AEM
Journals.ASM.org

Molecular Characterization of the EhaG and UpaG Trimeric Autotransporter Proteins from Pathogenic *Escherichia coli*

Makrina Totsika,[a] Timothy J. Wells,[a] Christophe Beloin,[b,c] Jaione Valle,[b]* Luke P. Allsopp,[a] Nathan P. King,[a] Jean-Marc Ghigo,[b,c] and Mark A. Schembri[a]

Australian Infectious Diseases Research Centre, School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, QLD, Australia[a]; Institut Pasteur, Unité de Génétique des Biofilms, Département de Microbiologie, Paris, France[b]; and CNRS, URA2172, Paris, France[c]

Trimeric autotransporter proteins (TAAs) are important virulence factors of many Gram-negative bacterial pathogens. A common feature of most TAAs is the ability to mediate adherence to eukaryotic cells or extracellular matrix (ECM) proteins via a cell surface-exposed passenger domain. Here we describe the characterization of EhaG, a TAA identified from enterohemorrhagic *Escherichia coli* (EHEC) O157:H7. EhaG is a positional orthologue of the recently characterized UpaG TAA from uropathogenic *E. coli* (UPEC). Similarly to UpaG, EhaG localized at the bacterial cell surface and promoted cell aggregation, biofilm formation, and adherence to a range of ECM proteins. However, the two orthologues display differential cellular binding: EhaG mediates specific adhesion to colorectal epithelial cells while UpaG promotes specific binding to bladder epithelial cells. The EhaG and UpaG TAAs contain extensive sequence divergence in their respective passenger domains that could account for these differences. Indeed, sequence analyses of UpaG and EhaG homologues from several *E. coli* genomes revealed grouping of the proteins in clades almost exclusively represented by distinct *E. coli* pathotypes. The expression of EhaG (in EHEC) and UpaG (in UPEC) was also investigated and shown to be significantly enhanced in an *hns* isogenic mutant, suggesting that H-NS acts as a negative regulator of both TAAs. Thus, while the EhaG and UpaG TAAs contain some conserved binding and regulatory features, they also possess important differences that correlate with the distinct pathogenic lifestyles of EHEC and UPEC.

Trimeric autotransporter adhesins (TAAs) are a subgroup of AT proteins that form a stable trimer on the bacterial cell surface (13). TAAs have been identified in a wide range of Gram-negative bacteria and, where characterized, are universally associated with virulence (31). TAAs are defined by the presence of a short 70- to 100-amino-acid C-terminal membrane anchor domain encompassing four β-sheets that form a trimer to create a full-sized β-barrel pore (50). This pore facilitates the translocation of the passenger domain to the cell surface (41, 50). This feature of TAAs is different from conventional AT proteins, which possess a translocation domain composed of around 300 amino acids that encompass 12 aliphatic β-pleated sheets that together form a β-barrel pore (4, 36).

The passenger domain of TAAs requires trimerization for stability and adhesive activity (12). Modeling of TAAs such as YadA from *Yersinia enterocolitica*, Hia from *Haemophilus influenzae*, and BadA from *Bartonella henselae* has revealed three distinct regions within the passenger domain: an N-terminal head, a neck, and a stalk (35, 51, 60). The N-terminal head structure differs between TAA proteins and is primarily involved in the adhesion properties of the protein (35, 60). The YadA head structure contains single-stranded, left-handed β-helices, the interface of which is formed by periodically occurring, conserved sequence motifs (35). These motifs can be found in the predicted head structures of many TAAs (51).

The head is connected to the stalk by a short, highly conserved sequence (the neck) which functions as an adapter between the large globular head and the narrow stalk domain. The neck is thought to act like a "safety pin," holding the three monomers together, partly explaining the stability of trimeric proteins (35). The stalk domain is repetitive, fibrous, and highly divergent in length with its primary function being to extend the head domain away from the surface of the bacterium (31). It can, however, confer other functional properties such as serum resistance (30, 41). The variable number of repeats in the stalk domain can lead to major differences in the size of TAAs. For example, BadA is more than 3,000 amino acids in length (40) while YadA is only 422 amino acids (27). Large TAAs like BadA have intermittent neck domains throughout the long stalk structure (31). The domain buildup of TAA proteins has been demonstrated from the crystal structure of EibD, an immunoglobulin binding TAA protein from *Escherichia coli* (30, 41).

Recently, a TAA from the uropathogenic *E. coli* (UPEC) strain CFT073 was identified (UpaG) that mediates adhesion to human bladder epithelial cells (56). UpaG promotes cell aggregation and biofilm formation on abiotic surfaces by CFT073 and various other UPEC strains as well as binding to the extracellular matrix (ECM) proteins fibronectin and laminin (56). Prevalence studies indicated that *upaG* is frequently associated with extraintestinal *E. coli* (ExPEC) strains (56). UpaG has also been identified as a potential protective antigen in ExPEC (20).

Enterohemorrhagic *Escherichia coli* (EHEC) is a pathogenic subclass of diarrheagenic *E. coli* (DEC). Here, we identify a TAA from *E. coli* O157:H7, EhaG, which is a positional orthologue of

Received 30 August 2011 Accepted 17 January 2012
Published ahead of print 27 January 2012
Address correspondence to Mark A. Schembri, m.schembri@uq.edu.au.
* Present address: Laboratory of Microbial Biofilms, Instituto de Agrobiotecnologia, Universidad Pública de Navarra-CSIC-Gobierno de Navarra, Pamplona, Spain.
Supplemental material for this article may be found at http://aem.asm.org/.

doi:10.1128/AEM.06680-11

UpaG but contains significant sequence divergence within the passenger-encoding domain. Cloning and expression of the *ehaG* gene from *E. coli* EDL933 revealed that the EhaG TAA possesses different functional properties than those of UpaG. These functional properties correlate with the distinct tissue tropism of EHEC and UPEC pathogens. While UpaG and EhaG displayed different functional characteristics, their expression in UPEC and EHEC is regulated in a common fashion by H-NS.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The following *E. coli* strains were used in this study: BL21(DE3) (Stratagene), MG1655, and MS427 (MG1655*flu*) (39); OS56 (MG1655*flu gfp*$^+$) (49); UPEC strain CFT073 (33); EHEC strain EDL933 (38); and CFT073*upaG* (56). Cells were routinely grown at 28°C or 37°C on solid or in liquid lysogeny broth (LB) medium (6), supplemented with the appropriate antibiotics: kanamycin (Kan; 100 μg/ml), chloramphenicol (Cam; 30 μg/ml), and ampicillin (Amp; 100 μg/ml). For growth under defined conditions, M63B1 supplemented with 0.4% glucose ($M63B1_{Glu}$) medium was used as indicated previously (43).

DNA manipulations and genetic techniques. DNA techniques were performed as previously described (43). 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 done using the DNeasy blood and tissue kit (Qiagen). Oligonucleotides were purchased from Sigma (Australia or France). All PCRs requiring proofreading were performed with the Expand high-fidelity polymerase system (Roche) as described by the manufacturer. Amplified products were sequenced to ensure fidelity of the PCR. DNA sequencing was performed using the ABI BigDye v3.1 kit (ABI) by the Australian Equine Genetics Research Centre, University of Queensland, Brisbane, Australia. Prevalence studies for the *upaG* and *ehaG* genes used *Taq* DNA polymerase, as described by the manufacturer (New England BioLabs), with the primers 144 (5′-AATACCCAGAGCATTACTAACCTG) and 145 (5′-ACCTTGT AATTTGTAGACCCAA).

Construction of plasmids. The *ehaG* gene was amplified by PCR from EHEC strain EDL933 using specific primers designed from the available genome sequence (130, 5′-CGCGCTCGAGATAATAAGGAACATTAAT GAACAAAATATTTAAAG, and 131, 5′-CGCGCAAGCTTTTACCACT GAATACCGGCACCG). The PCR product was digested with XhoI (forward primer) and HindIII/EcoRI (reverse primer) and ligated to XhoI-HindIII/EcoRI-digested plasmid pBAD/*Myc*-HisA. The resultant plasmid (pOMS01) was then digested with EcoRI and ligated with a correspondingly digested kanamycin resistance-encoding gene cassette to give rise to plasmid pEhaG-kan (pOMS15). Resistance to kanamycin was required to facilitate transformation of this plasmid into the *flu*-negative, *gfp*-positive *E. coli* K-12 strain OS56. The *upaG* gene from UPEC strain CFT073 has been described previously (56). Both genes were cloned using the same strategy, with expression of *ehaG*/*upaG* under the control of the arabinose-inducible *araBAD* promoter (25). Neither gene was cloned as a fusion to the 6×His-encoding sequence of pBAD/*Myc*-HisA.

Construction of mutants. In order to mutate the *ehaG* gene in EDL933 and create a *lacZ* reporter transcriptional fusion in CFT073, we used homologous recombination mediated by λ-red recombinase and either a one-step PCR procedure with 50-bp homology arms for recombination or a three-step PCR procedure with 500-bp homology arms for recombination (8, 15, 16, 53). The primers used to disrupt the *ehaG* gene in EDL933 were 843 (5′-ACAGCTAAAGAGTCAACTGG), 844 (5′-CCA TGAGGCGGCGACGTATCC), 845 (5′-GAAGCAGCTCCAGCCTACA CTAATGATGCTCGCTATTCCTTG), and 846 (5′-CTAAGGAGGATAT TCATATGTGATCCATTAAGTTAGTGTGACTAAGG). The mutation was confirmed using primers 789 (5′-AGGACGCCCGCCATAAACTG) and 790 (5′-GGTTTAACGGTTGTGGACAAC) and subsequent sequencing. A *upaG-lacZ* reporter transcriptional fusion in CFT073Δ*lac* was constructed using the same approach but employing primers upaG.lacZzeo.L-5 (CAGCTTCTGCGCTTATATCAAGGAATAGAGAGCAT CAATAATGACCATGATTACGGATTC) and upaG.lacZzeo.L-3 (CATCAGGCAATGTGGCGTTTACCCATTTGTTAATGGATGATCAG TCCTGCTCCTCGGCCAC). Mutants were confirmed via PCR and sequencing using primers upaG.ext-5 (AGGAATTCATCCTATGAACC) and upaG.ext-3 (TTATCGTTCGAACTGCTACTGTC). CFT073Δ*lac upaG*::*lacZ-zeo* mutants were screened after mutagenesis with the suicide plasmid pSC189 carrying the kanamycin-resistant Mariner transposon described previously (9, 14). Sequencing of the transposon in both directions enabled identification of the site of insertion and, hence, the gene disrupted. Mutation of the *hns* gene in EDL933 and CFT073 was performed as previously described (1).

Biofilm assays. Biofilm formation on polystyrene surfaces was monitored by using 96-well microtiter plates (Iwaki) essentially as previously described (47). Briefly, cells were grown for 18 h in LB (containing 0.2% arabinose for induction of AT-encoding genes) 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 590 nm (OD_{590}). Flow chamber experiments were performed as previously described (28, 46). Briefly, biofilms were allowed to form on glass surfaces in a multichannel flow system that permitted continuous monitoring of their structural characteristics. Flow cells were inoculated with standardized cultures with an OD_{600} of 0.02 pregrown overnight in M9 medium containing arabinose and kanamycin. Biofilm development was monitored by confocal scanning laser microscopy at 18 h postinoculation. For analysis of flow cell biofilms, z-stacks were analyzed using the COMSTAT software program (26).

Binding to ECM components. Bacterial binding to ECM components was performed in a microtiter plate enzyme-linked immunosorbent assay (ELISA) essentially as previously described (56). Microtiter plates (Maxisorp; Nunc) were coated overnight at 4°C with 2 μg of the following ECM proteins: collagen (types I to V), fibronectin, fibrinogen, laminin, elastin, heparan sulfate, human serum album, and bovine serum albumin (BSA) (Sigma-Aldrich) or the glycoproteins *N*-acetyl-D-galactosamine (NaGal), *N*-acetyl-D-glucosamine (NaGlu), or *N*-acetylneuraminic acid (NaNa). Wells were washed twice with TBS (137 mM NaCl, 10 mM Tris, pH 7.4) and then blocked with TBS-2% milk for 1 h. After being washed with TBS, 200 μl of washed and standardized (OD_{600} of 0.1) cultures of *E. coli* strain MS427pEhaG, MS427pUpaG, or MS427pBAD was added to 12 replicate wells per ECM component and the plates were incubated at 37°C for 2 h. After being washed to remove nonadherent bacteria, adherent cells were fixed with 4% paraformaldehyde (PFA), washed, incubated for 1 h with anti-*E. coli* serum (Meridian Life Sciences Inc.; catalogue no. B65001R) diluted 1:500 in 0.05% TBS-Tween–0.2% skim milk, washed, and incubated for 1 h with a secondary anti-rabbit antibody conjugated with horseradish peroxidase (diluted 1:1,000) (Sigma-Aldrich; catalogue no. A6154). After a final wash, adherent bacteria were detected by adding 150 μl of 0.3 mg/ml ABTS [2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (Sigma-Aldrich) in 0.1 M citric acid, pH 4.3, activated with 1 μl/ml 30% hydrogen peroxide] and the absorbance was measured at 405 nm.

Epithelial cell binding assays. T24 human bladder carcinoma cells and colorectal epithelial (Caco-2) cells were purchased from the American Type Culture Collection and cultured according to standard protocols. The adherence of MS427pEhaG and MS427pUpaG to cultured T24 or Caco-2 cells was examined essentially as previously described (7, 56). Briefly, wells were seeded with 1.5×10^5 epithelial cells in 24-well tissue culture plates and incubated at 37°C in 5% CO_2 until they were confluent. Overnight bacterial cultures were diluted and grown to an OD_{600} of 0.5 in LB with appropriate antibiotics (and 0.2% arabinose for *ehaG*/*upaG* induction). Bacterial cells were washed and added to triplicate epithelial cell monolayers at a multiplicity of infection of 10. Incubation was carried out for 1.5 h at 37°C in 5% CO_2. After incubation, monolayers were washed five times with phosphate-buffered saline (PBS) to remove nonadherent

bacteria. The remaining bacteria were then released by eukaryotic cell lysis with 0.1% Triton X-100. The number of adherent bacteria, as well as the inoculating dose, was determined by serial dilution and plating on LB agar. All experiments were performed in triplicate.

Coinfection of Caco-2 epithelial cell monolayers with fluorescently tagged MS427pEhaG and MS427pUpaG was carried out essentially as described above, but with the following modifications. Seeding of epithelial cells was performed in 4-chamber glass culture slides (BD Falcon). The two bacterial strains were mixed at a 1:1 ratio and incubated with the Caco-2 cell monolayers for 1.5 h at 37°C in 5% CO_2. Following incubation, monolayers were washed five times with PBS to remove nonadherent bacteria and the adherent bacteria were fixed with 4% paraformaldehyde (PFA). The PFA was then removed by washing with PBS, and adherent bacteria were visualized using fluorescence microscopy. Adherent bacteria were counted from 40 fields of view. In these experiments, the levels of EhaG and UpaG protein produced by MS427pEhaG and MS427pUpaG, respectively, were similar (data not shown).

Purification of 6×His-tagged EhaG′, antibody production, and immunoblotting. A 480-bp segment from the passenger-encoding domain of *ehaG* was amplified by PCR with primers 1707 (5′-TACTTCCAATCCAATGCAAACGCTATTGCGATAGGTGCTG) and 1708 (5′-TTATCCACTTCCAATGTTAAAGCTGTGAACCATTAATGG) from *E. coli* EDL933 genomic DNA; underlined nucleotides represent the ligation-independent cloning (LIC) overhangs used for insertion into plasmid pMCSG7 via LIC (17). The resultant plasmid (pEhaGTruncated) contained bp 652 to 1132 of *ehaG* fused to a 6×His-encoding sequence. *E. coli* BL21 was transformed with plasmid pEhaGTruncated and induced with isopropyl-β-D-thiogalactopyranoside (IPTG), and the resultant 6×His-tagged EhaG truncated protein (containing amino acids 218 to 378 of EhaG) was assessed by SDS-PAGE analysis as previously described (54). Polyclonal anti-EhaG serum was raised in rabbits by the Institute of Medical and Veterinary Sciences (South Australia). A polyclonal antiserum raised against UpaG has been described previously (56). 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). EhaG or UpaG rabbit polyclonal antiserum was used as primary antibody, 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.

β-Galactosidase assays. β-Galactosidase assays were performed essentially as previously described (32). Briefly, strains carrying *lacZ* fusions were grown on LB plates for 16 h and then inoculated into $M63B1_{Glu}$ minimal medium. After 16 to 18 h of growth, the culture was diluted in Z-buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 50 mM β-mercaptoethanol, 10 mM KCl, 1 mM $MgSO_4$, pH 7)–0.004% SDS, chloroform was added, and the samples were vortexed to permeabilize the cells. Samples were incubated at 28°C, and the reaction was initiated by the addition of *o*-nitrophenyl-β-D-galactopyranoside (ONPG). Reactions were stopped with the addition of sodium bicarbonate, and the enzymatic activity was assayed in quadruplicate for each strain by measuring the absorbance at 420 nm. Where required, β-galactosidase activity was also observed on LB agar plates containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal).

DNA curvature prediction and electrophoretic mobility shift assays. The *upaG* promoter region was analyzed *in silico* using Bendit, a program that enables the prediction of a curvature-propensity plot calculated with DNase I-based parameters (http://hydra.icgeb.trieste.it/dna/) (57). The curvature is calculated as a vector sum of dinucleotide geometries (roll, tilt, and twist angles) and expressed as degrees per helical turn (10.5°/helical turn = 1°/bp). Experimentally tested curved motifs produce curvature values of 5 to 25°/helical turn, whereas straight motifs give values below 5°/helical turn. The 250-bp *upaG* promoter region was amplified using primers upaG-250-5 (TTAGCAAAATGGCAGCAATT) and upaG+1-3 (TATTGATGCTCTCTATTCCT), and its intrinsic curvature was assessed by comparing its electrophoretic mobility with that of an unbent marker fragment (Promega 100-bp DNA ladder) on an 0.5% Tris-buffered EDTA (TBE)–7.5% PAGE gel at 4°C for retarded gel electrophoretic mobility.

Gel shift assays were performed essentially as previously described (5). A DNA mixture comprising the PCR-amplified *upaG* promoter region and TaqI-SspI-digested pBR322 at an equimolar ratio was incubated at room temperature for 15 min with increasing amounts of native purified H-NS protein (a gift from S. Rimsky) in 30 μl of reaction mixture containing 40 mM HEPES, pH 8, 60 mM potassium glutamate, 8 mM magnesium aspartate, 5 mM dithiothreitol, 10% glycerol, 0.1% octylphenoxypolyethoxyethanol, and 0.1 mg/ml BSA (H-NS binding buffer). DNA fragments and DNA-protein complexes were resolved by gel electrophoresis (0.5% TBE, 3% molecular screening agarose gel run at 50 V at 4°C) and visualized after staining with ethidium bromide.

RESULTS

Sequence variation in *E. coli* UpaG homologues correlates with strain pathotype. We used the translated amino acid sequence of *upaG* to probe all protein sequences encoded in 28 *E. coli* genomes available in the NCBI database (see Table S1 in the supplemental material). An intact gene encoding a putative TAA was found in 24 of the 28 genomes at the same genomic location as that of *upaG* in CFT073. Interestingly, three of the *E. coli* strains lacking the gene were commensal strains (ATCC 8739, K-12 MG1655, and W3110). Enteropathogenic *E. coli* (EPEC) strain E2348/69 contains a truncated gene due to a frameshift mutation. Multiple alignment of the 24 putative TAA protein sequences revealed that the signal sequence and translocation domain are highly conserved, while the passenger domain is highly variable in sequence and size (Fig. 1A). Phylogenetic analysis of the 24 translated full-length protein sequences revealed grouping according to strain pathotype, with sequences from DEC strains clading separately from those encoded in ExPEC strains (Fig. 1B). Proteins belonging to DEC strains shared high sequence identity (95%), while proteins encoded in ExPEC genomes shared 53% sequence identity, significantly higher than the overall 38% identity shared among the 24 *E. coli* TAA proteins examined.

***upaG* is highly prevalent among diarrheagenic *E. coli* strains.** We have previously demonstrated that *upaG* is commonly found in UPEC isolates (56). However, the primers used in our previous molecular screening were designed to amplify sequence from the passenger domain of *upaG* that is not conserved among the gene homologues found in the genome of DEC strains. In order to determine the prevalence of *upaG* homologues in DEC strains, we designed a new set of primers specific to a highly conserved region within the translocation domain and screened a collection of Shiga toxin-producing *E. coli* (STEC) isolates. We also screened a collection of UPEC isolates with the new set of primers. A product of the correct size was found in 92% (51/55) of STEC strains and 86.5% (64/74) of UPEC strains, indicating that *upaG* homologues are highly prevalent among pathogenic *E. coli* strains of intestinal as well as extraintestinal origin.

Comparative sequence analysis of UpaG and EhaG from prototypic strains UPEC CFT073 and EHEC EDL933, respectively. We hypothesized that the pathotype-specific grouping of UpaG homologues, due to the variable sequence of the function-encoding passenger domain, suggests that TAA proteins may have evolved specificity for their host environment. To test this tenet,

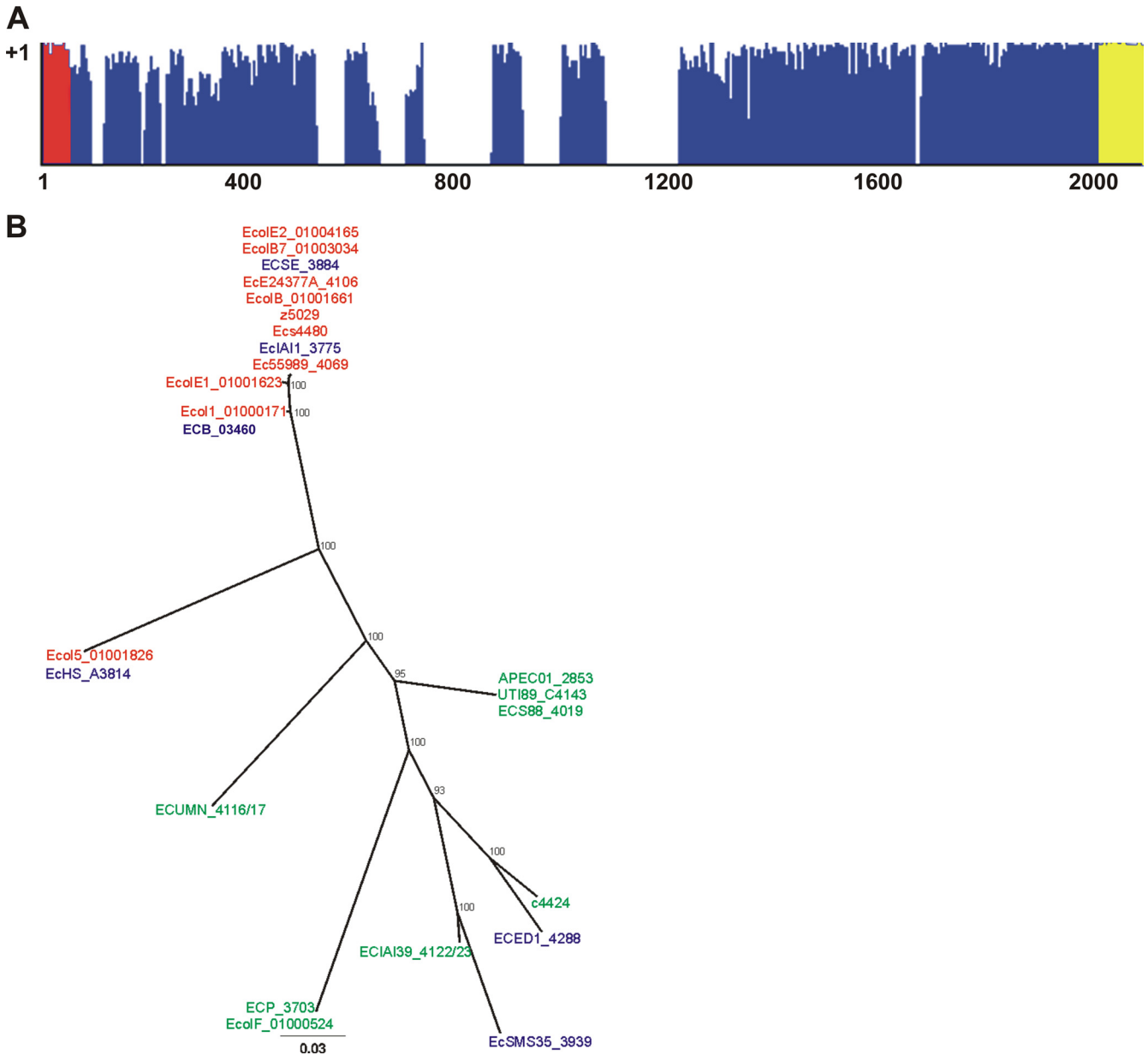

FIG 1 Sequence analysis of UpaG homologues from 24 *E. coli* genomes available on the NCBI database and listed in Table S1 in the supplemental material. (A) Similarity plot of 24 UpaG translated amino acid sequences aligned using Clustal W. High sequence conservation (indicated by shading on the *y* axis) was observed in the signal peptide (red) and translocator domain (yellow), while diversity in sequence and size was seen in the passenger domain (blue) of the proteins. (B) Unrooted phylogram of 24 UpaG translated amino acid sequences. Branch confidence levels are >90% and were determined from 1,000 bootstrap replicates of neighbor-joining trees calculated using PHYLIP (22). Taxon identifications represent the locus tags assigned to *upaG* orthologues in each *E. coli* genome sequence (see Table S1). Sequences from commensal and laboratory *E. coli* strains are indicated in blue, sequences from diarrheagenic *E. coli* (DEC) are indicated in red, and sequences from extraintestinal pathogenic *E. coli* (ExPEC) are indicated in green.

we selected one representative protein from the ExPEC group and one from the DEC clade to perform a detailed comparative sequence and functional analysis. UpaG from the prototypic UPEC strain CFT073 was chosen as the only functionally characterized TAA member from *E. coli* and was compared to its positional orthologue in the prototypic EHEC strain EDL933 (gene z5029). The z5029 gene was named *ehaG*, in a fashion consistent with previously characterized AT proteins from EHEC (58) while still acknowledging that it is a positional orthologue of *upaG* from UPEC. The *ehaG* gene (4,767 bp) encodes a protein similar to UpaG, containing an extended N-terminal signal sequence with predicted cleavage after amino acid 53, as well as an 89-amino-acid C-terminal translocation domain conserved among all TAA adhesins (Fig. 2A). Despite being positional orthologues, *upaG* and *ehaG* differ significantly in size and sequence (Fig. 2A). The predicted proteins are 1,778 and 1,588 amino acids, respectively, with the difference due to the presence of 190 additional amino acids in the passenger domain of UpaG (Fig. 2A). In fact, although their translocation domains are 100% identical, the two proteins share only 65% identity over the passenger domain. Sequence

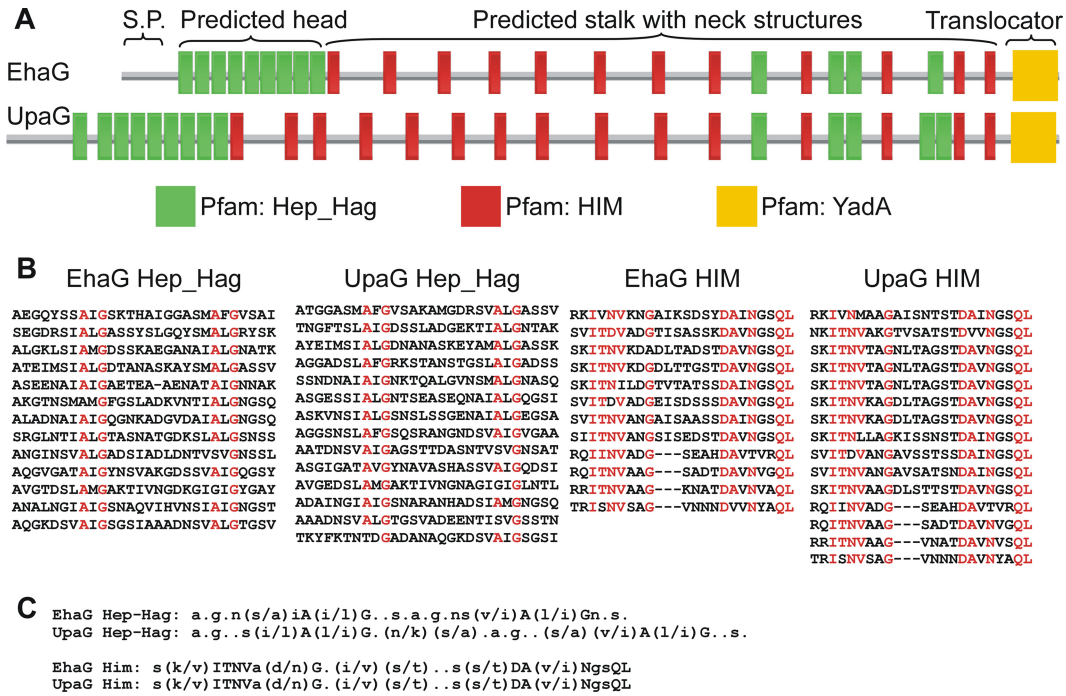

FIG 2 *In silico* analysis of UpaG and EhaG homologues. (A) Schematic illustration of the domain organization of UpaG and EhaG proteins. Indicated are the signal peptide (S.P.), translocator domain (Pfam: YadA), and the localization of the multiple invasin (Pfam: Hep_Hag) and hemagglutinin (Pfam: HIM) motifs within the passenger domain. (B) Multiple sequence alignment of the Hep_Hag repeats and HIM motifs of EhaG and UpaG. The domains were identified by sequence searches against the Pfam (23) and TIGRFAM databases hosted by the Comprehensive Microbial Resource (http://cmr.jcvi.org). Highly conserved residues are indicated in red. (C) Consensus sequence of the EhaG and UpaG Hep_Hag repeats and HIM motifs. Dots, no consensus; lowercase letters, some consensus; uppercase letters, high consensus.

analysis of the passenger domain of EhaG and UpaG against the Pfam and TIGRFAM databases showed that the proteins contain 13 and 14 Hep_Hag repeats, respectively. The majority of these 28-amino-acid-long repeats are found sequentially at the N-terminal end of the passenger domain (Fig. 2A). The head crystal structures of TAAs YadA and BadA are composed of such domains (35, 51), and it is likely that this region encodes the head of the EhaG and UpaG proteins (Fig. 2A). This is also supported by domain annotation for UpaG and EhaG proteins performed using the daTAA server (51) (data not shown). EhaG and UpaG also encompass 12 and 15 HIM motifs, respectively, dispersed along the passenger domain. This highly conserved 24-amino-acid motif is known to form the neck structure in YadA (35), a transition region between the globular head and the narrower stalk domain of the protein. Alignment of the multiple Hep_Hag and HIM sequences present in EhaG and UpaG (Fig. 2B) revealed that HIM sequences are well conserved between the two homologues, whereas the Hep_Hag sequences displayed some degree of variability (Fig. 2B and C).

EhaG from EHEC EDL933 mediates autoaggregation and biofilm formation. The significant sequence differences in the passenger domains of EhaG and UpaG prompted us to investigate whether EhaG possesses the same functional properties as does UpaG. For this purpose, the *ehaG* gene was amplified from the chromosome of EDL933 and cloned into pBAD/*Myc*-HisA under the control of the inducible *araBAD* promoter (25). The pEhaG vector was introduced into the previously described *E. coli flu* mutant strain MS427 (39), which is deficient in Ag43 expression and thus lacks the ability to autoaggregate and form a biofilm. Overexpression of EhaG in this background resulted in cell-cell aggregation from standing overnight cultures within 20 min, in contrast to the MS427pBAD control (Fig. 3A). We also tested the ability of EhaG to mediate biofilm formation in two distinct systems. In the polystyrene microtiter plate assay, MS427pEhaG formed a significant biofilm compared to the MS427pBAD control following induction with arabinose ($P < 0.001$; Fig. 3B). To compare the biofilm-forming capacities of EhaG and UpaG, we used the dynamic model of a continuous-flow chamber. Green fluorescent protein (GFP)-tagged MS427 cells (OS56) expressing *ehaG* from plasmid pEhaG or *upaG* from plasmid pUpaG were monitored for biofilm formation over 18 h using confocal laser scanning microscopy. In contrast to the OS56pBAD control, cells producing either EhaG or UpaG formed a strong biofilm across the entire surface of the chamber to a depth of approximately 15 μm (Fig. 3C).

EhaG and UpaG bind to ECM proteins. A common feature of TAAs is the ability to mediate binding to components of the ECM such as collagen and laminin. We therefore examined the ability of EhaG and UpaG to bind to a range of ECM proteins. The MS427pEhaG and MS427pUpaG strains both displayed binding to laminin, fibronectin, fibrinogen, and collagen types I, II, III, and V (Fig. 4). In contrast, no UpaG/EhaG-mediated binding was observed to type IV collagen, elastin, heparan sulfate, human serum albumin, bovine serum albumin, or the glycans *N*-acetyl-D-galactosamine (NaGal), *N*-acetyl-D-glucosamine (NaGlu), and *N*-acetylneuraminic acid (NaNa) (Fig. 4 and data not shown). Thus, EhaG and UpaG mediate binding to selected ECM proteins, and

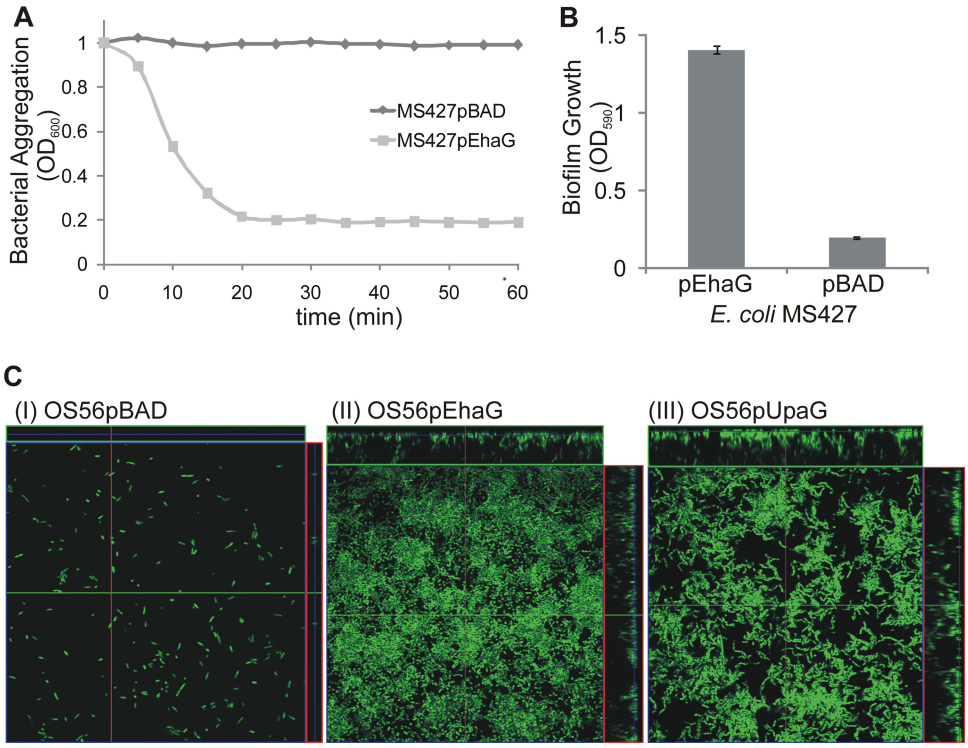

FIG 3 EhaG promotes autoaggregation and biofilm formation in *E. coli*. (A) Settling profile of liquid suspensions of *E. coli* strains MS427pBAD (vector control) and MS427pEhaG. Suspensions were prepared from overnight LB cultures supplemented with 0.2% arabinose normalized at an OD_{600} of 1. Bacterial autoaggregation is inversely proportional to the optical density of each suspension measured at 600 nm over a period of 60 min. (B) Biofilm formation by *E. coli* strains MS427pBAD (vector control) and MS427pEhaG following 18 h of culture in LB medium supplemented with 0.2% arabinose. Biofilm formation was examined in polystyrene 96-well microtiter plates using crystal violet staining. Bar charts show average absorbance measurements at 590 nm ± standard errors of the means from three independent experiments. (C) Fluorescent micrographs of continuous-flow biofilms formed in glass chambers 18 h after inoculation with GFP-labeled *E. coli* strains (I) OS56pBAD (vector control), (II) OS56pEhaG, and (III) OS56pUpaG. Micrographs show representative horizontal sections collected within each biofilm. Shown to the right and top of each individual panel are vertical sections representing the *yz* plane and the *xz* plane, respectively, at the positions indicated by the red and green lines.

the binding specificity to these ECM proteins is conserved between the two TAAs.

EhaG and UpaG mediate differential cellular adhesion. All TAAs characterized to date play a role in adhesion. We have previously demonstrated that UpaG enhances adhesion of CFT073 to T24 bladder epithelial cells (56). We therefore investigated if EhaG has a similar function; however, no increase was observed in the number of adherent bacteria recovered from T24 bladder epithelial cell monolayers infected with EhaG-producing MS427 cells compared to cells infected with MS427pBAD (Fig. 5A). Since EhaG is encoded by an enteric pathogen, we therefore tested its ability to mediate binding to intestinal epithelial cells. The number of adherent bacteria recovered after incubation of Caco-2 epithelial cell monolayers with MS427pEhaG was significantly higher than after incubation with the MS427pBAD control (Fig. 5A). The ability of EhaG to mediate bacterial adhesion to intestinal epithelial cells prompted us to investigate if UpaG can also mediate adhesion to Caco-2 cells. MS427pUpaG failed to adhere to Caco-2 epithelial cell monolayers at significantly higher numbers than did the MS427pBAD control (Fig. 5A). Upon coinfection of Caco-2 epithelial cell monolayers with a mixed (1:1) population of EhaG- and UpaG-producing MS427 cells that carried different fluorescent chromosomal markers (green for MS427pEhaG and red for MS427pUpaG), EhaG-producing cells adhered at 2.5-fold-greater numbers than did UpaG-producing cells (Fig. 5B and C). Reverse competition experiments using red MS427pEhaG and green MS427pUpaG cells produced a similar result (data not shown), ruling out a potential fluorophore-dependent bias in bacterial numbers.

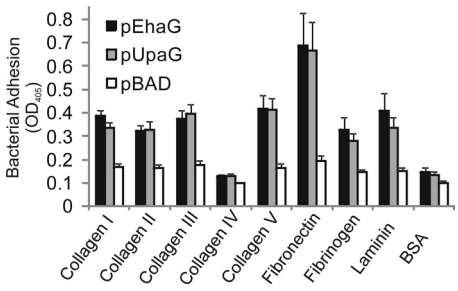

FIG 4 EhaG and UpaG mediate *E. coli* adherence to ECM proteins. ELISA-based assay demonstrating binding of *E. coli* MS427pEhaG (black bars), *E. coli* MS427pUpaG (gray bars), and *E. coli* MS427pBAD (white bars) to collagen (I to V), fibronectin, fibrinogen, and BSA. Results represent average absorbance readings at 405 nm + standard errors of the means from three independent experiments. The expression of EhaG and UpaG was induced with 0.2% arabinose.

Regulation of *upaG* expression by H-NS. We did not detect expression of UpaG or EhaG in protein samples prepared from *in*

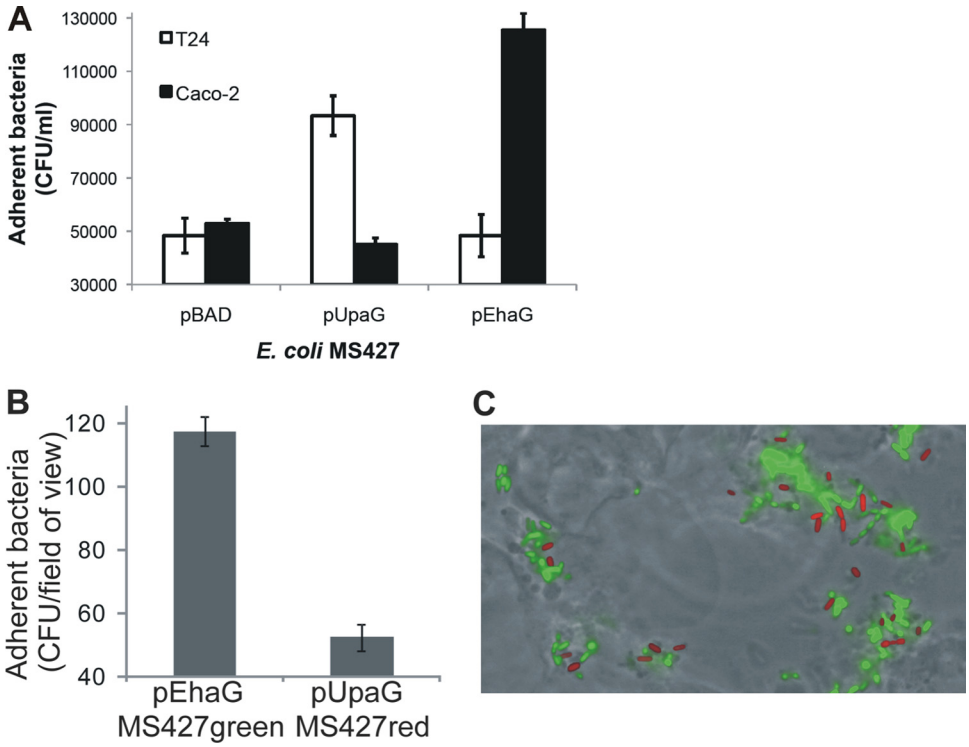

FIG 5 EhaG mediates *E. coli* adhesion to intestinal epithelial cells. (A) Efficiency of binding of *E. coli* MS427 containing pBAD, pEhaG, or pUpaG to T24 and Caco-2 epithelial cell monolayers. Bars represent the average numbers of adherent bacteria (CFU) per epithelial cell monolayer ± standard errors of the means. (B) Competitive adhesion to Caco-2 epithelial cell monolayers of mixed (1:1) bacterial inocula containing MS427green-pEhaG and MS427red-pUpaG. Bars represent average numbers of each bacterial strain (CFU) per field of view ± standard errors of the means. (C) Representative field of view of micrographs used to quantify the number of UpaG- and EhaG-producing adherent bacteria in the competitive Caco-2 epithelial cell adhesion assay. MS427pEhaG cells are tagged with GFP, and MS427pUpaG cells are tagged with red fluorescent protein. The expression of EhaG and UpaG was induced with 0.2% arabinose.

vitro-cultured UPEC CFT073 and EHEC EDL933, respectively, using Western blotting with specific polyclonal antiserum (data not shown) (56). To explore possible reasons for this lack of expression, we constructed a *lacZ* reporter fusion to the chromosomal promoter of *upaG* in CFT073 (CFT073Δ*lac_upaG*::*lacZ*) and subjected the strain to random Mariner transposon mutagenesis in order to identify negative regulators of *upaG* expression. Screening of 20,000 mutants identified two mutants with β-galactosidase activity. Both mutants carried transposon insertions in different positions of the gene encoding the global regulator H-NS (corresponding to amino acid positions 97 and 127 of the H-NS protein). To confirm the regulatory role of H-NS in expression of *upaG*, we constructed an *hns* deletion in CFT073Δ*lac_upaG*::*lacZ* and subsequently complemented this strain with a plasmid containing the *hns* gene downstream of the inducible *araBAD* promoter. Relative β-galactosidase activity was measured in $M63B1_{Glu}$-grown cultures of CFT073Δ*lac_upaG*::*lacZ*_Δ*hns* complemented with empty pBAD30 vector or pBAD30*hns* in the absence or presence of 0.2% arabinose. Deletion of *hns* in CFT073 derepressed *upaG* expression by 6-fold, which was reversed upon complementation with pBAD30*hns* but not the pBAD30 control under arabinose induction (Fig. 6A).

H-NS binds to the *upaG* promoter region. H-NS has an affinity for AT-rich, intrinsically curved double-stranded DNA. The 250-bp promoter region of *upaG* is 60.4% AT rich, and we examined its properties as well as its H-NS binding capacity using several approaches. First, we generated an *in silico* curvature-propensity plot calculated with DNase I-based parameters and showed that this 250-bp segment may adopt a curved conformation (Fig. 6B). Next, we demonstrated this curvature experimentally by examination of the PCR-amplified 250-bp fragment using polyacrylamide gel electrophoresis at 4°C. Using this method, which has been used previously to demonstrate DNA curvature (55, 59), the 250-bp *upaG* promoter region displayed a slightly retarded gel electrophoretic mobility compared to that of non-curved DNA standards (Fig. 6C). Finally, to demonstrate direct binding of H-NS to the 250-bp promoter region of *upaG*, we performed electrophoretic mobility shift assays. The 250-bp PCR product was mixed with TaqI-SspI-digested pBR322 DNA (which contains the *bla* promoter and has been previously shown to be bound by H-NS), incubated with increasing concentrations of purified H-NS protein, and subsequently visualized by gel electrophoresis. The 250-bp *upaG* promoter region and the fragment containing the *bla* promoter were retarded in mobility by the addition of 0.5 μM H-NS (Fig. 6D). The pBR322 fragments not containing the *bla* promoter were not influenced by H-NS at these concentrations, indicating that H-NS binds with specificity. These results suggest that H-NS binds to the regulatory region of *upaG* by recognizing a DNA region within 250 bp 5′ of the ATG translation start codon.

Mutation of the *hns* gene results in increased expression of *upaG* and *ehaG*. To confirm that H-NS acts as a repressor of *upaG* expression, we constructed an *hns* isogenic mutant in CFT073 and examined UpaG production by Western blot analysis. Loss of

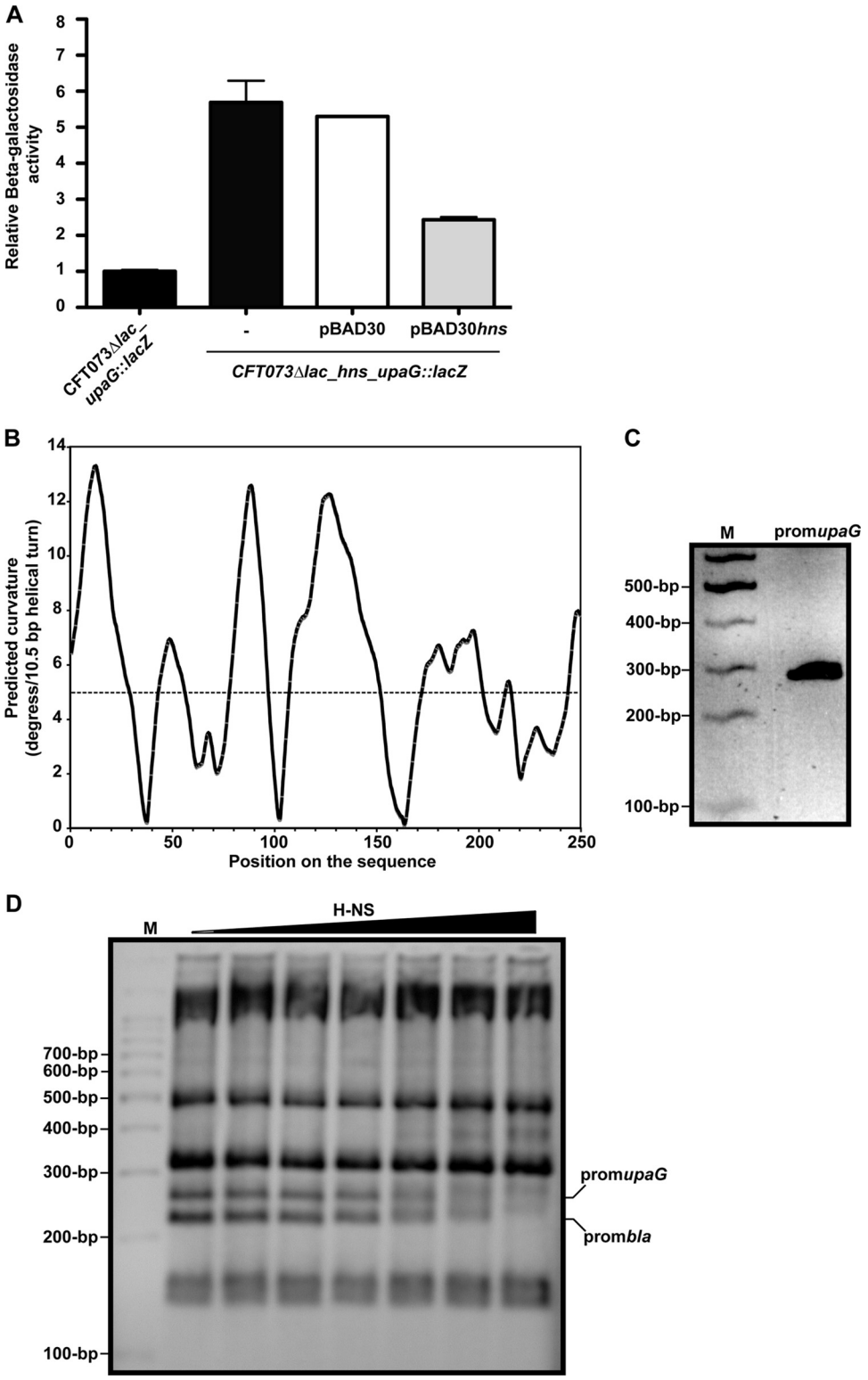

FIG 6 H-NS repression of *upaG* expression in UPEC CFT073. (A) β-Galactosidase activity of a chromosomal *upaG*::*lacZ* reporter fusion in CFT073Δ*lac* and CFT073Δ*lac_hns* alone, with pBAD30 (vector control) or pBAD30*hns* (complementation vector). pBAD30- and pBAD30*hns*-containing strains were grown in the presence of 0.2% arabinose. (B) Curvature plot of the *upaG* promoter sequence. (C) Electrophoretic mobility of the *upaG* promoter sequence. (D) Electrophoretic mobility shift assay of H-NS binding to the *upaG* and *bla* promoter sequences.

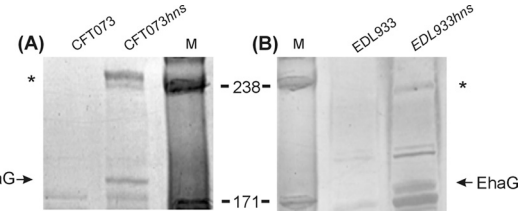

FIG 7 Production of UpaG (in CFT073) and EhaG (in EDL933) in an *hns* mutant background. Western blot analysis of whole-cell lysates prepared from *E. coli* CFT073 and CFT073*hns* employing a rabbit polyclonal anti-UpaG serum (A) and from *E. coli* EDL933 and EDL933*hns* using a rabbit polyclonal anti-EhaG serum (B). The monomeric UpaG and EhaG proteins are indicated; possible oligomeric forms of each protein are indicated by asterisks.

H-NS in CFT073 resulted in an increase in the production of UpaG that was detectable by Western blot analysis using an anti-UpaG specific serum (Fig. 7A). We also constructed an isogenic EDL933 *hns* mutant strain and tested for EhaG expression by Western blot analysis using an anti-EhaG specific serum. Similarly to CFT073, deletion of *hns* in EDL933 resulted in increased expression of the EhaG protein (Fig. 7B). Taken together, these results demonstrate that H-NS negatively regulates the expression of *upaG* in CFT073 and *ehaG* in EDL933.

DISCUSSION

TAAs are an important group of virulence factors in many Gram-negative pathogens. TAAs are translocated to the cell surface via the type V secretion pathway and adopt a trimeric conformation in the outer membrane. Three TAAs have been characterized from pathogenic *E. coli*, namely, UpaG, Saa, and the Eib group of proteins (37, 44, 56). Here we have characterized the EhaG protein, a TAA from enterohemorrhagic *E. coli* O157:H7. Like its positional orthologue UpaG from UPEC, EhaG mediates bacterial cell aggregation, biofilm formation, and adherence to ECM proteins. However, in contrast to UpaG (which mediates specific adherence to bladder epithelial cells), EhaG mediates specific adherence to intestinal epithelial cells.

EhaG and UpaG both possess a classical TAA domain structure consisting of an N-terminal signal sequence, a passenger domain, and a C-terminal translocation domain. The translocation domain is the most conserved feature of the TAA family, with the 70- to 100-amino-acid C-terminal region responsible for translocation of the passenger domain to the bacterial surface in all TAAs (41). We showed previously that the last 71 amino acids of UpaG, corresponding to the L1-β subdomain, represent the translocation unit (56). This region is completely conserved at the amino acid level in EhaG and UpaG.

In contrast to the translocation domain, the passenger domain of EhaG and UpaG exhibits extensive amino acid sequence variability (65% identity). Further analysis of the passenger domain revealed that the Hep_Hag motifs in the predicted head region were more variable than the HIM motifs in the stalk and neck structures. The head region has been found to be responsible for adhesion in other TAAs (31). Although the structural domains responsible for the binding properties of EhaG and UpaG have not been defined, it is likely that sequence differences in the head region account for the different EhaG and UpaG cell adherence phenotypes observed in this study.

Most trimeric AT proteins characterized to date display an adhesive activity mediating bacterial interactions with either host cells or ECM proteins (3, 10, 11, 24, 42, 45). Both EhaG and UpaG possessed several conserved features—they mediated cell aggregation; biofilm formation; and adhesion to laminin, fibronectin, fibrinogen, and collagen types I, II, III, and V. There are several features of both proteins that could account for these observations. The multiple Hep_Hag motifs in the head region, despite a degree of sequence variation, could mediate some of these functions. Alternatively, these properties may be associated with conserved structural features of both proteins rather than the presence of specific binding domains.

We found that the sequence divergence in the passenger domain of EhaG and UpaG corresponded strongly with diarrheagenic and extraintestinal *E. coli* pathotypes. In the case of DEC strains, EhaG is highly conserved and exhibits >95% amino acid identity. In contrast, the sequence of UpaG among ExPEC strains was more variable. The UpaG protein from UPEC CFT073 has previously been shown to mediate cell aggregation, biofilm formation, and adhesion to the ECM proteins laminin and fibronectin (56). The extensive sequence variation of UpaG suggests that there may be differences in these functional properties between UpaG variants from different UPEC/ExPEC strains. Thus, a more detailed analysis of the functions of different UpaG variants is required to properly assess its role in virulence.

Adherence of EHEC to the intestinal epithelium is essential for initiation of infection. The adherence of *E. coli* O157:H7 to Caco-2 cells has been shown to proceed in two stages: an initial diffuse adherence to epithelial cells followed by proliferation to develop microcolonies and intimate adherence (52). The intimate attachment and microcolony formation are thought to be mediated primarily by intimin-Tir interactions whereas the diffuse adherence to Caco-2 cells requires multiple factors encoded by the locus of enterocyte effacement (LEE) and additional chromosomal loci (52). Some of the proteins implicated in this initial adherence include ToxB and EspA (21, 52). The diffuse pattern of adherence mediated by EhaG to Caco-2 cells following expression in *E. coli* K-12 suggests that it may contribute to initial adherence of *E. coli* O157:H7 to intestinal epithelial cells. This remains to be demonstrated.

Immunodetection employing antibodies specific to EhaG and UpaG failed to detect these proteins in whole-cell extracts prepared from wild-type strains EDL933 and CFT073, respectively. We were also unable to detect expression of EhaG or UpaG from wild-type bacteria during interaction with cultured epithelial cells (data not shown). This prompted us to examine possible mechanisms by which the expression of *ehaG* and *upaG* could be repressed. We initially focused our analysis on *upaG* and identified a role for H-NS in its regulation. At the transcriptional level, a significant increase in *upaG* promoter activity was observed in a CFT073Δ*lac_upaG*::*lacZ-zeo hns* mutant. Consistent with this result, we also observed an increase in the expression of the UpaG protein by CFT073 in an *hns* mutant background. H-NS is a histone-like DNA binding protein that shows affinity for A-T-rich and bent nucleation sites on DNA (18, 19). Our data demonstrate that H-NS acts as a repressor of *upaG* transcription, most likely through direct binding to a region comprising the 250 bp upstream of the *upaG* open reading frame. This sequence shares 83% nucleotide sequence conservation with the corresponding region in *ehaG*; indeed, we also observed an increase in the expression of EhaG by EDL933 in an *hns* mutant background. Thus, the tran-

scription of *ehaG* and *upaG* is negatively regulated by H-NS. It is possible that the transcription of *ehaG* (in EHEC) and *upaG* (in UPEC) is coordinated with those of other H-NS-repressed genes. For example, mutation of the *hns* gene in UPEC strain 536 results in the derepression of multiple virulence factors, including alpha-hemolysin, iron uptake systems, and fimbriae (34), and several cryptic *E. coli* chaperone-usher fimbrial genes have also been shown to be repressed by H-NS in *E. coli* (29). The binding of H-NS to DNA is also modulated by temperature, with relief of repression for many genes observed above a threshold temperature of 32°C (19). Although we did not observe expression of EhaG/UpaG at 37°C in our experiments, it is possible that these proteins are expressed under specific conditions such as those during host infection.

In this study, we have examined the functional properties of the trimeric autotransporter EhaG from EHEC and compared its characteristics to UpaG from UPEC. Both proteins share several conserved features and yet also possess unique properties that may be associated with host tissue tropism and pathogenesis. Their common regulation by the global regulator H-NS suggests that their expression might be coordinated with other virulence factors, but the temporal and spatial patterns of UpaG and EhaG expression during disease remain to be elucidated. One property common to EhaG and UpaG is the ability to mediate biofilm formation. The role of biofilm formation in chronic bladder infection by UPEC (2) and environmental contamination of food by EHEC (48) has been documented. Indeed, given the recent outbreak of the STEC O104 strain in Germany, it will be important to thoroughly characterize the role of proteins such as EhaG in biofilm growth by pathogenic *E. coli*.

ACKNOWLEDGMENTS

We thank Sylvie Rimsky for providing the purified native H-NS protein.

This work was supported by grants from the Australian National Health and Medical Research Council (631654), the Australian Research Council (DP1097032), the University of Queensland (ECR grant to M.T.), the Institut Pasteur, the CNRS URA 2172, the Network of Excellence EuroPathoGenomics, and the European Community (LSHB-CT-2005-512061). M.A.S. is supported by an ARC Future Fellowship (FT100100662), and J.V. was a Marie-Curie Fellow.

REFERENCES

1. **Allsopp LP, et al.** 2012. Molecular characterization of UpaB and UpaC, two new autotransporter proteins of uropathogenic Escherichia coli CFT073. Infect. Immun. **80**:321–332.
2. **Anderson GG, et al.** 2003. Intracellular bacterial biofilm-like pods in urinary tract infections. Science **301**:105–107.
3. **Barenkamp SJ.** 1996. Immunization with high-molecular-weight adhesion proteins of nontypeable Haemophilus influenzae modifies experimental otitis media in chinchillas. Infect. Immun. **64**:1246–1251.
4. **Barnard TJ, Dautin N, Lukacik P, Bernstein HD, Buchanan SK.** 2007. Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. Nat. Struct. Mol. Biol. **14**:1214–1220.
5. **Beloin C, Dorman CJ.** 2003. An extended role for the nucleoid structuring protein H-NS in the virulence gene regulatory cascade of Shigella flexneri. Mol. Microbiol. **47**:825–838.
6. **Bertani G.** 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. J. Bacteriol. **62**:293–300.
7. **Bokil NJ, et al.** 2011. Intramacrophage survival of uropathogenic Escherichia coli: differences between diverse clinical isolates and between mouse and human macrophages. Immunobiology **216**:1164–1171.
8. **Chaveroche MK, Ghigo JM, d'Enfert C.** 2000. A rapid method for efficient gene replacement in the filamentous fungus Aspergillus nidulans. Nucleic Acids Res. **28**:E97.
9. **Chiang SL, Rubin EJ.** 2002. Construction of a mariner-based transposon for epitope-tagging and genomic targeting. Gene **296**:179–185.
10. **Comanducci M, et al.** 2002. NadA, a novel vaccine candidate of Neisseria meningitidis. J. Exp. Med. **195**:1445–1454.
11. **Cope LD, et al.** 1999. Characterization of the Moraxella catarrhalis uspA1 and uspA2 genes and their encoded products. J. Bacteriol. **181**:4026–4034.
12. **Cotter SE, Surana NK, Grass S, St. Geme JW, III.** 2006. Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. J. Bacteriol. **188**:5400–5407.
13. **Cotter SE, Surana NK, St. Geme JW, III.** 2005. Trimeric autotransporters: a distinct subfamily of autotransporter proteins. Trends Microbiol. **13**:199–205.
14. **Da Re S, Ghigo JM.** 2006. A CsgD-independent pathway for cellulose production and biofilm formation in Escherichia coli. J. Bacteriol. **188**:3073–3087.
15. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. **97**:6640–6645.
16. **Derbise A, Lesic B, Dacheux D, Ghigo JM, Carniel E.** 2003. A rapid and simple method for inactivating chromosomal genes in Yersinia. FEMS Immunol. Med. Microbiol. **38**:113–116.
17. **Donnelly MI, et al.** 2006. An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. Protein Expr. Purif. **47**:446–454.
18. **Dorman CJ.** 2007. H-NS, the genome sentinel. Nat. Rev. Microbiol. **5**:157–161.
19. **Dorman CJ.** 2004. H-NS: a universal regulator for a dynamic genome. Nat. Rev. Microbiol. **2**:391–400.
20. **Durant L, et al.** 2007. Identification of candidates for a subunit vaccine against extraintestinal pathogenic Escherichia coli. Infect. Immun. **75**:1916–1925.
21. **Ebel F, et al.** 1998. Initial binding of Shiga toxin-producing Escherichia coli to host cells and subsequent induction of actin rearrangements depend on filamentous EspA-containing surface appendages. Mol. Microbiol. **30**:147–161.
22. **Felsenstein J.** 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Department of Genome Science, University of Washington, Seattle, WA.
23. **Finn RD, et al.** 2010. The Pfam protein families database. Nucleic Acids Res. **38**:D211–D222.
24. **Geme JW, III, Cutter D.** 1995. Evidence that surface fibrils expressed by Haemophilus influenzae type b promote attachment to human epithelial cells. Mol. Microbiol. **15**:77–85.
25. **Guzman LM, Belin D, Carson MJ, Beckwith J.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. **177**:4121–4130.
26. **Heydorn A, et al.** 2000. Quantification of biofilm structures by the novel computer program COMSTAT. Microbiology **146**:2395–2407.
27. **Hoiczyk E, Roggenkamp A, Reichenbecher M, Lupas A, Heesemann J.** 2000. Structure and sequence analysis of Yersinia YadA and Moraxella UspAs reveal a novel class of adhesins. EMBO J. **19**:5989–5999.
28. **Kjaergaard K, Schembri MA, Ramos C, Molin S, Klemm P.** 2000. Antigen 43 facilitates formation of multispecies biofilms. Environ. Microbiol. **2**:695–702.
29. **Korea CG, Badouraly R, Prevost MC, Ghigo JM, Beloin C.** 2010. Escherichia coli K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. Environ. Microbiol. **12**:1957–1977.
30. **Leo JC, et al.** 2011. The structure of E. coli IgG-binding protein D suggests a general model for bending and binding in trimeric autotransporter adhesins. Structure **19**:1021–1030.
31. **Linke D, Riess T, Autenrieth IB, Lupas A, Kempf VA.** 2006. Trimeric autotransporter adhesins: variable structure, common function. Trends Microbiol. **14**:264–270.
32. **Miller JH.** 1992. A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
33. **Mobley HL, 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.
34. **Muller CM, et al.** 2006. Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic Escherichia coli. J. Bacteriol. **188**:5428–5438.

35. **Nummelin H, et al.** 2003. Structural studies of Yersinia adhesin YadA. Adv. Exp. Med. Biol. **529**:85–88.
36. **Oomen CJ, et al.** 2004. Structure of the translocator domain of a bacterial autotransporter. EMBO J. **23**:1257–1266.
37. **Paton AW, Srimanote P, Woodrow MC, Paton JC.** 2001. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic Escherichia coli strains that are virulent for humans. Infect. Immun. **69**:6999–7009.
38. **Perna NT, et al.** 2001. Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. Nature **409**:529–533.
39. **Reisner A, Haagensen JA, Schembri MA, Zechner EL, Molin S.** 2003. Development and maturation of Escherichia coli K-12 biofilms. Mol. Microbiol. **48**:933–946.
40. **Riess T, et al.** 2004. Bartonella adhesin a mediates a proangiogenic host cell response. J. Exp. Med. **200**:1267–1278.
41. **Roggenkamp A, et al.** 2003. Molecular analysis of transport and oligomerization of the Yersinia enterocolitica adhesin YadA. J. Bacteriol. **185**:3735–3744.
42. **Roggenkamp A, Neuberger HR, Flugel A, Schmoll T, Heesemann J.** 1995. Substitution of two histidine residues in YadA protein of Yersinia enterocolitica abrogates collagen binding, cell adherence and mouse virulence. Mol. Microbiol. **16**:1207–1219.
43. **Sambrook J, Fritsch EF, Maniatis T.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
44. **Sandt CH, Wang YD, Wilson RA, Hill CW.** 1997. Escherichia coli strains with nonimmune immunoglobulin-binding activity. Infect. Immun. **65**: 4572–4579.
45. **Scarselli M, et al.** 2006. Neisseria meningitidis NhhA is a multifunctional trimeric autotransporter adhesin. Mol. Microbiol. **61**:631–644.
46. **Schembri MA, Kjaergaard K, Klemm P.** 2003. Global gene expression in Escherichia coli biofilms. Mol. Microbiol. **48**:253–267.
47. **Schembri MA, Klemm P.** 2001. Biofilm formation in a hydrodynamic environment by novel fimh variants and ramifications for virulence. Infect. Immun. **69**:1322–1328.
48. **Seo KH, Frank JF.** 1999. Attachment of Escherichia coli O157:H7 to lettuce leaf surface and bacterial viability in response to chlorine treatment as demonstrated by using confocal scanning laser microscopy. J. Food Prot. **62**:3–9.
49. **Sherlock O, Schembri MA, Reisner A, Klemm P.** 2004. Novel roles for the AIDA adhesin from diarrheagenic Escherichia coli: cell aggregation and biofilm formation. J. Bacteriol. **186**:8058–8065.
50. **Surana NK, Cutter D, Barenkamp SJ, St. Geme JW, III.** 2004. The Haemophilus influenzae Hia autotransporter contains an unusually short trimeric translocator domain. J. Biol. Chem. **279**:14679–14685.
51. **Szczesny P, Lupas A.** 2008. Domain annotation of trimeric autotransporter adhesins—daTAA. Bioinformatics **24**:1251–1256.
52. **Tatsuno I, et al.** 2000. Isolation and characterization of mini-Tn5Km2 insertion mutants of enterohemorrhagic Escherichia coli O157:H7 deficient in adherence to Caco-2 cells. Infect. Immun. **68**:5943–5952.
53. **Ulett GC, et al.** 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.
54. **Ulett GC, Webb RI, Schembri MA.** 2006. Antigen-43-mediated autoaggregation impairs motility in Escherichia coli. Microbiology **152**:2101–2110.
55. **Ussery DW, Higgins CF, Bolshoy A.** 1999. Environmental influences on DNA curvature. J. Biomol. Struct. Dyn. **16**:811–823.
56. **Valle J, et al.** 2008. UpaG, a new member of the trimeric autotransporter family of adhesins in uropathogenic Escherichia coli. J. Bacteriol. **190**: 4147–4161.
57. **Vlahovicek K, Kajan L, Pongor S.** 2003. DNA analysis servers: plot.it, bend.it, model.it and IS. Nucleic Acids Res. **31**:3686–3687.
58. **Wells TJ, Tree JJ, Ulett GC, Schembri MA.** 2007. Autotransporter proteins: novel targets at the bacterial cell surface. FEMS Microbiol. Lett. **274**:163–172.
59. **Yamada H, Muramatsu S, Mizuno T.** 1990. An Escherichia coli protein that preferentially binds to sharply curved DNA. J. Biochem. **108**:420–425.
60. **Yeo HJ, et al.** 2004. Structural basis for host recognition by the Haemophilus influenzae Hia autotransporter. EMBO J. **23**:1245–1256.