JOURNAL OF BACTERIOLOGY, June 2011, p. 2981–2988 Vol. 193, No. 12
0021-9193/11/$12.00 doi:10.1128/JB.00033-11

A Unique Arabinose 5-Phosphate Isomerase Found within a Genomic Island Associated with the Uropathogenicity of *Escherichia coli* CFT073▽

Joshua A. Mosberg,[1]† Alejandra Yep,[1] Timothy C. Meredith,[1]‡ Sara Smith,[2] Pan-Fen Wang,[1] Tod P. Holler,[1] Harry L. T. Mobley,[2] and Ronald W. Woodard[1]*

Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1065,[1] and Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan 48109[2]

Received 7 January 2011/Accepted 31 March 2011

Previous studies showed that deletion of genes c3405 to c3410 from PAI-*metV*, a genomic island from *Escherichia coli* CFT073, results in a strain that fails to compete with wild-type CFT073 after a transurethral cochallenge in mice and is deficient in the ability to independently colonize the mouse kidney. Our analysis of c3405 to c3410 suggests that these genes constitute an operon with a role in the internalization and utilization of an unknown carbohydrate. This operon is not found in *E. coli* K-12 but is present in a small number of pathogenic *E. coli* and *Shigella boydii* strains. One of the genes, c3406, encodes a protein with significant homology to the sugar isomerase domain of arabinose 5-phosphate isomerases but lacking the tandem cystathionine beta-synthase domains found in the other arabinose 5-phosphate isomerases of *E. coli*. We prepared recombinant c3406 protein, found it to possess arabinose 5-phosphate isomerase activity, and characterized this activity in detail. We also constructed a c3406 deletion mutant of *E. coli* CFT073 and demonstrated that this deletion mutant was still able to compete with wild-type CFT073 in a transurethral cochallenge in mice and could colonize the mouse kidney. These results demonstrate that the presence of c3406 is not essential for a pathogenic phenotype.

Escherichia coli CFT073 is one of the most well-studied uropathogenic *E. coli* strains. An analysis of the transcriptome of CFT073 that compared growth in the mouse urinary tract to growth in rich medium (22) identified many genes upregulated during urinary tract infection (UTI). Highly upregulated genes included those encoding proteins from iron acquisition systems; those encoding proteins involved in capsular polysaccharide biosynthesis, lipopolysaccharide (LPS) biosynthesis, drug resistance, and microcin secretion; and those encoding subunits of type 1 fimbriae involved in adhesion. A significant number of genes of unknown function were also found to be upregulated during UTI. Subsequent analysis of the genome of CFT073 revealed 13 large genomic islands ranging in size from 32 to 123 kb (12). Three genomic island deletion mutants failed to compete with wild-type CFT073 following a transurethral cochallenge in mice, including a PAI-*metV* mutant. The deletion of PAI-*metV* also resulted in a statistically significant attenuation of CFT073's ability to independently colonize the mouse kidney. This attenuation was further localized to the cluster of genes including c3405 to c3410 (12).

Analysis of this gene cluster reveals that one of the genes, c3406, encodes a protein with significant homology to the sugar isomerase (SIS) domain of arabinose 5-phosphate (A5P) isomerases (see below). A5P isomerases catalyze the interconversion of D-ribulose 5-phosphate (Ru5P), a product of the pentose phosphate pathway, and A5P, an important intermediate in LPS and capsular polysaccharide biosynthesis (25). The genome of *E. coli* K-12 contains two genes encoding A5P isomerases. The *kdsD* gene, which encodes an A5P isomerase that catalyzes the committed step in 3-deoxy-D-*manno*-oculosonate (Kdo) biosynthesis (14), is found in the *yrb* gene cluster (*yrbABCDEFGHIK*; *kdsD* is also known as *yrbH*). The *gutQ* gene is found in the glucitol operon (*gutAEBDMRQ*) and encodes an A5P isomerase (15) that is involved in the metabolism of sorbitol. A third A5P isomerase is found in *E. coli* strains that express group 2 K antigens. These extraintestinal pathogens harbor the *kps* gene cluster (*kpsFEDUCS*), which expresses an A5P isomerase from the *kpsF* gene (13). Since *E. coli* CFT073 is an extraintestinal pathogen that expresses group 2 K antigens, it contains all three of these A5P isomerases in addition to c3406.

All three of these A5P isomerases, KdsD, GutQ, and KpsF, are composed of three protein domains: one SIS domain and two cystathionine β-synthetase (CBS) domains. SIS domains (2) are commonly found in aldose-ketose isomerases and are believed to be the portion of the protein responsible for catalysis. CBS domains are found in a diverse array of organisms, as part of proteins with widely varied functions (3), but their function in A5P isomerases is unknown. Since c3406 lacks these CBS domains (see below), confirmation of its A5P isomerase activity would make it unique among A5P isomerases. Intrigued by this and curious about the role this open reading frame might play in the pathogenesis of *E. coli* CFT073, we have characterized the A5P isomerase activity of isolated c3406 protein and examined whether c3406 is required for pathogenesis.

* Corresponding author. Mailing address: Department of Medicinal Chemistry, University of Michigan, Ann Arbor, MI 48109-1065. Phone: (734) 764-7366. Fax: (734) 647-8430. E-mail: rww@umich.edu.
† Present address: Department of Genetics, Harvard Medical School, Boston, MA 02115.
‡ Present address: Novartis Institute for Biomedical Research, Cambridge, MA 02139.
▽ Published ahead of print on 15 April 2011.

TABLE 1. Strains, plasmids, and primers used in this study

Strain, plasmid, or primer	Relevant characteristic(s)	Source or reference
E. coli		
BL21-DE3	Used for protein expression	Novagen
BW30270	*E. coli* K-12 MG1655 *rph*$^+$ *fnr*$^+$	CGSC[a]
CFT073	Blood isolate from a patient with acute pyelonephritis	16
TCM15	BW30270 (Δ*kdsD* Δ*gutQ*)	15
Δc3406	CFT073 (Δc3406) Kan^r	This study
XL1-Blue	Used for recombinant DNA methods	Stratagene
Plasmids		
pT7-7	Expression vector	23
pT7kdsD	*E. coli* K-12 *yrbH* inserted into NdeI/BamHI sites of pT7-7; Amp^r	14
pT7c3406	*E. coli* CFT073 c3406 inserted into NdeI/BamHI sites of pT7-7; Amp^r	This study
Primers		
c3406F	GGTGCTAGAATTCATATGAATAACACGGATCTTATCC	Invitrogen
c3406R	GAATTCGGATCCAAGTTATGAATTTGCTTTAATGCCG	Invitrogen
Δc3406F	GATCATGGATCCATGAATAACACGGATCTTATCCATCTCATTAAACATTTTATGCATAACGAGAC CGGTCAATTGGCTGGAG	Invitrogen
Δc3406R	GATCATCATATGTTATGAATTTGCTTTAATGCCGAGTGCACCTCCTGGATGATATAAGCCGAAAT ATCCTCCTTAGTTCC	Invitrogen

[a] CGSC, *E. coli* Genetic Stock Center (strain 7925).

MATERIALS AND METHODS

Materials. Thermal cycling was performed using primers synthesized by Invitrogen (Carlsbad, CA), the Failsafe PCR PreMix Selection kit from Epicentre Biotechnologies (Madison, WI), and an MJ Research PTC-200 Peltier thermal cycler. Enzymes required for cloning were purchased from New England BioLabs (Ipswich, MA). Plasmid DNA purification was performed with the Wizard kit from Promega (Madison, WI). DNA sequencing was performed by the University of Michigan Biomedical Resources Core Facility. Reagents, buffers, and biochemicals were purchased from Sigma-Aldrich (St, Louis, MO), Research Organics (Cleveland, OH), or Fisher Scientific (Pittsburgh, PA), except for the metal salts, which were "Puratronic grade" from Alfa Aesar (Ward Hill, MA).

Bacterial strains, plasmids, primers, and growth media. The bacterial strains, plasmids, and primers used in this study are described in Table 1. The genome of *E. coli* CFT073, a strain isolated from the blood of a patient admitted to the University of Maryland Medical System for the treatment of acute pyelonephritis (16), has been sequenced and annotated (26). *E. coli* TCM15 is a derivative of *E. coli* BW30270 in which the *kdsD* and *gutQ* genes were replaced with antibiotic resistance cassettes using the Lambda Red recombinase system (15). Strains were grown in M9 minimal medium (21), morpholinepropanesulfonic acid (MOPS) minimal medium (17), 2× YT medium, or LB medium (21). TCM15 cultures were additionally supplemented with glucose 6-phosphate (G6P; 10 μM) and A5P (15 to 50 μM).

Cloning, overexpression, and purification of c3406. The c3406 gene was amplified from *E. coli* CFT073 cells using a standard whole-cell PCR technique (9) and employing primers c3406F and c3406R (Table 1). The gel-purified PCR product was digested with NdeI and BamHI and then ligated into a similarly restricted, phosphatase treated pT7-7 expression vector (23) to form plasmid pT7c3406. The ligation mixture was used to transform chemically competent *E. coli* XL1-Blue cells; a transformant with the correct pT7c3406 plasmid was identified by restriction analysis and confirmed by DNA sequencing. Isolated pT7c3406 plasmid DNA was then used to transform chemically competent *E. coli* BL21(DE3) cells for protein expression studies. A fresh single colony was inoculated into 1 liter of 2× YT medium containing 100 mg/liter ampicillin and grown at 37°C with shaking at 200 rpm until the optical density at 600 nm (OD_{600}) reached approximately 0.9. The culture was then cooled to 22°C and induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM. After an additional 16 h of growth at 22°C, cells were harvested by centrifugation (10 min, 6,000 × *g*), resuspended in 60 ml of buffer A (20 mM Tris-HCl, pH 8.0, 1 mM dithiothreitol), sonicated, and clarified by centrifugation (30 min, 18,000 × *g*). The cell extract was loaded onto a MonoQ anion-exchange column (Amersham Biosciences) and eluted with a linear gradient of NaCl (0 to 600 mM) in buffer A over 2 h. Fractions containing A5P isomerase activity were pooled and precipitated with 15% ammonium sulfate (saturated solution in buffer A). After clarification by centrifugation (45 min, 18,000 × *g*), the pellet was resuspended in buffer A and dialyzed against buffer A. The protein solution was concentrated to 5 mg/ml, filtered through a 0.22-μm membrane, aliquoted, and frozen at −80°C. The c3406 protein produced was estimated to be 95% pure using SDS-PAGE analysis.

Molecular weight determinations. The subunit mass of c3406 was determined using matrix-assisted laser-desorption ionization mass spectrometry, performed by the University of Michigan Protein Structure Facility on a VESTEC-2000 instrument using a sinapinic acid matrix. The native molecular mass of c3406 was estimated by native gel electrophoresis (7). Samples of c3406 protein were run on native gels of 8, 10, 12, and 14% total acrylamide (%T) with an acrylamide/bisacrylamide ratio of 29:1. Standards included bovine serum albumin (BSA) dimer (132.4 kDa), BSA monomer (66.2 kDa), chicken egg white ovalbumin (45 kDa), and soybean trypsin inhibitor (21.5 kDa). Relative mobility (R_f) was determined for each protein in each gel. Log R_f versus %T was plotted for each of the proteins and fitted by linear regression, using SigmaPlot 9.0, to yield the retardation coefficient (K_r). A standard curve was obtained by plotting the −log K_r of the standards versus their log molecular weight and fitting with linear regression. The native molecular weight of c3406 was estimated by determining its K_r as described above and then matching the −log K_r of c3406 with the standard curve.

Metal content analysis. Untreated enzyme samples were prepared for metal content analysis by 48 h of dialysis at 4°C against 2 liters of metal-free buffer containing 20 mM Tris · HCl, pH 7.5. EDTA-treated enzyme samples were prepared by incubating c3406 protein in the presence of 10 mM EDTA for 2 h at 4°C, followed by desalting with an HR 10/10 fast desalting column (Amersham Biosciences) eluted with 20 mM Tris · HCl, pH 7.5. The divalent metal contents of both the untreated and EDTA-treated samples were determined using high-resolution inductively coupled plasma mass spectrometry on a Finnigan MAT Element instrument at the University of Michigan Department of Geology.

Assay of A5P isomerase activity. The discontinuous cysteine-carbazole colorimetric assay (6) adapted to a 96-well microplate format (14) was utilized to measure A5P isomerase activity in both catalytic directions. Briefly, a 96-well microplate containing 25 μl of a solution of c3406 protein in buffer B {100 mM 1,2-bis[Tris(hydroxymethyl)-methylamino]propane (BTP), pH 6.6, 2 mM EDTA} in each assay well was incubated at 37°C for 2 min using a Peltier thermal cycler. Each reaction was initiated by the addition of 25 μl of various concentrations of an A5P or Ru5P solution in buffer B at 37°C. After each reaction proceeded at 37°C for its specified time, it was quenched by the addition of 50 μl of 12.5 M H_2SO_4. Finally, the reaction mixtures were transferred to a flat-bottom 96-well assay plate containing 250 μl of cysteine-carbazole developing solution per well and incubated at room temperature for 1 h (A5P synthesis direction) or 3 h (Ru5P synthesis direction) before recording absorbances at 540 nm. All plates contained internal Ru5P standards and appropriate controls in triplicate.

Effects of divalent metals and pH on the A5P isomerase activity of c3406. To determine the effects of various divalent metals on enzyme activity, samples of c3406 protein diluted in buffer containing 100 mM BTP, pH 6.6, were incubated for 30 min at 4°C with 20 μM either a divalent metal salt or EDTA. Activity was then assessed as outlined above (final concentrations: 300 nM c3406, 50 mM BTP, 10 mM A5P, and 10 μM metal salt or EDTA) in triplicate with a 3-min reaction time and a 3-h color development time. The metal salts tested included $MgCl_2$, $CaCl_2$, $BaSO_4$, $MnCl_2$, $FeSO_4$, $CoCl_2$, $NiCl_2$, $CuSO_4$, $ZnSO_4$, $CdCl_2$, and $HgCl_2$.

To determine the effect of pH on enzyme activity, samples of c3406 were diluted in 200 mM buffer solutions that contained 2 mM EDTA and were adjusted (at 37°C) to pH levels between 5.0 and 8.0 in increments of 0.25 pH unit. The buffer 2-(*N*-morpholino)ethanesulfonic acid (MES) was used from pH 5.0 to pH 5.75, and the buffer BTP was used from pH 6.0 to pH 8.0. Activity was then assessed in triplicate using the standard assay conditions with a 3-min reaction time and a 3-h color development time. Final assay concentrations were 400 nM c3406, 100 mM buffer, 10 mM A5P, and 1 mM EDTA.

Substrate specificity. In order to determine whether c3406 protein isomerizes various other sugars and sugar phosphates, 25-μl aliquots of c3406 diluted in buffer B were assayed with 25 μl of 20 mM solutions containing one of the potential alternative aldose substrates D-ribose 5-phosphate, D-arabinose, D-G6P, D-mannose 6-phosphate, or D-glucosamine 6-phosphate. The final assay concentrations were 300 nM c3406, 50 mM BTP, 10 mM substrate, and 1 mM EDTA. Each reaction, performed in triplicate, was allowed to proceed for 10 min at 37°C, quenched, and then developed as previously described. Controls lacking enzyme, also in triplicate, were performed for each potential substrate. The colorimetric cysteine-carbazole assay detects both pentuloses and hexuloses (6) with the lower limit of detection estimated to be less than 1% of ketose formation.

Determination of kinetic parameters. Individual conditions were assayed in triplicate using the A5P isomerase activity assay described above. The enzyme c3406 was diluted in buffer B such that the final assay concentration was 300 nM. A reaction time of 2 min was used, ensuring that the reaction was quenched when less than 10% of the substrate had been consumed. Substrate concentrations ranged from approximately 0.1 to 5 K_m. Assay plates, including the appropriate controls in triplicate, were developed for 1 or 3 h when using Ru5P or A5P as the substrate, respectively. Values of K_m and k_{cat} were determined by fitting the averaged assay results (initial rates) to the Michaelis-Menten equation using nonlinear least-squares regression.

Equilibrium constant (K_{eq}) determination. Solutions containing 3 μM c3406 protein in 50 mM BTP (pH 6.6, adjusted at 37°C), 1 mM EDTA, and 10% (vol/vol) D_2O were incubated with either A5P or Ru5P at a 5 mM final concentration. Reactions were allowed to proceed at 37°C and were monitored periodically by ^{31}P nuclear magnetic resonance (NMR) analysis using a Bruker Avance DRX-300 instrument with WALTZ16 proton decoupling. Once both reactions demonstrated stable peak ratios, the spectra were acquired (64 scans). A 10-s delay time, which is greater than three times the T1 relaxation parameter for A5P and Ru5P, was used during the acquisition to ensure the complete relaxation of the phosphorus nucleus, thereby allowing the direct comparison of integrated peaks. K_{eq} is reported in the direction of Ru5P product formation from the A5P substrate ([Ru5P]/[A5P]).

Complementation of TCM15 with pT7c3406. Electrocompetent TCM15 cells were prepared by growing the cells in LB medium until early log phase (A_{600}, ~0.5), spinning them down, and then washing them three times in ice-cold 10% glycerol. Electrocompetent TCM15 cells were transformed with pT7c3406 and then plated on LB agar supplemented with 100 μg/ml ampicillin. A fresh transformant was grown overnight at 37°C in MOPS minimal medium (21) supplemented with 100 μg/ml ampicillin, 0.2% glycerol, 15 μM A5P, and 10 μM G6P. The overnight culture was diluted 1:50 into fresh MOPS minimal medium supplemented with 100 μg/ml ampicillin and 0.2% glycerol. Under these conditions, expression of c3406 occurs as a result of the leaky nature of the T7 promoter in *E. coli*.

LPS analysis. The nature and amount of LPS from *E. coli* BW30270, TCM15, TCM15/pT7kdsD, and TCM15/pT7c3406 were determined by SDS-PAGE analysis. Overnight cultures were grown in MOPS minimal medium with 0.2% glycerol as the sole carbon source and with supplements as described in the section on bacterial strains, plasmids, primers, and growth media. Overnight cultures were diluted 1:20 in fresh medium and shaken for 6 h at 37°C. Samples of *E. coli* BW30270, TCM15, TCM15/pT7kdsD, and TCM15/pT7c3406 were prepared from equal numbers of cells, based upon OD_{600}. Aliquots of 1 to 5 ml cells were pelleted, washed twice with phosphate-buffered saline (PBS), resuspended in sample buffer (50 mM Tris-HCl, pH 6.5, 2% SDS, 4% 2-mercaptoethanol, 10% glycerol, 5 mg/ml proteinase K), and incubated for 1 h at 56°C. Just before loading, samples were heated for 5 min at 95°C. SDS-PAGE analysis was performed with a 5% stacking layer and a 15% running gel at 120 V for 1.5 h at ambient temperature and visualized by silver staining (8).

Induction of the sorbitol operon. Glucose-sorbitol diauxic growth experiments were performed as described before (15). Briefly, strains were grown in M9 minimal medium containing 2 mM glucose and 2 mM sorbitol (D-glucitol) as the dual carbon sources. Cell density was monitored as a function of time by measuring absorbance at 600 nm. Cells exhausted the supply of glucose before reaching stationary phase and either ceased growth (no complementation of the phenotype caused by *gutQ*) or resumed growth (complementation of the phenotype caused by *gutQ*) after a lag phase.

Chromosomal knockout of c3406. *E. coli* CFT073 Δc3406 was constructed from *E. coli* CFT073 using the Lambda Red recombinase system (11) and a kanamycin resistance cassette amplified with primers Δc3406F and Δc3406R (Table 1). This procedure replaced chromosomal c3406 with a kanamycin resistance cassette. *E. coli* CFT073 Δc3406 cells were grown in either LB medium supplemented with 50 μg/ml kanamycin or MOPS minimal medium supplemented with 50 μg/ml kanamycin and 0.2% glycerol.

Murine model of ascending UTI. A CBA/J mouse model of ascending UTI was used to assess the virulence of a mixture of *E. coli* CFT073 and the chromosomal knockout CFT073 Δc3406. Inocula were prepared as follows. Overnight cultures of the wild type and the deletion mutant were centrifuged at 4,000 × *g* for 10 min, resuspended in PBS to an OD_{600} of 4, mixed in a 1:1 ratio, and used to deliver a total of 2×10^8 CFU per mouse. To determine the input number of CFU/ml of each strain, dilutions of each inoculum were plated on LB agar plates, containing kanamycin when required, using an Autoplate 4000 (Spiral Biotech). Bacterial counts were determined using a Q-Count colony counter and accompanying software (Spiral Biotech). Female 6- to 8-week-old CBA/J mice (Jackson Laboratory, Bar Harbor, ME) were anesthetized with 100 mg ketamine and 10 mg xylazine per kg body weight and inoculated transurethrally with 50 μl of bacterial suspension per mouse. A sterile polyethylene catheter connected to an infusion pump was used to deliver the inocula over a 30-s period. At 48 h postinoculation, urine was collected and the mice were sacrificed. Their bladder and kidneys were aseptically removed, weighed, and homogenized in 3 ml sterile PBS. Homogenized tissue samples were plated onto LB plates with or without kanamycin (as required) to determine the output number of CFU/g of tissue for each strain. Wild-type bacterial counts were obtained by subtracting the number of CFU/g on the kanamycin-containing plates from the number of CFU/g on the plain LB plates. The lower limit of detection of this assay is 10^2 CFU/g of tissue. All animal protocols were approved by the University Committee on Use and Care of Animals at the University of Michigan Medical School.

RESULTS

The c3406 protein from *E. coli* CFT073 has A5P isomerase activity *in vitro*. A BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) of all completed genomes was conducted using genes c3405 to c3410 from *E. coli* strain CFT073. Genes c3405 to c3410 appear to behave as an operon, and close homologs of these genes are found in a small number of pathogenic *E. coli* and *Shigella boydii* strains (Fig. 1). The genes are annotated as a putative 2-hydroxyacid dehydrogenase (c3405), an SIS (c3406), a beta-cystathionase (c3407), a "PTS system, maltose and glucose-specific IIABC component" (c3408), and an antiterminator (c3409). The short open reading frame c3410 is listed simply as a "hypothetical protein". Having studied A5P isomerases in *E. coli* for many years, we were intrigued by gene c3406 because it encodes a protein with significant identity to the three known A5P isomerases from *E. coli* which are encoded by the *kdsD* (or *yrbH*), *gutQ*, and *kpsF* genes. An alignment of the c3406 protein with the three known A5P isomerases from *E. coli* CFT073 is shown in Fig. 2. The c3406 protein is much shorter than the other A5P isomerases. It consists solely of a SIS domain and lacks the tandem CBS domains found at the C termini of the other three *E. coli* A5P isomerases. When aligned individually with the SIS domains of the other three A5P isomerases, the c3406 protein shares 48.7%

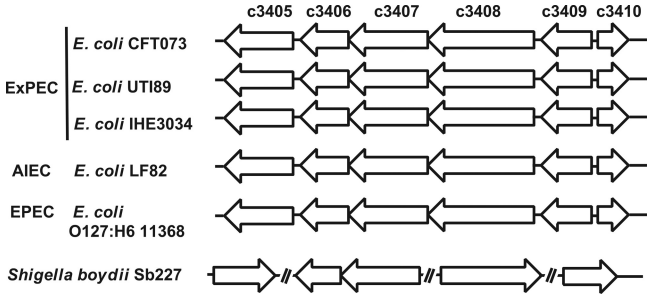

FIG. 1. Examples of the 16 species of *E. coli* or *Shigella* that have the c3405-to-c3410 operon. ExPEC, extraintestinal pathogenic *E. coli*. CFT073 and UTI89 cause UTI; IHE3034 is associated with neonatal meningitis. AIEC, adherent invasive *E. coli*. LF82 is associated with Crohn's disease. EPEC, enteropathogenic *E. coli*. Strain 0127:H6 11368 causes diarrhea; *S. boydii* causes dysentery.

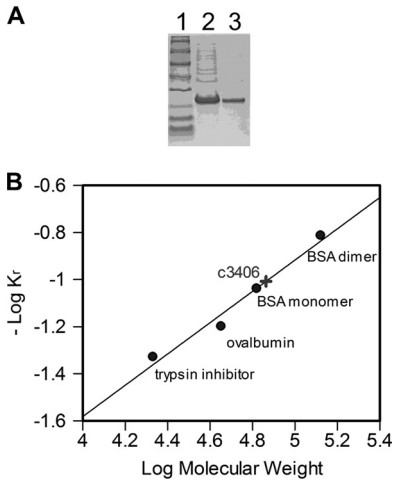

FIG. 3. c3406 molecular weight determination by PAGE. (A) SDS-PAGE of c3406. Lane 1: Bio-Rad molecular weight standards; lanes 2 and 3, recombinant c3406. (B) Standard curve (Ferguson plot) from native molecular weight determination by native PAGE.

identity with KdsD, 37.0% identity with GutQ, and 39.3% identity with KpsF, making the sequence of c3406 most similar to the SIS domain of KdsD. To confirm that the c3406 protein possesses A5P isomerase activity, it was overexpressed in *E. coli* and purified to substantial homogeneity. Preliminary experiments demonstrated that purified c3406 possessed A5P isomerase activity, so it was subjected to more careful physical and biochemical analysis.

Recombinant c3406 protein migrated at approximately 21 kDa on denaturing SDS-PAGE (Fig. 3A), as expected, and its identity was verified by mass spectrometry (Table 2). The observed subunit mass was within 0.04% of the expected subunit mass (20,888 Da). Quaternary structure was probed by native PAGE. The native c3406 protein migrated with an apparent molecular mass of 73 kDa (Fig. 3B), or 3.5 times its monomer molecular weight. Since *E. coli* KdsD (14) and KpsF (13) are tetramers, we think it reasonable to suggest that c3406 is also a tetramer in solution, although we cannot rule out a trimeric structure based upon these data alone.

A metal content analysis of the isolated c3406 protein revealed 0.72 equivalent of Zn per subunit and trace amounts of other metals. We tested the effects of several different metal ions on the activity of c3406 (Fig. 4). Incubation with divalent Zn, Cd, or Hg salt reduced its activity to less than 25% of the untreated rate, showing that these metals inhibit enzymatic activity. Addition of divalent Mg, Ca, Ba, Mn, Fe, Co, Ni, or Cu salt, on the other hand, increased its activity by about 2-fold, suggesting that these metals do not inhibit enzymatic activity to any significant degree but that they do displace the inhibitory Zn^{2+} ions present in c3406 before treatment. Consistent with this hypothesis, removal of all metal ions via EDTA treatment increased the activity of c3406 by approximately 2.5-fold. These results are essentially identical to those obtained with KdsD (14) and KpsF (13). They demonstrate that c3406, like other *E. coli* A5P isomerases, does not require a metal ion for catalysis.

A pH-rate profile revealed a narrow pH optimum at 6.6 (Fig. 5), with only trace A5P isomerase activity observed below pH 5.5 or above pH 7.5. A detailed kinetic analysis of the A5P

```
KdsD  MSHVELQPGFD------FQQAGKEVLAIERECLAELDQYIN-QNFTLACEKMFWCKGKVVVMGMGKSGHIGRKMAATFASTGTPS 78
GutQ  MSEALLN-------------AGRQTLMLELQEASHLPERLG-DDFVRAANIILHCEGKVVVSGIGKSGHIGKKIAATLASTGTPA 71
KpsF  MSERHLPDDQSSTIDPYLITSVRQTLAEQSAALQNLSKQLDSGQYQRVLNLIMNCKGHVILSGMGKSGHVGRKISATLASTGTPS 85
c3406 MNNTDLIHLIK------HFMH-NELKAVEEVIDSPLSEFAN-LIKVLQS-----CQGKVVFIGVGKSGIIARKLAATFASTGTPS 72
                    SIS DOMAIN

KdsD  FFVHPGEAAHGDLGMVTPQDVVIAISNSGESSEITALIPVLKRLHVPLICITGCPESSMARAADVHLCVKVAKEACPLGLAPTSS 165
GutQ  FFVHPAEALHGDLGMIESRDVMLFISYSGGAKELDLIIPRLEDKSIALLAMTGKPTSPLGLAAKAVLDISVEREACPMHLAPTSS 158
KpsF  FFIHPAEAFHGDLGMITPYDLLILISASGETDEILKLVPSLKNFGNRIIAITNNGNSTLAKNADAVLELHMANETCPNNLAPTTS 172
c3406 FFVHGTEAVHGDLGMVAKDDVVILISNSGETAEILATLPSLKKMGNYLISFTRSHHSSLAISCDLSVEIPVKSEADNLGLAPSCS 159
      SIS DOMAIN

KdsD  TTATLVMGDALAVALLKARGFTAEDFALSHPGGALGRKLLLRVNDIMHTGDEIPHVKKTASLRDALLEVTRKNLGMTVICDDNMM 263
GutQ  TVNTLMMGDALAMAVMQARGFNEEDFARSHPAGALGARLLNKVHHLMRRDDAIPQVALTASVMDAMLELSRTGLGLVAVCDDQRL 236
KpsF  TTLTMAIGDALAIAMIHQRKFMPNDFARYHPGGSLGRRLLTRVADVMQHD--VPAVQLDASFKTVIQRITSGCQGMVMVEDAEGG 268
c3406 STVVLVVGDAVALALSELKKFTRADFGLYHPGGALG----IKANS---------------------------------------- 198
      SIS DOMAIN   CBS DOMAINS

KdsD  IEGIFTDGDLRRVFDMGVDVRRLSIADVMTPGGIRVRPGILAVEALNLMQSRHITSVMVADG-DHLLGVLHMHDLLRAGVV  328
GutQ  VKGVFTDGDLRR-WLVGGGALTTPVNEAMTVGGTTLQSQSRAIDAKEILMKRKITAAPVVDENGKLTGAINLQDFYQAGII  285
KpsF  LAGIITDGDLRRFMEKEGSLTSATAAQMMTREPLTLPEDTMIIEAEEKMQKHRVSTLLVTNKANKVTGLVRIFD-------  298
c3406 ----------------------------------------------------------------------------------  198
      CBS DOMAINS
```

FIG. 2. Alignment of the four A5P isomerases from *E. coli* CFT073.

TABLE 2. Comparison of the biochemical properties of *E. coli* KdsD, GutQ, KpsF, and c3406

Protein	k_{cat} (A5P to Ru5P, s^{-1})	K_m (A5P, mM)	k_{cat}/K_m (A5P, M^{-1} s^{-1})	k_{cat} (Ru5P to A5P, s^{-1})	K_m (Ru5P, mM)	k_{cat}/K_m (Ru5P, M^{-1} s^{-1})	K_{eq}	Optimum pH	Subunit mass (Da)
KdsD[a]	157 ± 4[d]	0.61 ± 0.06	2.6×10^5	255 ± 16	0.35 ± 0.08	7.3×10^5	0.50 ± 0.06	8.4	35,084
GutQ[b]	218 ± 4	1.2 ± 0.1	1.8×10^5	242 ± 11	0.64 ± 0.08	3.8×10^5	0.47	8.25	33,909
KpsF[c]	15 ± 1	0.57 ± 0.04	2.6×10^4	19 ± 2	0.30 ± 0.03	6.3×10^4	0.48 ± 0.02	7.75	35,447
c3406	16.8 ± 0.2	1.92 ± 0.05	8.8×10^3	10.5 ± 0.8	0.70 ± 0.12	1.5×10^4	0.52	6.6	20,880

[a] Data from reference 14.
[b] Data from reference 15.
[c] Data from reference 13.
[d] Values are averages and standard deviations.

isomerase activity of c3406 at pH 6.6, in the presence of 2 mM EDTA, is presented in Table 2, alongside literature values for the KdsD and GutQ of *E. coli* K-12 and the KpsF from *E. coli* CFT073. The biochemical properties of c3406 are distinct from those of KdsD, GutQ, and KpsF but most closely resemble those of KpsF. The c3406 protein is the poorest enzyme among the four, with both low k_{cat} values and high substrate K_m values. Suspecting that there might be a relaxed substrate specificity, we tested c3406 for the ability to catalyze the isomerization of a series of potential alternate substrates, including D-ribose 5-phosphate, D-arabinose, D-G6P, D-mannose 6-phosphate, and D-glucosamine 6-phosphate. None of these alternative substrates was turned over, suggesting that the activity of c3406 is specific to the interconversion of Ru5P and A5P. Finally, the interconversion of Ru5P and A5P was allowed to reach equilibrium, after which the relative equilibrium concentrations of A5P and Ru5P were estimated by ^{31}P NMR analysis. The apparent Ru5P-A5P equilibrium established by c3406 was essentially the same as those established by the other *E. coli* A5P isomerases, as expected (Table 2).

The c3406 protein functions as an A5P isomerase *in vivo*. Deletion of either *kdsD* or *gutQ* produces a strain that is still able to survive on minimal medium. Deletion of both genes, however, produced a strain (TCM15) that requires both A5P (to support LPS production) and G6P (to induce the transport system for A5P) to survive and grow (15). To determine the ability of the c3406 gene to complement the double knockout, TCM15 cells were transformed with expression plasmid pT7-7 containing c3406 to produce strain TCM15/pT7c3406. A positive control containing *kdsD* (TCM15/pT7kdsD) and a negative control containing the empty vector (TCM15/pT7-7) were also made. These three strains were streaked onto agar plates containing either MOPS minimal medium and glycerol or MOPS minimal medium and glycerol supplemented with A5P and G6P (Fig. 6). Both TCM15/pT7c3406 and TCM15/pT7kdsD were able to grow in the MOPS minimal medium and glycerol alone, while TCM15/pT7-7 required MOPS minimal medium and glycerol supplemented with A5P and G6P to grow. Separate experiments demonstrated that TCM15/pT7c3406 could grow in glycerol-supplemented minimal medium at a rate comparable to that of either wild-type BW30270 or TCM15/pT7kdsD (data not shown). The LPS was extracted from BW30270, TCM15/pT7kdsD, and TCM15/pT7c3406 cultures grown on glycerol-supplemented minimal medium and analyzed by SDS-PAGE. Similar levels of LPS were produced by BW30270, TCM15/pT7kdsD, and TCM15/pT7c3406 (data not shown), suggesting that c3406 has A5P isomerase activity sufficient to support normal synthesis of LPS *in vivo*.

c3406 sustains activation of the sorbitol operon. The *gutQ* gene, which encodes the A5P isomerase GutQ, is part of an operon that allows *E. coli* to utilize sorbitol as a carbon source (20). Cells grown in medium containing both glucose and sorbitol preferentially utilize glucose but can continue to grow on sorbitol, after a short induction phase, after the glucose has been exhausted. Deletion of *gutQ* alone does not change this phenotype, nor does deletion of *kdsD* alone (15). If both *gutQ* and *kdsD* are deleted, however, the level of induction of the *gut* operon by sorbitol is sensitive to the concentration of A5P added to the growth medium (15). If a low concentration (5

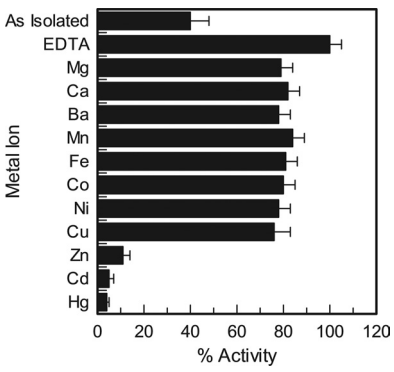

FIG. 4. Effects of divalent metals on the activity of c3406. The enzyme was incubated with no additive (As Isolated), EDTA, or a divalent metal ion as described in Materials and Methods and then assayed for activity. The activity of the EDTA-treated sample was assigned a value of 100%.

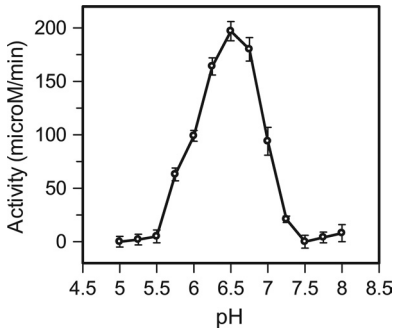

FIG. 5. pH-rate profile of c3406. The rate of A5P isomerization to Ru5P was measured at 37°C in a series of different pH environments. Experimental details can be found in Materials and Methods.

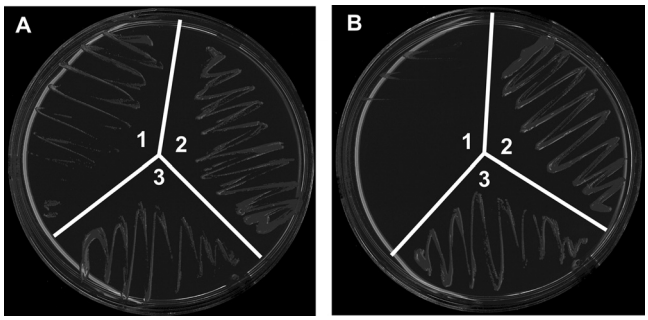

FIG. 6. Complementation of A5P auxotroph TCM15 on agar plates. (A) MOPS minimal medium plus 0.2% glycerol, 15 μM A5P, and 10 μM G6P. (B) MOPS minimal medium plus 0.2% glycerol. In both panels, the wedges contain TCM15/pT7-7 (wedge 1), TCM15/pT7KdsD (wedge 2), and TCM15/pT7c3406 (wedge 3).

μM) of A5P is added to the growth medium, the induction of the *gut* operon by sorbitol, measured as mRNA levels or GutD activity in cell lysates, is greatly diminished. To evaluate the ability of c3406 to induce the utilization of sorbitol, dual substrate growth curves were recorded for TCM15/pT7c3406, the positive control TCM15/pT7kdsD, and BW30270. Although the lag time for induction was reproducibly longer for TCM15/pT7c3406 cells than for TCM15/pT7kdsD or BW30270 cells, TCM15/pT7c3406 was able to resume growth on sorbitol after the glucose in the medium had been consumed (Fig. 7).

Deletion of c3406 does not inhibit the growth of *E. coli* CFT073 or its colonization of the mouse kidney and bladder. The gene c3406 is included in the genomic island PAI-*metV* within the genome of *E. coli* CFT073. Deletion of this genomic island significantly attenuated CFT073 colonization of the mouse bladder and kidney. Deletion of just the c3405-to-c3410 region of PAI-*metV* generated similar results, suggesting that this group of genes is essential for the pathogenic phenotype (12). We deleted the single gene c3406 from the genome of *E. coli* CFT073, which produced a strain (Δc3406, Table 1) that showed growth rates in minimal medium similar to those of the wild-type CFT073 strain (data not shown). We then tested the ability of strain Δc3406 to colonize the mouse urinary tract by cochallenge with wild-type CFT073. The results showed no significant differences between CFT073 and Δc3406 in either

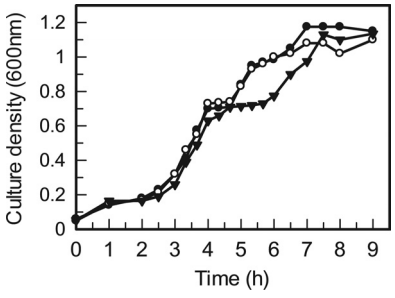

FIG. 7. Glucose-sorbitol diauxic growth experiments. Growth in minimal medium supplemented with glucose and sorbitol allowed growth to an OD_{600} of approximately 0.7, followed by a lag phase, after which growth on sorbitol allowed the OD_{600} to reach approximately 1.2. Closed circles, BW30270; open circles, TCM15/pT7kdsD; closed triangles, TCM15/pT7c3406.

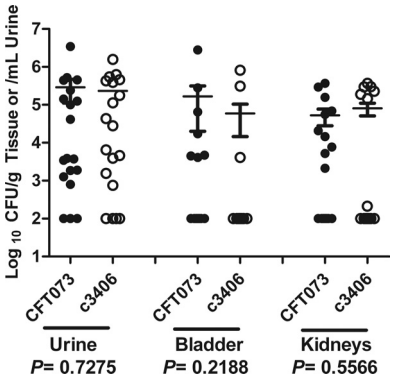

FIG. 8. Competitive mixed-infection experiment. CBA/J mice (n = 20) were challenged with a 1:1 mixture of wild-type CFT073 and CFT073 Δc3406. Data points represent numbers of bacterial CFU/g of tissue from the indicated organs of individual mice or numbers of CFU/ml of urine.

cell numbers in the urine or the ability to colonize the bladder and kidneys (Fig. 8).

DISCUSSION

The identification of a small operon associated with uropathogenicity in *E. coli* CFT073 (12) led to the identification of a group of similar operons in several pathogenic *E. coli* and *Shigella* strains (Fig. 1). Each of these operons contains (annotations are from the CFT073 genome, GenBank accession no. AE 014075) a putative 2-hydroxyacid dehydrogenase (c3405), a putative phosphosugar isomerase (c3406), a putative beta-cystathionase (c3407), a putative phosphotransferase (PTS) system component (c3408), and a putative antiterminator (c3409). The specific function of this operon is unknown, but the fact that it encodes a putative integral membrane protein (c3408) typically found in carbohydrate PTS systems (19) suggests that it may have a role in the internalization and utilization of a carbohydrate.

We decided to study c3406, the putative phosphosugar isomerase contained within this operon, because it has significant homology to the SIS domains of the *E. coli* A5P isomerases that we have studied (KdsD, GutQ, and KpsF) but is unique in that it lacks their C-terminal tandem CBS domains. We have shown that c3406 catalyzes the interconversion of A5P and ribulose 5-phosphate, supporting the annotation of c3406 as an A5P isomerase. Detailed biochemical characterization of c3406 shows (Table 2) that it has relatively high substrate K_m values, like those of GutQ, but relatively low k_{cat} values, like those of KpsF. The efficiency of c3406 as an A5P isomerase is low, since its k_{cat}/K_m value is 4-fold lower than that of the next most efficient A5P isomerase, KpsF, and 30- to 50-fold lower than that of the most efficient A5P isomerase, KdsD. The pH-rate optimum of c3406 is also at the extreme for an A5P isomerase—a full pH unit lower than its nearest biochemical neighbor, KpsF. The pH-activity profile of c3406 is also narrow, like that of KpsF. Taken together, these data demonstrate that c3406 is a unique A5P isomerase with biochemical properties most similar to those of KpsF.

Expressed from a leaky T7 promoter in *E. coli*, c3406 is capable of supporting LPS biosynthesis at a rate sufficient for

growth, a role normally ascribed to KdsD. It is also capable of supporting a shift from glucose utilization to sorbitol utilization, a role normally ascribed to GutQ. The ability of c3406 to participate in these pathways despite its lack of tandem CBS domains deepens the mystery surrounding the function of the CBS domains in A5P isomerases that contain them. CBS domains have been shown to bind adenine-containing nucleotides and serve as regulatory units in diverse contexts (10). Normally occurring in pairs, they affect the multimerization and sorting of proteins, ligand binding, and other properties, depending upon the protein to which they are fused (10). A negative regulatory role has been proposed for the tandem CBS domains of *Yersinia pestis* YrbH (24), which is a CBS domain-containing A5P isomerase, based upon an increase in biofilm-forming activity when these domains are removed.

The c3406 protein shares significant sequence identity (34%) with a putative phosphosugar isomerase from *Bacteroides fragilis* NCTC 9343. An X-ray crystal structure of this protein, which also lacks CBS domains, has been deposited in the Protein Data Bank (PDB; ID 3ETN), but it has not been described in the literature. *B. fragilis* is an anaerobic, Gram-negative organism. Its genome (4) encodes proteins sharing significant identity with the other enzymes of the Kdo biosynthesis and utilization pathway (KdsA, KdsC, KdsB, and WaaA) but lacks a gene encoding a typical CBS domain-containing A5P isomerase (a KdsD). Since *B. fragilis* makes LPS, albeit an atypical LPS (5), one might expect that this putative phosphosugar isomerase would have sufficient A5P isomerase activity to support LPS biosynthesis. Our complementation studies with c3406 support this expectation. It is interesting that the crystal structure of this putative A5P isomerase from *B. fragilis* contains a bound molecule of CMP-Kdo, with the CMP moiety protruding from the surface of the protein. It is not clear how this molecule came to be part of this structure, but the fact that the Kdo moiety binds to the isolated SIS domain suggests to us that the CBS domains of other A5P isomerases may fulfill a regulatory role by helping bind the cytidine portion of a molecule like CMP-Kdo. It seems reasonable to further suggest that the roles played by the CBS domains of KdsD, GutQ, and KpsF might be somewhat different from one another, much as the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthases AroF, AroG, and AroH, which coexist in *E. coli* strains and catalyze the same reaction, are regulated by phenylalanine, tyrosine, and tryptophan, respectively (1). Further research into the roles of the CBS domains of A5P isomerases is needed to address these questions.

The c3406 protein also shares significant sequence identity (31%) with a second putative phosphosugar isomerase, from *Listeria monocytogenes*, that has not been described in the literature but for which a crystal structure has been deposited (PDB ID 3FXA). *L. monocytogenes* is a Gram-positive facultative anaerobe. Its genome (Genome database entry NC_002973) (18) does not encode proteins with significant identity to the enzymes of Kdo biosynthesis. The putative phosphosugar isomerase in *L. monocytogenes* is sandwiched between two genes (LMOf2365_0530 and LMOf2365_0532) encoding PTS components. Our observation that c3406 is an A5P isomerase, the significant sequence identity between c3406 and this *L. monocytogenes* protein, and the fact that the genes encoding these proteins are both adjacent to PTS components suggest that this *L. monocytogenes* protein is also an A5P isomerase. The precise role of an A5P isomerase in the utilization of a specific exogenous sugar has not been established, although GutQ has been known for some time to be involved in the utilization of sorbitol (15). We cannot exclude the possibility that the greater lag time observed for c3406 in the glucose-sorbitol diauxic growth experiment is related to the absence of CBS domains in this protein. However, since the switch to sorbitol utilization can be supported by either a CBS domain-containing A5P isomerase (GutQ, KdsD) or a CBS domain-free A5P isomerase (c3406), our working hypothesis is that the switch to sorbitol utilization requires an intracellular concentration of A5P that is greater than the concentration needed to sustain LPS biosynthesis. In this model, the greater lag time associated with c3406 is related to its poorer efficiency as an A5P isomerase. Additional research into the role of A5P isomerization in these carbohydrate utilization systems is warranted.

Finally, we were curious about whether deletion of c3406 from the genome of *E. coli* CFT073 would attenuate this organism's pathogenicity. Our experiments have demonstrated that c3406 is not solely accountable for the phenotype observed when the entire operon is deleted. It is possible that the deletion of c3406 can be complemented by the presence of the other three A5P isomerases of *E. coli* CFT073. Alternatively, the pathogenic phenotype may result from an accumulation of factors that increase fitness in the pathogenic state, to which c3406 makes only a small contribution. The biological function of the c3405-to-c3410 operon and its role in pathogenesis remain to be fully elucidated.

ACKNOWLEDGMENTS

This work was supported by National Institute of Health grant AI-061531 (to R.W.W.).

We thank Ted Houston of the University of Michigan Department of Geology at the W. M. Keck Elemental Geochemistry Laboratory for measuring metal contents.

REFERENCES

1. **Ahmad, S., B. Rightmaier, and R. A. Jensen.** 1986. Evolution of the regulatory isozymes of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase present in the *Escherichia coli* genealogy. J. Bacteriol. **165:**146–154.
2. **Bateman, A.** 1999. The SIS domain: a phosphosugar-binding domain. Trends Biochem. Sci. **24:**94–95.
3. **Bateman, A.** 1997. The structure of a domain common to archaebacteria and the homocystinuria disease protein. Trends Biochem. Sci. **22:**12–13.
4. **Cerdeño-Tárraga, A. M., et al.** 2005. Extensive DNA inversions in the B. fragilis genome control variable gene expression. Science **307:**1463–1465.
5. **Delahooke, D. M., G. R. Barclay, and I. R. Poxton.** 1995. A re-appraisal of the biological activity of Bacteroides LPS. J. Med. Microbiol. **42:**102–112.
6. **Dische, Z., and E. Borenfreund.** 1951. A new spectrophotometric method for the detection and determination of keto sugars and trioses. J. Biol. Chem. **192:**583–587.
7. **Gallagher, S. R.** 1 May 2001, posting date. One-dimensional electrophoresis using nondenaturing conditions. Curr. Protocols Protein Sci. **Chapter 10:** Unit 10.3. doi:10.1002/0471140864.ps1003s00.
8. **Hitchcock, P. J., and T. M. Brown.** 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J. Bacteriol. **154:**269–277.
9. **Hofmann, M. A., and D. A. Brian.** 1991. Sequencing PCR DNA amplified directly from a bacterial colony. Biotechniques **11:**30–31.
10. **Ignoul, S., and J. Eggermont.** 2005. CBS domains: structure, function, and pathology in human proteins. Am. J. Physiol. Cell Physiol. **289:**C1369–C1378.
11. **Lee, D. J., et al.** 2009. Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. BMC Microbiol. **9:**252.
12. **Lloyd, A. L., T. A. Henderson, P. D. Vigil, and H. L. T. Mobley.** 2009. Genomic islands of uropathogenic *Escherichia coli* contribute to virulence. J. Bacteriol. **91:**3469–3481.

13. **Meredith, T. C., and R. W. Woodard.** 2006. Characterization of Escherichia coli D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. Biochem. J. **395:**427–432.
14. **Meredith, T. C., and R. W. Woodard.** 2003. Escherichia coli YrbH is a D-arabinose 5-phosphate isomerase. J. Biol. Chem. **278:**32771–32777.
15. **Meredith, T. C., and R. W. Woodard.** 2005. Identification of GutQ from *Escherichia coli* as a D-arabinose 5-phosphate isomerase. J. Bacteriol. **187:** 6936–6942.
16. **Mobley, H. L. T., et al.** 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.
17. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119:**736–747.
18. **Nelson, K. E., et al.** 2004. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen Listeria monocytogenes reveal new insights into the core genome components of this species. Nucleic Acids Res. **32:**2386–2395.
19. **Peterkofsky, A., G. Wang, and Y.-J. Seok.** 2006. Parallel PTS systems. Arch. Biochem. Biophys. **453:**101–107.
20. **Postma, P. W., J. W. Lengeler, and G. R. Jacobson.** 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. **57:**543–594.
21. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed., vol. III. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
22. **Snyder, J. A., et al.** 2004. Transcriptome of uropathogenic *Escherichia coli* during urinary tract infection. Infect. Immun. **72:**6373–6381.
23. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. U. S. A. **82:**1074–1078.
24. **Tan, L., and C. Darby.** 2006. Yersinia pestis YrbH is a multifunctional protein required for both 3-deoxy-D-manno-oct-2-ulosonic acid biosynthesis and biofilm formation. Mol. Microbiol. **61:**861–870.
25. **Tzeng, Y.-L., et al.** 2002. KpsF is the arabinose-5-phosphate isomerase required for 3-deoxy-D-manno-octulosonic acid biosynthesis and for both lipooligosaccharide assembly and capsular polysaccharide expression in Neisseria meningitidis. J. Biol. Chem. **277:**24103–24113.
26. **Welch, R. A., et al.** 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli* . Proc. Natl. Acad. Sci. USA **99:**17020–17024.