the murine urinary tract.** A) Mice were transurethrally inoculated with $10^8$ CFU of either CFT073 or EFC5. 48 hours post-inoculation, bladders and kidneys were removed and CFU/g of tissue was determined. Each dot represents CFU load for one mouse. B) and C) show combined results for 10 *tosA*-negative and 7 *tosA*-positive strains for bladder and kidney tissue, respectively. Each dot represents one mouse (N=44 and 48 for *tosA*-positive and *tosA*-negative, respectively).

One *tosA*-positive (ECOR 57) and one *tosA*-negative strain (ECOR 70) isolated from gorillas were compared for the ability to colonize in the murine model described above. A second pair of *tosA*-positive and *tosA*-negative fecal isolates from orangutans (ECOR 52 and ECOR 7) was also compared. Both *tosA*-positive isolates colonized both bladder and kidney, while only ECOR 7, a *tosA*-negative strain isolated from an orangutan, colonized the bladder and kidneys (data not shown). An additional *tosA*-positive fecal isolate collected from a lion (ECOR 58) was tested and found to colonize both bladder and kidneys in the murine model. ECOR 58 is unusual as this strain only contained three additional genes tested for in this study (*fimA*, *picU*, and *iroN*). These results demonstrate that the *tosA* is a marker for strains that can successfully colonize the murine model including fecal *E. coli* isolated from other animals as well as for isolates that contained relatively few of the genes that have been well studied in the context of pathogenesis in the urinary tract.

***tosA*-positive strains comprise a distinct subset of the B2 lineage of *E. coli.*** A multiplex assay (19) to place *E. coli* into one of the four main ECOR phylogenetic group assignments, A, B1, B2, or D, was applied to each strain in this study (Table 3). *E. coli* fecal isolates were frequently placed within group A (42.9% of isolates), but this group less commonly contained isolates cultured from the urinary tract; for example, only 7.7% of strains isolated from cases of pyelonephritis localized to group A. On the other hand, 70% of isolates from cases of cystitis and pyelonephritis were found to be members of group B2, which was the most common group among isolates from the urinary tract. In contrast to this 63 of the 64 (98.4%) *tosA*-positive strains in this study

were determined to reside in the B2 phylogenetic group of *E. coli*. A single *tosA*-positive isolate (ECOR 58) localized to the B1 phylogenetic group. The hypothesis that these isolates comprised a closely related subset of UPEC was then examined further using sequence analysis

The majority of *tosA*-positive isolates (8 of 12, 66.6%) examined by multilocus sequence typing (MLST) belonged to a single sequence type complex, ST complex 73. In addition to the ECOR strain collection, the strains examined in the murine model were assigned to a sequence type (ST) using standard protocols (131), and this collection of sequences was used to build a dendrogram illustrating the relationship of these groups to each other (Figure 3-3). The ECOR strain collection provided a framework of sequence types associated with each of the four main *E. coli* groups as well as a source of several groups of *E. coli* previously determined to have undergone extensive genetic recombination. The only isolate that contained *tosA* outside of the B2 group (ECOR 58) was previously determined to belong to a group of *E. coli* that appear to have undergone extensive recombination with other groups of *E. coli* (131). Further examination of the MLST database showed that the majority of strains in ST complex 73 (29 of 48, 60.4%) are UPEC isolates and were collected in the setting of a human UTI (131). Based on the MLST and ECOR typing, *tosA* appears to be a marker for a group of strains that are enriched for UPEC and present in a few outlier groups that may have acquired this gene by extensive genetic recombination.

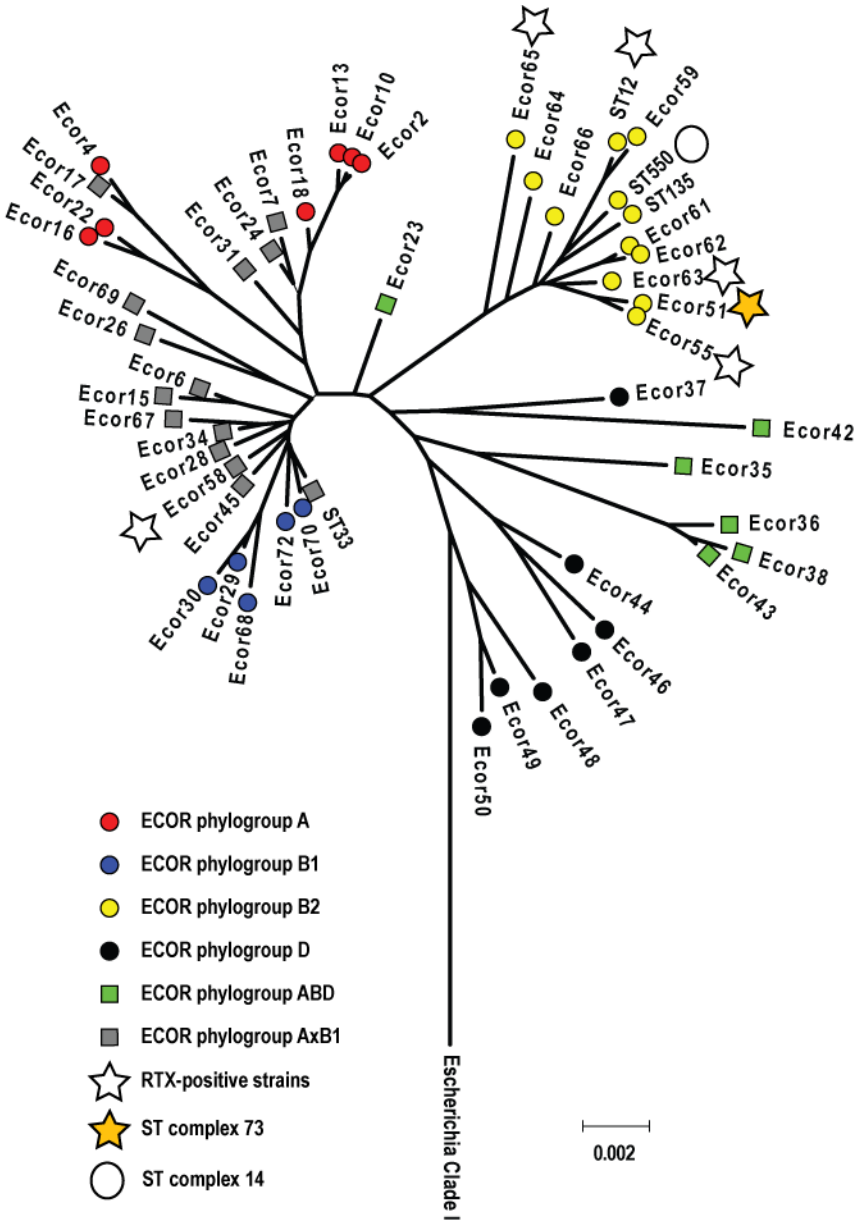



**Figure 3-3. Phylogenetic relationship of *tosA*-positive strains.** Representative *tosA*-positive strains included in this study are marked with a star. The four main groups of *E. coli* (A, B1, B2, and D) and two groups of recombinants (ABD and AxB1) were assigned according to Wirth *et al*. (131). MLST complexes 73 and 14 are indicated next to representative isolates from each group. Sequences from Escherichia clade I isolates that do not group with phylogenetic groups A, B1, B2, D, or the hybrid groups AxB and ABD, were used to root the dendrogram (120).

**Testing of three models for assessing pathogenic potential.** While we have discovered that the presence of *tosA* predicts virulence regardless of clinical source, can we also use the full virulence and fitness gene prevalence data set to differentiate strains by clinical source alone. Fecal isolates of *E. coli* on average had significantly fewer virulence genes (mean = 4.2) than isolates from the urinary tract (mean = 7.5) (Table 4). One simple model that could explain the difference in the distribution of the genes under study is that isolates that caused the most severe clinical disease contain more genes and mechanisms linked to pathogenesis than those that cause milder clinical cases of disease or those that were isolated from healthy individuals. This model was tested by averaging the number of genes detected as present in a group of strains and compared across all groups using a one-way ANOVA (Figure 3-4A). The global test statistic was significant ($p<0.001$), however, pair-wise testing was only significant for fecal isolates compared to the other groups, and for comparing pyelonephritis isolates with isolates from patients with compromised urinary tract anatomy. This indicates that fecal isolates can be differentiated from strains of *E. coli* isolated from the urinary tract based on the average number of virulence-associated genes present, but that this is insufficient to fully differentiate between urinary tract isolates from specific clinical scenarios of different severity. Isolates that were presumed to have limited pathogenic potential, such as ABU isolates, were indistinguishable from isolates known to be highly pathogenic and elicit strong symptoms from the host, such as pyelonephritis isolates, based on this test. While it is possible that ABU isolates are true pathogens and the lack of symptoms is the result of host innate immune defects in the infected individual (51),

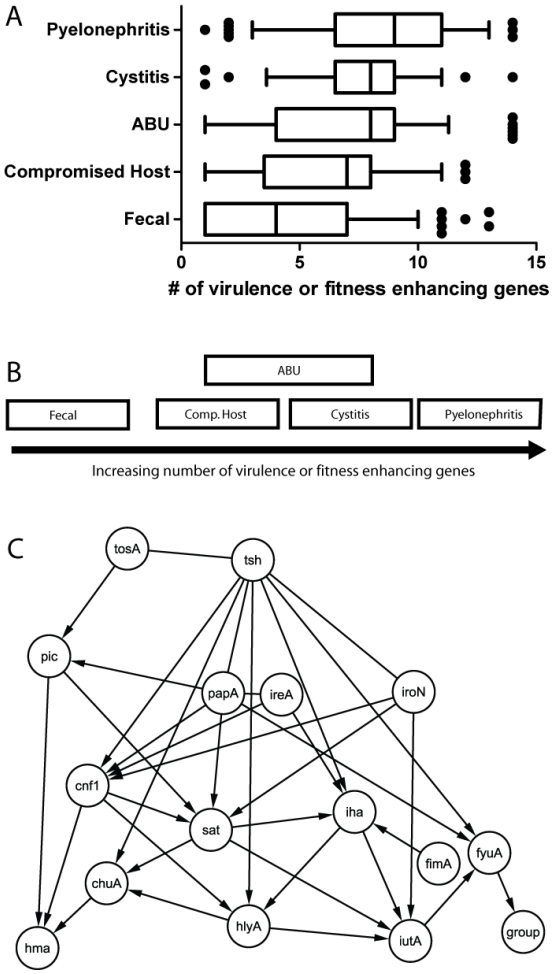



**Figure 3-4. Models demonstrating differences and similarities between *E. coli* isolates from different clinical settings.** A) Under the first model, groups are placed according to the average number of virulence factors as described in a box and whisker plot where the mean is indicated by a vertical line; the 25$^{th}$ and 75$^{th}$ percentiles are indicated by the left and right border of the box; and the 10$^{th}$ and 90$^{th}$ percentiles are indicated by the whiskers. Dots represent outliers in the analysis. B) The third model used multivariate ANOVA of the full presence or absence data for each gene to test for differences in gene prevalence rates between all five groups of isolates. Groups shown to overlap cannot be readily differentiated based on the data collected in this study, while those that do not overlap where found to contain different combinations of the genes under study. C) Bayesian model of unique connections between virulence and fitness factor genes. The presence or absence of each gene in each of 314 strains was modeled to identify a consensus network of the genes. Knowledge of the presence or absence of a given gene predicts the presence or absence of another gene. Arrows indicate directed edges linking two variables together in the network.

it is also possible that mutations exist within virulence gene operons, rendering their products inactive (70).

A second model that could explain the distribution of the genetic mechanisms of virulence, based on classes of virulence determinants present in the urinary tract, was explored. UPEC are genetically heterogeneous and often exhibit multiple molecular mechanisms for causing disease, such as secreting a series of protein toxins (89)or synthesizing multiple acquisition systems for the uptake of iron and iron-containing compounds (47). The second model assumed that success in the urinary tract and the production of symptoms in the host was the sum of three activities the production of large secreted protein toxins or RTX family member (*cnf1, hlyA, tosA*), the production of secreted proteolytic autotransporter proteins (*sat, pic, tsh*), and the acquisition of iron (*chuA, hma, iutA, iroN, fyuA, iha, ireA*). Under this model the isolates from different clinical groups may be differentiated by comparing the potential of the strains to achieve each of these three activities (data not shown). After averaging the number of genes in each category per strain and adjusting for the mean and standard deviation of each category, a multivariate version of an ANOVA design was used to seek differences between the clinical groupings. Similar to the first model, although the global test was significant ($p<0.001$), pairwise testing only clearly differentiated fecal isolates from the other clinical isolates. The urine-source groups were statistically indistinguishable, regardless of disease severity or host compromise status.

The third model, which was able to differentiate between groups of urinary tract isolates, assumed that each genetic mechanism of survival and pathogenesis (defined by the presence of distinct virulence or fitness genes) confers a unique advantage in the host. Under this model, the level of success of a strain of *E. coli* in different patient populations and sites within the urinary tract would be partly determined by the unique contribution of each genetic mechanism of pathogenesis. While this study could not assess the entire collection of known virulence factors described in the literature, the 15 genes assayed by PCR provided sufficient information to allow a multivariate ANOVA model using the full presence or absence data for each gene to test for differences in gene prevalence rates between all five groups of isolates. Similar to the first two models the global test was significant ($p<0.001$) and fecal isolates were significantly different from other strain collections. However, significant differences were also observed among *E. coli* isolated from patients with different clinical scenarios. Under this model, all groups of urinary tract isolates were distinguished from each other except for an overlap between ABU isolates and isolates from patients with compromised urinary tract anatomy and between ABU isolates and cystitis isolates. Figure 3-4B presents a visual representation of the results of the ANOVA and MANOVA testing and demonstrates that, even after allowing each of the 15 genes studied to count as separate variables, urinary tract isolates of *E. coli* from the four different classes of clinical severity exhibited more similarities to other urinary tract isolates than to fecal isolates.

***tosA*-positive isolates are significantly different from *tosA*-negative strains in all three statistical models tested**. Due to the high similarity noted among the gene prevalence data, the three models described above were retested with the strains grouped by the presence or absence of *tosA*. Because only two groups (*tosA*-positive and *tosA*-negative) were being tested, the original ANOVA was replaced by a student's *t*-test for evaluating the average number of genes presence. Both multivariate tests were conducted as before, and in each case *tosA*-positive isolates were significantly different from *tosA*-negative isolates ($p<0.01$).

**Bayesian network modeling reveals extensive connections between virulence and fitness genes.** Finally, to seek relationships between the combinations of virulence genes present in specific strains, Bayesian network modeling was employed. This technique allows the influence of the presence of a gene on the presence of the other genes to be assessed visually without the need for hundreds of individual pair-wise statistical comparisons. A consensus network compiled from the best scoring models discovered numerous connections between genes (Figure 3-4C). For example, three genes *tosA*, *tsh*, and *iroN* are at the top of the hierarchical structure within the consensus network. These genes appear to influence or closely associate with the presence or absence of the other genes. This model illustrates that only the yersiniabactin receptor (*fyuA*) was directly linked to a particular clinical setting ("group"). Indeed, *fyuA* is less common in fecal isolates (47.3%), but steadily increases in presence as the severity of the disease increases, reaching a maximum prevalence of 93.5% in pyelonephritis isolates (Table 2). Our model also indicates that *fyuA* is closely

associated and possibly influenced by three other genes *tsh*, *papA*, and *iutA*. Table 2 indicates that these three genes have a similar pattern to *fyuA*. However, based on the available binary data, our modeling predicted that *fyuA* has strong direct association with the clinical setting.

**Discussion**

Virulence-associated gene content among *E. coli* isolated from the urinary tract demonstrated more similarities to other UPEC strains than to fecal isolates. While a general trend of increasing number of virulence factors was observed in strains isolated from the more severe clinical settings (Table 4, Figure 3-4A), each group of isolates could only be differentiated by multivariate analysis of the entire virulence gene prevalence dataset (Figure 3-4B). In contrast, strains that carried *tosA* were easily differentiated from *tosA*-negative strains by several measures. *tosA*, a putative RTX family member gene (77), co-occurred on average with more than 10 of the other 14 genes assayed for in this study and thus its presence was predictive of highly virulent strains (Tables 2 and 3). This gene represents a marker for *E. coli* belonging to the B2 phylogenetic group and is enriched in sequence type complex 73, a group of *E. coli* enriched for UPEC isolates including the prototype human pyelonephritis strain CFT073 (131) (Figure 3-3).

UPEC strains reside in the gastrointestinal tract and thus can occasionally be isolated among commensal strains cultured from fecal samples (33, 86, 137). Indeed, the presence of *tosA* predicted successful colonization of all but one of seven *tosA*-positive fecal *E. coli* isolates tested in the murine model of an ascending UTI (Figure 3-2B and 2C). ECOR 51, the only *tosA*-positive fecal isolate that did not colonize the animal model lacked the gene *fimA*, the major structural subunit of type 1 fimbriae. While type 1 fimbriae are critical for the success of UPEC in the murine model (9, 20, 42), recent

studies have questioned the importance of this factor in human infections (46, 74) and raise the possibility that this isolate could potentially colonize the human urinary tract. While it was beyond the scope of this study to identify a specific role for this gene in bacterial pathogenesis, the results presented here indicate that success of *tosA*-positive strains may rely in part on the co-occurrence of known virulence factors of UPEC. With the exception of *papA* and *ireA*, the highest prevalence of the 12 other genes included in this study among all groups of isolates was observed among *tosA*-positive strains (Table 2). However, the observation that ECOR 58, a fecal *E. coli* isolate that lacked many of the best characterized UPEC virulence factors, demonstrates that *tosA* can nevertheless serve as a marker both for UPEC strains that are predicted to cause disease through well known mechanisms as well as for strains that may contain unique and as yet uncharacterized approaches to colonizing the host urinary tract.

One goal of the study was to determine whether we could differentiate strains based on clinical source and virulence and fitness gene content. First, it was straightforward to statistically separate fecal strains from strains isolated from the urinary tract using the three models (Figure 3-4). However, separating isolates by specific clinical populations proved more difficult. We found seemingly minimal differences in the numbers of virulence or fitness genes among UPEC isolates from different patient populations (Figure 3-4A). Using model 3, however, these differences were sufficient to differentiate between strains from patients with complicated UTI (compromised host urinary tract anatomy) and those from uncomplicated symptomatic infections (cystitis and pyelonephritis) (Figure 3-4B).

ABU strains in this study, on the other hand, failed to coalesce into a discrete group and instead revealed a spectrum of isolates that ranged from strains that contained only 1 of the 15 genes assayed to strains that contained 14 of these same genes (Figure 3-4A). These strains, by our analysis, closely resemble fecal commensal strains that failed to colonize the murine model (that is, low virulence potential) and those strains that more closely resemble human cystitis isolates (higher virulence potential), respectively. That these strains colonize but elicit no symptoms could be explained in at least three ways. a) Low virulence strains, unrelated to typical UPEC strains, could colonize a healthy individual if there were additional as yet undetermined colonization genes present (69, 139). Lack of virulence factors would not damage the host and thus cause no symptoms. b) Strains with a high number of virulence genes could contain mutations within these genes that are nevertheless present and detectable by PCR (70, 97). *E. coli* strain 83972, for example, may represent such a strain that contains virulence genes but also carries mutations within some of these genes (70, 96). c) Truly high virulence strains with intact virulence genes could colonize a host with defects in innate immunity who cannot respond properly to infection (51). The broad overlap of ABU isolates with strains isolated from hosts with compromised urinary tract anatomy and the overlap with cystitis isolates observed in this study could be explained by a combination of these three scenarios.

Bayesian network analysis is a powerful statistical method that infers potential relationships among a set of variables and providing an alternative to analyzing hundreds of individual pair-wise comparisons (53, 54) . In the current study, such a

network indicates that the knowledge of the presence or absence of a given gene can predict the presence or absence of another gene. Before conducting this analysis, we predicted that genes contained in the same pathogenicity associated island (PAI), large stretches of DNA thought to be acquired *en masse* through horizontal gene transfer, would be linked together in the consensus network (Figure 3-4C). For example, $PAI_{CFT073}$-*pheV* contains *papA, sat*, *iha*, *iutA and hlyA*; these genes are all serially connected in the consensus network. What was not expected was that nearly every gene included in the study would demonstrate some level of connectivity. As an example, the gene *tsh*, which encodes an autotransporter expressed both during *in vivo* and *in vitro* growth of CFT073 (52), is not associated with any known PAI (Figure 3-1B). Despite this, *tsh* contains direct links to 8 other genes in the consensus network. This result indicates that gene association between established UPEC virulence factors may not be confined to those genes contained in the same PAI. In addition, gene content of a PAI, the location of additional copies of the same genes outside of the archetypal PAI, and differences in number of PAIs have been previously documented between strains of UPEC (15, 126). The Bayesian network results (Figure 3-4C) are consistent with the hypothesis that such genetic reassortments of virulence factors are common among UPEC.

Given the previous observations that Bayesian network analysis is sufficiently robust to screen out noise in large gene expression arrays (34, 53), it is surprising that only the yersiniabactin receptor gene *fyuA*, and not *tosA* or other genes, was determined to directly influence the probability that a given strain originated in a

particular clinical group. Indeed, our Bayesian network modeling found that the presence and absence of *fyuA* was directly associated or influenced by three other genes (*tsh*, *papA*, and *iutA*), and the presence of *tosA* appears to play a leading role in predicting the presence of many of the other genes (Figure 3-4C). This was somewhat surprising since *tosA* was found in only one-fourth of strains. Taken together, this result may suggest that distinct pathotypes of UPEC exist, but identifying these groups will require knowledge about the diverse mechanisms of pathogenesis found within each of the different subpopulations of UPEC.

We define a UPEC strain as one that infects the human urinary tract. Our tools to define such an isolate include virulence and fitness gene content, colonization of the murine urinary tract, and original clinical source. We observed here that clinical source alone is not necessarily predictive of UPEC status; nor is the presence of a single gene. Not all isolates that have the potential to colonize the urinary tract, for example, carry the *tosA* gene. Only 24.3% (28/115) of cystitis and pyelonephritis isolates, in the current study, contain this gene. These isolates are most likely to resemble ST complex 73 strains such as CFT073, but other groups of UPEC exist. Strains such as *E. coli* fecal control 24 (EFC24), one of the *tosA*-negative strains investigated in this study that successfully colonized the bladder of infected mice, represent one of the groups that together comprise the remaining 75% of UPEC isolates. EFC24 was determined in this study to belong to ST complex 14, a group of B2 isolates highly enriched for UPEC. In fact, 10 of the 13 (76.9%) isolates that make up ST complex 14 were isolated from the human urinary tract (131). While many of the *tosA*-negative fecal strains that were

successful in the murine model appear to belong to larger groups of urinary tract isolates, none of the genes included in this study could differentiate these isolates from those *tosA*-negative isolates that failed to colonize the host urinary tract. Additional studies will be necessary to identify a reliable marker for groups of UPEC that lack the *tosA* gene.

Despite these limitations, the presence or absence of *tosA* provides useful information in assessing the pathogenic potential of a given *E. coli* isolate. As this gene is found in one quarter of UPEC strains that cause symptomatic infections in normal human hosts, assaying for this one gene would enable larger studies that have previously not been feasible, such as studies that compare UPEC carriage rates in patient populations with different rates of urinary tract infections. When combined with other virulence and fitness genes such as *hma, picU,* and *fyuA*, the urovirulence potential of these isolates could be accurately predicted. Given the heterogeneity in pathogenetic mechanisms observed among UPEC, such an understanding may prove critical to future efforts to develop novel therapeutic strategies to combat this widespread human pathogen.

**Table 3-1. *E. coli* isolates used in this study**.

| Clinical group | No. of isolates[a] | Source of collection[a] |
|---|---|---|
| **Fecal** | 91 | (83, 91) |
| **Compromised Host** | 41 | (80, 123) |
| **Asymptomatic Bacteriuria** | 68 | (57, 91) |
| **Cystitis** | 37 | (46, 91, 113) |
| **Pyelonephritis** | 78 | (91, 122) |

[a]number of total isolates of each category is given along with the reference to the publication that describes the source of the isolates.

**Table 3-2. Prevelance of virulence associated genes in among *E. coli* isolated from different clinical settings.**

| | adhesins | | toxins | | | | | | outer membrane iron receptors | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *fimA* | *papA* | *hlyA* | *cnf1* | *tosA* | *sat* | *pic* | *tsh* | *chuA* | *hma* | *iutA* | *iroN* | *fyuA* | *iha* | *ireA* |
| fecal | 98.9 | 26.4 | 13.2 | 7.7 | 11.0 | 19.8 | 12.1 | 20.9 | 39.6 | 26.4 | 31.9 | 20.9 | 47.3 | 31.9 | 14.3 |
| CHA | 97.6 | 51.2 | 19.5 | 4.9 | 19.5 | 41.5 | 12.2 | 48.8 | 75.6 | 34.1 | 58.5 | 26.8 | 78.0 | 46.3 | 24.4 |
| ABU | 98.5 | 54.4 | 35.3 | 33.8 | 26.5 | 33.8 | 23.5 | 63.2 | 73.5 | 51.5 | 45.6 | 41.2 | 83.8 | 35.3 | 16.2 |
| cystitis | 97.3 | 67.6 | 48.6 | 32.4 | 16.2 | 32.4 | 18.9 | 54.1 | 83.8 | 81.1 | 43.2 | 48.6 | 89.2 | 37.8 | 10.8 |
| pyelonephritis | 98.7 | 63.6 | 48.1 | 13.0 | 28.6 | 51.9 | 29.9 | 67.5 | 89.6 | 61.0 | 66.2 | 41.6 | 93.5 | 55.8 | 27.3 |
| *tosA*-positive strains | 98.4 | 59.4 | 70.3 | 37.5 | 100.0 | 67.2 | 89.1 | 96.9 | 98.4 | 92.2 | 67.2 | 59.4 | 98.4 | 67.2 | 20.3 |
| *tosA*-negative strains | 98.4 | 47.2 | 21.6 | 12.0 | 0.0 | 26.8 | 2.0 | 36.8 | 61.6 | 36.4 | 43.2 | 28.0 | 69.6 | 34.4 | 18.4 |

Gene prevalence expressed as % of isolates that scored positive on multiplex PCR. CHA - compromised host urinary tract isolates, ABU - asymptomatic bacteriuria isolates.

**Table 3-3. Characteristics of isolates marked by the presence of a single virulence or fitness gene.**

| Gene | Positive[a] | Negative[a] | ΔVF[b] | Avg Δ in %gene prevalence[c] | *t*-test bladder[d] | *t*-test kidney[d] |
|---|---|---|---|---|---|---|
| ***tosA*** | 11.2 | 5.4 | 5.9 | 34.7 | <0.0001 | 0.0071 |
| ***picU*** | 11.2 | 5.4 | 5.8 | 34.6 | <0.0001 | NS |
| ***hlyA*** | 10.1 | 4.9 | 5.2 | 30.2 | NS | NS |
| ***cnf1*** | 10.1 | 5.8 | 4.3 | 23.7 | NS | NS |
| ***sat*** | 9.5 | 5.0 | 4.5 | 25.2 | NS | NS |
| ***hma*** | 9.4 | 4.0 | 5.4 | 31.6 | <0.0001 | 0.0074 |
| ***tsh*** | 9.4 | 3.9 | 5.5 | 32.1 | <0.0001 | NS |
| ***iroN*** | 9.0 | 5.3 | 3.7 | 19.5 | NS | NS |
| ***Iha*** | 9.0 | 4.9 | 4.1 | 22.1 | NS | NS |
| ***papA*** | 9.0 | 4.2 | 4.8 | 26.9 | 0.014 | NS |
| ***ireA*** | 8.7 | 6.1 | 2.7 | 11.8 | 0.0002[e] | 0.0155[e] |
| ***iutA*** | 8.7 | 4.6 | 4.1 | 22.1 | 0.0405[e] | NS |

| | | | | | | |
|---|---|---|---|---|---|---|
| ***chuA*** | 8.4 | 2.4 | 6.0 | 35.4 | 0.0003 | NS |
| ***fyuA*** | 8.1 | 1.7 | 6.4 | 38.6 | 0.0017 | NS |
| ***fimA*** | 6.6 | 6.6 | 0.0 | -7.5 | - | - |

[a]average number of the 15 virulence factors in isolates that contain (positive) or lack (negative) the indicated gene

[b]$\Delta$VF, the difference in number of virulence factors between positive and negative

[c]% change in gene prevalence of all 15 genes is averaged for each group.

[d]data in Figure 3-3B and 3-3C were grouped according to the presence or absence of each gene and reanalyzed; the *p* value of each group that was found to be significant is reported. NS, not significant.

[e]these genes were found to be negatively correlated with CFU/g of tissue. The presence of *ireA* was found previously (63)to negatively correlate with lethality in a murine model of peritonitis.

**Table 3-4. Average number of virulence associated genes and ECOR group membership**.

| Clinical group | Average number virulence genes | ECOR phylogenetic group (% of isolates)[a] | | | |
|---|---|---|---|---|---|
| | | A | B1 | B2 | D |
| **Fecal** | 4.2 | 42.9 | 18.7 | 20.9 | 17.6 |
| **Compromised Host** | 6.4 | 17.1 | 7.3 | 51.2 | 24.4 |
| **Asymptomatic Bacteriuria** | 7.2 | 15.9 | 11.1 | 58.7 | 14.3 |
| **Cystitis** | 7.6 | 16.7 | 0.0 | 71.4 | 11.9 |
| **Pyelonephritis** | 8.4 | 7.7 | 2.6 | 70.5 | 19.2 |
| | | | | | |
| ***tosA*-Positive Strains** | 11.2 | 0.0 | 1.6 | 98.4 | 0.0 |
| ***tosA*-Negative Strains** | 5.4 | 27.6 | 11.2 | 39.2 | 22.0 |

[a]ECOR group membership expressed as % of isolates that belong to the four main ECOR phylogenetic groups

## CHAPTER 4

## THE REPEAT-IN-TOXIN (RTX) FAMILY MEMBER TOSA MEDIATES ADHERENCE OF UROPATHOGENIC *ESCHERICHIA COLI* AND SURVIVAL DURING BACTEREMIA

### Introduction

Repeats in toxin (RTX) proteins are widespread among gram-negative bacteria, with more than 1000 family members detected in a survey of genome sequences from 251 bacterial species (76). Two common features present in known RTX family members are a characteristic glycine- and aspartate-rich repeat near the C terminus of the protein and a conserved type 1 secretion system (T1SS) that exports the protein out of the cell into the extracellular environment, bypassing the periplasmic space (76, 125). The model protein for the RTX family member is α-hemolysin, which inserts into host membranes and forms pores that allow an influx of $Ca^{2+}$ into host cells, altering host physiology or leading to cell death (125). Additional RTX family members have displayed a wide array of functions in bacterial pathogens; secreted proteases (81, 104) and lipases (27, 133), crosslinkers of cellular actin that cause host cell rounding (100), and surface associated coats of protein that form the bacterial S-layer (92, 108) to name a few. However, most RTX family members remain uncharacterized. Given the widespread distribution and diverse roles that known family members

contribute to bacterial pathogenesis, the identification and characterization of novel RTX family members remains an important area of research.

One of the best characterized RTX family members, α-hemolysin, has been shown to enhance host damage in the urinary tract during an *Escherichia coli*-mediated infection (110). In addition, this protein contributes to disseminated infections (15, 44), such as those that result from an ascending urinary tract infection (2, 18, 40). Despite the importance of this toxin, not all uropathogenic *Escherichia coli* (UPEC) produce this toxin (14, 59). Additional uncharacterized RTX family members have been detected in the genomic sequences of uropathogens (76), raising the possibility that these proteins might serve as alternatives to α-hemolysin.

The human pyelonephritis/urosepsis isolate CFT073 contains two annotated repeat-in-toxin family members (RTX): *hlyA* encoding the prototype of the family, α-hemolysin, and *tosA* (originally annotated as *upxA*) (126). *In silico* analysis has previously indicated that *tosA* shares little homology with characterized RTX family members (93) and may represent a novel virulence factor of UPEC. Subsequent analysis of the region surrounding *tosA* revealed that the gene is part of a pathogenicity associated island (PAI) and appears to be linked to genes encoding a type one secretion system (T1SS) composed of a *tolC* homolog and homologs of *hlyB* and *hlyD* (93). Previous work (119) revealed that *tosA* is only expressed in the host environment and contributes significantly to the success of this strain of UPEC in colonizing an animal model of an ascending UTI. However, it remains unclear whether this novel RTX protein

functions in a similar way as α-hemolysin, or if it confers a different set of advantages to the bacterium.

To explore this question, we characterize the expression of the genes in the region of the *tosA* locus, and identify areas in the host during an animal model of ascending UTI that allow TosA expression. We then explore the function of TosA protein using *in vitro* tissue culture systems and *in vivo* models of UTI, bacteremia, and sepsis to cover the full range of environments UPEC encounter during the natural course of infection. Finally, we explore the use of purified TosA protein in an experimental animal vaccine model to protect against urosepsis. These results indicate that while the *tos* locus shares many features in common with the α-hemolsyin locus, the TosA protein appears to play a distinct role in UPEC pathogenesis.

## Materials and Methods

### Bacterial strains and plasmids

The UPEC isolate CFT073 was isolated from blood and urine cultures of a human patient with pyelonephritis (122). An isogenic strain lacking the gene tosA was already created by our group for use in a previous study (77). A plasmid containing the gene for GFP under the constitutive em7 promoter, pGEN mut 3.1 (Ref), was electroporated into WT CFT073 cells for use in the murine model of ascending UTI. The pBAD-*tosA* plasmid was created by PCR amplifying the *tosA* gene using the following conditions: 1ug CFT073 genomic DNA was mixed with 4µl 2.5mM dNTP mixture, 5µl 5X GC Phusion buffer (NEB), 2µl 10µM primer mixture, 1U of Phusion polymerase in 0.25µl volume (NEB), and 12.75µl distilled water. PCR reactions were cycled according to the following protocol: 1) 98°C for 30 seconds, 2) 98°C for 30 seconds, 3) 64°C for 30 seconds, 4) 72°C for 4 minutes, 5) repeat steps 2-4 30 times, 6) 72°C for 5 minutes. PCR products were ligated into TOPO Blunt PCR cloning vector (Invitrogen) according to manufacture instructions. A clone containing the correct insert was selected and plasmid was purified using the Qiagen plasmid mini prep kit and cut with BglII and EcoRI restriction enzymes (NEB). The DNA fragment containing the tosA sequence was purified after agarose gel electrophoresis and ligated overnight with a BglII/EcoRI (NEB) cut pBAD/Myc-His A vector (Invitrogen) using T4 DNA ligaes (NEB) according to manufacture instructions. Ligation products were electroporated into TOP10 *E. coli* cells (Invitrogen) and resulting colonies were screened for protein expression after growth in LB + 10 mM L-arabinose at 37°C for 3 hours (referred to as inducing conditions).

*araB*$_P$-*tosC* construct of CFT073 was constructed using the lambda red recombinase system (23) (Figure 4-3). Two PCR products were created; one included a cassette conferring kanamycin resistance amplified from the pKD4 plasmid, and the other from the pKD46 plasmid included the arabinose inducible promoter. These two products were created such that the 5' end of the pKD4 fragment and the 3' end of the pKD46 product contained regions of homology to the genomic region upstram of *tosC*. The 3' end of the pKD4 product contained a linking primer that was the reverse complement of a linking primer carried on the 5' end of the pKD46 product. Each of the two PCR products were created in separate PCR products according to PCR protocol outlined for pBAD-*tosA* vector creation with the following modification: extension step was modified to 60°C for minutes. The resulting PCR products were gel purified as above and combined in the same PCR to create a chimeric PCR product containing the two products linked end-to-end in a PCR reaction as described above with the one modification. Ten thermal cycles (steps 2-4) were conducted without any primers to combine the two products. After ten cycles, the cycling was paused, and PCR primers corresponding to the 5' end of the pKD4 and 3' end of PKD46 fragments were added and cycling was continued for an additional 30 cycles to amplify the combined product. Chimeric PCR products were purified using the DNA Clean & Concentrator kit (Zymo Research). Purified PCR products were electroporated into CFT073 cells containing the intact pKD46 plasmid and standard protocols were followed to complete the lambda red recombination procedure (23). Clones were checked for protein production after growth in the arabinose inducing conditions described above.

**RT-PCR and qPCR**

RNA was isolated from bacteria collected from expelled urine of infected mice according to the protocols of Vigil *et al.* Briefly, 3 groups of 5 CBA/J female mice each were grouped together and transurethrally infected with $10^8$ CFU of WT CFT073. Starting at 5 hours post infection, urine was collected from infected mice three times daily for three days. The urine was centrifuged at 13,000 RPM to pellet bacteria, supernatant was removed, and the 20µl of RNA protect (Qiagen) was added to each pellet. Pellets were frozen at -80°C until later use. RNA was extracted using RNeasy mini kit (Qiagen) and cDNA was created with SuperScript II Reverse Transciptase kit (Invitrogen), according to manufacturer instructions.

RT PCR was performed on the same cDNA and no reverse transcriptase control material using primers that span from the 3' end of one gene to the 5' end of another gene according to the above protocol and PCR products were visualized on a 1.2% agarose gel following electrophoresis, stained with ethidium bromide, and visualized on a ChemiDoc imaging system (BioRad).

**Purification of TosA, cell fractionation and generation of antiserum**

pBAD-*tosA* containing TOP10 cells were grown in LB at 37°C in shaking incubator until early exponential phase. At this point, L-arabinose was added to a final concentration of 10mM and the cultures were allowed to grow for an additional 3 hours. Cells were pelleted by centrifugation and frozen at -80C until purification was conducted. Pellets were thawed in 3ml PBS and passed three times through a French cell press. The cell lysates were centrifuged at 112,000xg for 30 minutes to pellet

unlysed cells and cell membrane material. The clarified supernatant was filtered through a 0.2µM syringe filter (Millipore). Lysate was then immediately loaded onto a HiPrep Sephacryl-300 26/60 gel filtration column (GE Healthcare) and the column was run with PBS + 0.5M NaCl + 0.25M urea. Collected fractions were analyzed by SDS-PAGE for TosA presence. Pooled fractions were concentrated using Centricon Plus-70 filter units with 100kDA cutoff and washed 2X with 70ml DPBS.

Cell fractionation was perfomed on pBAD-*tosA* TOP10 cellls and CFT073 $araB_P$-*tosC* cells grown in LB under arabinose inducing conditions, lysed, centrifuged briefly at 2,000xg to remove unlysed cells, and supernatant was centrifuged at 112,000xg for 30 minutes. Supernatant, representing soluble cytoplasmic/periplasmic proteins was carefully removed and pellets, representing membrane fractions, were resuspended in either DPBS + 1% Triton X-100 or DPBS + 2% Sarkosyl at room temperate for 30 minutes. Resuspended pellets were centrifuged at 112,000xg for 30 minutes, and supernatant (soluble membrane proteins) and pellets (insoluble membrane fraction) were saved. Filtered culture supernatant was concentrated on Amicon centrifugal filter units with 10,000 MWC pore size (Millipore). Concentrated culture supernatant, cytoplasmic/periplasmic, soluble and insoluble membrane fractions equivalent to the quantity present in 50µl culture volume were analyzed by SDS-PAGE stained with Imperial Protein Stain (Thermo Scientific) and imaged on a ChemiDoc imaging system (BioRad).

TosA for generation of α-TosA antiserum was prepared by separating cell lysate from pBAD-*tosA* preps prepared as described above using SDS-PAGE. Bands of ~250kDa

in size, corresponding to TosA, were excised from multiple gels, pooled, and sent to Rockland Immuochemical for production of α-TosA polyclonal antiserum in rabbits. Affinity purification of the resulting α-TosA antiserum was performed by crosslinking purified TosA to Ultrlink Biosupport resin (Thermo Scientific) according to standard protocols.

**Animal models of infection**

All mouse studies were approved by the University of Michigan Committee on the Use and Care of Animals. A previously described murine model of ascending UTI (48) was utilized as described below for immunocytochemistry experiments. A recently developed murine model of bacteremia (109) was used for non-lethal competition studies as described. A zebra fish model of ExPEC pathogenesis as described by Wiles *et al.* was used as a lethal sepsis model. Briefly, bacteria were prepared as described above for the murine model and 1nl of a suspension contain 1000 CFU was microinjected into the pericardial space or into the dorsal aorta of 48 hour post hatch zebrafish embryos. Co-challenge was conducted using equal mixtures of both strains and injecting 1000 CFU total in each body site. For forced TosA expression, WT and *araB*$_P$-*tosC* CFT073 were grown in LB + 10mM L-arabinose for 3 hours prior to infection. Infection proceeded as described above, but zebrafish were kept in water supplemented with 50mM L-arabinose during the course of the experiment. Lethality was assessed at regular intervals for 2 days following infection. CFU counts were determined by homogenizing infected fish embryos in 500µl sterile PBS and plating serial dilutions on MacConkey agar.

**Cell adherence and intact bladder adherence assays**

Adherence to cell lines grown *in vitro* was conducted using the following cell lines: Hs 769.T, UM-UC-3, MM55.k, Vero, and HEK 293 (ATTC) grown in DMEM + 10% fetal bovine serum, and Sv Huc1 grown in F12K + 10% FBS and RT4 grown in McCoy's 5a + 10% FBS (Invitrogen). All cell lines were cultured at 37°C and 5% $CO_2$. Cells were grown in sterile 6 well plates until 90-95% confluence and washed with 1ml sterile DPBS. WT and *araB*$_P$-*tosC* CFT073 were grown in LB + 10mM L-arabinose for 3 hours prior to infection, O.D.$_{600}$ reading taken and diluted in sterile DPBS, and used to infect washed mammalian cells at an MOI of 0.5-1. The plates were centrifuged at 500xg for 5 minutes and incubated at 37°C and 5% $CO_2$ for 10 minutes. Cells were washed twice with 1ml sterile PBS 2X and then lifted off in 1ml EDTA/0.25% Trypsin solution (Invitrogen). Serial dilutions of cell suspensions were plated on LB agar plates and CFU counts obtained after overnight growth at 37°C. The input dilution of bacteria was also plated to determine CFU count of the inoculums.

A modification of the murine model of ascending UTI described above was created to test adherence of UPEC on intact bladder epithelium. Female C57/Bl6 mice were transurethrally inoculated with $2x10^6$ CFU of either WT or *araB*$_P$-*tosC* CFT073 grown under the inducing conditions described above in a total volume of 25µl. 30 minutes after infection mice were euthanized and bladders were removed and cut in half with a sterile scalpel blade. Bladders were washed in 1ml sterile DPBS for 5 minutes on a rotating microcentrifuge mixer under low speed, transferred to new tubes and washed a second time to remove non-adherent bacteria. Washed bladders were

transferred to 3ml sterile PBS, homogenized, and plated on LB agar plates for CFU determination. Recovered bacteria were compared to CFU counts of the prepared inoculum.

**Immunogold-TEM**

Immunogold labeling of intact bacterium was carried out by growing WT CFT073, pBAD-*tosA*containg TOP10, and *araB*$_P$-*tosC* cells in LB + 10mM arabinose for 3 hours until late exponential phase of growth. 10µl of cell suspension was spotted on nickel coated formvar/carbon film nickel coated TEM grids (EMS). After 15 minutes, the liquid was wicked away with filter paper and blocked in 10µl of DPBS + 5% goat normal serum + 5% BSA for 15 minutes. Blocking solution was exchanged with 10µl of α-TosA antiserum diluted 1:250 in incubation solution, containing DPBS + 5% goat serum + 0.1% cold water fish skin gelatin + 0.1% BSA-c (EMS). After 15 minutes excess fluid was wicked away with filter paper and exchanged for 10µl incubation solution for 5 minutes. The wash was repeated and then exchanged with 10µl goat anti-rabbit IgG conjugated with 10nm gold particles (EMS) diluted 1:250 in incubation solution. After 15 minutes grids were washed twice with incubation solution and twice with distilled water. Grids were air dried and imaged on a Philips CM-100 transmission electron microscope. Grids for negative staining were incubated with 10µl of 1% uranyl acetate for 3 minutes.

**immunoflourescent imaging**

Female C57/Bl6 mice were transurethrally infected with $10^8$ CFU of pGEN mut3.1 containg CFT073 as described above. 48 hours post infection mice were euthanized and bladder, kidneys, and spleen was removed from each mouse and fixed in 10% buffered

formalin for 4 hours. Organs were then processed through a sucrose gradient and frozen in OCT (Andwin Scientific). 10µm thick frozen sections were cut from each organ.

Tissue sections were blocked in DPBS + 5% goat normal serum + 5% donkey normal serum for 15 minutes. α-TosA antiserum and chicken α-GFP antiserum (Abcam) was each diluted 1:500 in blocking solution was incubated for 30 minutes, followed by 2 washes in DPBS for 10 minutes. Anti-rabbit IgG Dylight 549 and anti-chicken IgY Dylight 488 (Jackson ImmunoResearch) was diluted 1:1000 and 1:500, respectively, in blocking solution and 1U of Alexa Fluor 647 phalloidin conjugate was added. Tissue sections were incubated in this solution for 30 minutes, then washed 2 times in DPBS. Sections were mounted with ProLong Gold antifade solution (Invitrogen) and imaged on a Zeiss LSM-510 META confocal laser scanning microscope.

**Vaccination model**

Purified TosA was diluted with DPBS + 10% Imject Alum (Thermo Scientific) and administered via subcutaneous inject to deliver 100ug protein in 100µl volume. Control mice were given a solution of DPBS + 10% alum. On the same day retro-orbital eye bleeds were collected from each animal and serum was separated and stored at -20°C until further analysis. At one week and two weeks after primary vaccination, animals were boosted with 25ug protein in 100µl DPBS + 5% alum. Control mice were vaccinated with DPBS + 5% alum. One week after the second boost, three weeks after primary vaccination, retro-orbital eye bleeds were collected and animals were challenged with WT CFT073. UTI challenge was performed according to the

transurethral model and bacteremia challenge was performed according to the protocols outlined above.

**Data analysis**

Statistical tests were carried out in the Prism statistical software package (Graphpad Software). Cell adherence data were analyzed by student's *t* test. Co-challenge data were analyzed by calculating the $\log_{10}$ transformation of competitive index (mutant/WT) and performing a Wilcoxon rank sum test with a hypothetical median of 0. Survival curves were analyzed by Mantel-Cox test.

## Results

***tosA* locus of *E. coli* CFT073 includes the genes for a RTX family member, a type 1 secretion system, and putative regulators.** The gene content and organization of the 15kb region of the CFT073 genome containing *tosA* (Figure 4-1A) is contained within the $PAI_{CFT073\text{-}aspV}$ pathogenicity-associated island (77, 78, 126). This genomic island had previously been implicated in enhancing fitness of UPEC in a murine model of an ascending UTI (77), by significantly contributing to the virulence of UPEC strain CFT073 in both bladder and kidney tissue of infected mice (119). The type 1 secretion system (T1SS) genes, *tosCBD,* were identified in a previous *in silico* analysis (93) as components of a conserved T1SS apparatus predicted to mediate the export of an RTX family member. *tosC* is predicted to encode a homolog of the outer membrane protein *tolC*; *tosB* and *tosD* are predicted to form the inner membrane ABC transporter and membrane fusion protein of the T1SS, respectively. Consistent with this predicted mode of export, TosA does not carry an identifiable cleavable leader peptide. Two additional open reading frames (ORF) designated c0364 and c0365 follow the putative RTX family member *tosA* in the same orientation. In addition, our analysis uncovered a previously unannotated ORF (designated *tosR*) directly upstream of the start codon of *tosC* (52 nt separate the two genes); the ORF has homology to *papB*, a regulator of the P fimbrial operon (8). The *tos* locus is flanked by fragments of genes predicted to encode remnants of a transposon system possibly indicating that this region was acquired by horizontal gene transfer.

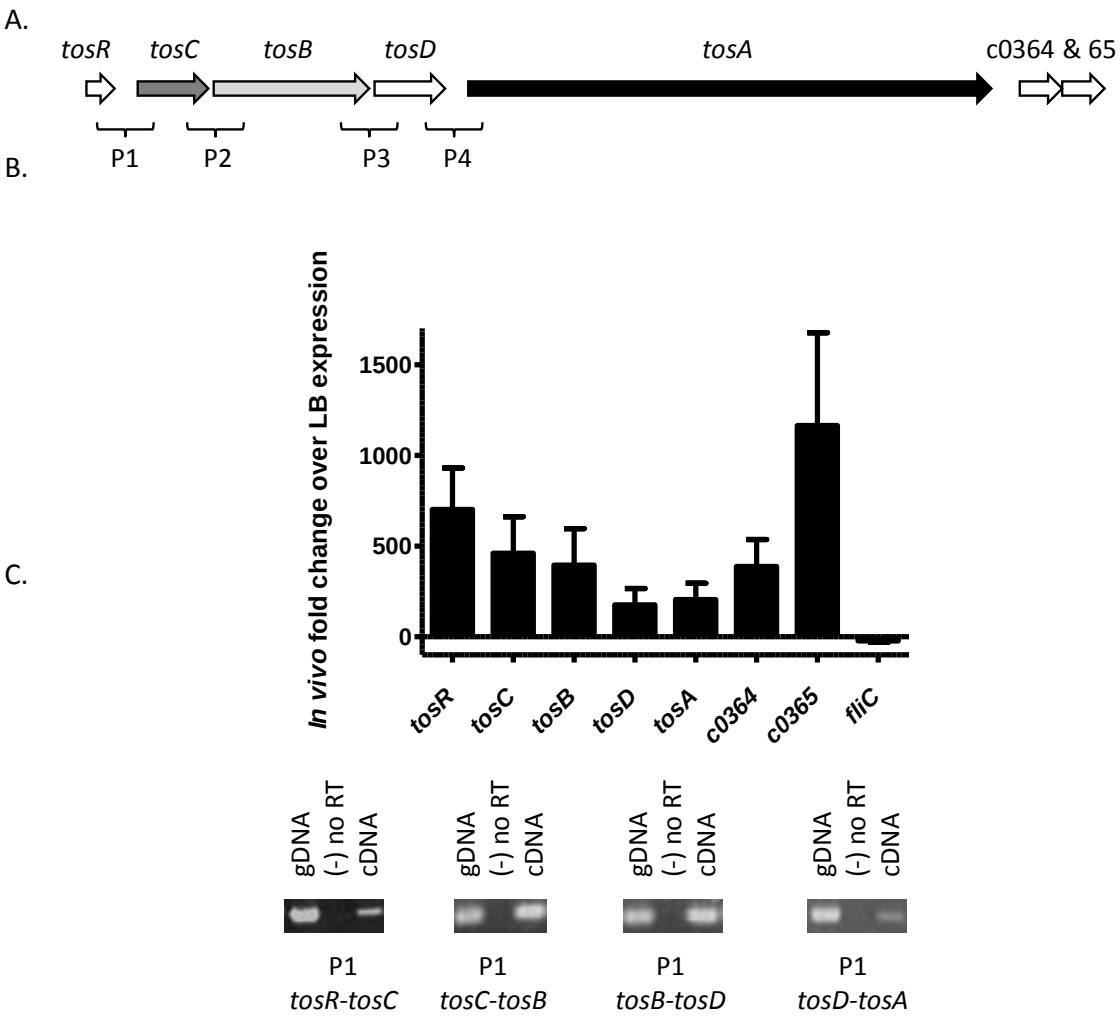



**Figure 4-1. Putative *tos* operon structure. A.** Schematic showing each ORF in the 15kb region of *tosA*. Area denoted by brackets below genes, marked P1-4, illustrate amplicons that span between adjacent genes shown in panel C. **B.** qPCR results using bacterial RNA collected from mice transurethrally infected with WT CFT073. Five mice were pooled for each of three replicates normalized to *gapA* expression and shown as fold change over growth in LB at 37°C shaking with aeration. **C.** RT PCR of amplicons spanning the region between the two genes denotate below each section. gDNA – genomic CFT073 DNA; (-) no RT – no reverse transcriptase control; cDNA – cDNA prepared as in panel B.

All ORFs in the *tosA* locus are expressed at significant levels *in vivo*, whereas their expression is undetectable during *in vitro* culture (Figure 4-1B). qPCR conducted on bacteria voided in the urine of UPEC-infected mice had previously determined that *tosA* expression was limited to the *in vivo* environment (119). To extend this analysis, a third set of probes was used to assess gene expression of all *tos* locus ORFs illustrated in Figure 4-1A. These data are consistent with our previous study and suggests a possible operon structure for these genes.

RT-PCR of amplicons using primers that span the region from the end of one gene to the beginning of the adjacent gene (Figure 4-1A) indicate that the genes *tosR* through *tosA* are transcribed as part of a continuous transcript (Figure 4-1C). qPCR was sufficiently sensitive to detect expression of c0364 and c0365 (Figure 4-1B), however, RT-PCR was not able to detect these genes. Thus there is not enough evidence at present to include c0364 and c0365 in the operon. This result is consistent with an operon structure consisting of *tosRCBDA*, but additional experiments are required to confirm this result and define the full extent of the *tos* operon.

***tosA* expression is limited to the host *in vivo* environment.** *E. coli* progress through the host urinary tract in an ascending manner (2, 18, 40), encountering a variety of environments in the process. To determine whether the *tosA* expression observed in Figure 4-1 was limited to a particular site in the urinary tract, we utilized a murine model of an ascending UTI. 48 hours after transurethral inoculation, bladder, kidney and spleens were removed and processed for immunofluorescence microscopy. A plasmid

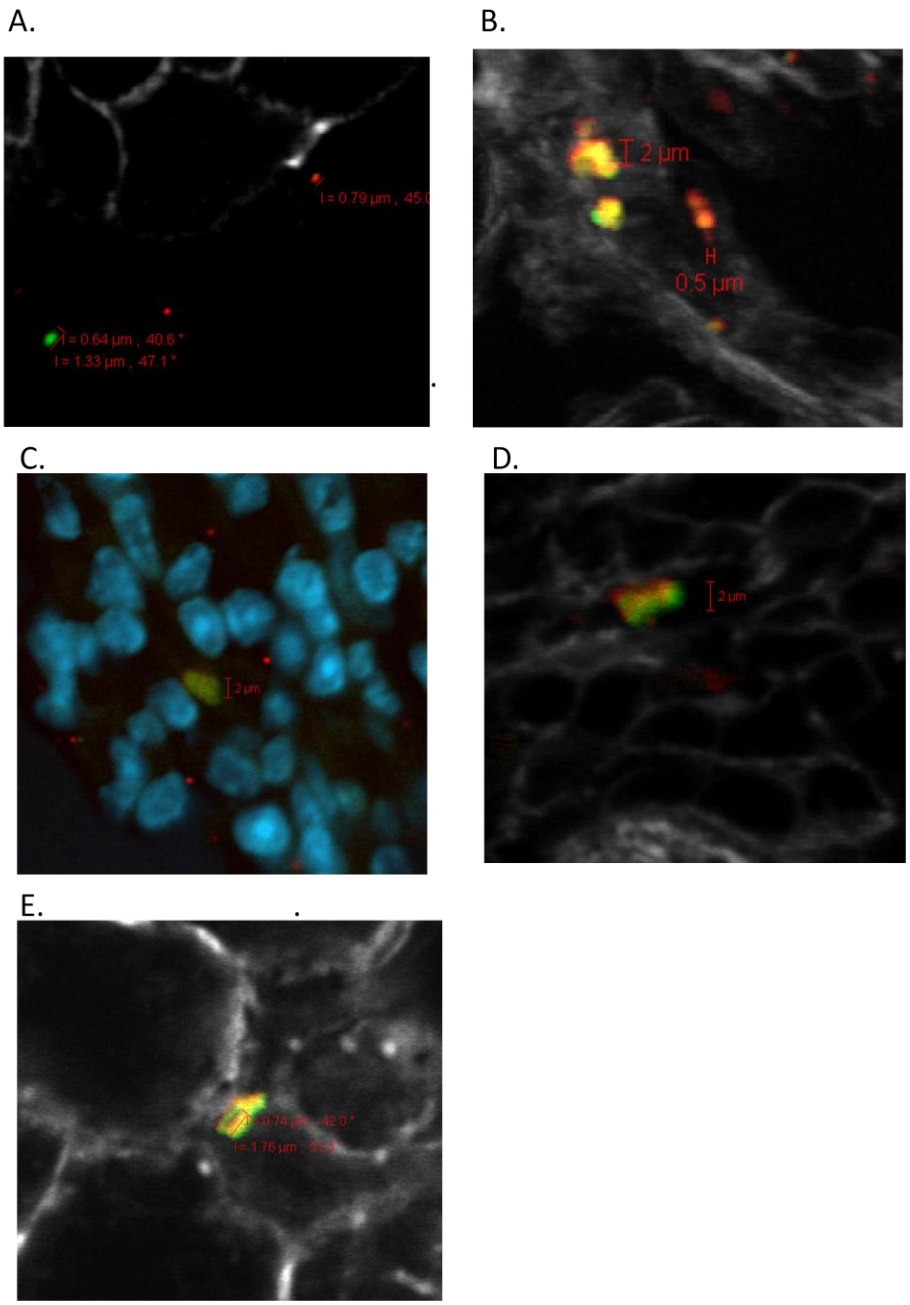



**Figure 4-2. TosA is expressed in infected bladder, kidney, spleen, and liver.** Female C57/Bl6 mice transurethrally (**A-C**) infected or infected via tail vein injection (**D-E**) with CFT073 pGEN mut3.1. 48 hours post infect bladders (**A**), kidneys (**B**), and spleens (**C**), and at 24 hours post infection spleens (**D**), and livers (**E**) were removed and process for immunofluorescence microscopy. α-GFP staining shown in green, α-TosA in red, phalloidin staining in white, and DAPI in blue. Measurements of individual bacteria denoted by scale bars.

that constitutively expresses GFP was utilized to mark wild-type CFT073 bacteria and TosA-antiserum stained bacteria expressing TosA. Tissue sections of bladder, kidney, and spleen (Figure 4-2A, B, and C, respectively) demonstrate that UPEC express TosA protein in each organ 48 hours post-inoculation. Additionally, in separate experiments, our murine bacteremia model reveals that TosA is expressed at 24 hours post inoculation in spleen and liver tissue (Figure 4-2D and E, respectively), two sites where UPEC was previously determined to display enhanced fitness and survival over non-pathogenic *E. coli* (109). The results of these two experiments suggest that TosA expression is important for UPEC during both ascending UTI and disseminated infections.

**Genetic manipulation of strain CFT073 allows *in vitro* expression of TosA.** Because *tosA* expression is limited to the *in vivo* environment (119), two constructs allowing inducible expression of *tosA* were engineered to allow additional experimentation. A plasmid-based system (pBAD-*tosA*) that carries the cloned *tosA* gene under an arabinose-inducible promoter, but which lacks the T1SS genes and accessory gene content illustrated in Figure 4-1A, allows *in vitro* TosA production and purification (Figure 4-3B). However, the TosA produced remains confined to the soluble cytoplasmic fraction of bacteria (Figure 4-4A).

To construct a strain that expresses *tosA in vitro* and localizes the protein to its wild-type location, we used the lambda red recombinase system and a hybrid PCR product containing an arabinose-inducible promoter to recombine into the

chromosome of CFT073, placing the *araB* promoter sequence upstream of the start codon of *tosC* (Figure 4-3A). Upon arabinose induction, this *araB*$_P$-*tosC* construct produces TosA protein that is detectable using an anti-TosA polyclonal antiserum. In contrast, wild-type CFT073 cultured under identical conditions does not produce detectable levels of the protein (Figure 4-3C).

**TosA protein localizes to the outer surface of *E. coli*.** When both arabinose-induced *E. coli* CFT073 (pBAD-*tosA*) and the *araB*$_P$-*tosC* construct were subjected to cell fractionation into supernatant, soluble and membrane fractions, TosA localized to different cellular fractions. The soluble fraction contained the majority of the detectable TosA when expressed from pBAD-*tosA*, while the membrane fraction contained the majority of this protein when induced by arabinose in the *araB*$_P$-*tosC* construct (Figure 4-4A). An equivalent volume of concentrated, filtered culture supernatant from both cultures indicated that, in contrast to the majority of known RTX family members (76), TosA is not fully secreted from the bacterium. In arabinose-induced *araB*$_P$-*tosC* bacteria, which contain the intact putative *tos* T1SS, TosA exhibited differential solubility in two common detergents used for cellular fractionation experiments and was extracted efficiently into 2% sarkosyl fraction, but remained insoluble in the presence of 1% Triton X-100 PBS. This suggested that TosA localized to the membrane fraction, an observation confirmed by enzyme function analysis of the soluble and insoluble fractions (data not shown), but it could not differentiate between inner membrane or outer membrane localization. The results of the fractionation

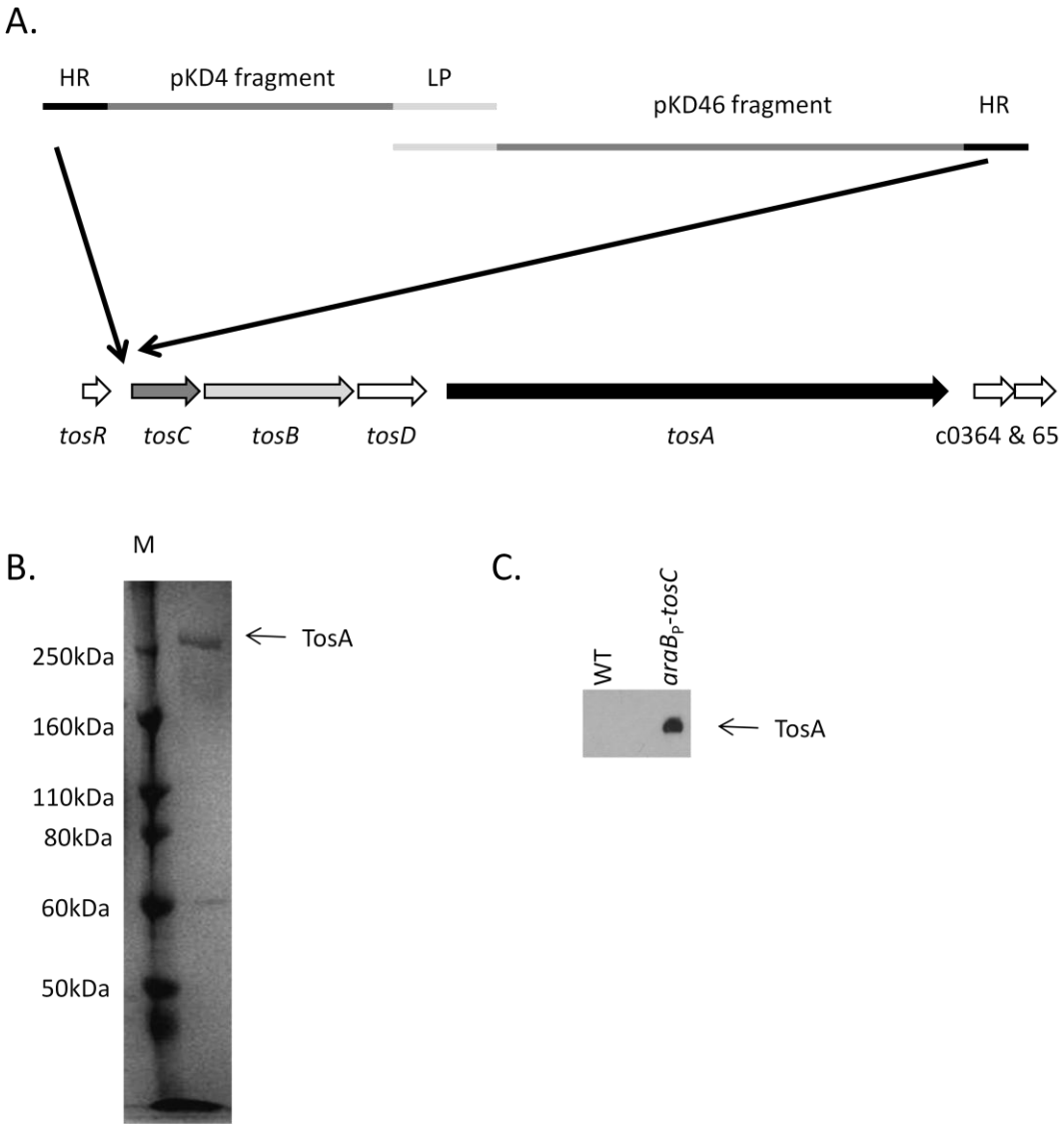



**Figure 4-3. *In vitro* TosA expression constructs. A.** Schematic illustrating construction of *araB*$_P$-*tosC* CFT073. A chimeric PCR construct consisting of two plasmid fragments generated from pKD4 and pKD46, containing a kanR conferring gene and an arabinose inducible promoter, respectively, were amplified with primers that contain homology to the region between *tosR* and *tosC* (HR). A second set of primers, one the reverse complement of the other, were used as linking primers to combine the two PCR products together in a third PCR reaction (LP). **B.** TosA purified from pBAD-*tosA* construct purified by gel filtration chromatography. 0.8µg protein loaded, M – size standard marker. **C.** Western blot using polyclonal α-TosA antiserum of WT and *araB*$_P$-*tosC* CFT073 cells grown under arabinose inducing conditions. 25ul of late exponential culture was loaded per lane; TosA marks the position just above the 250kDa size standard.

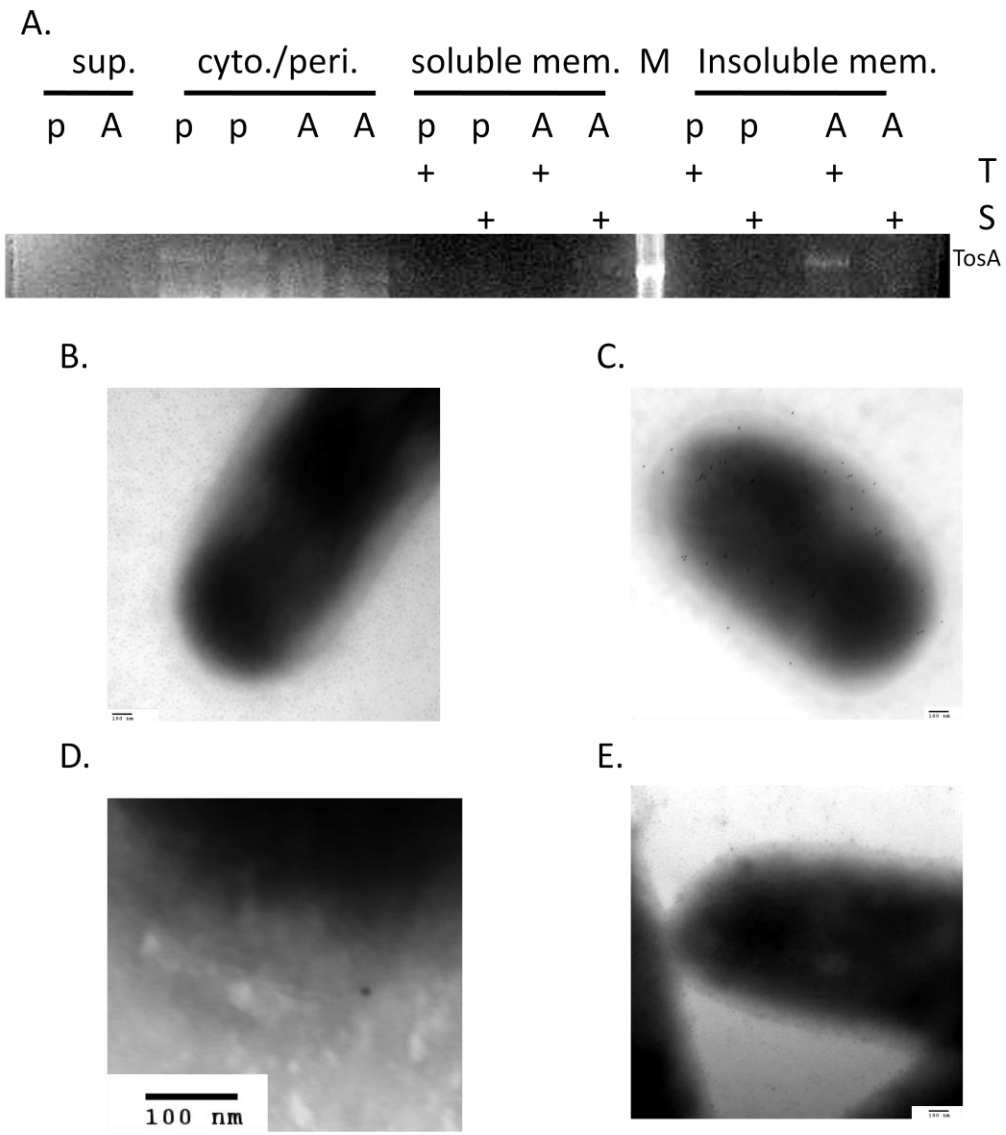



**Figure 4-4. TosA localizes to the outer membrane. A.** Coomassie blue stained gel of cell fractionation results. The equivalent of 50ul of culture of pBAD-*tosA* (p) and *araB*$_P$-*tosC* CFT073 (A) grown uder arabinose inducing conditions was loader per lane. Sup. – culture supernatant; cyto./peri. – soluble cytoplasmic and periplasmic fraction; Membrane fractions solublized in either DPBS + 1% Triton X-100 (T) or DPBS + 2% sarkosyl (S). **B-E** Immunogold-TEM micrographs of α-TosA stained bacteria. **B.** Δ*tosA* CFT073 do not show any immunogold particles. **C.** *araB*$_P$-*tosC* CFT073 marked with numerous immunogolg particles. **D.** *araB*$_P$-*tosC* CFT073 negatively stained with uranyl acetate show immunogold particles on the surface of the bacterial outer membrane. **E.** pBAD-*tosA* TOP10 cells show no immunogold staining. All images acquired at 64,000X magnification; scale bares denote 100nm.

experiment were also confirmed with immunoflouresence microscopy on pBAD-*tosA* containing cells and *araB*$_P$-*tosC* CFT073 cultured under inducing conditions (Figure 4-S1).

Immunogold labeling of arabinose-induced *araB*$_P$-*tosC* bacteria expressing TosA, imaged by transmission electron microscopy, localizes TosA to the outer surface of the bacterium. When *araB*$_P$-*tosC* and WT CFT073 are cultured *in vitro* under arabinose-inducing conditions, only bacteria expressing TosA are marked with immunogold particles, while WT CFT073 remains unlabeled (Figure 4-4B and C). Negative staining of these cells reveals that the immunogold particles marking TosA localize to the surface of the bacterium (Figure 4-4D). Finally, pBAD-*tosA* cells cultured under inducing conditions show no immunogold staining (Figure 4-4E), confirming that TosA requires a specific transport mechanism, not found in the TOP10 *E. coli* strain that carries the pBAD plasmid, to be exported out of the cytoplasm. While the specific transport system is assumed to be encoded by *tosCBD,* additional experiments will be required to confirm this.

***tosA* expression enhances adherence to the epithelial cells that line the urinary tract.** Given the surface localization of TosA, we reasoned that this protein may mediate adherence to epithelial cells that line the host urinary tract. When WT CFT073 and *araB*$_P$-*tosC* bacteria, both cultured in arabinose, were incubated with a murine kidney epithelial cell line (MM55.K) *in vitro*, 5.4% WT bacteria were found to adhere to the epithelial cells after washing with PBS. Bacteria expressing TosA were observed to adhere under these conditions at nearly three times this level (15.4%) (Figure 4-5A),

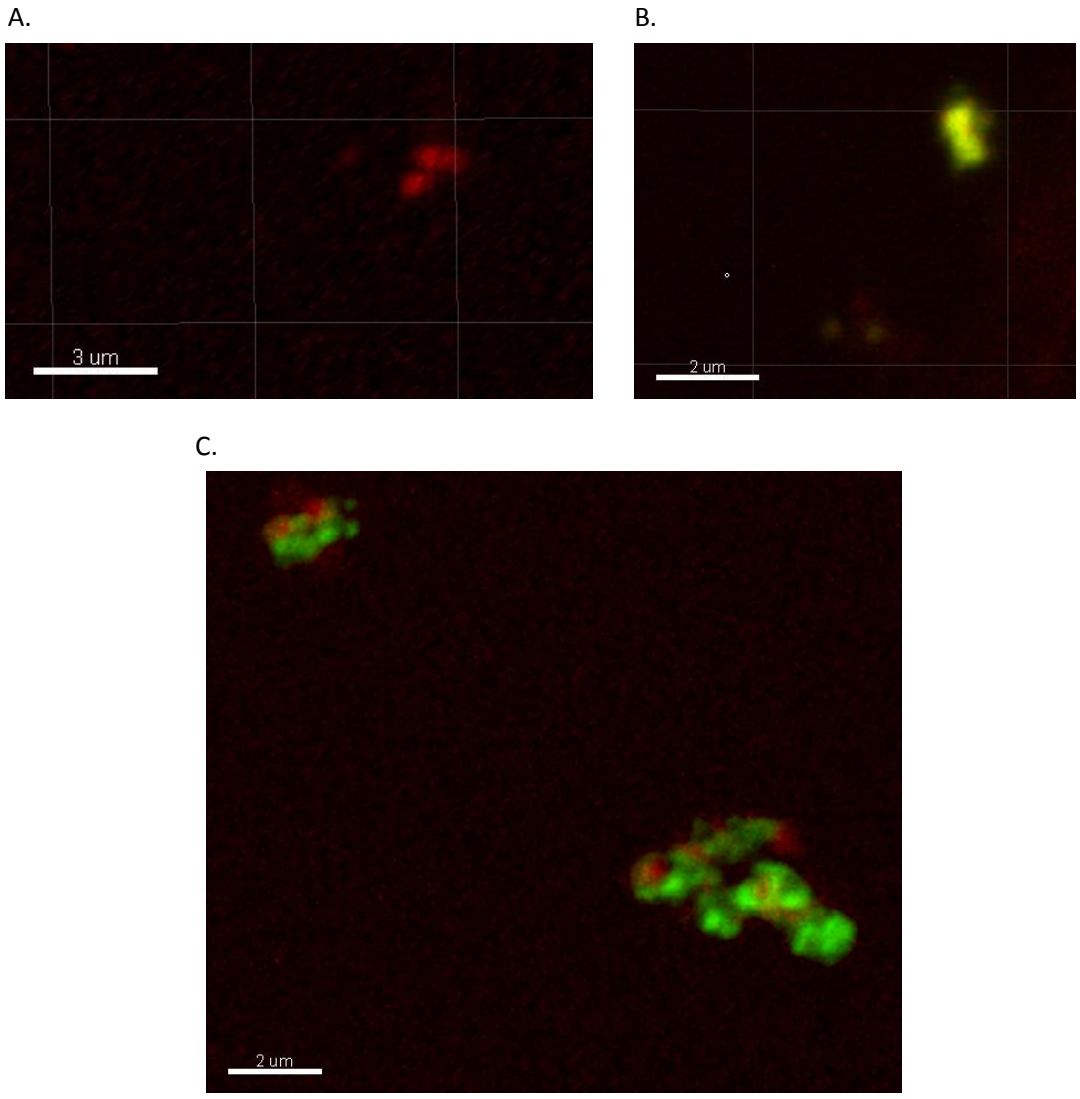



**Figure 4-S1. TosA localization with and without the linked T1SS.** WT CFT073 (**A**), pBAD-tosA TOP10 (**B**), and *araB*$_P$*-tosC* CFT073 (**C**) bacteria were grown in arabinose inducing conditions and spotted on glass slides. Cells were permeabilized with 0.2% Triton-X 100 and stained with propidium iodide (red) and α-TosA antiserum (green). Images were acquired as Z-stacks on a confocal microscope and individual slices are shown in each panel. Co-localization of nucleic acids and TosA appears as yellow staining.

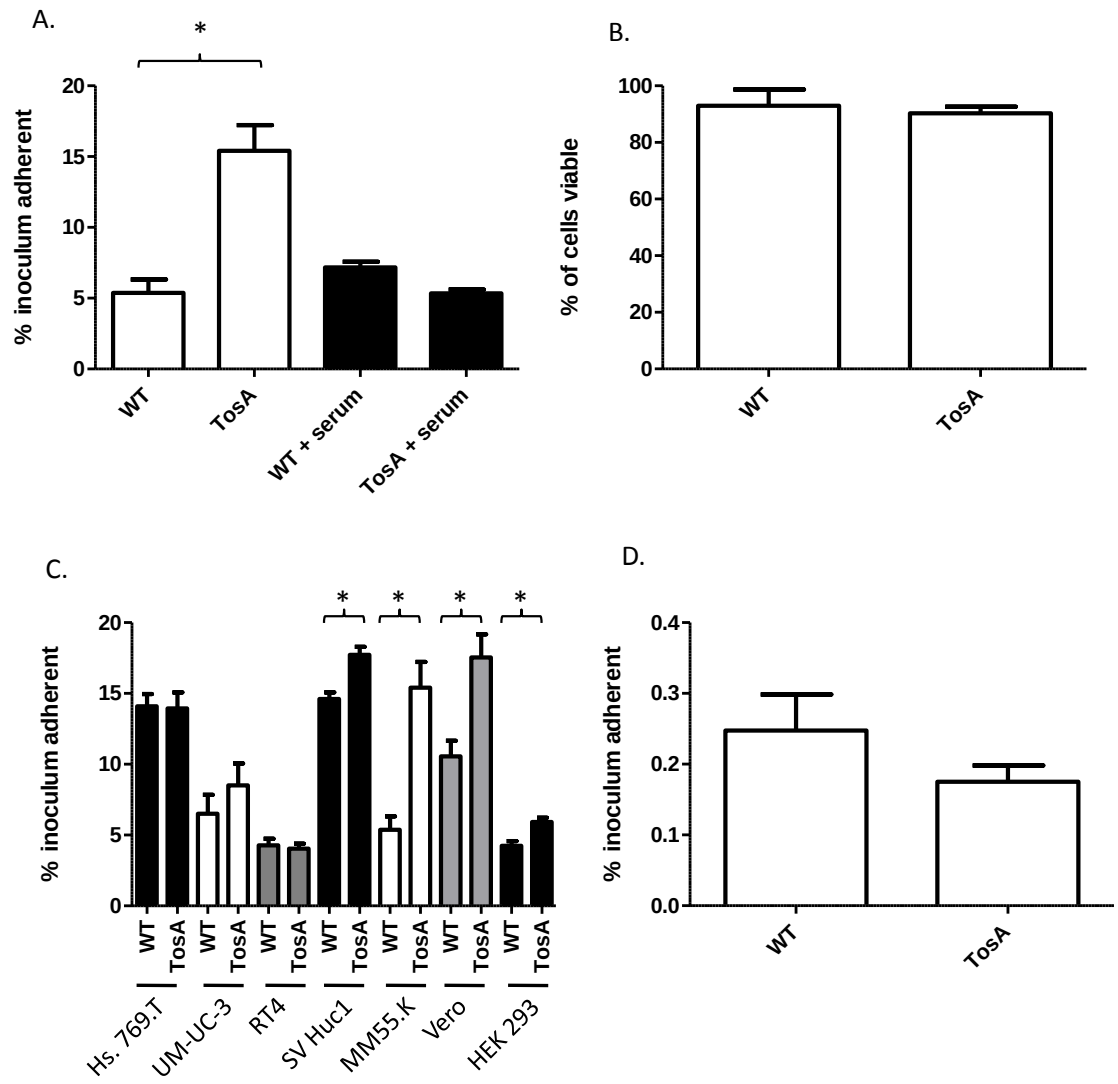



**Figure 4-5. TosA mediates adherence to epithelial cells of the upper urinary tract. A.** and **C.** WT CFT073 (WT) and *araB*$_P$-*tosC* CFT073 (TosA) grown under arabinose inducing conditions were used to test adherence to confluent monolayer of cells *in vitro* at an MOI of 0.5-1. After incubation and washing, cells were lifted off and plated for CFU determination and graphed as % of adherent cells recovered. A**.** MM55.K cell adherence assay was repeated after cells were pretreated for 5 minutes with α-TosA antiserum. **B.** Serum pretreatment was repeated and plated for CFU determination. Viability was assessed as % of input bacteria that were recovered after serum treatment. **D.** *In vivo* adherence assay conducted on female C57/Bl6 mice transurethrally inoculated with WT and TosA expressing CFT073. After 30 minutes, bladders were remove, washed, and CFU/organ determined. % of input cells that remained adherent after treatment graphed as in panels **A.** and **B.** * $p<0.05$ calculated by student's t-test.

suggesting that this novel RTX family member mediates adherence to epithelial cells that line the host urinary tract.

The adherence assay was repeated with WT and TosA-expressing bacteria that had been pre-incubated for 5 min with α-TosA antiserum. The results of the adherence assay showed an ablation of the TosA-mediated increase in adherence back to WT levels, but had little impact on WT bacteria treated with the same antiserum (Figure 4-5A). Given the lack of reactivity against WT bacteria cultured under identical conditions (Figure 4-3C), only bacteria expressing TosA were expected to be impacted by the pre-treatment. To confirm that bacteria treated with α-TosA antiserum were viable and that this result was not the result of complement-mediated lysis of TosA-producing bacteria, bacterial suspensions were pretreated with α-TosA antiserum and plated to assess viability. No significant difference was observed in viability between the number of WT bacteria and bacteria expressing TosA (Figure 4-5B). This result confirmed that the increase in adhesion was specific to the production of TosA protein.

The adherence assay described above was repeated with different cell lines that together represent the epithelial cell types that line the entire host urinary tract. UPEC adhered strongly to cell lines derived from a human transitional cell carcinoma of the urethra (Hs 769.T) and at lower levels in two cell lines derived from a human transitional cell carcinoma and a transitional cell papilloma of the bladder (UM-UC-3 and RT4, respectively). However, no differences were observed in adherence levels between WT and *araB*$_P$-*tosC* cells expressing TosA (Figure 4-5C). WT CFT073, a prototype human

pyelonephritis/urosepsis isolate (126), was able to adhere to cells derived from the epithelial lining of the lower urinary tract efficiently without the need for TosA.

In contrast, bacteria expressing TosA adhered in higher numbers than WT CFT073 to all tested cell lines derived from the upper urinary tract. A transformed transitional cell epithelial line derived from a normal adult human specimen (SV Huc 1) allowed WT bacteria to adhere in high numbers; however, a significant increase was observed when TosA was expressed. Epithelial cells derived from embryonic human kidney tissue (HEK 293) showed a modest increase of adhesion with TosA expression and an even greater increase was observed for a transformed adult primate kidney epithelial cell line (Vero) (Figure 4-5C).

To confirm that the result obtained in celazls lines under *in vitro* conditions was relevant during the course of infection, a similar adherence assay was developed by modifying an established murine model of ascending UTI (48). Thirty minutes after transurethral inoculation of either WT or *araB*$_P$-*tosC* bacteria induced with arabinose, mice were euthanized and bladders were removed and cut in half and weighed. Bladder tissue was subjected to PBS washing similar to the *in vitro* adherence assays to remove non-adherent bacteria. The number of bacteria that remained adherent (Figure 4-5D) demonstrated no significant difference between adherence levels between both strains of bacteria. Taken together with the previous results, this indicates that TosA can mediate adherence to the epithelial cell lining of the upper urinary tract but does not contribute significantly to adherence of UPEC in the urethra or bladders of infected hosts.

***tosA* enhances fitness in two animal models of bacteremia and sepsis.** Because *tosA* was discovered in a urosepsis isolate and is expressed in spleen and livers during bacteremia (Figure 4-2D and E), the possibility that the protein has a secondary role during bacteremia and sepsis was investigated using two different animal models of extra-intestinal pathogenic *Escherichia coli* (ExPEC) pathogenesis. First, a Δ*tosA* strain that had been previously shown to have reduced fitness in the murine model of ascending UTI during co-challenge (77) and during independent challenge (119) was tested in our recently developed murine model of bacteremia (109). Our group developed this model to assess the impact of known and putative UPEC virulence factors in the success of bacteria that spread from the kidneys to distant organ sites. We have found that many genes that confer an advantage in the urinary tract also confer a fitness advantage in this model. Using this model, CFT073 Δ*tosA* was recovered in significantly lower CFUs in both spleen and liver tissue 24 hours post infection (P=0.023 and P=0.008, respectively) (Figure 4-6), suggesting a second role for this gene during bacteremia.

A second animal model using zebrafish to study ExPEC pathogenesis during sepsis (130) was also employed to study the same Δ*tosA* strain. In the zebrafish model, WT bacteria were recovered at equal levels as Δ*tosA* bacteria during co-challenge in the

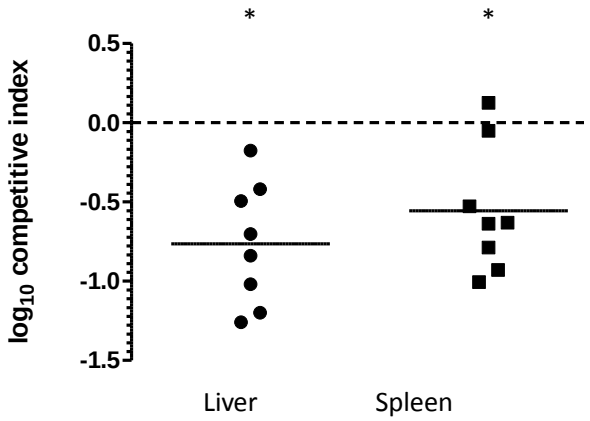



**Figure 4-6. TosA enhances fitness during bactermia.** A murine model of bacteremia was used to compete WT CFT073 against Δ*tosA* CFT073. An equal mixture of both strains was diluted to deliver $10^6$ CFU in 100ul, which was injected in the tail vein of female C57/Bl6 mice. 24 hours post infection liver and spleen tissue was removed and CFU of each strain was determined. The log10 competitive index (mutant/WT) is shown with lines representing median values. * $p<0.05$.

bloodstream and when injected into the pericardial space (Figure 4-7A). During independent challenge in the bloodstream, identical survival curves were obtained for the two strains, indicating that in this model system *tosA* does not confer an advantage to bacterial survival (Figure 4-7B).

The discrepancy between the two animal models was resolved by staining zebrafish infected with WT CFT073 for TosA and imaging by immunofluorescence microscopy. While in mice, this strain expresses TosA in the spleen following transurethral inoculation and in the spleen and liver following intravenous injection (Figure 4-2), no detectable TosA protein was observed in bacteria either in the pericardial space of the fish (Figure 4-7C) or that had disseminated throughout the circulatory system (data not shown). However, in fish that had been infected with *araB*$_P$-*tosC,* bacteria produced TosA protein if the fish were kept in water supplemented with arabinose (data not shown) and allowed this model system to be utilized with this strain of CFT073.

When induced *araB*$_P$-*tosC* bacteria were inoculated directly into the bloodstream of zebrafish that were maintained in water supplemented with arabinose, almost 80% of infected fish died by 18 hr post inoculation (Figure 4-7D). Fish infected with the same strain, but not treated with arabinose took nearly 36 hours to reach this level of mortality ($P<0.01$). When zebrafish were kept in arabinose-supplemented water and mock infected with PBS + arabinose more than 80% of the fish survived to 48 hours,

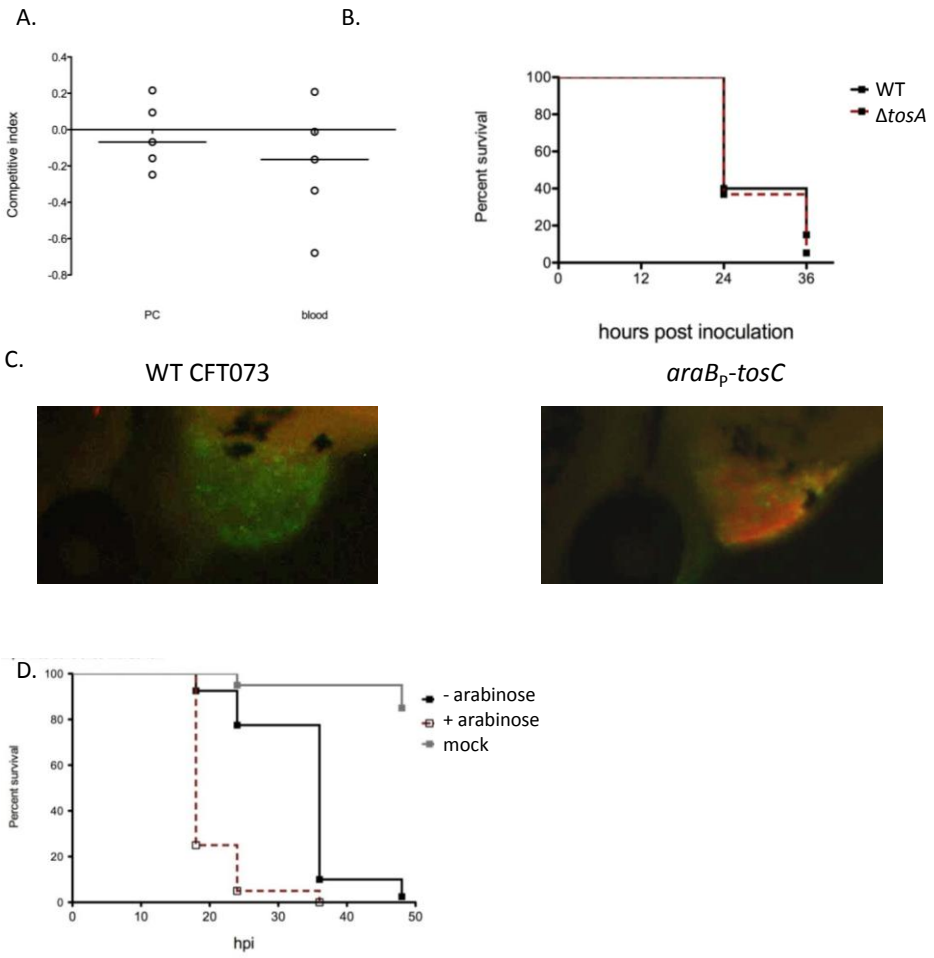



**Figure 4-7. TosA enhances lethality in a zebrafish model of ExPEC pathogenesis.** 48 hours post hatch zebrafish embryos were microinjected with bacteria into either the pericardial space around the developing heart or into the dorsal aorta to initiate bacteremia and sepsis. **A.** WT and Δ*tosA* CFT073 were mixed in equal proportions and inoculated into pericardial space (PC) or dorsal aorta (blood) with 1000 CFU in 1nl. 16 hours post infection embryos were homogenized and CFU load of each strain was determined. Competitive index (WT/mutant) is shown. **B.** Survival curve of zebrafish embryo's injected in the dorsal aorta with 1000 CFU either WT or Δ*tosA*. **C.** Immunofluorescence microscopy of pericardial space of embryos injected with WT CFT073 or *araB*$_P$-*tosC* bacteria grown under arabinose inducing conditions. Fish were kept in water supplemented with arabinose and 16 hours post infect fixed and processed for imaging. α-*E. coli* antibody staining marks bacteria green and α-TosA antiserum is displayed in red. **D.** Survival curves of fish infected with *araB*$_P$-*tosC* construct incubated with or without arabinose induction. Mock infected embryos were incubated in the presence of arabinose.

indicating that the arabinose treatment *per se* was not contributing to the enhanced killing kinetics observed with TosA expression.

**Vaccination against TosA confers protection against urosepsis caused by a human pyelonephritis UPEC isolate.** In the final set of experiments presented in this study, we determine that TosA protein contributes to the ability of UPEC to spread from the kidneys to distant organ sites. TosA was purified from the pBAD-*tosA* system, mixed with alum as adjuvant and administered to mice by subcutaneous injections following the vaccination schedule outlined in Figure 4-8A. Vaccinated mice were transurethrally inoculated with WT CFT073 and CFUs in bladder, kidney, and spleen were quantified and compared to CFUs of mock-vaccinated mice that only received PBS + alum injections. Vaccination had no impact on bacterial survival in the urinary tract, however, a significant decrease in CFU/g spleen tissue was observed at 48 hours post infection (Figure 4-8B). The vaccination protocol was repeated and mice were challenged by intravenous injection and CFU levels in spleen and liver were determined 24 hours post inoculation (Figure 4-8C). In this model, however, no protection was observed among vaccinated mice (data not shown), suggesting that the protection observed in the previous vaccination was specific to bacteria that ascended the urinary tract and later disseminated into the bloodstream.

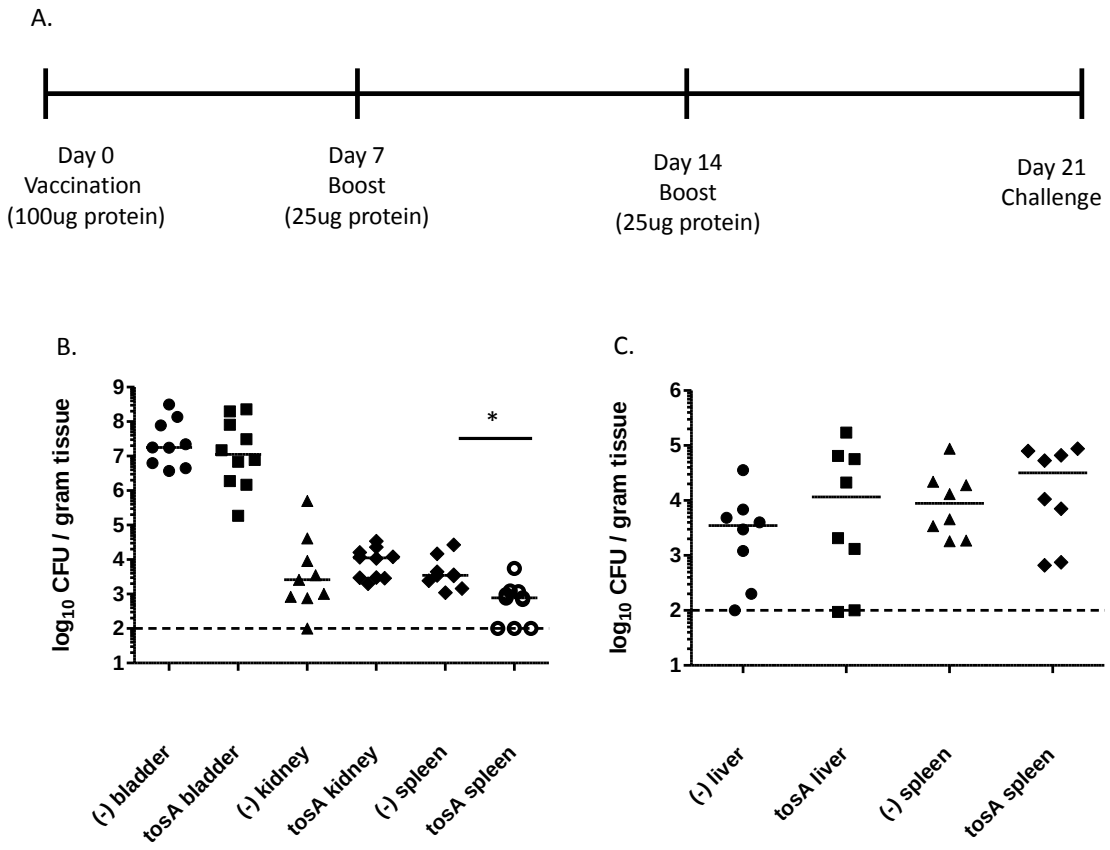


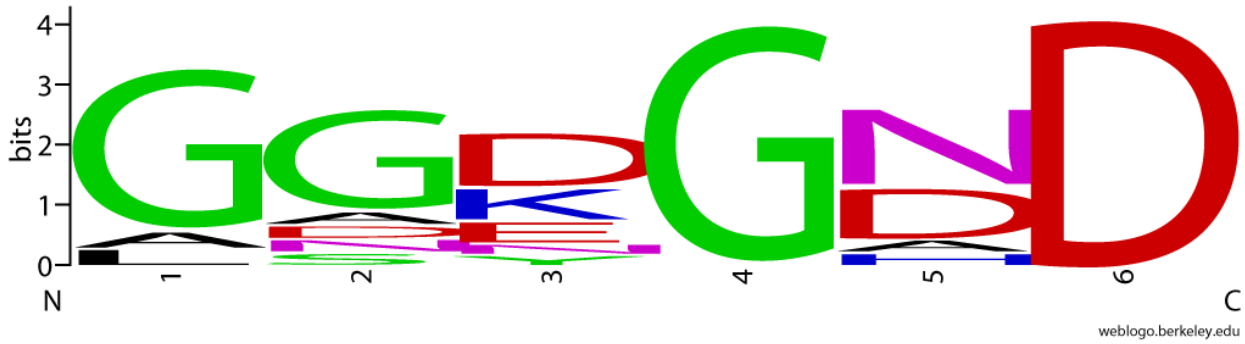

**Figure 4-8. TosA vaccination protects against urosepsis. A.** Purified TosA was utilized in an experimental vaccine model according to the following schedule: female C57/Bl6 mice were vaccinated with 100ug protein, then boosted one and two weeks later with 25ug. Three weeks after primary vaccination mice were challenged with WT CFT073 via transurethral infection **B.** or tail vein injection **C.** CFU/g tissue was determined for each organ site after 48 hours of infection (**B**) or 24 hours post infection (**C**) and compared to mice mock vaccinated with PBS. * p=0.004.

**Discussion**

The gene sequence of *tosA* had previously been annotated as belonging to a putative RTX family of exoproteins based on its association with the conserved T1SS (93) and the presence of characteristic glycine- and aspartate-rich repeats in the C-terminus of the protein (76). Our group has independently identified 9 repeats (Figure 4-S2) that match closely to the GGXGXD consensus RTX sequence (76) in the C-terminus of the protein. Additionally, a novel repeat was discovered in *tosA* consisting of 5 large tandem repeats ~1kb in length (Figure 4-S3).

The strict regulation of *tosA* expression (only *in vivo* induction) appears to be unique among RTX family members. Little is known about the regulation of RTX operons beyond the regulatory mechanisms defined for α-hemolysin (26, 50, 67). CFT073 displays hemolytic activity during *in vitro* growth (83), but does not express detectable levels of TosA (Figure 4-3C). This suggests that different regulation mechanisms operate on the two RTX systems present in this strain. Furthermore, the well defined RTX operons, conferring diverse functions ranging from the pore forming toxins to secreted proteases, do not appear to contain genes that potentially regulate gene expression (76). The presence of *tosR*, a gene with homology to the *papB* family of fimbrial gene expression regulators (58, 134), immediately upstream of a novel UPEC adherence factor gene remains unexplained. One attractive hypothesis is that UPEC, which are known to utilize *papB* family members to coordinate expression of the multitude of fimbriae encoded in the genomes of UPEC (56, 75, 135), adapted the same network of regulators to allow *tosA* expression to respond in concert

A.




B.


**Figure 4-S2. TosA contains characteristic RTX repeats.** DNA sequences of *hlyA* and *tosA* from the CFT073 genome were searched for the RTX repeat structure GGXGXD using PATTINPROT software program set to find regions with 75% identity to the consensus RTX sequence (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_pattinprot.html). Twelve repeats were found for *hlyA* (**A**) and 9 for *tosA* (**B**). The repeats were aligned and used to create sequence logos using WebLogo software (http://weblogo.berkeley.edu/logo.cgi). Height of letters at each position indicates the level of conservation at that amino acid position.

A.


B.


C.

GSVSVELPGDANKGDTVDVTFEDEKGGKHTVTLEKGDNGWTSSDPTLIPDSTGDKATIPADNVKDN
SEVTGVAKDPSGNESDPSTVTSKTDGVADAPVLSIPEVTDGYANADELKDGLQAEVTLPAGTAEGA
VITLTVTRPDKTTENVTHTVTKDEVAAGKVSMDIPKDAVIDGQNSVSVTLTQGSNPAKPGNVVDFA
ADTQIPGDTDGDGATDATPVVAIPEAADGVNAEELKDGVQTEVTVPKGSAAGDTLTLTVTKPDGTT
DTVEHTLTADEVAAGKADVTIPADKVTADGQYSVTAEITDPAGNTSGQGQPADFAVDTVAPSAPVL
KAEDD

**Figure 4-S3. TosA contains a novel repeat structure. A.** The Cft073 genome annotation originally described *tosA* (originally termed *upxA*) as a 5,325nt long ORF encoding a putative protein 1,610 amino acids long (predicted MW 164kDa). Cloning and expression of *tosA* indicate the gene is ~8,000nt long and encodes a protein slightly larger than 250kDa (Figure 3). Sequence analysis identified two tandem repeats in the nucleotide sequence, depicted in panel **A** as gray bars underneath a schematic of *tosA* which shows the repeats relative location close to the 5' of the gene and N terminus of the predicted protein. Restriction enzyme mapping of full *tosA* PCR products is consistent with five of these repeats in a tandem array (data not shown). **C.** Amino acid sequence of the tandem repeats is 335 residues long.

with changes in expression of other adherence factors. Further work will be required to determine whether *tosR* is a papB family member and determine its effects on the expression of other adherence factors. Additionally, while RNA transcripts corresponding to ORFs c0364 and c0365 were detected, no function has yet been assigned to these sequences.

The surface localization of TosA also sets this RTX family member apart from other RTX proteins. Only the RTX members that comprise the S-layer of certain bacteria, notably the motility-associated RTX of *Cyanobacteria* (13) and the RTX surface coat of *Campylobacter spp*. that protects against immune system attack (92), are known to remain associated with the bacterium. All other known RTX members are fully secreted into the external milieu. However, many RTX proteins can bind to host cell membranes, many with selectivity (115). The only adaptation that would be required to convert such a receptor-specific RTX into an adherence factor would be the incorporation of the protein into the outer membrane. Previous *in silico* analysis identified a putative transmembrane spanning region in the N-terminus of TosA (93), however, additional study is required to determine if this region allows TosA to embed in the outer membrane.

A possible secondary role for TosA of mediating survival during bacteremia and urosepsis is supported by three lines of evidence. First, a Δ*tosA* strain of CFT073 demonstrated reduced fitness when competed against WT bacteria in our murine model of bacteremia (Figure 4-6). This model has determined that many UPEC virulence factors, including many non-toxins, enhance fitness in the host (109). Second, the

protection observed in TosA-vaccinated mice from the final stage of an ascending UTI (Figure 4-8A), urosepsis, may indicate that TosA plays a role, or at least is expressed during, the transition from the urinary tract to the bloodstream. Finally, the enhanced lethality of TosA-expressing bacteria in the zebrafish sepsis model (Figure 4-7E) argues for an important role during disseminated infections. However, these data do not specifically address whether the protein has a direct role in host toxicity. Several attempts by our group to observe such a role, including incubating TosA-expressing bacteria or purified TosA with kidney epithelial cells *in vitro*, have failed to provide any evidence for a direct role in cytotoxicity (data not shown). Given the results at hand, TosA may only be enhancing host damage and survival indirectly by allowing bacteria expressing the protein to bind to target cells, enhancing delivery of other secreted products.

Finally, the experimental vaccination results presented in Figure 4-8 raise the possibility that a novel UTI/urosepsis therapeutic could be developed to target TosA. This model utilizes an earlier end point than the transurethral UTI model (24 *vs* 48 hr). It is possible that the host immune system has not had sufficient time to mount an effective response in the short time frame. It is also possible that the difference was due to the manner in which bacteria entered the blood stream in between the two models impacted the results. In the bacteremia model, the bacteria were grown *in vitro*, a condition were tosA is not expressed (119), and injected all at once. In contrast, it is expected that movement from the urinary tract into the bloodstream occurred throughout the course of the infection, and given the results of this study, the bacteria

in this model were already expressing TosA. This may have presented the immune system of vaccinated animals with two very different challenges and could explain the difference in the two vaccination experiments.

Regardless of the differences observed in the two vaccination models, two additional observations suggest that further development of anti-TosA therapeutics may prove beneficial. In a previous study, our group identified that bacteria flushed out of the urinary tract in voided urine from women with UTI show detectable transcript levels of *tosA* (46), indicating that the restricted expression of *tosA* may include the human urinary tract. Finally, in a recent survey conducted by our group of urinary tract and fecal isolates of *E. coli*, we discovered that *tosA* is present in ~25% of uncomplicated UTI isolates, but rare in fecal isolates. 98% of strains that carry *tosA* were part of the B2 phylogenetic lineage of *E. coli*, and most belonged to a subgroup that were enriched for highly pathogenic *E. coli* (Vigil *et a.l*, under review). Thus, while not all UPEC express this protein, TosA may still be an attractive target for the development of multivalent vaccines to combat uncomplicated urinary tract infections and the severe sequelae that can result.

# CHAPTER 5

## CONCLUSIONS AND FUTURE DIRECTIONS

The major conclusions of each study are detailed in the preceding chapters and only a brief outline of these will be presented here. Of all the genes present in the genome of UPEC that contribute to UTI pathogenesis, *tosA* stands out as one that is very specifically expressed in the host urinary tract. It contributes significantly to the success of CFT073 in the murine model of ascending infection and in a murine model of bacteremia. This gene can serve as a marker of a specific subset of ExPEC isolates, B2 isolates in general and specifically ST complex 73 isolates. Regardless of the strain collection or host species, *tosA* presence predicts success of fecal isolates of *E. coli* in the murine model of ascending UTI, a result that few other UPEC virulence markers can match. The genetic loci containing *tosA* contains all three components of the T1SS employed by RTX family members and includes a newly discovered member of the PapB family of transcriptional regulators, suggesting co-regulation of TosA expression with fimbriae expression. TosA was found to mediate adherence specifically to the epithelial cells that line the upper urinary tract, but not the lower urinary tract. A possible secondary role of TosA mediating, or at least enhancing, UPEC toxicity during bactermia and sepsis was discovered. Finally, experimental vaccines targeting TosA protected against urosepsis during an ascending UTI.

In addition, several issues of broader scope will be considered with particular emphasis placed on the questions raised by these studies that remain unanswered.

**The role of TosA in UTI and urosepsis**

While a clear role for TosA has been established in mediating adhesion to the epithelial lining of the ureters and especially the kidney, this raises several questions. Multiple adherence mechanisms can be found in typical UPEC strains, with more than two dozen characterized and putative fimbrial and non-fimbrial adherence systems contained in the CFT073 genome (126). The advantage gained by having so many seemingly redundant systems remains a mystery, although one possible explanation will be offered later in this section.

In addition, the role of TosA adherence during disseminated infections is difficult to understand. Given the results of Chapter 4, it would appear that TosA-mediated adherence is specific to particular cell types. But unlike the species specificity defined for the cytotoxic RTX members (115), TosA-mediated adherence was observed in tissue culture lines derived from humans, non-human primates, and rodents. Additional cell types covering a wide range of tissues will have to be assessed in the adherence assay in order to determine whether the adherence phenotype covers an even broader range of host environments. And what the exact contribution of adherence in any form to bacteremia and sepsis, seen not just in TosA but also in type 1 and P fimbriae (109), confers to pathogenesis outside of the urinary tract remains an open question.

**Molecular mechanisms of adhesion and toxicity**

Work conducted on multiple RTX family members, including the well known cytotoxic members, suggests that RTX members bind to specific host cell receptors (76, 125). Given the fact that TosA-expressing bacteria were unable to bind to many of the cell types tested, it is unlikely that the protein exerts its mechanism of action by inserting into the lipid bilayer of the host membrane without recognition of a specific host factor. The identification of what that factor is and what the impact this has on the host cell will likely be the subject of future studies.

On a related note, the mechanism of enhanced fitness in a murine model of bacteremia and enhanced lethality in the zebrafish model is completely unknown at this point. While the MARTX family has been shown to encode multiple functions in a single protein (100), no direct role for TosA has been established in mediating toxicity to the epithelial cells lining the upper urinary tract. It is possible that TosA may exhibit different functions on different cell types, so the additional work proposed above to continue *in vitro* tests using a wider array of tissue culture systems may also prove useful in addressing this hypothesis. An alternative is that the role may be indirect, perhaps by allowing intimate adherence of TosA-expressing bacteria enhancingthe delivery of additional bacterial toxins. Studies using strains deficient in both TosA and well characterized UPEC toxins may help clarify whether TosA serves a direct toxic effect on host cells.

**The regulation network of UPEC adherence factors**

The expression pattern of TosA was striking, showing no detectable protein synthesis under several *in vitro* culture conditions, while *in vivo* expression was detected

in bladder, kidney, spleen, and liver tissue. The regulatory mechanisms that restrict TosA expression to the host environment could provide many insights into bacterial pathogenesis and could potentially shed light on the many putative RTX systems that remain uncharacterized. Many of these systems, especially those in other uropathogens, may be regulated by similar mechanisms.

An attractive hypothesis is that TosA expression may be tied into the complex regulatory network that UPEC utilize to coordinate expression of the chaperone-usher fimbrial operons. The best characterized member, PapB, binds to high affinity sites that enhance expression of the *pap* operons (encoding P fimbriae), but high protein levels allow binding of low affinity repressor sequences that limit expression (30). In addition, PapB can repress the expression of F1C fimbriae (75) and both PapB and the related family member FocB can repress the off-to-on conversion of the invertible element that controls type 1 fimbrial expression (55, 135). Discovery of the *papB* family member *tosR* suggests that expression of TosA may also result in repression of other adherence factors, perhaps limiting the expression of fimbriae while TosA is being expressed. This kind of coordinated expression may prove UPEC a mechanism to avoid the adaptive immune system response of the host. Fimbriae are readily recognized by the host immune response, and TosA was originally characterized as an antigenic protein recognized by the murine immune system. Switching between many different antigenic proteins with similar functions could potentially be of sufficient benefit to the bacterium to explain the presence of so many different adherence systems with overlapping function in the same genome. Work to characterize the regulatory mechanisms that

control TosA expression may benefit from building on what is already known about the PapB family of transcriptional regulators.

**Future development of UPEC vaccnines**

As outlined in Chapter 1, the development of RTX-based vaccines is of great interest in the field of bacterial pathogenesis. The preliminary vaccine results presented in Chapter 4 are encouraging and suggest that further development of a TosA based vaccine will be of great benefit. Even if TosA never makes it into a clinical vaccine, the experiments conducted during vaccine development are likely to illuminate many aspects of TosA and RTX biology. For example, the results of our vaccine trials suggest that TosA may be more important for the process that UPEC use to move from the lumen of the kidney tubules across the capillary endothelium into the bloodstream.

Finally, an interesting corollary was noted in the *A. pleuropneumoniae* RTX vaccine development. While isolates of this species are known to produce at least four RTX toxin members, one of these, ApxIV, is known to be only expressed during *in vivo* growth and would thus be absent from the whole cell vaccine currently in use (102). This RTX-based vaccine showed enhanced protection over a heat-killed whole cell vaccine, suggesting that other vaccines that target RTX proteins made specifically in the host environment may have great therapeutic potential. In addition, this observation suggests that while many properties of TosA may seem novel when compared to known RTX members, it may serve as a prototype for the many uncharacterized RTX members, of which at least 1000 are known, and an even larger number have yet to be discovered.

**REFERENCES**

1. **Aboderin, O. A., A. R. Abdu, B. W. Odetoyin, and A. Lamikanra.** 2009. Antimicrobial resistance in Escherichia coli strains from urinary tract infections. J Natl Med Assoc **101:**1268-73.
2. **Al-Hasan, M. N., J. E. Eckel-Passow, and L. M. Baddour.** 2010. Bacteremia complicating gram-negative urinary tract infections: a population-based study. J Infect **60:**278-85.
3. **Alteri, C. J., E. C. Hagan, K. E. Sivick, S. N. Smith, and H. L. Mobley.** 2009. Mucosal immunization with iron receptor antigens protects against urinary tract infection. PLoS Pathog **5:**e1000586.
4. **Alteri, C. J., and H. L. Mobley.** 2007. Quantitative profile of the uropathogenic *Escherichia coli* outer membrane proteome during growth in human urine. Infect Immun **75:**2679-88.
5. **Alteri, C. J., S. N. Smith, and H. L. Mobley.** 2009. Fitness of *Escherichia coli* during urinary tract infection requires gluconeogenesis and the TCA cycle. PLoS Pathog **5:**e1000448.
6. **Andersen, C., C. Hughes, and V. Koronakis.** 2001. Protein export and drug efflux through bacterial channel-tunnels. Curr Opin Cell Biol **13:**412-6.
7. **Bacheller, C. D., and J. M. Bernstein.** 1997. Urinary tract infections. Med Clin North Am **81:**719-30.
8. **Baga, M., M. Goransson, S. Normark, and B. E. Uhlin.** 1985. Transcriptional activation of a *pap* pilus virulence operon from uropathogenic *Escherichia coli*. EMBO J **4:**3887-93.
9. **Bahrani-Mougeot, F. K., E. L. Buckles, C. V. Lockatell, J. R. Hebel, D. E. Johnson, C. M. Tang, and M. S. Donnenberg.** 2002. Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic Escherichia coli virulence determinants in the murine urinary tract. Mol Microbiol **45:**1079-93.
10. **Baumann, U., S. Wu, K. M. Flaherty, and D. B. McKay.** 1993. Three-dimensional structure of the alkaline protease of Pseudomonas aeruginosa: a two-domain protein with a calcium binding parallel beta roll motif. EMBO J **12:**3357-64.
11. **Boardman, B. K., and K. J. Satchell.** 2004. Vibrio cholerae strains with mutations in an atypical type I secretion system accumulate RTX toxin intracellularly. J Bacteriol **186:**8137-43.
12. **Bours, P. H., R. Polak, A. I. Hoepelman, E. Delgado, A. Jarquin, and A. J. Matute.** 2010. Increasing resistance in community-acquired urinary tract infections in Latin America, five years after the implementation of national therapeutic guidelines. Int J Infect Dis **14:**e770-4.
13. **Brahamsha, B.** 1996. An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium Synechococcus. Proc Natl Acad Sci U S A **93:**6504-9.
14. **Brooks, H. J., F. O'Grady, M. A. McSherry, and W. R. Cattell.** 1980. Uropathogenic properties of Escherichia coli in recurrent urinary-tract infection. J Med Microbiol **13:**57-68.
15. **Brzuszkiewicz, E., H. Bruggemann, H. Liesegang, M. Emmerth, T. Olschlager, G. Nagy, K. Albermann, C. Wagner, C. Buchrieser, L. Emody, G. Gottschalk, J. Hacker, and U. Dobrindt.** 2006. How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic Escherichia coli strains. Proc Natl Acad Sci U S A **103:**12879-84.
16. **Chakupurakal, R., M. Ahmed, D. N. Sobithadevi, S. Chinnappan, and T. Reynolds.** 2010. Urinary tract pathogens and resistance pattern. J Clin Pathol **63:**652-4.

17. **Chambers, S. T., and C. M. Kunin.** 1987. Isolation of glycine betaine and proline betaine from human urine. Assessment of their role as osmoprotective agents for bacteria and the kidney. J Clin Invest **79:**731-7.
18. **Chin, B. S., M. S. Kim, S. H. Han, S. Y. Shin, H. K. Choi, Y. T. Chae, S. J. Jin, J. H. Baek, J. Y. Choi, Y. G. Song, C. O. Kim, and J. M. Kim.** 2011. Risk factors of all-cause in-hospital mortality among Korean elderly bacteremic urinary tract infection (UTI) patients. Arch Gerontol Geriatr **52:**e50-5.
19. **Clermont, O., S. Bonacorsi, and E. Bingen.** 2000. Rapid and simple determination of the Escherichia coli phylogenetic group. Appl Environ Microbiol **66:**4555-8.
20. **Connell, I., W. Agace, P. Klemm, M. Schembri, S. Marild, and C. Svanborg.** 1996. Type 1 fimbrial expression enhances Escherichia coli virulence for the urinary tract. Proc Natl Acad Sci U S A **93:**9827-32.
21. **Culham, D. E., C. Dalgado, C. L. Gyles, D. Mamelak, S. MacLellan, and J. M. Wood.** 1998. Osmoregulatory transporter ProP influences colonization of the urinary tract by Escherichia coli. Microbiology **144 ( Pt 1):**91-102.
22. **Culham, D. E., A. Lu, M. Jishage, K. A. Krogfelt, A. Ishihama, and J. M. Wood.** 2001. The osmotic stress response and virulence in pyelonephritis isolates of Escherichia coli: contributions of RpoS, ProP, ProU and other systems. Microbiology **147:**1657-70.
23. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-5.
24. **Deb, D. K., P. Dahiya, K. K. Srivastava, R. Srivastava, and B. S. Srivastava.** 2002. Selective identification of new therapeutic targets of Mycobacterium tuberculosis by IVIAT approach. Tuberculosis (Edinb) **82:**175-82.
25. **Dielubanza, E. J., and A. J. Schaeffer.** 2011. Urinary tract infections in women. Med Clin North Am **95:**27-41.
26. **Dobrindt, U., L. Emody, I. Gentschev, W. Goebel, and J. Hacker.** 2002. Efficient expression of the alpha-haemolysin determinant in the uropathogenic Escherichia coli strain 536 requires the leuX-encoded tRNA(5)(Leu). Mol Genet Genomics **267:**370-9.
27. **Duong, F., C. Soscia, A. Lazdunski, and M. Murgier.** 1994. The Pseudomonas fluorescens lipase has a C-terminal secretion signal and is secreted by a three-component bacterial ABC-exporter system. Mol Microbiol **11:**1117-26.
28. **Economou, A., W. D. Hamilton, A. W. Johnston, and J. A. Downie.** 1990. The Rhizobium nodulation gene nodO encodes a Ca2(+)-binding protein that is exported without N-terminal cleavage and is homologous to haemolysin and related proteins. EMBO J **9:**349-54.
29. **Felmlee, T., S. Pellett, and R. Welch.** 1985. Nucleotide sequence of an Escherichia coli chromosomal hemolysin. J Bacteriol **163:**94-105.
30. **Forsman, K., M. Goransson, and B. E. Uhlin.** 1989. Autoregulation and multiple DNA interactions by a transcriptional regulatory protein in E. coli pili biogenesis. EMBO J **8:**1271-7.
31. **Foxman, B.** 2002. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am J Med **113 Suppl 1A:**5S-13S.
32. **Foxman, B., R. Barlow, H. D'Arcy, B. Gillespie, and J. D. Sobel.** 2000. Urinary tract infection: self-reported incidence and associated costs. Ann Epidemiol **10:**509-15.
33. **Foxman, B., S. D. Manning, P. Tallman, R. Bauer, L. Zhang, J. S. Koopman, B. Gillespie, J. D. Sobel, and C. F. Marrs.** 2002. Uropathogenic Escherichia coli are more likely than commensal E. coli to be shared between heterosexual sex partners. Am J Epidemiol **156:**1133-40.

34. **Friedman, N., M. Linial, I. Nachman, and D. Pe'er.** 2000. Using Bayesian networks to analyze expression data. J Comput Biol **7:**601-20.
35. **Fullner, K. J., and J. J. Mekalanos.** 2000. In vivo covalent cross-linking of cellular actin by the Vibrio cholerae RTX toxin. EMBO J **19:**5315-23.
36. **Gadeberg, O. V., and I. Orskov.** 1984. In vitro cytotoxic effect of alpha-hemolytic Escherichia coli on human blood granulocytes. Infect Immun **45:**255-60.
37. **Gentschev, I., J. Hess, and W. Goebel.** 1990. Change in the cellular localization of alkaline phosphatase by alteration of its carboxy-terminal sequence. Mol Gen Genet **222:**211-6.
38. **Glaser, P., H. Sakamoto, J. Bellalou, A. Ullmann, and A. Danchin.** 1988. Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of Bordetella pertussis. EMBO J **7:**3997-4004.
39. **Goebel, W., and J. Hedgpeth.** 1982. Cloning and functional characterization of the plasmid-encoded hemolysin determinant of Escherichia coli. J Bacteriol **151:**1290-8.
40. **Gransden, W. R., S. J. Eykyn, I. Phillips, and B. Rowe.** 1990. Bacteremia due to Escherichia coli: a study of 861 episodes. Rev Infect Dis **12:**1008-18.
41. **Griebling, T. L.** 2005. Urologic diseases in America project: trends in resource use for urinary tract infections in women. J Urol **173:**1281-7.
42. **Gunther, N. W. t., J. A. Snyder, V. Lockatell, I. Blomfield, D. E. Johnson, and H. L. Mobley.** 2002. Assessment of virulence of uropathogenic Escherichia coli type 1 fimbrial mutants in which the invertible element is phase-locked on or off. Infect Immun **70:**3344-54.
43. **Haardt, M., B. Kempf, E. Faatz, and E. Bremer.** 1995. The osmoprotectant proline betaine is a major substrate for the binding-protein-dependent transport system ProU of Escherichia coli K-12. Mol Gen Genet **246:**783-6.
44. **Hacker, J., C. Hughes, H. Hof, and W. Goebel.** 1983. Cloned hemolysin genes from Escherichia coli that cause urinary tract infection determine different levels of toxicity in mice. Infect Immun **42:**57-63.
45. **Haga, Y., S. Ogino, S. Ohashi, T. Ajito, K. Hashimoto, and T. Sawada.** 1997. Protective efficacy of an affinity-purified hemolysin vaccine against experimental swine pleuropneumonia. J Vet Med Sci **59:**115-20.
46. **Hagan, E. C., A. L. Lloyd, D. A. Rasko, G. J. Faerber, and H. L. Mobley.** 2010. Escherichia coli global gene expression in urine from women with urinary tract infection. PLoS Pathog **6:**e1001187.
47. **Hagan, E. C., and H. L. Mobley.** 2007. Uropathogenic Escherichia coli outer membrane antigens expressed during urinary tract infection. Infect Immun **75:**3941-9.
48. **Hagberg, L., I. Engberg, R. Freter, J. Lam, S. Olling, and C. Svanborg Eden.** 1983. Ascending, unobstructed urinary tract infection in mice caused by pyelonephritogenic Escherichia coli of human origin. Infect Immun **40:**273-83.
49. **Hang, L., M. John, M. Asaduzzaman, E. A. Bridges, C. Vanderspurt, T. J. Kirn, R. K. Taylor, J. D. Hillman, A. Progulske-Fox, M. Handfield, E. T. Ryan, and S. B. Calderwood.** 2003. Use of in vivo-induced antigen technology (IVIAT) to identify genes uniquely expressed during human infection with Vibrio cholerae. Proc Natl Acad Sci U S A **100:**8508-13.
50. **Haugen, B. J., S. Pellett, P. Redford, H. L. Hamilton, P. L. Roesch, and R. A. Welch.** 2007. In vivo gene expression analysis identifies genes required for enhanced colonization of the mouse urinary tract by uropathogenic Escherichia coli strain CFT073 dsdA. Infect Immun **75:**278-89.

51. **Hawn, T. R., D. Scholes, H. Wang, S. S. Li, A. E. Stapleton, M. Janer, A. Aderem, W. E. Stamm, L. P. Zhao, and T. M. Hooton.** 2009. Genetic variation of the human urinary tract innate immune response and asymptomatic bacteriuria in women. PLoS One **4:**e8300.
52. **Heimer, S. R., D. A. Rasko, C. V. Lockatell, D. E. Johnson, and H. L. Mobley.** 2004. Autotransporter genes pic and tsh are associated with Escherichia coli strains that cause acute pyelonephritis and are expressed during urinary tract infection. Infect Immun **72:**593-7.
53. **Hodges, A. P., D. Dai, Z. Xiang, P. Woolf, C. Xi, and Y. He.** 2010. Bayesian network expansion identifies new ROS and biofilm regulators. PLoS One **5:**e9513.
54. **Hodges, A. P., P. Woolf, and Y. He.** 2010. BN+1 Bayesian network expansion for identifying molecular pathway elements. Commun Integr Biol **3:**549-54.
55. **Holden, N., B. Uhlin, and D. Gally.** 2001. PapB paralogues and their effect on the phase variation of type 1 fimbriae in Escherichia coli. Mol Microbiol **42:**319-30.
56. **Holden, N. J., B. E. Uhlin, and D. L. Gally.** 2001. PapB paralogues and their effect on the phase variation of type 1 fimbriae in Escherichia coli. Mol Microbiol **42:**319-30.
57. **Hooton, T. M., D. Scholes, A. E. Stapleton, P. L. Roberts, C. Winter, K. Gupta, M. Samadpour, and W. E. Stamm.** 2000. A prospective study of asymptomatic bacteriuria in sexually active young women. N Engl J Med **343:**992-7.
58. **Hultdin, U. W., S. Lindberg, C. Grundstrom, S. Huang, B. E. Uhlin, and A. E. Sauer-Eriksson.** 2010. Structure of FocB--a member of a family of transcription factors regulating fimbrial adhesin expression in uropathogenic Escherichia coli. FEBS J **277:**3368-81.
59. **Ikaheimo, R., A. Siitonen, U. Karkkainen, P. Kuosmanen, and P. H. Makela.** 1993. Characteristics of Escherichia coli in acute community-acquired cystitis of adult women. Scand J Infect Dis **25:**705-12.
60. **Issartel, J. P., V. Koronakis, and C. Hughes.** 1991. Activation of Escherichia coli prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. Nature **351:**759-61.
61. **John, M., I. T. Kudva, R. W. Griffin, A. W. Dodson, B. McManus, B. Krastins, D. Sarracino, A. Progulske-Fox, J. D. Hillman, M. Handfield, P. I. Tarr, and S. B. Calderwood.** 2005. Use of in vivo-induced antigen technology for identification of Escherichia coli O157:H7 proteins expressed during human infection. Infect Immun **73:**2665-79.
62. **Johnson, J. R., C. Clabots, and M. A. Kuskowski.** 2008. Multiple-host sharing, long-term persistence, and virulence of Escherichia coli clones from human and animal household members. J Clin Microbiol **46:**4078-82.
63. **Johnson, J. R., O. Clermont, M. Menard, M. A. Kuskowski, B. Picard, and E. Denamur.** 2006. Experimental mouse lethality of Escherichia coli isolates, in relation to accessory traits, phylogenetic group, and ecological source. J Infect Dis **194:**1141-50.
64. **Johnson, J. R., N. Kaster, M. A. Kuskowski, and G. V. Ling.** 2003. Identification of urovirulence traits in Escherichia coli by comparison of urinary and rectal E. coli isolates from dogs with urinary tract infection. J Clin Microbiol **41:**337-45.
65. **Johnson, J. R., and T. A. Russo.** 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) Escherichia coli. Int J Med Microbiol **295:**383-404.
66. **Jorgensen, S. E., P. F. Mulcahy, G. K. Wu, and C. F. Louis.** 1983. Calcium accumulation in human and sheep erythrocytes that is induced by Escherichia coli hemolysin. Toxicon **21:**717-27.

67. **Juarez, A., J. M. Nieto, A. Prenafeta, E. Miquelay, C. Balsalobre, M. Carrascal, and C. Madrid.** 2000. Interaction of the nucleoid-associated proteins Hha and H-NS to modulate expression of the hemolysin operon in Escherichia coli. Adv Exp Med Biol **485:**127-31.
68. **Keane, W. F., R. Welch, G. Gekker, and P. K. Peterson.** 1987. Mechanism of Escherichia coli alpha-hemolysin-induced injury to isolated renal tubular cells. Am J Pathol **126:**350-7.
69. **Klemm, P., V. Hancock, and M. Schembri.** 2007. Mellowing out: adaptation to commensalism by Escherichia coli asymptomatic bacteriuria strain 83972. Infect Immun **75:**3688-95.
70. **Klemm, P., V. Roos, G. C. Ulett, C. Svanborg, and M. A. Schembri.** 2006. Molecular characterization of the Escherichia coli asymptomatic bacteriuria strain 83972: the taming of a pathogen. Infect Immun **74:**781-5.
71. **Kumar, M., F. G. Khan, S. Sharma, R. Kumar, J. Faujdar, R. Sharma, D. S. Chauhan, R. Singh, S. K. Magotra, and I. A. Khan.** 2011. Identification of Mycobacterium tuberculosis genes preferentially expressed during human infection. Microb Pathog **50:**31-8.
72. **Leblond-Francillard, M., J. L. Gaillard, and P. Berche.** 1989. Loss of catalase activity in Tn1545-induced mutants does not reduce growth of Listeria monocytogenes in vivo. Infect Immun **57:**2569-73.
73. **Leeds, J. A., and R. A. Welch.** 1996. RfaH enhances elongation of Escherichia coli hlyCABD mRNA. J Bacteriol **178:**1850-7.
74. **Lim, J. K., N. W. t. Gunther, H. Zhao, D. E. Johnson, S. K. Keay, and H. L. Mobley.** 1998. In vivo phase variation of Escherichia coli type 1 fimbrial genes in women with urinary tract infection. Infect Immun **66:**3303-10.
75. **Lindberg, S., Y. Xia, B. Sonden, M. Goransson, J. Hacker, and B. E. Uhlin.** 2008. Regulatory Interactions among adhesin gene systems of uropathogenic Escherichia coli. Infect Immun **76:**771-80.
76. **Linhartova, I., L. Bumba, J. Masin, M. Basler, R. Osicka, J. Kamanova, K. Prochazkova, I. Adkins, J. Hejnova-Holubova, L. Sadilkova, J. Morova, and P. Sebo.** 2010. RTX proteins: a highly diverse family secreted by a common mechanism. FEMS Microbiol Rev **34:**1076-112.
77. **Lloyd, A. L., T. A. Henderson, P. D. Vigil, and H. L. Mobley.** 2009. Genomic islands of uropathogenic Escherichia coli contribute to virulence. J Bacteriol **191:**3469-81.
78. **Lloyd, A. L., D. A. Rasko, and H. L. Mobley.** 2007. Defining genomic islands and uropathogen-specific genes in uropathogenic Escherichia coli. J Bacteriol **189:**3532-46.
79. **Lo, R. Y., C. A. Strathdee, and P. E. Shewen.** 1987. Nucleotide sequence of the leukotoxin genes of Pasteurella haemolytica A1. Infect Immun **55:**1987-96.
80. **Lomberg, H., M. Hellstrom, U. Jodal, and C. Svanborg Eden.** 1989. Secretor state and renal scarring in girls with recurrent pyelonephritis. FEMS Microbiol Immunol **1:**371-5.
81. **Loomes, L. M., M. A. Kerr, and B. W. Senior.** 1993. The cleavage of immunoglobulin G in vitro and in vivo by a proteinase secreted by the urinary tract pathogen Proteus mirabilis. J Med Microbiol **39:**225-32.
82. **Marrs, C. F., L. Zhang, and B. Foxman.** 2005. Escherichia coli mediated urinary tract infections: are there distinct uropathogenic E. coli (UPEC) pathotypes? FEMS Microbiol Lett **252:**183-90.
83. **Mobley, H., D. Green, A. Trifillis, D. Johnson, G. Chippendale, C. Lockatell, B. Jones, and J. Warren.** 1990. Pyelonephritogenic Escherichia coli and killing of cultured human

renal proximal tubular epithelial cells: role of hemolysin in some strains. Infect Immun **58:**1281-9.

84. **Mobley, H. L., D. M. Green, A. L. Trifillis, D. E. Johnson, G. R. Chippendale, C. V. Lockatell, B. D. Jones, and J. W. Warren.** 1990. Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. Infect Immun **58:**1281-9.
85. **Mobley, H. L. T., M.S. Donnenberg, and E.C. Hagan.** 2009. Uropathogenic Escherichia coli. *In* R. C. I. A. Böck, J. B. Kaper, P. D. Karp, F. C. Neidhardt, T. Nyström, J. M. Slauch, C. L. Squires, and D. Ussery (ed.), EcoSal—Escherichia coli and Salmonella: Cellular and Molecular Biology. ASM Press, Washington, D.C.
86. **Moreno, E., A. Andreu, C. Pigrau, M. Kuskowski, J. Johnson, and G. Prats.** 2008. Relationship between Escherichia coli strains causing acute cystitis in women and the fecal E. coli population of the host. J Clin Microbiol **46:**2529-34.
87. **Murray, A. C., M. A. Kuskowski, and J. R. Johnson.** 2004. Virulence factors predict Escherichia coli colonization patterns among human and animal household members. Ann Intern Med **140:**848-9.
88. **Nagy, G., A. Altenhoefer, O. Knapp, E. Maier, U. Dobrindt, G. Blum-Oehler, R. Benz, L. Emody, and J. Hacker.** 2006. Both alpha-haemolysin determinants contribute to full virulence of uropathogenic Escherichia coli strain 536. Microbes Infect **8:**2006-12.
89. **Nielubowicz, G. R., and H. L. Mobley.** 2010. Host-pathogen interactions in urinary tract infection. Nat Rev Urol **7:**430-41.
90. **O'Hanley, P., G. Lalonde, and G. Ji.** 1991. Alpha-hemolysin contributes to the pathogenicity of piliated digalactoside-binding Escherichia coli in the kidney: efficacy of an alpha-hemolysin vaccine in preventing renal injury in the BALB/c mouse model of pyelonephritis. Infect Immun **59:**1153-61.
91. **Ochman, H., and R. Selander.** 1984. Standard reference strains of Escherichia coli from natural populations. J Bacteriol **157:**690-3.
92. **Okuda, K., T. Kigure, S. Yamada, T. Kaneko, K. Ishihara, T. Miura, T. Kato, and I. Takazoe.** 1997. Role for the S-layer of Campylobacter rectus ATCC33238 in complement mediated killing and phagocytic killing by leukocytes from guinea pig and human peripheral blood. Oral Dis **3:**113-20.
93. **Parham, N. J., S. J. Pollard, R. R. Chaudhuri, S. A. Beatson, M. Desvaux, M. A. Russell, J. Ruiz, A. Fivian, J. Vila, and I. R. Henderson.** 2005. Prevalence of pathogenicity island IICFT073 genes among extraintestinal clinical isolates of Escherichia coli. J Clin Microbiol **43:**2425-34.
94. **Pearson, M. M., M. Sebaihia, C. Churcher, M. A. Quail, A. S. Seshasayee, N. M. Luscombe, Z. Abdellah, C. Arrosmith, B. Atkin, T. Chillingworth, H. Hauser, K. Jagels, S. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, D. Walker, S. Whithead, N. R. Thomson, P. N. Rather, J. Parkhill, and H. L. Mobley.** 2008. Complete genome sequence of uropathogenic Proteus mirabilis, a master of both adherence and motility. J Bacteriol **190:**4027-37.
95. **Rollins, S. M., A. Peppercorn, L. Hang, J. D. Hillman, S. B. Calderwood, M. Handfield, and E. T. Ryan.** 2005. In vivo induced antigen technology (IVIAT). Cell Microbiol **7:**1-9.
96. **Roos, V., M. Schembri, G. Ulett, and P. Klemm.** 2006. Asymptomatic bacteriuria Escherichia coli strain 83972 carries mutations in the foc locus and is unable to express F1C fimbriae. Microbiology **152:**1799-806.

97. **Roos, V., M. A. Schembri, G. C. Ulett, and P. Klemm.** 2006. Asymptomatic bacteriuria Escherichia coli strain 83972 carries mutations in the foc locus and is unable to express F1C fimbriae. Microbiology **152:**1799-806.
98. **Russo, T. A., and J. R. Johnson.** 2000. Proposal for a new inclusive designation for extraintestinal pathogenic isolates of Escherichia coli: ExPEC. J Infect Dis **181:**1753-4.
99. **Salim, K. Y., D. G. Cvitkovitch, P. Chang, D. J. Bast, M. Handfield, J. D. Hillman, and J. C. de Azavedo.** 2005. Identification of group A Streptococcus antigenic determinants upregulated in vivo. Infect Immun **73:**6026-38.
100. **Satchell, K.** 2007. MARTX, multifunctional autoprocessing repeats-in-toxin toxins. Infect Immun **75:**5079-84.
101. **Seah, J. N., J. Frey, and J. Kwang.** 2002. The N-terminal domain of RTX toxin ApxI of Actinobacillus pleuropneumoniae elicits protective immunity in mice. Infect Immun **70:**6464-7.
102. **Shao, M., Y. Wang, C. Wang, Y. Guo, Y. Peng, J. Liu, G. Li, H. Liu, and S. Liu.** 2010. Evaluation of multicomponent recombinant vaccines against Actinobacillus pleuropneumoniae in mice. Acta Vet Scand **52:**52.
103. **Sheahan, K. L., C. L. Cordero, and K. J. Satchell.** 2007. Autoprocessing of the Vibrio cholerae RTX toxin by the cysteine protease domain. EMBO J **26:**2552-61.
104. **Shibuya, Y., T. Yamamoto, T. Morimoto, N. Nishino, T. Kambara, and H. Okabe.** 1991. Pseudomonas aeruginosa alkaline proteinase might share a biological function with plasmin. Biochim Biophys Acta **1077:**316-24.
105. **Sivick, K. E., and H. L. Mobley.** 2009. An "omics" approach to uropathogenic Escherichia coli vaccinology. Trends Microbiol **17:**431-2.
106. **Sivick, K. E., and H. L. Mobley.** 2010. Waging war against uropathogenic Escherichia coli: winning back the urinary tract. Infect Immun **78:**568-85.
107. **Sivick, K. E., M. A. Schaller, S. N. Smith, and H. L. Mobley.** 2010. The innate immune response to uropathogenic Escherichia coli involves IL-17A in a murine model of urinary tract infection. J Immunol **184:**2065-75.
108. **Sleytr, U. B., and T. J. Beveridge.** 1999. Bacterial S-layers. Trends Microbiol **7:**253-60.
109. **Smith, S. N., E. C. Hagan, M. C. Lane, and H. L. Mobley.** 2010. Dissemination and Systemic Colonization of Uropathogenic Escherichia coli in a Murine Model of Bacteremia. MBio **1**.
110. **Smith, Y. C., S. B. Rasmussen, K. K. Grande, R. M. Conran, and A. D. O'Brien.** 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. Infect Immun **76:**2978-90.
111. **Soler-Garcia, A. A., and A. E. Jerse.** 2007. Neisseria gonorrhoeae catalase is not required for experimental genital tract infection despite the induction of a localized neutrophil response. Infect Immun **75:**2225-33.
112. **Stanley, P., V. Koronakis, and C. Hughes.** 1991. Mutational analysis supports a role for multiple structural features in the C-terminal secretion signal of Escherichia coli haemolysin. Mol Microbiol **5:**2391-403.
113. **Stapleton, A., S. Moseley, and W. E. Stamm.** 1991. Urovirulence determinants in Escherichia coli isolates causing first-episode and recurrent cystitis in women. J Infect Dis **163:**773-9.
114. **Strathdee, C. A., and R. Y. Lo.** 1987. Extensive homology between the leukotoxin of Pasteurella haemolytica A1 and the alpha-hemolysin of Escherichia coli. Infect Immun **55:**3233-6.

115. **Taichman, N. S., D. L. Simpson, S. Sakurada, M. Cranfield, J. DiRienzo, and J. Slots.** 1987. Comparative studies on the biology of Actinobacillus actinomycetemcomitans leukotoxin in primates. Oral Microbiol Immunol **2:**97-104.
116. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol **24:**1596-9.
117. **Trifillis, A. L., M. S. Donnenberg, X. Cui, R. G. Russell, S. J. Utsalo, H. L. Mobley, and J. W. Warren.** 1994. Binding to and killing of human renal epithelial cells by hemolytic P-fimbriated *E. coli*. Kidney Int **46:**1083-91.
118. **van den Bosch, H., and J. Frey.** 2003. Interference of outer membrane protein PalA with protective immunity against Actinobacillus pleuropneumoniae infections in vaccinated pigs. Vaccine **21:**3601-7.
119. **Vigil, P. D., C. J. Alteri, and H. L. Mobley.** 2010. Identification of in vivo induced antigens including a RTX family
exoprotein required for uropathogenic Escherichia coli virulence. Infect Immun **In press**.
120. **Walk, S. T., E. W. Alm, D. M. Gordon, J. L. Ram, G. A. Toranzos, J. M. Tiedje, and T. S. Whittam.** 2009. Cryptic lineages of the genus Escherichia. Appl Environ Microbiol **75:**6534-44.
121. **Warren, J. (ed.).** 1996. Urinary Tract Infections: Molecular Pathogenesis and Clinical Management. ASM Press, Washington D.C.
122. **Warren, J., E. Miller, B. Fitzpatrick, D. DiFranco, E. Caplan, J. Tenney, and W. Anthony.** 1983. A randomized, controlled trial of cefoperazone vs. cefamandole-tobramycin in the treatment of putative, severe infections with gram-negative bacilli. Rev Infect Dis **5 Suppl 1:**S173-80.
123. **Warren, J. W., J. H. Tenney, J. M. Hoopes, H. L. Muncie, and W. C. Anthony.** 1982. A prospective microbiologic study of bacteriuria in patients with chronic indwelling urethral catheters. J Infect Dis **146:**719-23.
124. **Wassif, C., D. Cheek, and R. Belas.** 1995. Molecular analysis of a metalloprotease from Proteus mirabilis. J Bacteriol **177:**5790-8.
125. **Welch, R. A.** 1991. Pore-forming cytolysins of gram-negative bacteria. Mol Microbiol **5:**521-8.
126. **Welch, R. A., V. Burland, G. Plunkett, 3rd, P. Redford, P. Roesch, D. Rasko, E. L. Buckles, S. R. Liou, A. Boutin, J. Hackett, D. Stroud, G. F. Mayhew, D. J. Rose, S. Zhou, D. C. Schwartz, N. T. Perna, H. L. Mobley, M. S. Donnenberg, and F. R. Blattner.** 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. Proc Natl Acad Sci U S A **99:**17020-4.
127. **Welch, R. A., E. P. Dellinger, B. Minshew, and S. Falkow.** 1981. Haemolysin contributes to virulence of extra-intestinal *E. coli* infections. Nature **294:**665-7.
128. **Welch, R. A., and S. Falkow.** 1984. Characterization of Escherichia coli hemolysins conferring quantitative differences in virulence. Infect Immun **43:**156-60.
129. **Welch, R. A., R. Hull, and S. Falkow.** 1983. Molecular cloning and physical characterization of a chromosomal hemolysin from Escherichia coli. Infect Immun **42:**178-86.
130. **Wiles, T. J., J. M. Bower, M. J. Redd, and M. A. Mulvey.** 2009. Use of zebrafish to probe the divergent virulence potentials and toxin requirements of extraintestinal pathogenic Escherichia coli. PLoS Pathog **5:**e1000697.
131. **Wirth, T., D. Falush, R. Lan, F. Colles, P. Mensa, L. H. Wieler, H. Karch, P. R. Reeves, M. C. Maiden, H. Ochman, and M. Achtman.** 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol Microbiol **60:**1136-51.

132. **Wood, J. M.** 1999. Osmosensing by bacteria: signals and membrane-based sensors. Microbiol Mol Biol Rev **63:**230-62.
133. **Woods, R. G., M. Burger, C. A. Beven, and I. R. Beacham.** 2001. The aprX-lipA operon of Pseudomonas fluorescens B52: a molecular analysis of metalloprotease and lipase production. Microbiology **147:**345-54.
134. **Xia, Y., K. Forsman, J. Jass, and B. E. Uhlin.** 1998. Oligomeric interaction of the PapB transcriptional regulator with the upstream activating region of pili adhesin gene promoters in Escherichia coli. Mol Microbiol **30:**513-23.
135. **Xia, Y., D. Gally, K. Forsman-Semb, and B. E. Uhlin.** 2000. Regulatory cross-talk between adhesin operons in Escherichia coli: inhibition of type 1 fimbriae expression by the PapB protein. EMBO J **19:**1450-7.
136. **Yamamoto, S.** 2007. Molecular epidemiology of uropathogenic Escherichia coli. J Infect Chemother **13:**68-73.
137. **Yamamoto, S., T. Tsukamoto, A. Terai, H. Kurazono, Y. Takeda, and O. Yoshida.** 1997. Genetic evidence supporting the fecal-perineal-urethral hypothesis in cystitis caused by Escherichia coli. J Urol **157:**1127-9.
138. **Young, J., and I. B. Holland.** 1999. ABC transporters: bacterial exporters-revisited five years on. Biochim Biophys Acta **1461:**177-200.
139. **Zdziarski, J., C. Svanborg, B. Wullt, J. Hacker, and U. Dobrindt.** 2008. Molecular basis of commensalism in the urinary tract: low virulence or virulence attenuation? Infect Immun **76:**695-703.