UpaH Is a Newly Identified Autotransporter Protein That Contributes to Biofilm Formation and Bladder Colonization by Uropathogenic *Escherichia coli* CFT073[∇]

Luke P. Allsopp, Makrina Totsika, Jai J. Tree,# Glen C. Ulett,† Amanda N. Mabbett, Timothy J. Wells, Bostjan Kobe, Scott A. Beatson, and Mark A. Schembri*

School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane QLD 4072, Australia

Escherichia coli is the primary cause of urinary tract infection (UTI) in the developed world. The major factors associated with virulence of uropathogenic *E. coli* (UPEC) are fimbrial adhesins, which mediate specific attachment to host receptors and trigger innate host responses. Another group of adhesins is represented by the autotransporter (AT) subgroup of proteins. In this study, we identified a new AT-encoding gene, termed *upaH*, present in a 6.5-kb unannotated intergenic region in the genome of the prototypic UPEC strain CFT073. Cloning and sequencing of the *upaH* gene from CFT073 revealed an intact 8.535-kb coding region, contrary to the published genome sequence. The *upaH* gene was widely distributed among a large collection of UPEC isolates as well as the *E. coli* Reference (ECOR) strain collection. Bioinformatic analyses suggest β-helix as the predominant structure in the large N-terminal passenger (α) domain and a 12-strand β-barrel for the C-terminal β-domain of UpaH. We demonstrated that UpaH is expressed at the cell surface of CFT073 and promotes biofilm formation. In the mouse UTI model, deletion of the *upaH* gene in CFT073 and in two other UPEC strains did not significantly affect colonization of the bladder in single-challenge experiments. However, in competitive colonization experiments, CFT073 significantly outcompeted its *upaH* isogenic mutant strain in urine and the bladder.

Urinary tract infections (UTIs) are among the most common infectious diseases of humans. It is estimated that 40 to 50% of adult healthy women have experienced at least one UTI episode in their lifetime, and there is a tendency for these infections to become chronic due to a high rate of recurrence. Almost all patients with an indwelling urinary catheter for 30 days or longer develop catheter-associated UTI, which accounts for 40% of all nosocomial infections (13).

Uropathogenic *Escherichia coli* (UPEC) is the cause of the majority (>80%) of UTIs in humans and one of the most common sources of Gram-negative bacteremia in hospitalized patients. The ability of UPEC to colonize the urinary tract and cause disease involves adhesins (e.g., type 1 and P fimbriae), toxins (e.g., hemolysin), and iron acquisition systems that utilize siderophores (e.g., enterobactin, salmochelin, and aerobactin) (31, 60). Adherence to the urinary tract epithelium enables bacteria to resist the hydrodynamic forces of urine flow, to trigger host and bacterial cell signaling pathways, and to establish infection. Among adhesins, P and type 1 fimbriae correlate strongly with uropathogenesis and mediate binding to specific receptors in the urinary tract (7, 39, 47, 54, 62, 63). Both P and type 1 fimbriae recognize their receptor targets by virtue of organelle tip-located adhesins, PapG and FimH, respectively (27, 33).

In addition to fimbrial adhesins, a number of autotransporter (AT) proteins associated with virulence have been characterized for UPEC. These include the secreted toxin Sat (17, 18), the phase-variable outer membrane protein antigen 43 (Ag43) (57), and the surface-located trimeric AT protein UpaG (59). AT proteins are unique in that their sequence is sufficient to direct their transport across the bacterial membrane system and final routing of a variable passenger (α) domain to the cell surface. AT proteins are generally highly similar with respect to the structure of the translocation (β) domain that assists the transport of the α-domain across the outer membrane, whereas they differ substantially in their α-domain, which determines the unique functional characteristics of AT proteins (23, 24). Once at the bacterial surface, the α-domain may be processed and released into the extracellular surroundings (e.g., Pet and EspP), or it may be cleaved but remain in contact with the cell surface via noncovalent interactions with the β-domain (e.g., Ag43 and AIDA). Thus, AT proteins have diverse functions, ranging from cell-associated adhesins to secreted toxins (24).

Ten putative AT-encoding genes have been identified in the sequenced genome of the prototype UPEC strain CFT073 (45). In this study, we identified and characterized a new AT-encoding gene from CFT073, which we have named *upaH*. The *upaH* gene was not identified in the original analysis of the CFT073 genome sequence due to a sequence misassembly within a highly repetitive region. We show that *upaH* encodes a large cell surface-located AT protein that contributes to biofilm formation and colonization of the mouse urinary tract in coinfection experiments.

* Corresponding author. Mailing address: School of Chemistry and Molecular Biosciences, Building 76, University of Queensland, Brisbane QLD 4072, Australia. Phone: 617 33653306. Fax: 617 33654699. E-mail: m.schembri@uq.edu.au.

Present address: Immunity and Infection Division, the Roslin Institute and R(D)SVS, Chancellor's Building, University of Edinburgh, Edinburgh EH16 4SB, United Kingdom.

† Present address: Centre for Medicine and Oral Health, Griffith University, Gold Coast Campus, Queensland 4222, Australia.

∇ Published ahead of print on 9 February 2010.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristic(s)	Reference
E. coli strains		
BL21	F^- *ompT hsdS*(r_B^- m_B^-) *dcm gal*	Stratagene
CFT073	Wild-type UPEC isolate	38
CFT073^{cam}	CFT073*lacZ::cam-gfp*; Cam^r	This study
CFT073*upaH*	CFT073*upaH*::*kan*; Kan^r	This study
CFT073PcL*upaH*	CFT073PcL*upaH*; constitutively expressed UpaH	This study
M161	Wild-type UPEC isolate	This study
M161*upaH*	M161*upaH*::*kan*	This study
M357	Wild-type UPEC isolate	This study
M357*upaH*	M357*upaH*::*kan*	This study
MS427	MG1655*flu*	46
OS56	MG1655*flu*; GFP$^+$; Amp^r	53
Plasmids		
pUpaH^{Truncated}	747-bp portion of *upaH* gene from CFT073 was amplified with primers UpaHTF and UpaHTR, digested with NcoI-HindIII, and ligated into NcoI-HindIII-digested pBAD/*Myc*-HisB; Amp^r	This study
pUpaH	*upaH* gene from CFT073 was amplified with primers 520 and 521, digested with XhoI-HindIII, and ligated into XhoI-HindIII-digested pBAD/*Myc*-HisA-kan; Amp^r Kan^r	This study
pSG10	Chloramphenicol resistance gene from pKD3 inserted into HindIII site of *gfp* expression plasmid pKEN2	This study

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The strains and plasmids used in this study are listed in Table 1. *E. coli* strains used to study the prevalence of the *upaH* gene have been described previously (35) or were part of the *E. coli* Reference (ECOR) collection (42). Cells were routinely grown at 37°C on solid or in liquid Luria broth (LB) medium (5) supplemented with the appropriate antibiotics: kanamycin (Kan) (50 μg/ml), ampicillin (Amp) (100 μg/ml), or chloramphenicol (Cam) (30 μg/ml). For growth under defined conditions, M9 minimal medium supplemented with 0.2% glucose (49) was used as indicated. The plasmid pUpaH contains the *upaH* gene from CFT073 and was created by PCR amplification using the primers 520 (5′-CGCGCTCGAGATAATAAGGAATT ATTATGCAAAGGAAAACTCTATTGTCG) and 521 (5′-CGCGCAAGCTT TTAGAATGTTTTCTTGAAGT), digestion with XhoI-HindIII, and insertion into XhoI-HindIII-digested pBAD/*Myc*-HisA. A final modification of the plasmid was made by inserting a kanamycin resistance-encoding gene into the HindIII site. Resistance to kanamycin was required to facilitate transformation of these plasmids into the *flu*-negative *gfp*-positive K-12 strain OS56. In pUpaH, expression of *upaH* is under the control of the arabinose-inducible *araBAD* promoter (19). Mutants containing deletions in the *upaH* gene were constructed by λ-red-mediated homologous recombination as previously described (9). Primers with 50-bp homology extensions (718, 5′-GCTTACCCTCTGAATGCTCCT CTGTAGTCGAACAGAACCTGATGCCATGGGTGTAGGCTGGAGCTG CTTC; and 719, 5′-CTGTCCACCGCCGTTTCGCTGCCGCTGTTCATGCTC AGGTTAAAGTCACCCATATGAATATCCTCCTTAG) were employed to amplify the kanamycin resistance cassette used to replace *upaH*. Mutants were confirmed via PCR and sequencing using the primers 716 (5′-GCAAAGGAA AACTCTATTGTCG) and 717 (5′-GGCTGTTGAAACCACCACC). The *upaH* promoter insertion mutant (CFT073PcL*upaH*) was constructed using a strategy described previously (8, 57) and the following primers: upstream region (587, 5′-AGAACATTACTTCTTTGTGC; and 586, 5′-GTGAGAATTACTAACTTG AGCGAATCAAATGCGTCACATAGATATTTAT), downstream region (584, 5′-CATTAAAGAGGAGAAAGGTACCGCATGCAAAGGAAAACTCTATT GTCGGCCT; and 585, 5′-TCGGTATAAGTAAAAACTGA), PcL promoter and kanamycin resistance cassette (588, 5′-CTCTGGCAAGCGCCTCGATTA CTG; and 589, 5′-CATCACCTTCACCCTCTCCACTGAC); verification (592, 5′-GCTTGAGACCCGTAGCGATC; 591, 5′-CTTCCTCGTGCTTTACGGTA TCG; 593, 5′-AGAGAAATAGTCCCGTCAGC; and 590, 5′-CAGAGCAGCC GATTGTCTGTTG) (8, 57). CFT073PcL*upaH* was confirmed by PCR amplification and subsequent DNA sequencing. CFT073^{cam} was constructed by inserting the chloramphenicol resistance gene cassette from pKD3 at the *lacZ* locus using primers with 50-bp homology extensions (1547, 5′-TGATCATTAACTATCCGCTGGATG ACCAGGATGCCATTGCTGTGGAAGCTTAGGGCGAATTGACATTGTG; and 1606, 5′-CCGCGCCAATAACAGAATTATCCCCGATGGTGACGCCTGG ATTAATAACCCTGGCAGTTCCCTACTCTCG). The mutant was confirmed via PCR and sequencing with the following primers: upstream region (1551, 5′-CCGG TGTCTCTTATCAGACC; and 1177, 5′-TCACCTTCACCCTCTCCACT), down-

stream region (1638, 5′-CTGGAGTGAATACCACGACG; and 1635, 5′-TCCCG GTCGTTTATTTCGCG).

DNA manipulations and genetic techniques. DNA techniques were performed as previously described (49). Isolation of plasmid DNA was carried out using the QIAprep spin miniprep kit (Qiagen). Restriction endonucleases were used according to the manufacturer's specifications (New England Biolabs). Chromosomal DNA purification was made using the GenomicPrep cell and tissue DNA isolation kit (Amersham Pharmacia Biotech Inc.). Oligonucleotides were purchased from Sigma, Australia. The *upaH* gene was amplified with the Expand long-template polymerase system (Roche) as described by the manufacturer. DNA sequencing was performed using the ABI Big Dye ver3.1 kit (ABI) by the Australian Genome Research Facility, Brisbane, Australia. Prevalence studies for the *upaH* gene used Nova*Taq* DNA polymerase, as described by the manufacturer (Novagen), with the primers 524 (5′-AGTGAAGGGGCAAAAAC CTT) and 525 (5′-TGAAACCACCACCATTCTGA). The following program was used: 94°C for 15 s, 50°C for 15 s, and 72°C for 1 min (25×). The orientation of the type 1 fimbrial switch promoter was assessed as described previously (55).

Sequencing of *upaH*. *In vitro* transposon mutagenesis of the *upaH* gene cloned into pUpaH was performed with the Epicentre EZ::Tn5<KAN-2> transposon mutagenesis kit as described by the manufacturer. A library of overlapping transposon mutants was generated by mapping the transposon insertion site relative to the start site of *upaH*. A primer targeted to the end of the transposon was used to sequence into *upaH*.

RT-PCR. For the detection of *upaH* mRNA by reverse transcriptase PCR (RT-PCR), urine from six infected mice was collected just prior to euthanasia, pooled, and added directly to a two-times volume of RNAprotect (Qiagen). Total bacterial RNA was isolated and purified using an RNeasy minikit (Qiagen). After purification, RNA was treated with RNase-free DNase I (Novagen) to remove contaminating DNA and repurified using Qiagen RNeasy columns. RNA samples were quantified spectrophotometrically at 260 nm and additionally checked by gel electrophoresis. RNA was converted to cDNA using Superscript III first-strand synthesis supermix as described by the manufacturer (Invitrogen Life Technologies). cDNA was used directly as a template for PCR using primers specific for *upaH* (524 and 525). A negative control using the original RNA was run in parallel to confirm the absence of contaminating DNA.

Purification of truncated UpaH, antibody production, and immunoblotting. A 747-bp segment from the passenger-encoding domain of *upaH* was amplified with the primers UpaHTF (5′-GGCCATGGAAATCGCGGGTAATAACGCG GTGGTGAACCAGGA) and UpaHTR (5′-GGAAGCTTGGCAAATATACT GATGAGTT), digested with NcoI/HindIII, and ligated into NcoI/HindIII-digested pBAD/*Myc*-HisB. The resultant plasmid (pUpaHTruncated) contained the base pairs 5050 to 5796 of *upaH* fused to a 6×His-encoding sequence. *E. coli* BL21 was transformed with the plasmid pUpaHTruncated, induced with arabinose, and the resultant 6×His-tagged protein was purified using the Qiagen nickel-nitrilotriacetic acid (Ni-NTA) spin kit according to the manufacturer's instructions. Protein purity was assessed by SDS-PAGE analysis as previously

described (58). Polyclonal anti-UpaH serum was raised in rabbits by the Institute of Medical and Veterinary Sciences (South Australia), following a standard protocol (http://www.imvs.sa.gov.au/). For immunoblotting, whole-cell lysates were subjected to SDS-PAGE using NuPAGE Novex 3 to 8% Tris-acetate precast gels with NuPAGE Tris-acetate running buffer and subsequently transferred to polyvinylidene difluoride (PVDF) microporous membrane filters using the iBlot dry blotting system as described by the manufacturer (Invitrogen). Serum raised against the passenger subunit of UpaH was used as primary serum, and the secondary antibody was alkaline phosphatase-conjugated anti-rabbit IgG. Sigma Fast 5-bromo-4-chloro-3-indolylphosphate–nitroblue tetrazolium (BCIP/NBT) was used as the substrate in the detection process.

Biofilm assays. Biofilm formation on polystyrene surfaces was monitored by using 96-well microtiter plates (BD Falcon) essentially as previously described (52). Briefly, cells were grown for 24 h in glucose M9 minimal medium (containing 0.1% arabinose for induction of *upaH* gene expression) at 37°C, washed to remove unbound cells, and stained with crystal violet. Quantification of bound cells was performed by addition of acetone-ethanol (20:80) and measurement of the dissolved crystal violet at an optical density of 595 nm. Flow chamber experiments were performed as previously described (30, 51). Briefly, biofilms were allowed to form on glass surfaces in a multichannel flow system (4 ml/h) that permitted online monitoring of community structures. Flow cells were inoculated with optical density at 600 nm (OD_{600})-standardized cultures pregrown overnight in glucose M9 minimal medium. For *E. coli* OS56 biofilms, kanamycin was added as required for maintenance of plasmids. CFT073, CFT073*upaH*, and CFT073*upaH*PcL strains were tagged by introduction of the *gfp*-containing plasmid pSG10, and biofilms were grown as described above but in the presence of chloramphenicol. Biofilm development was monitored by confocal scanning laser microscopy at 24 h or 40 h after inoculation. For analysis of the flow cell biofilms, 16 z-stacks were collected for each strain and analyzed by using the COMSTAT software program (26). All experiments were performed in triplicate.

Microscopy and image analysis. An anti-UpaH specific serum was used for immunofluorescence microscopy and immunogold electron microscopy as previously described (56, 59). Microscopic observation of biofilms and image acquisition were performed on a scanning confocal laser microscope (LSM510 Meta; Zeiss) equipped with detectors and filters for monitoring of green fluorescent protein (GFP). Vertical cross sections through the biofilms were visualized by using the Zeiss LSM image examiner. Images were further processed for display by using Photoshop software (Adobe, Mountain view, CA).

Mouse model of UTI. Female C57BL/6 mice (8 to 10 weeks) were purchased from the University of Queensland Animal Facility and housed in sterile cages with *ad libitum* access to sterile water. The mouse model of UTI described previously was used for this study (48). Briefly, an inoculum of 20 μl, containing 5×10^8 CFU of bacteria in phosphate-buffered saline (PBS) containing 0.1% India ink, was instilled directly into the bladder using a 1-ml tuberculin syringe attached to the catheter. Urine was collected from each mouse at 18 h after inoculation for quantitative colony counts. Mice were euthanized at 18 h after challenge by cervical dislocation; bladders and kidneys were then excised aseptically, weighed, and homogenized in PBS. Bladder and kidney homogenates were serially diluted in PBS and plated on LB agar for colony counts. Data are expressed as the median number of CFU per 0.1 g of bladder tissue or ml of urine. All animal experiments were repeated at least twice. Competitive mixed-infection assays were performed as described above except that mice were inoculated with a mixture of 2.5×10^8 CFU of CFT073[cam] and 2.5×10^8 CFU of CFT073*upaH* grown in the absence of antibiotics. The two strains were differentiated by their resistance to chloramphenicol (CFT073[cam]) and kanamycin (CFT073*upaH*), respectively, to enable differential colony counts of the recovered bacteria. Median bacterial CFU were compared using either the nonparametric Mann-Whitney test for the independent challenges or the nonparametric Wilcoxon matched-pairs test for cochallenge infections within the InStat software package (Graphpad, San Diego, CA). *P* values of <0.05 were considered significant.

Computational sequence and structure analysis. Amino acid sequence searches were performed using the BlastP and PSI-BLAST software programs (1) as implemented on the NCBI (National Center for Biotechnology Information) website. Sequence repeats were analyzed using the RADAR (22), REPRO (16), and XSTREAM (41) software programs. Secondary structure was predicted using PSI-PRED software program (v2.6) (6). Fold recognition was performed using the I-TASSER (65) and mGenThreader software (6). The three-dimensional structures were modeled using the program Modeler 9v5 (12). The BetaWrapPro (37) server was used for the detection of beta-helix motifs. Structure figures were prepared using the PyMol software program (DeLano Scientific LLC).

Nucleotide sequence accession number. The complete nucleotide sequence of *upaH* has been deposited in the NCBI database (Bankit number 1182914).

RESULTS

Identification of *upaH*, a new AT-encoding gene in the genome sequence of UPEC CFT073. We performed a bioinformatic analysis of the *E. coli* CFT073 genome to search for novel putative adhesins and identified a 6.5-kb unannotated intergenic region that contains a disrupted copy of a previously unidentified AT-encoding gene (Fig. 1A). At the corresponding position on the *E. coli* K-12 MG1655 genome, there is a large pseudogene (8622 nucleotides), called *ydbA*, that has been truncated by insertion sequences (Fig. 1A). Were it not for these insertion sequences, *ydbA* would encode a full-length AT protein. An intact positional orthologue is also present in the genome sequence of UPEC strain UTI89, although it has been misannotated as a major facilitator superfamily-type transporter (Fig. 1A). In the published genome sequence of *E. coli* CFT073, a frameshift is reported in the *ydbA* gene; however, this may be a sequencing artifact since there is a section of uncertainty (CAAGGCGGTYGTSAAYAAA) in this region. Close examination of the *E. coli* CFT073 genome suggests that the repeat region of *ydbA* was misassembled during genome closure and an intact AT-encoding gene may actually be present. In accordance with the terminology adopted for other AT-encoding genes of UPEC (45), we have termed the *ydbA* gene *upaH*. PCR amplification of the entire *upaH* gene from *E. coli* CFT073 revealed a product that was approximately 2 kb larger than that expected based on the published genome sequence, confirming an error in genome assembly. In order to sequence the *upaH* gene, the *upaH* PCR product was cloned as a transcriptional fusion downstream of the tightly regulated *araBAD* promoter in the pBAD/*Myc*-HisA expression vector to generate the plasmid pUpaH. *In vitro* transposon mutagenesis was employed to generate a library of overlapping transposon mutants, which enabled sequencing across the repetitive region of the gene. This revealed that, in contrast to the published genome sequence, the *E. coli* CFT073 *upaH* gene is an intact coding region of 8.535 kb in length.

Bioinformatic analysis of UpaH predicts it is an AT protein. Analysis of the translated *upaH* sequence using the Conserved Domain Database (CDD) search (36) detected the AT β-domain (translocation domain) in the C-terminal region of the protein (Fig. 1B). Sequence searches against the Protein Data Bank revealed no significant similarities even when searching with a position-specific scoring matrix (PSSM) based on PSI-BLAST searches of the nonredundant protein sequences database. The known crystal structures of AT β-domains include NalP from *Neisseria meningitidis* (43) and EspP from *E. coli* (3). The similarity of both of these structures with the β-domain of UpaH could be detected by the fold recognition algorithm I-TASSER (65). Alignment of the UpaH and NalP β-domain sequences with ClustalW (34) shows 15% identity. We modeled the structure of this UpaH domain (residues 2528 to 2845) with the software program Modeler (12) using the NalP structure as a template (Fig. 1C). The model predicts a central helix (residues 2530 to 2556) winding through a 12-strand β-barrel. The periplasmic β-barrel loops are short and well defined in the model; on the other hand, the extracellular loops
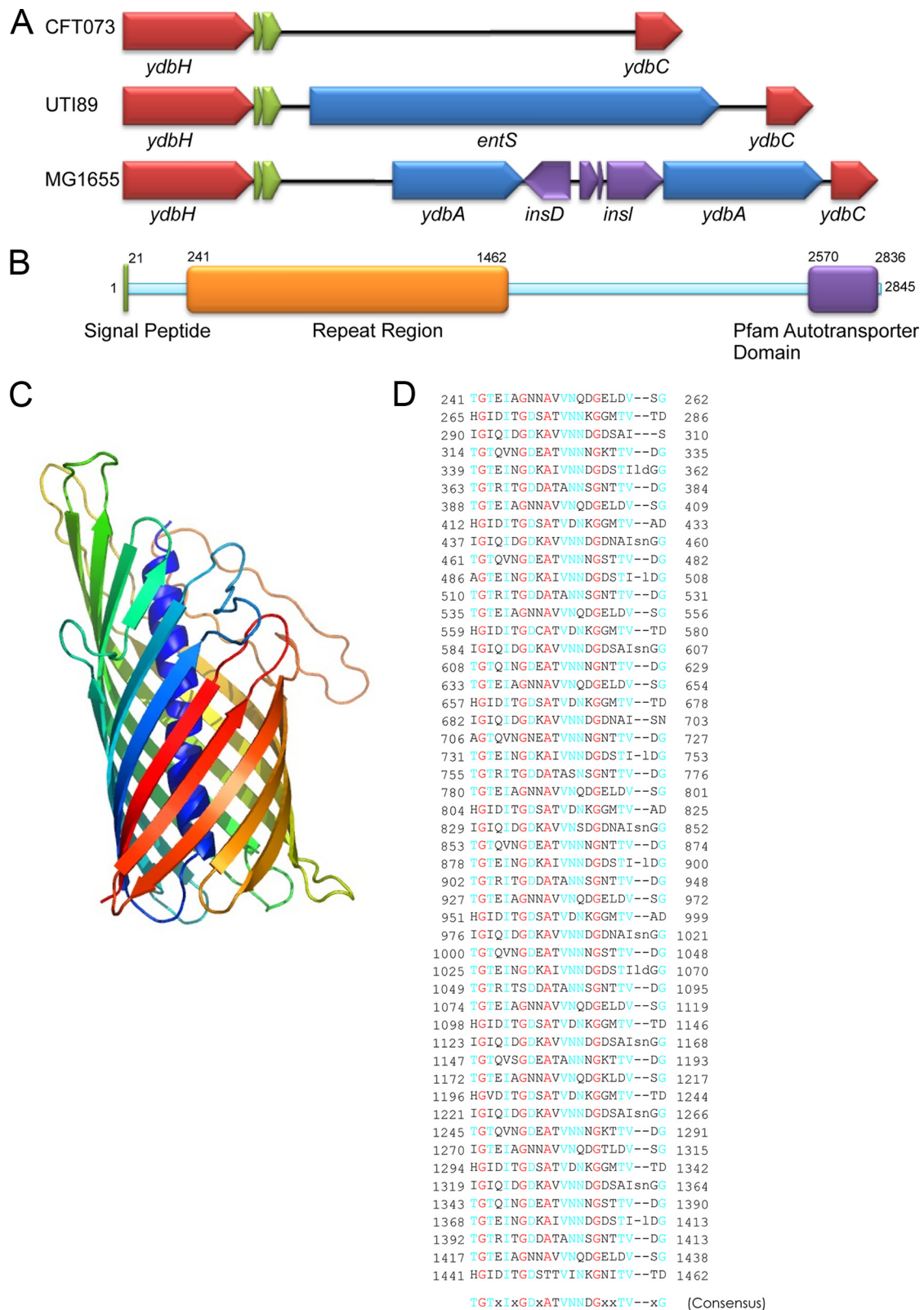
FIG. 1. (A) Physical map indicating the genomic location of *upaH* in *E. coli* CFT073, UTI89, and MG1655. (B) Schematic illustration of the domain organization of UpaH from *E. coli* CFT073. Indicated are the signal peptide, the repeat region within the passenger domain, and the transmembrane β-domain. Numbers indicate amino acid positions of region boundaries. (C) Predicted structure of the β-domain of UpaH. Shown is a ribbon diagram of the structure of the β-domain (residues 2528 to 2845) obtained by homology modeling based on the crystal structure of the NalP β-domain (43), colored blue (N terminus) through green and yellow to red (C terminus). (D) Amino acid sequence of the repeats present in UpaH. Amino acids in red indicate ≤98% conservation; amino acids in blue indicate ≤55% conservation.

are longer and show more variability in different models created by Modeler, suggesting they are less well defined in the model (in particular the loops connecting β-barrel strands 1 and 2 [residues 2573 to 2579], 3 and 4 [residues 2616 to 2624], 5 and 6 [residues 2673 to 2679], and 9 and 10 [residues 2783 to 2802]). The large passenger domain of UpaH contains ~50 imperfect tandem sequence repeats (Fig. 1D). The smallest repeating unit alternates in length between 24 and 25 amino acids, and larger-length repeats are evident that are multiples of 24 to 25 (98, 147, 196, etc). It appears that the sequence of this domain may have evolved through a number of sequential duplications involving fragments containing variable numbers of repeats. The predominant secondary structure predicted in this domain is β-structure, and a repeating pattern of three β-sheets per sequence repeat can be observed.

The *upaH* gene is highly prevalent among *E. coli* isolates. The widespread occurrence of AT-encoding genes in *E. coli* prompted us to investigate the prevalence of *upaH* in UPEC isolates from our laboratory collection, as well as the 72 strains from the *E. coli* reference (ECOR) collection. For this purpose, we employed primers designed to amplify the region encompassing nucleotides 8083 to 8410 of *upaH*. We found that the correct-size base pair product was amplified from 90/118 (76%) of the UPEC strains (15/21 cystitis strains, 28/31 pyelonephritis strains, and 47/66 asymptomatic bacteriuria [ABU] strains), as well as 65/72 (90%) of the ECOR strains. The presence of *upaH* as assessed by PCR was also determined for these strains according to their phylogenetic group and revealed the following prevalence pattern: group A, 30/35 (86%); B1, 17/20 (85%); B2, 80/95 (84%); and D, 23/33 (70%). These results suggest that the *upaH* gene is widely distributed among *E. coli* strains.

The *upaH* gene encodes a large protein that is located at the bacterial cell surface. To demonstrate expression of the UpaH protein, plasmid pUpaH was transformed into the previously described *E. coli* K-12 *flu* mutant strain MS427 (53). This strain is unable to mediate the classical cell aggregation and biofilm phenotypes associated with Ag43 expression (46). UpaH expression was detected using a rabbit polyclonal antiserum targeting a region of UpaH comprising amino acids 1424 to 1673 of the N-terminal passenger domain. Western blot analysis employing anti-UpaH antibodies detected the UpaH protein in whole-cell lysates prepared from *E. coli* MS427(pUpaH) following induction with arabinose. The size of the UpaH protein was approximately 460 kDa, significantly larger than its expected molecular mass of 290 kDa (Fig. 2A). Such a disparity in mass has been observed previously for proteins containing repeated acidic sequences, including the AT protein ShdA from *Salmonella enterica* (29). UpaH is also highly acidic, containing 376/2,845 (13.2%) acidic amino acids (glutamate and aspartate). The overexpression of UpaH also resulted in multiple smaller bands presumed to be breakdown products (Fig. 2A).

Our bioinformatic analysis suggested that *upaH* encodes an AT protein and thus UpaH may localize to the outer membrane. To test this tenet, we performed immunofluorescence microscopy (Fig. 2B). UpaH antiserum readily reacted with *E. coli* MS427(pUpaH) following induction with arabinose during growth in LB broth, demonstrating that the N-terminal region of UpaH was effectively translocated to the cell surface. No reaction was seen with *E. coli* MS427(pBAD) cells. The surface
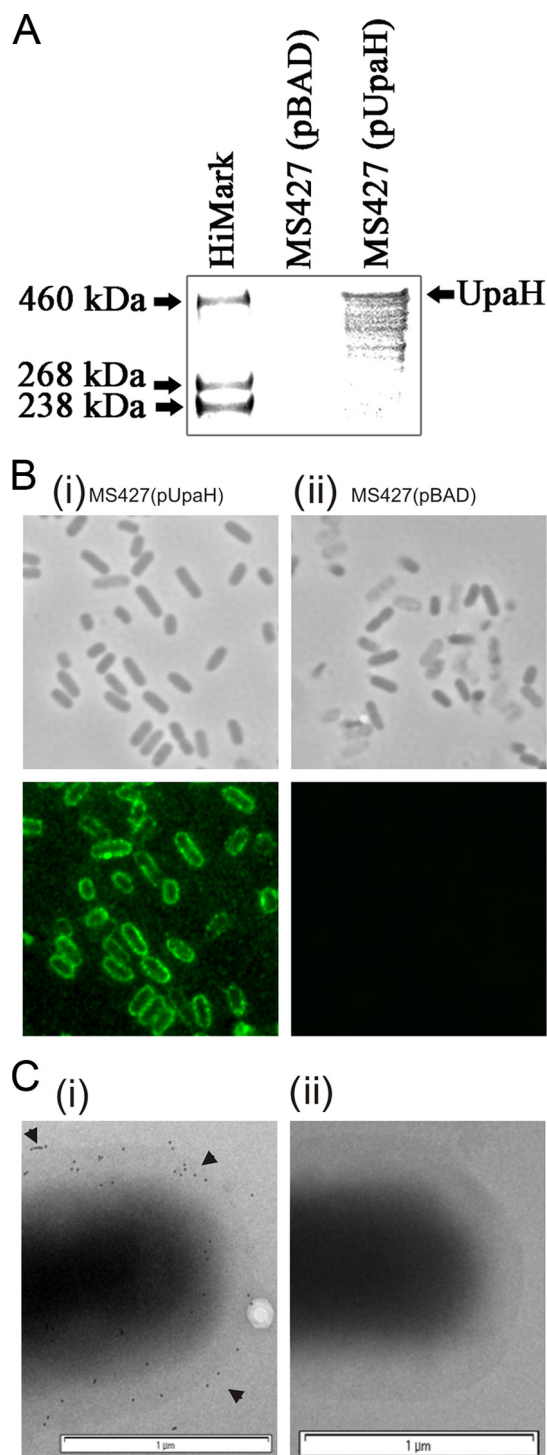

FIG. 2. (A) Western blot analysis of UpaH performed using whole-cell lysates prepared from *E. coli* MS427(pBAD) and MS427(pUpaH) grown in the presence of 0.1% arabinose. (B) Phase-contrast and immunofluorescence microscopy employing UpaH-specific antiserum against cells of *E. coli* strains MS427(pUpaH) (i) or MS427(pBAD) (ii) grown in the presence of arabinose. Overnight cultures were fixed and incubated with anti-UpaH serum followed by incubation with goat anti-rabbit IgG coupled to Alexa Fluor 488. (C) Immunogold electron microscopy of *E. coli* MS427(pUpaH) (i) and MS427(pBAD) (ii) grown in the presence of arabinose and labeled with anti-UpaH serum followed by protein A-gold (10 nm) conjugate. Gold particles were observed on the surface of *E. coli* MS427(pUpaH) but not MS427(pBAD).

location of UpaH was confirmed using immunogold labeling and electron microscopy (Fig. 2C). *E. coli* MS427(pUpaH) cells displayed abundant surface labeling with gold particles following incubation with UpaH antiserum compared to results for *E. coli* MS427(pBAD) cells. Thus, the *upaH* gene is functional and encodes a large protein located at the cell surface in *E. coli*.

Phenotypic properties of UpaH. The presence of multiple repeats within the predicted passenger-encoding domain of UpaH suggested it may possess some functional properties similar to those of the well-characterized Ag43 and AIDA-I AT proteins. However, we saw no evidence of UpaH-mediated cell aggregation in *E. coli* MS427 harboring pUpaH (data not shown). UpaH overexpression also failed to mediate adherence to T24 bladder epithelial cells or to the extracellular matrix proteins collagen I, collagen III, collagen IV, fibronectin, and laminin (data not shown). To determine whether the UpaH protein promotes biofilm formation, we examined the effect of UpaH overexpression in *E. coli* MS427 (or the *gfp*-tagged derivative strain OS56) using static and dynamic biofilm assays. First, we employed a static nontreated polystyrene microtiter plate model and showed that UpaH promotes biofilm formation after growth in M9 minimal medium (Fig. 3A). Next, we tested the ability of UpaH to promote biofilm formation under dynamic conditions using the continuous-flow-chamber model, which permits monitoring of the bacterial distribution within an evolving biofilm at the single-cell level due to the combination of GFP-tagged cells and scanning confocal laser microscopy. UpaH promoted strong biofilm growth under these conditions and produced a structure with a depth of approximately 10 μm (Fig. 3B).

The UpaH protein is expressed in UPEC isolates. To determine whether UpaH is expressed in *E. coli* CFT073, we constructed a *upaH* deletion mutant employing λ-red-mediated homologous recombination of linear DNA (referred to as CFT073*upaH*). Examination of whole-cell lysates prepared from these isogenic strains following growth in LB broth by SDS-PAGE and Western blotting demonstrated expression of UpaH in *E. coli* CFT073 but not in CFT073*upaH* (Fig. 4A). Furthermore, complementation of CFT073*upaH* with plasmid pUpaH restored the ability of this strain to produce UpaH (albeit with evidence of protein degradation) (Fig. 4A). We also randomly selected 16 UPEC strains positive in the *upaH* PCR screen and tested them for the ability to express the UpaH protein following growth in LB by Western blot analysis employing UpaH-specific antiserum. A band corresponding to the UpaH protein was identified in all of these strains (data not shown). To confirm that this was in fact UpaH, we deleted the *upaH* gene in two strains (i.e., M161 and M357) using λ-red-mediated homologous recombination. Whole-cell lysates prepared from *E. coli* M161*upaH* and M357*upaH* failed to react with the UpaH-specific antiserum (Fig. 4B). Thus, UpaH is a newly identified AT protein expressed by *E. coli* CFT073 and other UPEC isolates. We also tested for surface localization of UpaH on *E. coli* CFT073 using immunofluorescence microscopy; however, we could not detect any expression. This was not caused by a defect in the ability to transport UpaH to the outer membrane, since surface localization of UpaH was detected in CFT073*upaH*(pUpaH) and a UpaH overexpression strain constructed by inserting a chromosomally located con-
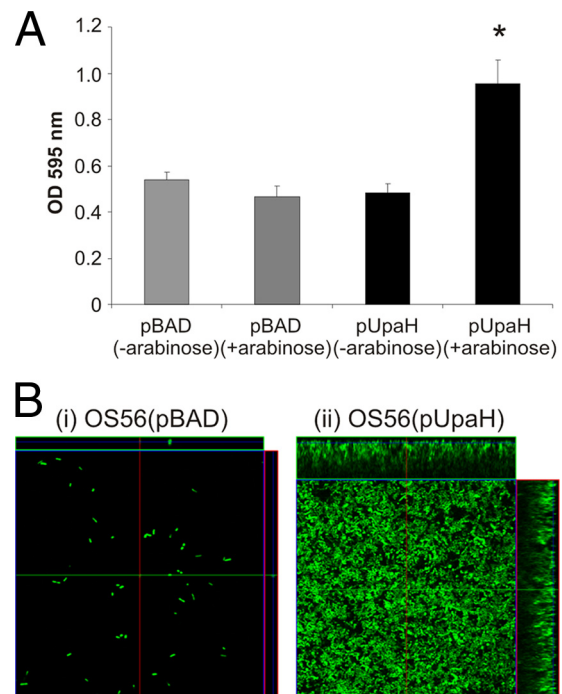

FIG. 3. UpaH mediates biofilm formation. (A) Biofilm formation in polystyrene microtiter plates by *E. coli* MS427(pBAD) and MS427(pUpaH) following static growth in the absence (−) or presence (+) of 0.1% arabinose. Biofilm growth was quantified by solubilization of crystal violet-stained cells with ethanol-acetone and determination of the absorbance at 595 nm. Results represent averages of a minimum of 8 replicates ± SEM. Significant biofilm growth was observed for *E. coli* MS427(pUpaH) following induction of UpaH expression with arabinose (*, $P < 0.001$ as determined by analysis of variance [NOVA]). (B) Dynamic-flow-chamber assay examining biofilm formation of *E. coli* OS56(pBAD) and OS56(pUpaH) grown in the presence of 0.1% arabinose. Biofilm development was monitored by confocal scanning laser microscopy after 24 h. The images are representative horizontal sections collected within each biofilm and vertical sections (to the right of and below each larger panel, representing the yz plane and the xz plane, respectively) at the positions indicated by the white lines. *E. coli* MS427(pUpaH) produced a significant biofilm in this assay.

stitutive promoter upstream of the *upaH* coding region (CFT073PcL*upaH*) (Fig. 4C).

Deletion of *upaH* reduces biofilm formation by UPEC strain CFT073. The contribution of UpaH to biofilm formation by *E. coli* CFT073 was assessed, employing both microtiter plate and flow-cell-based assays. In the microtiter plate assay, *E. coli* CFT073 produced a significantly increased biofilm compared to *E. coli* CFT073*upaH* (Fig. 5A). We also examined biofilm formation by the *E. coli* M161/M161*upaH* and M357/M357*upaH* isogenic strain sets. *E. coli* M161 and M357 did not form a strong biofilm in this assay; there was no significant difference in the biofilm formed by *E. coli* M161*upaH* and M357*upaH* (Fig. 5A). To characterize the function of UpaH in *E. coli* CFT073 further, we tested the ability of *E. coli* CFT073, CFT073*upaH*, and CFT073PcL*upaH* to promote biofilm formation in a continuous-flow-chamber system. In this assay, CFT073*upaH* formed a biofilm with a lower biovolume, substratum coverage, and mean thickness than those of the biofilms formed by *E. coli* CFT073 and CFT073PcL*upaH* ($P =$
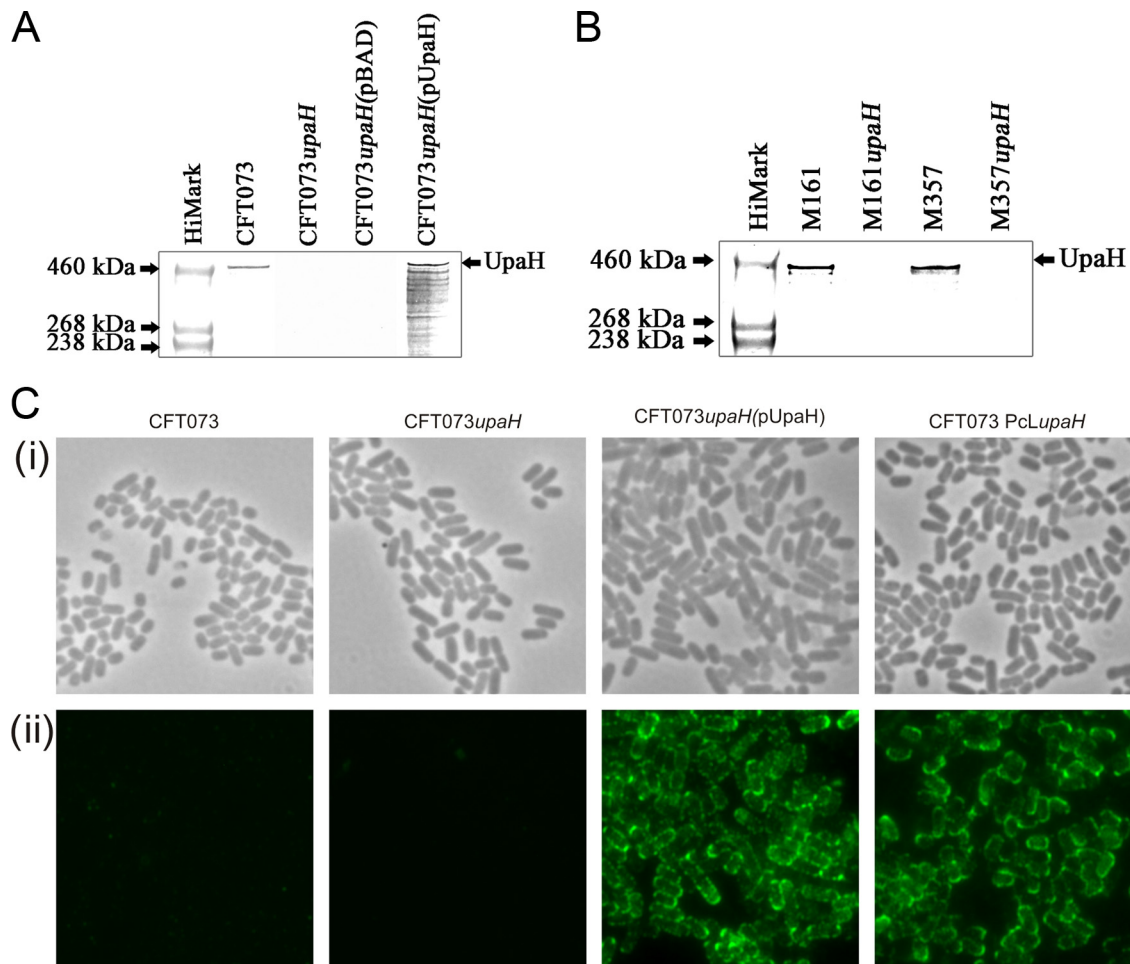
FIG. 4. (A) Western blot analysis of UpaH performed using whole-cell lysates prepared from *E. coli* CFT073, CFT073*upaH*, CFT073*upaH*(pBAD), and CFT073*upaH*(pUpaH). A band corresponding to UpaH was detected in *E. coli* CFT073 and CFT073*upaH*(pUpaH) but not in *E. coli* CFT073*upaH* and CFT073*upaH*(pBAD). (B) Western blot analysis of UpaH performed using whole-cell lysates prepared from *E. coli* M161, M161*upaH*, M357, and M357*upaH*. A band corresponding to UpaH was detected in *E. coli* M161 and M357 but not in *E. coli* M161*upaH* or M357*upaH*. (C) Phase-contrast (i) or immunofluorescence (ii) microscopy employing UpaH-specific antiserum against cells of *E. coli* CFT073, CFT073*upaH*, CFT073*upaH*(pUpaH), or CFT073Pc*LupaH*. Overnight cultures were fixed and incubated with anti-UpaH serum, followed by incubation with goat anti-rabbit IgG coupled to Alexa Fluor 488. A positive reaction indicating surface localization of UpaH was detected only for CFT073*upaH*(pUpaH) and CFT073Pc*LupaH*.

0.008, $P = 0.008$, and $P = 0.012$, respectively) (Fig. 5B and C). Taken together, these results demonstrate that UpaH contributes to biofilm formation by CFT073.

The *upaH* gene does not contribute to colonization of the mouse bladder in single-infection experiments. To study the role of UpaH in virulence, we examined the ability of CFT073 and CFT073*upaH* to survive in the mouse urinary tract. There was no significant difference in the number of CFT073 or CFT073*upaH* cells recovered from the bladders and urine of mice following infection (data not shown). No significant colonization of the kidneys was observed for either strain; this is consistent with previous data from our laboratory using C57BL/6 mice (57, 59). To examine if this was a strain-specific effect, we also compared *E. coli* M161/M161*upaH* and M357/M357*upaH* for their ability to colonize the mouse urinary tract as described above. There was no significant difference in colonization of the bladder or recovery from urine for the parent and *upaH* mutant of both isogenic strain sets (data not shown);

no significant colonization of the kidneys was observed for any of the strains.

The *upaH* gene is expressed during colonization of the mouse urinary tract. The above result prompted us to examine whether *upaH* is transcribed during infection of C57BL/6 mice by CFT073 using RT-PCR. RNA was extracted from pooled mouse urine collected 18 h postinoculation and converted to cDNA. A PCR using *upaH*-specific primers revealed the presence of an amplification product from the cDNA samples, while no PCR product was obtained from the RNA sample prior to cDNA synthesis (Fig. 6A). Thus, the *upaH* gene located on the *E. coli* CFT073 chromosome is transcribed during colonization of the mouse urinary tract.

Deletion of *upaH* from CFT073 reduces colonization of the mouse bladder in a mixed-infection experiment. For a more sensitive method to test the role of *upaH* in colonization of the mouse urinary tract, we performed competitive-colonization experiments where we coinoculated mice with the CFT073 and
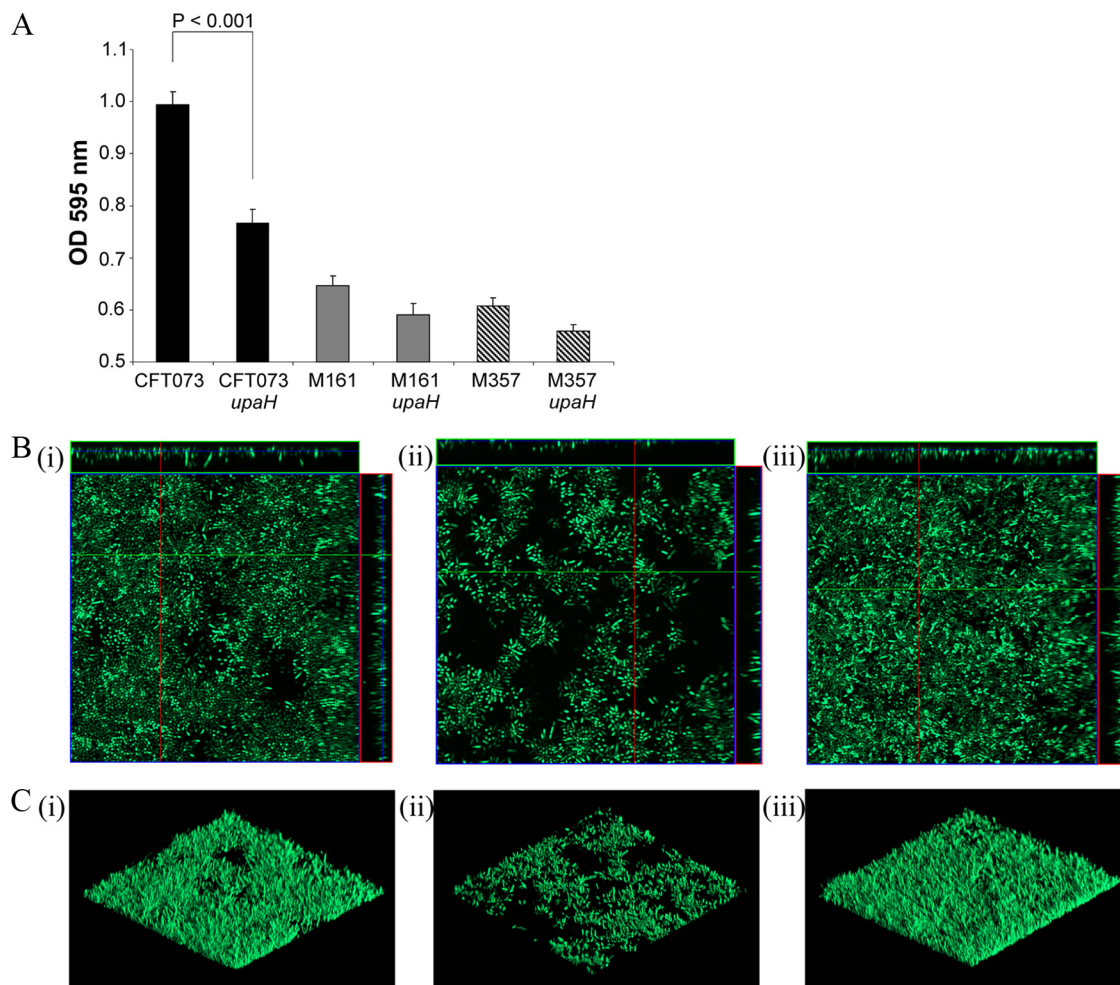
FIG. 5. (A) Biofilm formation in polystyrene microtiter plates by three sets of isogenic strains with respect to *upaH*: *E. coli* CFT073 and CFT073*upaH*; *E. coli* M161 and M161*upaH*; and *E. coli* M357 and M357*upaH*. Results represent averages of OD_{595} readings of at least 4 independent experiments ± SEM. There was no difference in the growth of wild-type and *upaH* mutant strains. (B) Dynamic-flow-chamber assay examining biofilm formation of *E. coli* CFT073 (i), *E. coli* CFT073*upaH* (ii), and *E. coli* CFT073PcL*upaH* (iii). Biofilm development was monitored by confocal scanning laser microscopy after 40 h. The images are representative horizontal sections collected within each biofilm and vertical sections (to the right of and below each larger panel, representing the *yz* plane and the *xz* plane, respectively) at the positions indicated by the white lines. (C) Three-dimensional image demonstrating the structure of the biofilm formed by *E. coli* CFT073 (i), *E. coli* CFT073*upaH* (ii), and *E. coli* CFT073PcL*upaH* (iii).

CFT073*upaH* strains in a 1:1 ratio. These experiments were performed using a derivative of CFT073 tagged at the *lacZ* locus with a chloramphenicol resistance gene cassette (i.e., CFT073^{cam}) and CFT073*upaH* (which contains a kanamycin resistance gene cassette in *upaH*). The two strains had an identical growth rate in pooled human urine (and LB broth) and did not display any difference with respect to type 1 fimbria production (as assessed by yeast cell agglutination and a *fim* switch orientation PCR), hemolysis of red blood cells, and motility (data not shown). We found that CFT073^{cam} significantly outcompeted CFT073*upaH* at 18 h in urine ($P < 0.001$) and the bladder ($P < 0.001$) (Fig. 6B). In contrast, we did not observe a significant difference in colonization using mixed-infection experiments employing M161/M161*upaH* or M357/M357*upaH* (data not shown). These results suggest that the ability to express UpaH confers a competitive advantage for

the survival of CFT073 in the mouse urinary tract but that this effect is not specific to all strains of UPEC.

DISCUSSION

UPEC produces a range of fimbrial and nonfimbrial adhesins that play a role in virulence and contribute to persistent infection of the urinary tract. Some UPEC fimbrial adhesins, including type 1 and P fimbriae, have been well characterized with respect to their expression, regulation, receptor-binding target, and role in virulence. Of the nonfimbrial adhesins, the AT family of proteins represents a novel group of virulence factors because of their role in adhesion, invasion, and biofilm formation in other organisms. Although UPEC strains possess multiple AT-encoding genes, very little is known about their function, and only Ag43, UpaG, and Sat have been shown to
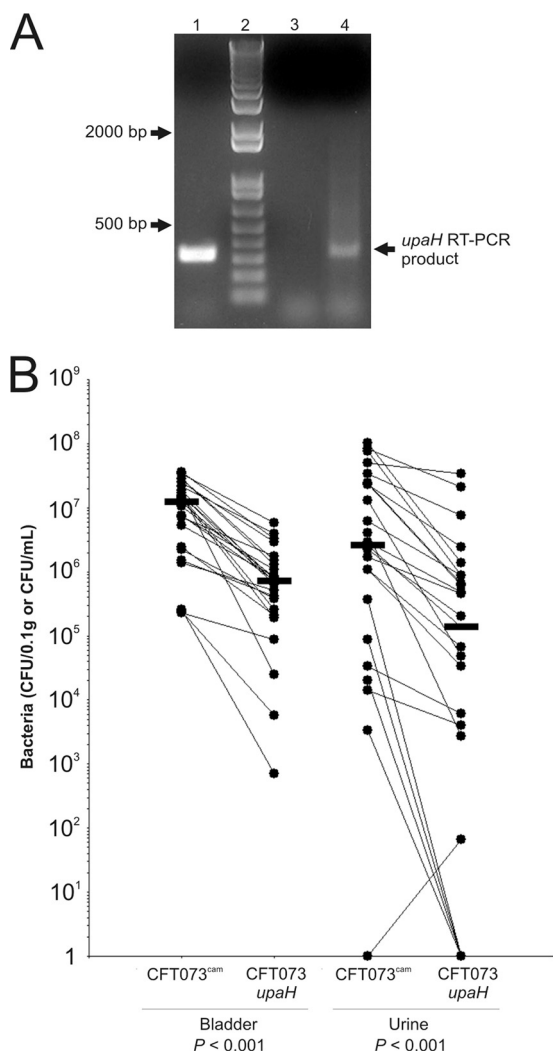
FIG. 6. (A) RT-PCR analysis of *upaH* transcription by CFT073 during colonization of the mouse urinary tract. Urine from six mice infected with CFT073 was collected and pooled at 18 h postinoculation and added directly into a 2× volume of RNAprotect (Qiagen). CFT073 cells were concentrated by centrifugation, and total RNA was extracted. Lane 1, *upaH*-specific PCR product (328 bp) obtained from genomic DNA (positive control); lane 2, 1-Kb Plus DNA ladder (Invitrogen); lane 3, PCR employing RNA prior to cDNA synthesis as a template (negative control); lane 4, *upaH*-specific PCR product (328 bp) obtained using the cDNA as the template. (B) Competitive mixed-infection experiment. C57BL/6 mice ($n = 24$) were challenged with a 1:1 mixture of *E. coli* CFT073[cam] and CFT073*upaH*. Total CFU were enumerated on selective medium to differentiate between CFT073[cam] (chloramphenicol resistant) and CFT073*upaH* (kanamycin resistant). Each symbol represents the total number of CFU for an individual mouse per 0.1 g of bladder tissue or the total number of CFU per ml urine. Lines connect values for the same mouse. Horizontal bars represent the median values. No significant colonization of the kidneys was observed.

play a role in virulence. Here we identify and characterize UpaH as a new AT protein of UPEC and demonstrate that it contributes to biofilm formation and colonization of the mouse bladder in a mixed-infection assay.

The *upaH* gene is not annotated in the genome sequence of *E. coli* CFT073 due to a misassembly within a region of highly repetitive DNA. In *E. coli* K-12, a large pseudogene (8,622 bp), called *ydbA*, that has been truncated by insertion sequences is located at the corresponding position on the chromosome. Positional orthologues of *ydbA* are also discernible in several other pathogenic and nonpathogenic *E. coli* strains, although they are not always intact (e.g., *E. coli* O157:H7 EDL933 z2322/z2323 represent a split-gene version of *ydbA*). In the UPEC strain UTI89 genome, *ydbA* has been misannotated as encoding the "EntS/YbdA major facilitator superfamily (MFS) transporter." This unfortunate error likely arose because the EntS MFS enterobactin transporter described by Furrer et al. (14) was encoded by the gene *ybdA*. Similarity searches with the true EntS sequence (gi|2506613) clearly demonstrate that the UTI89 EntS protein and the original EntS/YbdA protein are not related proteins. The combination of events that led to the misannotation of *upaH* emphasizes the need for caution when assembling highly repetitive regions of genome sequence data, especially since in some cases these repeats may occur within genes that encode cell surface proteins. The "genome rot" caused by the UTI89 EntS error (already propagated to three recent *E. coli* genome annotations) also highlights the need for caution when predicting function based on annotation alone, even when the annotation has been manually curated.

The passenger domain of UpaH contains 50 imperfect repeats, each of which comprises 24 to 25 amino acid residues. The solved structures of passenger domains from other AT proteins, including pertactin from *Bordetella pertussis* (10), hemoglobin protease from *E. coli* (44), and pore-forming toxin from *Helicobacter pylori* (15), all contain the β-helix motif. The β-helix structure contains a superhelical arrangement of structural repeats, each consisting of three β-strands in a triangular arrangement (64). These repeats pack into a cylindrical solenoid structure, with the β-strands forming parallel β-sheets along the cylinder. The repeating sequence pattern in the UpaH passenger domain is consistent with structural repeats found in a β-helix. Despite the lack of detectable sequence similarities, we therefore propose that the bulk of the passenger domain would correspond to a β-helix structure. Several other pieces of evidence support this prediction: (i) the program BetaWrapPro (37), which specializes in identifying β-helix motifs, identifies several such motifs in UpaH; (ii) the fold recognition programs, such as I-TASSER (65) and mGenThreader (6), suggest the sequence of the UpaH passenger domain is most compatible with the structures of AT passenger domains and other β-helix proteins; and (iii) most AT passenger domains have been suggested previously to contain the β-helix structure, presumably to facilitate efficient secretion across the outer membrane and subsequent folding and to present the functional determinant at a distance from the bacterial surface (28).

Short repeat sequences have been described within the passenger domain of two other surface-located AT proteins from *E. coli*, namely, AIDA-I (4) and Ag43 (25). AIDA-I contains 21 repeats built up of 19 to 20 amino acid residues followed, after a short interruption, by an additional 10 repeats. The repeat structure of Ag43 is less well defined and more variable in length (12 to 20 amino acids). Repeat sequences often form solenoid structures with frequent roles in protein-protein interactions (32) and possibly ligand binding (61), both of which are consistent with the surface location of UpaH (and AIDA-I and Ag43) and the ability of these proteins to mediate biofilm formation. Other possible functions of short repeat sequences

may be to increase flexibility, alter the length, incorporate repetitive domain architecture, and mediate antigenic variation via duplication or deletion. The precise biological role of these short repeat sequences in UpaH and their level of conservation in UPEC remain to be determined.

In the murine UTI model, UPEC colonization of the bladder is associated with the formation of intracellular bacterial communities (IBCs) within superficial umbrella cells (2). IBCs possess biofilm-like properties and are comprised of compact cell clusters encased in a polysaccharide matrix. IBC formation creates a quiescent state that may be associated with long-term persistence in the bladder (11, 40). Two cell surface factors associated with biofilm formation, type 1 fimbriae, and Ag43 are expressed by cells within IBCs (2). We observed a significant difference in biofilm formation between CFT073 and CFT073upaH. However, this difference did not translate into a significant role for UpaH in the mouse urinary tract in single-infection experiments. To investigate this further, we also constructed a upaH deletion in two other clinical isolates from our laboratory collection that expressed UpaH (i.e., M161 and M357). There was no significant difference in bladder colonization or biofilm growth for M161/M161upaH or M357/M357upaH. Despite these results, we did observe a competitive advantage for CFT073 for colonization of the urinary tract in mixed-infection experiments with CFT073upaH. It is possible that UpaH and other AT proteins (e.g., Ag43 and UpaG) possess a high degree of functional redundancy with respect to the biofilm phenotype. The apparent strain-specific effect of UpaH in CFT073 versus M161/M357 may be associated with their different biofilm formation properties. Other surface factors differentially expressed between CFT073, M161, and M357 (e.g., fimbriae, flagella, and capsule) could also interfere with the function of UpaH, as has been reported previously for Ag43 (21, 50, 58). Thus, our in vivo analysis of the function of UpaH suggests it may play a modulatory role in colonization of the mouse urinary tract by CFT073 and that this contribution could occur in concert with other AT proteins (e.g., Ag43 and UpaG) that possess similar biofilm formation properties.

In a recent study, Hagan and Mobley identified 23 outer membrane proteins from E. coli CFT073 that elicit a humoral response in the mouse UTI model (20). This included porins, transporters, and receptors (including siderophore receptors). There are several possible reasons for the failure to detect UpaH in this screen. First, it is possible that as a result of its very large size, the UpaH protein did not migrate properly into the gel. We note that despite an apparent molecular mass of 290 kDa, UpaH migrated at approximately 460 kDa. Second, UpaH may be expressed in vivo at low levels. Although we were able to detect the presence of a upaH transcript by RT-PCR, the exact level of UpaH expression in vivo remains to be determined. Finally, it is possible that UpaH does not elicit a strong humoral response. However, the highly repetitive sequence of the passenger domain and its similarity to other AT proteins would suggest otherwise. The UpaH homologues present in the published genomes exhibit some variation in size, which may reflect the natural variation in the number of repeats. Alternatively, these genomes may contain incorrect assemblies due to the challenges of sequencing repetitive regions.

REFERENCES

1. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.
2. Anderson, G. G., J. J. Palermo, J. D. Schilling, R. Roth, J. Heuser, and S. J. Hultgren. 2003. Intracellular bacterial biofilm-like pods in urinary tract infections. Science 301:105–107.
3. Barnard, T. J., N. Dautin, P. Lukacik, H. D. Bernstein, and S. K. Buchanan. 2007. Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. Nat. Struct. Mol. Biol. 14:1214–1220.
4. Benz, I., and M. A. Schmidt. 1992. AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic Escherichia coli strain 2787 (O126:H27), is synthesized via a precursor molecule. Mol. Microbiol. 6:1539–1546.
5. Bertani, G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. J. Bacteriol. 62:293–300.
6. Bryson, K., L. J. McGuffin, R. L. Marsden, J. J. Ward, J. S. Sodhi, and D. T. Jones. 2005. Protein structure prediction servers at University College London. Nucleic Acids Res. 33:W36–W38.
7. 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. 93:9827–9832.
8. Da Re, S., B. Le Quere, J. M. Ghigo, and C. Beloin. 2007. Tight modulation of Escherichia coli bacterial biofilm formation through controlled expression of adhesion factors. Appl. Environ. Microbiol. 73:3391–4303.
9. 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. 97:6640–6645.
10. Emsley, P., I. G. Charles, N. F. Fairweather, and N. W. Isaacs. 1996. Structure of Bordetella pertussis virulence factor P.69 pertactin. Nature 381:90–92.
11. Eto, D. S., J. L. Sundsbak, and M. A. Mulvey. 2006. Actin-gated intracellular growth and resurgence of uropathogenic Escherichia coli. Cell Microbiol. 8:704–717.
12. Fiser, A., and A. Sali. 2003. Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol. 374:461–491.
13. Foxman, B. 2002. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am. J. Med. 1139(Suppl. 1A):5S–13S.
14. Furrer, J. L., D. N. Sanders, I. G. Hook-Barnard, and M. A. McIntosh. 2002. Export of the siderophore enterobactin in Escherichia coli: involvement of a 43 kDa membrane exporter. Mol. Microbiol. 44:1225–1234.
15. Gangwer, K. A., D. J. Mushrush, D. L. Stauff, B. Spiller, M. S. McClain, T. L. Cover, and D. B. Lacy. 2007. Crystal structure of the Helicobacter pylori vacuolating toxin p55 domain. Proc. Natl. Acad. Sci. U. S. A. 104:16293–16298.
16. George, R. A., and J. Heringa. 2000. The REPRO server: finding protein internal sequence repeats through the Web. Trends Biochem. Sci. 25:515–517.
17. Guyer, D. M., I. R. Henderson, J. P. Nataro, and H. L. Mobley. 2000. Identification of Sat, an autotransporter toxin produced by uropathogenic Escherichia coli. Mol. Microbiol. 38:53–66.
18. Guyer, D. M., S. Radulovic, F. E. Jones, and H. L. Mobley. 2002. Sat, the secreted autotransporter toxin of uropathogenic Escherichia coli, is a vacuolating cytotoxin for bladder and kidney epithelial cells. Infect. Immun. 70:4539–4546.
19. Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD promoter. J. Bacteriol. 177:4121–4130.
20. Hagan, E. C., and H. L. Mobley. 2007. Uropathogenic Escherichia coli outer membrane antigens expressed during urinary tract infection. Infect. Immun. 75:3941–3949.
21. Hasman, H., T. Chakraborty, and P. Klemm. 1999. Antigen 43-mediated autoaggregation of Escherichia coli is blocked by fimbriation. J. Bacteriol. 181:4834–4841.
22. Heger, A., and L. Holm. 2000. Rapid automatic detection and alignment of repeats in protein sequences. Proteins 41:224–237.
23. Henderson, I. R., and J. P. Nataro. 2001. Virulence functions of autotransporter proteins. Infect. Immun. 69:1231–1243.
24. Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez, and D. Ala'Aldeen. 2004. Type V protein secretion pathway: the autotransporter story. Microbiol. Mol. Biol. Rev. 68:692–744.
25. Henderson, I. R., and P. Owen. 1999. The major phase-variable outer membrane protein of Escherichia coli structurally resembles the immunoglobulin A1 protease class of exported protein and is regulated by a novel mechanism involving Dam and OxyR. J. Bacteriol. 181:2132–2141.

26. **Heydorn, A., A. T. Nielsen, M. Hentzer, C. Sternberg, M. Givskov, B. K. Ersboll, and S. Molin.** 2000. Quantification of biofilm structures by the novel computer program COMSTAT. Microbiology **146**(Pt 10)**:**2395–2407.
27. **Jones, C. H., J. S. Pinkner, R. Roth, J. Heuser, A. V. Nicholes, S. N. Abraham, and S. J. Hultgren.** 1995. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc. Natl. Acad. Sci. U. S. A. **92:**2081–2085.
28. **Junker, M., C. C. Schuster, A. V. McDonnell, K. A. Sorg, M. C. Finn, B. Berger, and P. L. Clark.** 2006. Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. Proc. Natl. Acad. Sci. U. S. A. **103:**4918–4923.
29. **Kingsley, R. A., D. Abi Ghanem, N. Puebla-Osorio, A. M. Keestra, L. Bergh-man, and A. J. Baumler.** 2004. Fibronectin binding to the *Salmonella enterica* serotype Typhimurium ShdA autotransporter protein is inhibited by a monoclonal antibody recognizing the A3 repeat. J. Bacteriol. **186:**4931–4939.
30. **Kjaergaard, K., M. A. Schembri, C. Ramos, S. Molin, and P. Klemm.** 2000. Antigen 43 facilitates formation of multispecies biofilms. Environ. Microbiol. **2:**695–702.
31. **Klemm, P., and M. A. Schembri.** 2000. Bacterial adhesins: function and structure. Int. J. Med. Microbiol. **290:**27–35.
32. **Kobe, B., and A. V. Kajava.** 2000. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. Trends Biochem. Sci. **25:**509–515.
33. **Kuehn, M. J., J. Heuser, S. Normark, and S. J. Hultgren.** 1992. P pili in uropathogenic E. coli are composite fibres with distinct fibrillar adhesive tips. Nature **356:**252–255.
34. **Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947–2948.
35. **Mabbett, A. N., G. C. Ulett, R. E. Watts, J. J. Tree, M. Totsika, C. L. Ong, J. M. Wood, W. Monaghan, D. F. Looke, G. R. Nimmo, C. Svanborg, and M. A. Schembri.** 2009. Virulence properties of asymptomatic bacteriuria Escherichia coli. Int. J. Med. Microbiol. **299:**53–63.
36. **Marchler-Bauer, A., J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. De-Weese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, S. Lu, G. H. Marchler, M. Mullokandov, J. S. Song, A. Tasneem, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, and S. H. Bryant.** 2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. **37:**D205–D210.
37. **McDonnell, A. V., M. Menke, N. Palmer, J. King, L. Cowen, and B. Berger.** 2006. Fold recognition and accurate sequence-structure alignment of sequences directing beta-sheet proteins. Proteins **63:**976–985.
38. **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:**1281–1289.
39. **Mulvey, M. A., Y. S. Lopez-Boado, C. L. Wilson, R. Roth, W. C. Parks, J. Heuser, and S. J. Hultgren.** 1998. Induction and evasion of host defenses by type 1-piliated uropathogenic Escherichia coli. Science **282:**1494–1497.
40. **Mysorekar, I. U., and S. J. Hultgren.** 2006. Mechanisms of uropathogenic *Escherichia coli* persistence and eradication from the urinary tract. Proc. Natl. Acad. Sci. U. S. A. **103:**14170–14175.
41. **Newman, A. M., and J. B. Cooper.** 2007. XSTREAM: a practical algorithm for identification and architecture modeling of tandem repeats in protein sequences. BMC Bioinformatics **8:**382.
42. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of Escherichia coli from natural populations. J. Bacteriol. **157:**690–693.
43. **Oomen, C. J., P. van Ulsen, P. van Gelder, M. Feijen, J. Tommassen, and P. Gros.** 2004. Structure of the translocator domain of a bacterial autotransporter. EMBO J. **23:**1257–1266.
44. **Otto, B. R., R. Sijbrandi, J. Luirink, B. Oudega, J. G. Heddle, K. Mizutani, S. Y. Park, and J. R. Tame.** 2005. Crystal structure of hemoglobin protease, a heme binding autotransporter protein from pathogenic *Escherichia coli*. J. Biol. Chem. **280:**17339–17345.
45. **Parham, N. J., U. Srinivasan, B. Desvaux, B. Foxman, C. F. Marrs, and I. R. Henderson.** 2004. PicU, a second serine protease autotransporter of uropathogenic Escherichia coli. FEMS Microbiol. Lett. **230:**73–83.
46. **Reisner, A., J. A. Haagensen, M. A. Schembri, E. L. Zechner, and S. Molin.** 2003. Development and maturation of *Escherichia coli* K-12 biofilms. Mol. Microbiol. **48:**933–946.
47. **Roberts, J. A., B. I. Marklund, D. Ilver, D. Haslam, M. B. Kaack, G. Baskin, M. Louis, R. Mollby, J. Winberg, and S. Normark.** 1994. The Gal(alpha 1-4)Gal-specific tip adhesin of Escherichia coli P-fimbriae is needed for pyelonephritis to occur in the normal urinary tract. Proc. Natl. Acad. Sci. U. S. A. **91:**11889–11893.
48. **Roos, V., G. C. Ulett, M. A. Schembri, and P. Klemm.** 2006. The asymptomatic bacteriuria Escherichia coli strain 83972 outcompetes uropathogenic E. coli strains in human urine. Infect. Immun. **74:**615–624.
49. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
50. **Schembri, M. A., D. Dalsgaard, and P. Klemm.** 2004. Capsule shields the function of short bacterial adhesins. J. Bacteriol. **186:**1249–1257.
51. **Schembri, M. A., and K. Kjaergaard, and P. Klemm.** 2003. Global gene expression in *Escherichia coli* biofilms. Mol. Microbiol. **48:**253–267.
52. **Schembri, M. A., and P. Klemm.** 2001. Biofilm formation in a hydrodynamic environment by novel fimh variants and ramifications for virulence. Infect. Immun. **69:**1322–1328.
53. **Sherlock, O., M. A. Schembri, A. Reisner, and P. Klemm.** 2004. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. J. Bacteriol. **186:**8058–8065.
54. **Svanborg-Eden, C., L. Hagberg, R. Hull, S. Hull, K. E. Magnusson, and L. Ohman.** 1987. Bacterial virulence versus host resistance in the urinary tracts of mice. Infect. Immun. **55:**1224–1232.
55. **Tree, J. J., G. C. Ulett, C. L. Ong, D. J. Trott, A. G. McEwan, and M. A. Schembri.** 2008. Trade-off between iron uptake and protection against oxidative stress: deletion of cueO promotes uropathogenic Escherichia coli virulence in a mouse model of urinary tract infection. J. Bacteriol. **190:**6909–6912.
56. **Ulett, G. C., A. N. Mabbett, K. C. Fung, R. I. Webb, and M. A. Schembri.** 2007. The role of F9 fimbriae of uropathogenic Escherichia coli in biofilm formation. Microbiology **153:**2321–2331.
57. **Ulett, G. C., J. Valle, C. Beloin, O. Sherlock, J. M. Ghigo, and M. A. Schembri.** 2007. Functional analysis of antigen 43 in uropathogenic Escherichia coli reveals a role in long-term persistence in the urinary tract. Infect. Immun. **75:**3233–3244.
58. **Ulett, G. C., R. I. Webb, and M. A. Schembri.** 2006. Antigen-43-mediated autoaggregation impairs motility in *Escherichia coli*. Microbiology **152:**2101–2110.
59. **Valle, J., A. N. Mabbett, G. C. Ulett, A. Toledo-Arana, K. Wecker, M. Totsika, M. A. Schembri, J.-M. Ghigo, and C. Beloin.** 2008. UpaG, a new member of the trimeric autotransporter family of adhesins in uropathogenic Escherichia coli. J. Bacteriol. **190:**4147–4161.
60. **Wiles, T. J., R. R. Kulesus, and M. A. Mulvey.** 2008. Origins and virulence mechanisms of uropathogenic Escherichia coli. Exp. Mol. Pathol. **85:**11–19.
61. **Wren, B. W., R. R. Russell, and S. Tabaqchali.** 1991. Antigenic cross-reactivity and functional inhibition by antibodies to *Clostridium difficile* toxin A, *Streptococcus mutans* glucan-binding protein, and a synthetic peptide. Infect. Immun. **59:**3151–3155.
62. **Wu, X. R., T. T. Sun, and J. J. Medina.** 1996. In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: relation to urinary tract infections. Proc. Natl. Acad. Sci. U. S. A. **93:**9630–9635.
63. **Wullt, B., G. Bergsten, H. Connell, P. Rollano, N. Gebretsadik, R. Hull, and C. Svanborg.** 2000. P fimbriae enhance the early establishment of Escherichia coli in the human urinary tract. Mol. Microbiol. **38:**456–464.
64. **Yoder, M. D., N. T. Keen, and F. Jurnak.** 1993. New domain motif: the structure of pectate lyase C, a secreted plant virulence factor. Science **260:**1503–1507.
65. **Zhang, Y.** 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics **9:**40.

Editor: J. B. Bliska