Conserved FimH mutations in the global *Escherichia coli* ST131 multi-drug resistant lineage weaken interdomain interactions and alter adhesin function

Jilong Qin [a], Katie A. Wilson [b], Sohinee Sarkar [a], Begoña Heras [c], Megan L. O'Mara [b], Makrina Totsika [a,*]

[a] *Centre for Immunology and Infection Control, School of Biomedical Sciences, Queensland University of Technology, Brisbane, QLD 4006, Australia*
[b] *Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia*
[c] *Department of Biochemistry and Chemistry, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC 3086, Australia*

ARTICLE INFO

ABSTRACT

The binding of the type 1 fimbrial adhesin FimH to mannosylated receptors is allosterically regulated to enhance the fitness of uropathogenic *Escherichia coli* (UPEC) during urinary tract infection (UTI). Mutations in the two FimH domains (pilin and lectin) located outside the mannose binding pocket have been shown to influence mannose binding affinity, yet the details of the allostery mechanism are not fully elucidated. Here we characterised different FimH conformational states (termed low-affinity tense and high-affinity relaxed conformations) of natural FimH variants using molecular dynamics (MD) simulation techniques and report key structural dynamics differences between them. The clinically dominant FimH30 variant from the pandemic multidrug resistant *E. coli* ST131 lineage contains an R166H mutation that weakens FimH interdomain interactions and allows enhanced mannose interactions with pre-existing high-affinity relaxed conformations. When expressed in an isogenic ST131 strain background, FimH30 mediated high human cell adhesion and invasion, and enhanced biofilm formation over other variants. Collectively, our computational and experimental findings support a model of FimH protein allostery that is mediated by shifts in the pre-existing conformational equilibrium of FimH, additional to the sequential step-wise process of structural perturbations transmitted from one site to another within the protein. Importantly, it is the first study to shed light into how natural mutations in a clinically dominant FimH variant influence the protein's conformational landscape optimising its function for ST131 fitness at intestinal and extraintestinal niches.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creative-commons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Uropathogenic *Escherichia coli* (UPEC) is the most common Gram-negative bacterial pathogen causing bacteremia and urinary tract infections (UTIs) [1]. UPEC express several adhesive hair-like appendages on their cell surface, with type 1 fimbriae (or pili) being the best characterised and shown to mediate UPEC colonisation of both the bladder and more recently the colon epithelium in mouse models [2–5]. Adhesion to mannosylated host cell receptors is mediated by the type 1 fimbriae tip adhesin FimH. FimH consists of two domains connected by a 5 amino acid linker: an *N*-terminal lectin domain (FimH$_{LD}$, 1–155) located at the distal end of the fimbriae, and a C-terminal pilin domain (FimH$_{PD}$, 161–279) which

accommodates a donor strand from FimG, the previous subunit in the type 1 fimbrial structure (Fig. 1A). The FimH$_{LD}$ contains a mannose binding pocket of three conserved loops [6–8], which recognises mannosylated host surface molecules as receptors, including the uroplakin Ia glycoprotein that is abundantly expressed on the surface of human uroepithelial cells [9].

A two-domain architecture allows FimH to modulate its catch bond host cell adhesion properties, whereby FimH affinity for mannosylated host cell receptors is enhanced by domain separation under shear force [10,11]. Indeed, the movement of red blood cells (RBC) bound to a surface coated with *E. coli* was stalled under increased flow conditions [12,13]. Several FimH structural studies support this tenet. For example, purified FimH in complex with the type 1 chaperone FimC, as well as purified FimH$_{LD}$ or an engineered FimH mutant (A188D) with high mannose-binding affinity, all have disrupted FimH$_{LD}$-FimH$_{PD}$ interdomain interactions [14]. In contrast, purified FimH complemented with the FimG donor

---

* Corresponding author.
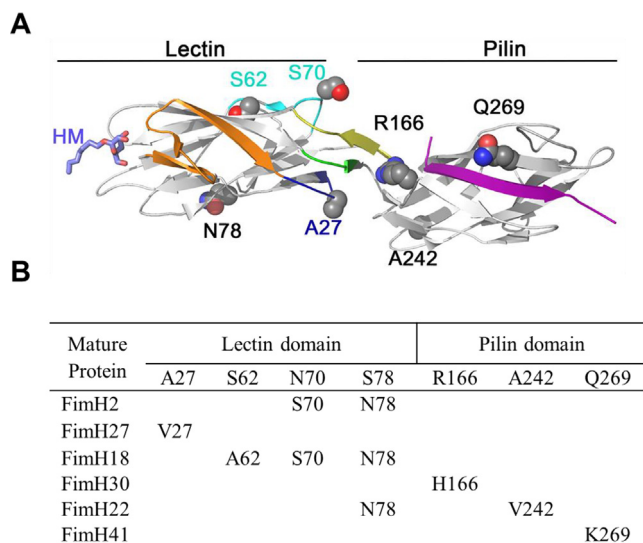*E-mail address:* makrina.totsika@qut.edu.au (M. Totsika).

Fig. 1. Mutations in natural FimH variants examined in this study. Mutations in FimH variants investigated as part of this study located in the FimH lectin and pilin domains (outside the conserved mannose binding pocket), showing their structural location (grey spheres) in the FimH2 structure (PDB 4XOE) (A) and amino acid substitutions (B). Clamp segment (residues 2–22, orange); swing segment (residues 23–32, blue); bulge helix segment (residues 60–72, cyan); insertion loop (residues 112–124, yellow); linker (residues 150–155, green); DsG donor peptide (purple); and mannose (purple) are highlighted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Mature	Lectin domain				Pilin domain		
Protein	A27	S62	N70	S78	R166	A242	Q269
FimH2			S70	N78			
FimH27	V27						
FimH18		A62	S70	N78			
FimH30					H166		
FimH22				N78		V242	
FimH41							K269

peptide (FimH·DsG) with a low affinity for mannose showed close interdomain interactions [10]. In addition, recent structures of FimH complemented with the non-native donor peptide from FimF (DsF) [10] or the engineered high-affinity FimH mutant A27V/V163A [15] also captured FimH in the elongated (domain-separated) conformation. Both of these recent structures showed a nearly identical FimH$_{LD}$ conformation but had distinctive bend angles between the two domains [15]. Further, biochemical and biophysical analyses of FimH in solution revealed that the protein can be found simultaneously in the associated conformation (termed tense or T state) and in multiple elongated conformations (termed relaxed or R state) prior to mannose binding [15]. The authors posited that the equilibrium between pre-existing FimH conformational ensembles is subject to evolutionary selection pressure that optimally adapts the adhesive strength of this adhesin to suit the pathogen's lifestyle. While this tenet is biologically plausible it has not been explored with natural FimH variants.

While FimH is ubiquitously present in *E. coli,* it has been documented to display extensive allelic variation with over 60 alleles reported to date [7,16–20]. Despite FimH allelic variation having been shown to affect the protein's function, its impact on protein structure and allostery has not been widely investigated and current understanding remains limited to just a handful of FimH variants [21]. In fact, all wild-type FimH structures available in the Protein Data Bank (PDB) to date, are representative of the FimH2 and FimH27 sequence variants [10,22]. While both contain pathoadaptive mutations known to affect FimH cellular function [6], they do not represent the wide variety of FimH variants found in *E. coli*, especially those associated with clinically relevant lineages. Here, we studied the impact of natural variation in *fimH* alleles from globally dominant UPEC sequence type 131 (ST131) lineages (FimH30-Clade C; FimH22-Clade B; FimH41-Clade A, [23]) on FimH protein allostery at the atomic level and adhesin function during host-pathogen interactions. Using advanced computational approaches, we distinguish FimH R state conformations

from the domain-separated T state conformation under shear force. We show that the FimH30 variant associated with the multidrug resistant (MDR) ST131 Clade C forms fewer interdomain interactions in the T state, whereas when in R states it displays longer lasting interactions with mannose. When expressed in an isogenic ST131 strain background, the FimH30 variant confers enhanced intestinal cell adhesion and bladder cell invasion and mediates high biofilm formation. Our findings support the notion that the R166H mutation in FimH30$_{PD}$ favors a conformation lanscape that allows this variant to optimally adapt its virulence function for cell adhesion, invasion and biofilm formation.

2. Results

2.1. Mannose-bound FimH variants maintain stable domain conformations in both T and R states with large interdomain motion only observed in the R state

FimH-mannose binding was previously shown to be affected by engineered mutations outside of the mannose binding pocket that alter local interaction networks and influence the conformational equilibrium of FimH [14,15]. To investigate this in naturally occurring FimH sequences from diverse *E. coli* strains, we selected three FimH variants, each closely associated with the three main ST131 clades (FimH30 – Clade C; FimH22 – Clade B; and FimH41 – Clade A). These were compared to the extensively characterised FimH18 variant from UPEC strain UTI89 [24] and the only two structurally solved variants: FimH2 and FimH27 [21], which are commonly, but not exclusively, encoded by commensal *E. coli* strains. In comparison to previously characterised FimH variants (FimH2, FimH27 and FimH18), the three ST131 clade-defining variants have non-synonymous mutations located outside the mannose binding site and predominantly in the pilin domain (R166H, S78N/A242V and Q269K for FimH30, FimH22 and FimH41, respectively; Fig. 1).

To explore if these non-synonymous mutations outside of the mannose binding pocket impact on FimH protein structure and dynamics, we performed 300 ns of nonbiased MD simulations in triplicate for each FimH variant. The previously reported structures of FimH·DsG bound with mannose in both the T state (FimH2, PDB ID: 4XOE) and the R state (FimH27, PDB ID: 4XOB) were used to initiate the simulations as they represent the first step of FimH-mannose binding at different pre-existing conformations (T or R; Fig. S1). Irrespective of the variant, all simulations initiated from the T state (PDB ID 4XOE, Fig. 2A) showed a maximum backbone root-mean-square deviation (RMSD) of 1.6 ± 0.4 Å with respect to the starting conformation of the entire protein (Fig. 2A), or 1.1 ± 0.4 Å for an individual domain (Fig. 2B-C), indicative of normal thermal motion (Fig. 2D). However, all simulations initiated from an R state (PDB ID 4XOB, Fig. 2A) showed elevated backbone RMSDs with respect to the starting structure, ranging between 5.8 ± 3.9 Å for FimH2 to 10.1 ± 3.6 Å for FimH27, indicative of large-scale conformational changes (Fig. 2E). On calculation of the RMSD of each domain for simulations based on the R state, the maximum RMSD is 1.3 ± 0.2 Å (Fig. 2B-C), confirming that each individual domain maintains a stable conformation throughout the simulation and the large overall RMSD is due to the relative motion of the domains with respect to one another, rather than intra-domain conformational changes. Consistent with the difference in overall RMSD between the two states, a smaller radius of gyration is also observed for the FimH variants in the T state (24–28 Å) compared to the R state (20–32 Å, Fig. S2). Collectively, all mannose-bound FimH variants simulated either from an initial T or an R state maintained their conformational states, with the T state showing a constrained conformation, while R states sampling multiple conformations.
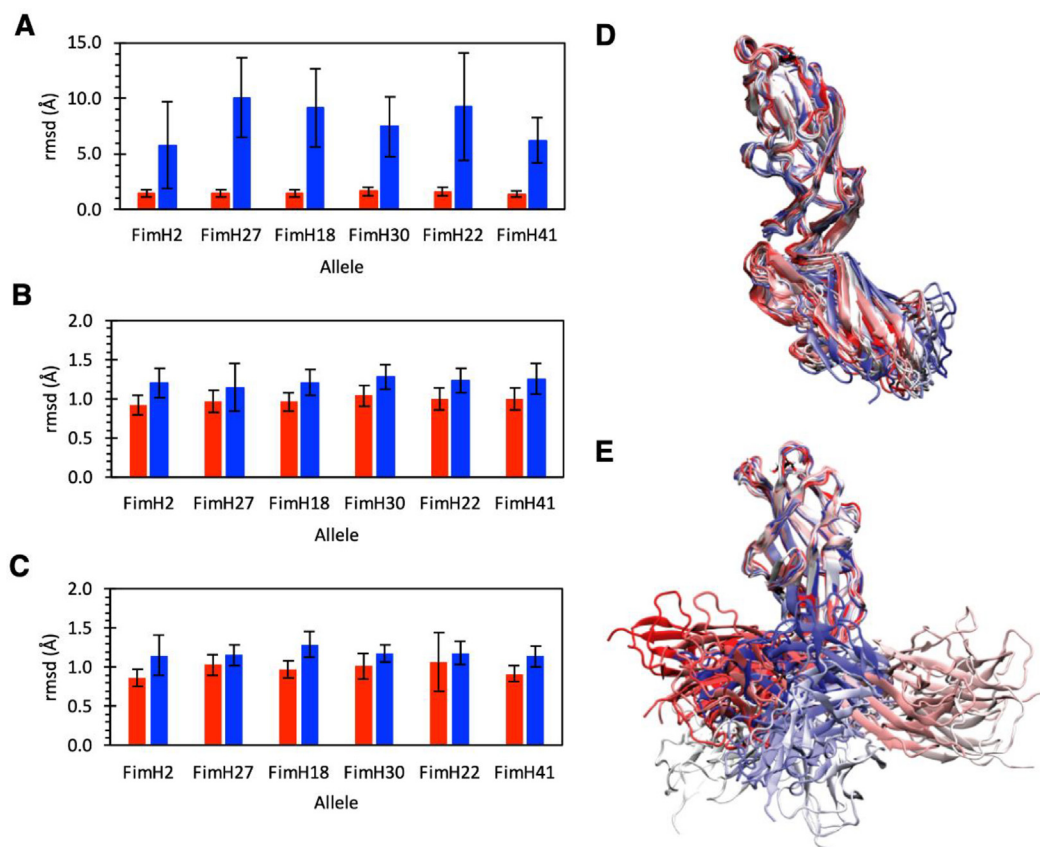
Fig. 2. Average domain motion for FimH in the T and R states. Average backbone RMSD for FimH variants simulated from the crystallographic T state (red, PDB 4XOE, FimH2) or R state (blue, PDB 4XOB, FimH27) calculated for (A) the entire protein, (B) pilin domain, or (C) lectin domain, with respect to the starting geometry. Standard deviation is shown as an error bar. Superposition of the relative conformation of FimH at 3 ns intervals between the start (red) and the end (blue) of the 300 ns MD simulation for the (D) T state or (E) R state. FimH30 is shown as a representative example. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. FimH30 shows decreased interdomain interactions in the T state and maintains persistent mannose interactions in the R state

FimH recognises and binds to mannosylated receptors through its conserved mannose binding pocket located at the distal end of the lectin domain. We compared the hydrogen bonding frequencies between mannose and FimH variants simulated in both the R and T states. The mannose hydrogen bonding pattern is well maintained between all modelled FimH variants in both states. Specifically, the backbone of Asp47 and Phe1 and side chains of Asn135, Gln133 and Asp54 form 7 hydrogen bonds to mannose, coordinating it within the binding site (Fig. 3A & Table S2). FimH30 shows the most consistent increase in the occupancy of all 7 mannose hydrogen bonds in the R state relative to the T state (Fig. 3B & Table S2). The FimH27 and FimH18 variants also show an increase in mannose hydrogen bonding occupancies, while there is little difference between the mannose hydrogen bonding occupancy between the two states for FimH2, FimH22 and FimH41 (Fig. 3B & Table S2).

Since FimH interdomain interactions have been reported to negatively affect mannose binding [14,25], we next assessed hydrogen bonding frequencies between domains in each FimH variant. At any point of the simulation in the T state there are between 3 and 15 interdomain hydrogen bonds (Fig. 4A) with 8 of these being high occupancy (persisting for >70 % of the total simulation time; Table S3). In contrast, all FimH variants maintained a low number of interdomain hydrogen bonds in the R state conformations (Fig. 4B). In the case of FimH30 the R166H mutation

leads to complete loss of two high occupancy hydrogen bonds involving R166H at the interdomain interface in the T state (70–100 % occupancy in all other alleles, Fig. 4C & Tables S3 and S4), and also complete loss of the S114(O)-G159(NH) hydrogen bond present in all other variants in the R state (Table S3). Collectively, these data suggest that FimH30 has less interdomain interactions in both T and R state conformations and increased mannose binding occupancy in R state conformations, which is in agreement with reduced interdomain interactions leading to tighter mannose binding.

2.3. Applying force to FimH30 induces a domain-separated conformation that is different to R state conformations

Previous work has shown that shear force can facilitate the T to R state transition in FimH [10–12]. We next sought to use MD simulations to explore the molecular motion of FimH30 (the variant with fewer interdomain interactions in both T and R states) under shear force, by performing adaptive bias steered molecular dynamics (ASMD) starting from the T state conformation. Force was applied between the mannose ligand and the DsG to mimic the scenario whereby FimH30 at the fimbrial tip binds to mannosylated receptors under increased shear force. In the T state conformation, both the insertion loop and swing segment of the lectin domain are interacting with the pilin domain (Fig. 5A). Upon application of the force, the contacts between the swing segment and pilin domain are lost first (N33–P157 and N33–G159 hydrogen bonds) (Supplementary Video 1 and Figure S3). Subsequently, at
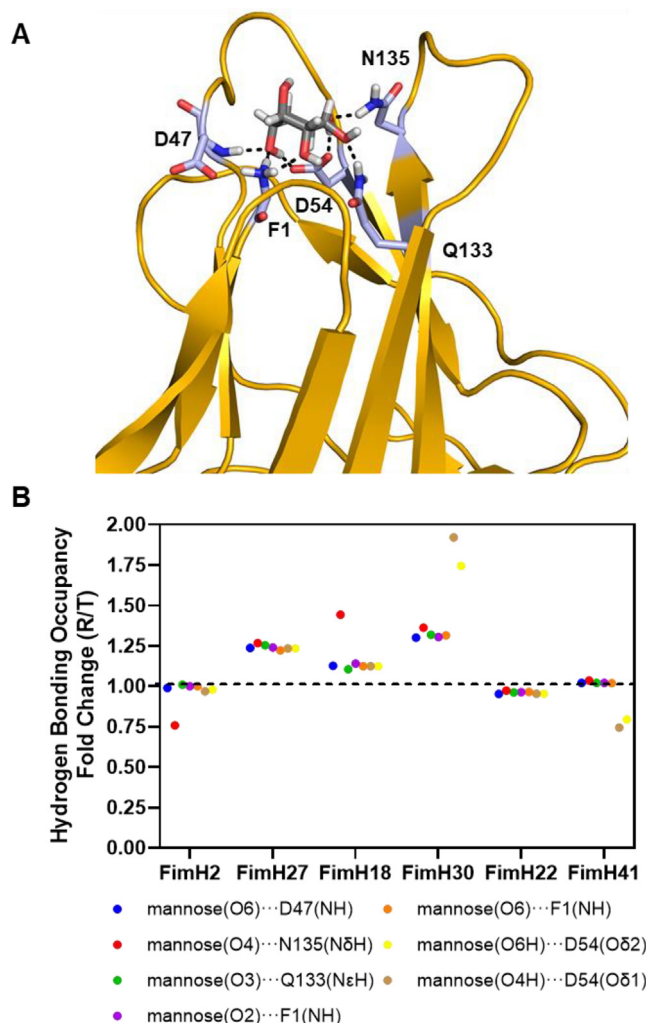
Fig. 3. Mannose hydrogen bonding interactions with FimH. (A) Representative structure from MD showing the hydrogen-bonding geometry for interactions between mannose and residues in the FimH30 mannose-binding pocket. (B) Fold change in hydrogen bond occupancies observed in simulations with each FimH variant in the R state over the T state, where a ratio of 1 (dotted line) indicates no changes in mannose binding stability and above 1 suggests higher mannose binding occupancy in the R state.

Fig. 4. FimH interdomain hydrogen bonding interactions. Number of hydrogen bonds formed between the lectin and pilin domains during simulations of the (A) T state and (B) R state. (C) Representative structure from MD showing the hydrogen bonding at R166H between FimH2 (top, right) and FimH30 (bottom, right) in T state.

backbone geometry of P26 in the non-biased simulations differs between the R state and T state conformations (Fig. 5D). These data suggest that the resulting FimH conformation induced by applying force on the T state under our experimental settings may differ from the domain-separated R state conformations.

high forces the insertion loop contacts with the pilin domain are broken. Many interactions occur between the insertion loop and pilin domain, with the main interactions seen between V112, S113, S114 and G116, and P157, T158, G159, G160 and C161. During this conformational change the angle between the two domains changes from ~ 150° to 180° (Fig. 5A). Interestingly, when the final frame of the pull is subsequently simulated in the absence of force, FimH30 collapses back to the original T state conformation (Fig. 5B). These results indicate that the domain-separated FimH30 obtained from the T state when simulated under shear force does not represent the FimH30 starting R state conformations used in our non-biased simulations. Indeed, a key difference between the T and R states is the backbone geometry of the P26 residue, which influences the conformation of the swing loop (where this residue resides), ultimately altering the conformation of the adjacent linker region and insertion loop in the interdomain region leading to the two distinct configurations (Fig. 5C & 5D, Fig. S1C). During the ASMD simulation separation of the FimH30 domains in the T state, in which a force was applied to separate the domains, the backbone geometry of P26 did not change (Fig. 5C). Notably, the
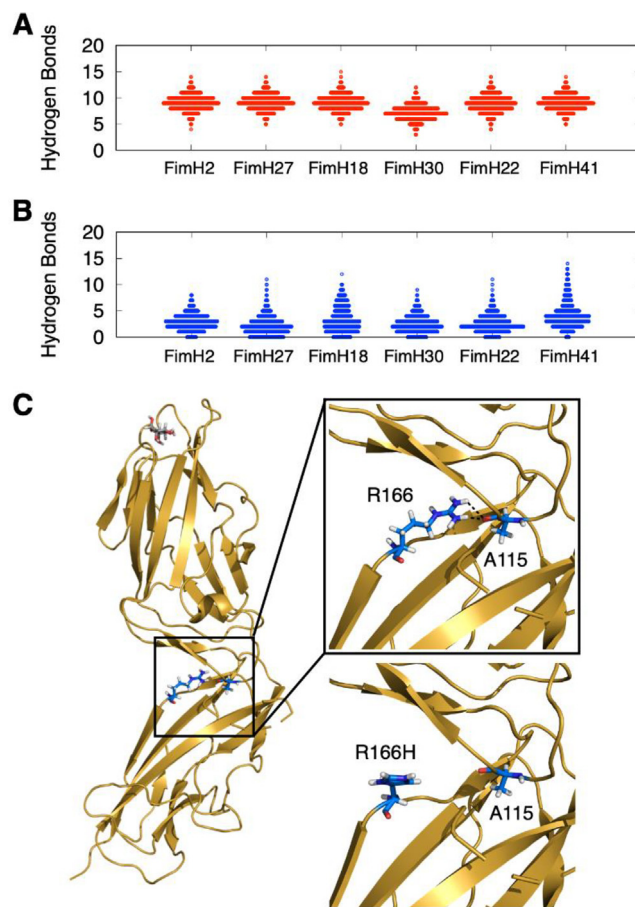
3. Construction of an isogenic set of *E. coli* ST131 EC958 strains expressing equivalent levels of distinct FimH variants

To study the functional consequences of differences in protein dynamics caused by these natural mutations in FimH, we performed allelic exchange and introduced the three prevalent *fimH* alleles from ST131 clades A (*fimH41*), B (*fimH22*), and C (*fimH30*), and well-studied alleles *fimH18* and *fimH27* into the ST131 *E. coli* EC958 strain background, by replacing the original *fimH30* gene with each of the natural alleles. This resulted in a set of 5 isogenic EC958 *fimH* mutant strains (referred as H30, H22, H41, H27 and H18 throughout this study). Due to an insertion within the type 1 fimbrial regulator gene *fimB*, the off-to-on switch frequency of type 1 fimbriae in ST131 *E. coli* EC958 is slow, but the production of fimbriae can be enriched under sequential aerobic static culture conditions [26,27]. To confirm equal production of type 1 fimbriae in the isogenic strains, we first compared the orientation of the *fim* switch (*fimS*; contains the *fim* promoter) in EC958 strains harbouring different *fimH* alleles. No difference in *fimS* 'off' and 'on' proportions was observed between strains, with over 90 percent of the
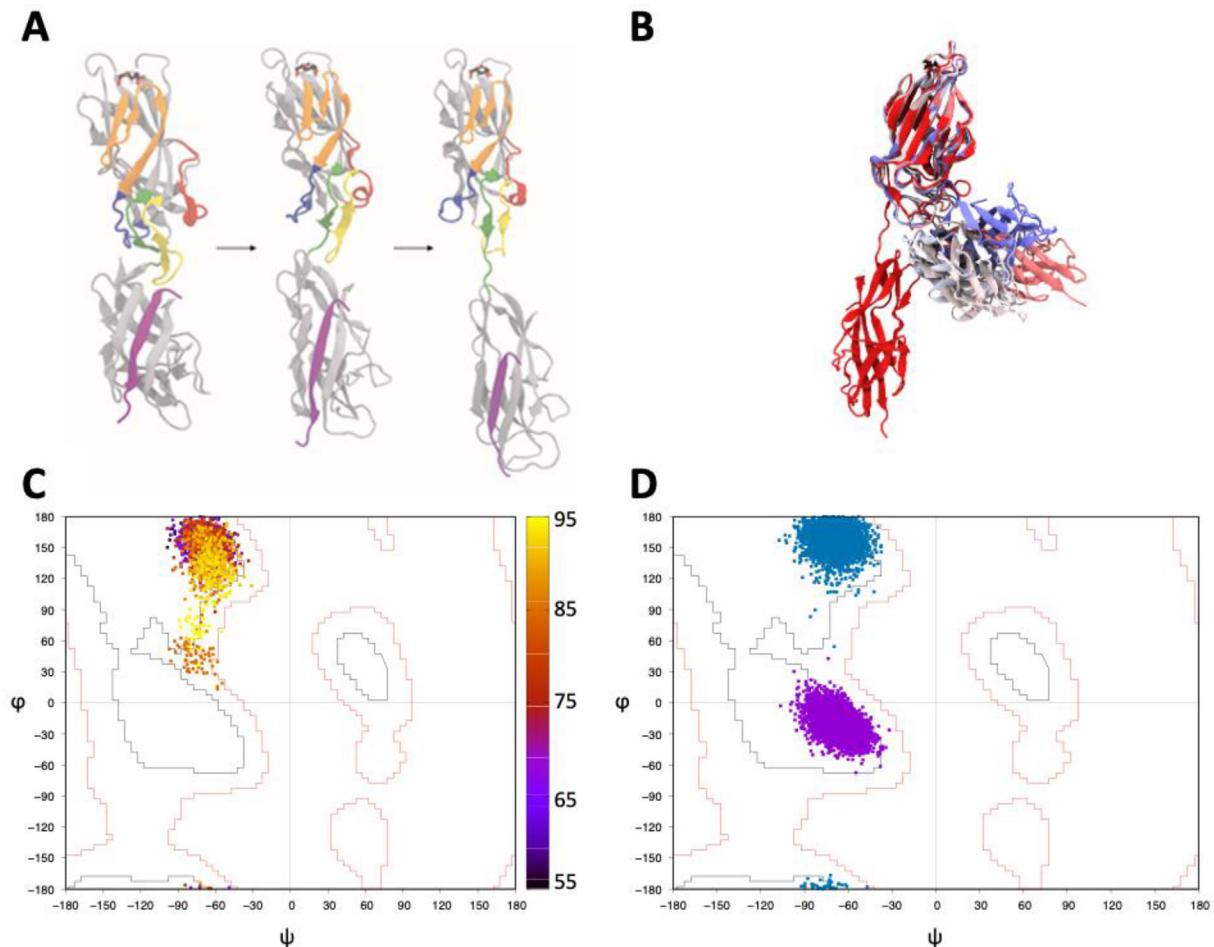
Fig. 5. Distinctive structural difference between T state conformation with domain separation under shear force and R state conformations in the absence of force. (A) Conformational change observed during the ASMD from the T state conformation (left) to loss of contacts between the swing segment and pilin domain (middle) and the fully separated conformation with the swing segment not interacting with the pilin domain. Regions of the protein including the swing segment (blue), DsG (purple), clamp segment (orange), bulge helix segment (red), insertion loop (yellow) and linker (green) are highlighted. (B) Superposition of the relative conformation of FimH30 at 50 ns intervals between the start (red) and the end (blue) of the 300 ns MD simulation of the FimH30 variant in the absence of force starting from the final frame of the ASMD pull. (C) Ramachandran plot of the ϕ and ψ dihedral angles of P26 over the ASMD simulation of FimH30 from a separation distance between mannose and the DsG of 55 (black) to 95 (yellow) Å. (D) Ramachandran plot of the ϕ and ψ dihedral angles of P26 from 300 ns unbiased simulation of FimH30 in the R-state (purple) and *T*-state (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

fimS-containing amplicons in the 'switch on' orientation, similar to that of EC958 wild-type (WT) cultured under identical *fim* enriching conditions (Fig. S3A). Equivalent expression of type 1 fimbriae across strains was also confirmed by immunoblotting using anti-FimA sera (Fig. S3B). Lastly, functional production of type 1 fimbriae on the surface of EC958 as assessed by yeast agglutination confirmed that the isogenic EC958 *fimH* mutants had similar levels of type 1 fimbriae to WT. FimH-mediated yeast cell agglutination by all strains was fully inhibited by the addition of methyl-α-mannopyranoside (Fig. S3C). Together, these data confirmed that the expression of different *fimH* alleles in ST131 *E. coli* EC958 was not detectably different to that of the *fimH30* encoding parent strain.

4. FimH30 mediates high levels of intestinal cell adhesion, bladder cell invasion and biofilm formation by ST131 *E. coli* EC958

We have previously shown that FimH30 in ST131 *E. coli* EC958 mediates human bladder epithelial cell adhesion and invasion [26], interactions with human intestinal epithelial cells [3] and biofilm

formation [28]. We therefore compared cell adhesion and invasion mediated by EC958 isogenic *fimH* mutants to T24 human bladder epithelial cells (Fig. 6A&B) and Caco-2 human colonic epithelial cells (Fig. 6 C&D). While swapping *fimH30* for other natural *fimH* alleles in EC958 did not affect bacterial adhesion to bladder cells (Fig. 6A), the isogenic H30 strain showed significantly higher cell invasion compared to H41, H27 and H18 (Fig. 6B). Intriguingly, while H30 displayed significantly higher adhesion to Caco-2 cells compared to H22 and H41, it showed similar intestinal cell invasion to all other strains, except H41 (Fig. 6D). It is also worth noting that all ST131 *fimH* mutants mediated overall higher adhesion levels to intestinal cells compared to bladder cells and higher levels of invasion into bladder over intestinal cells (Fig. 6E). Our results indicate that the *fimH30* allele could confer a potential advantage to EC958 for intestinal adhesion and bladder invasion over other *fimH* alleles from ST131 clade A and B. We also tested the isogenic *fimH* mutants for biofilm formation on abiotic surfaces (Fig. 6F). H30 along with H22 displayed significantly higher biofilm levels than H41, H27 and H18, which lends further support to the tenet that FimH30 is evolutionarily optimised for enhanced ST131 fitness across various niches.
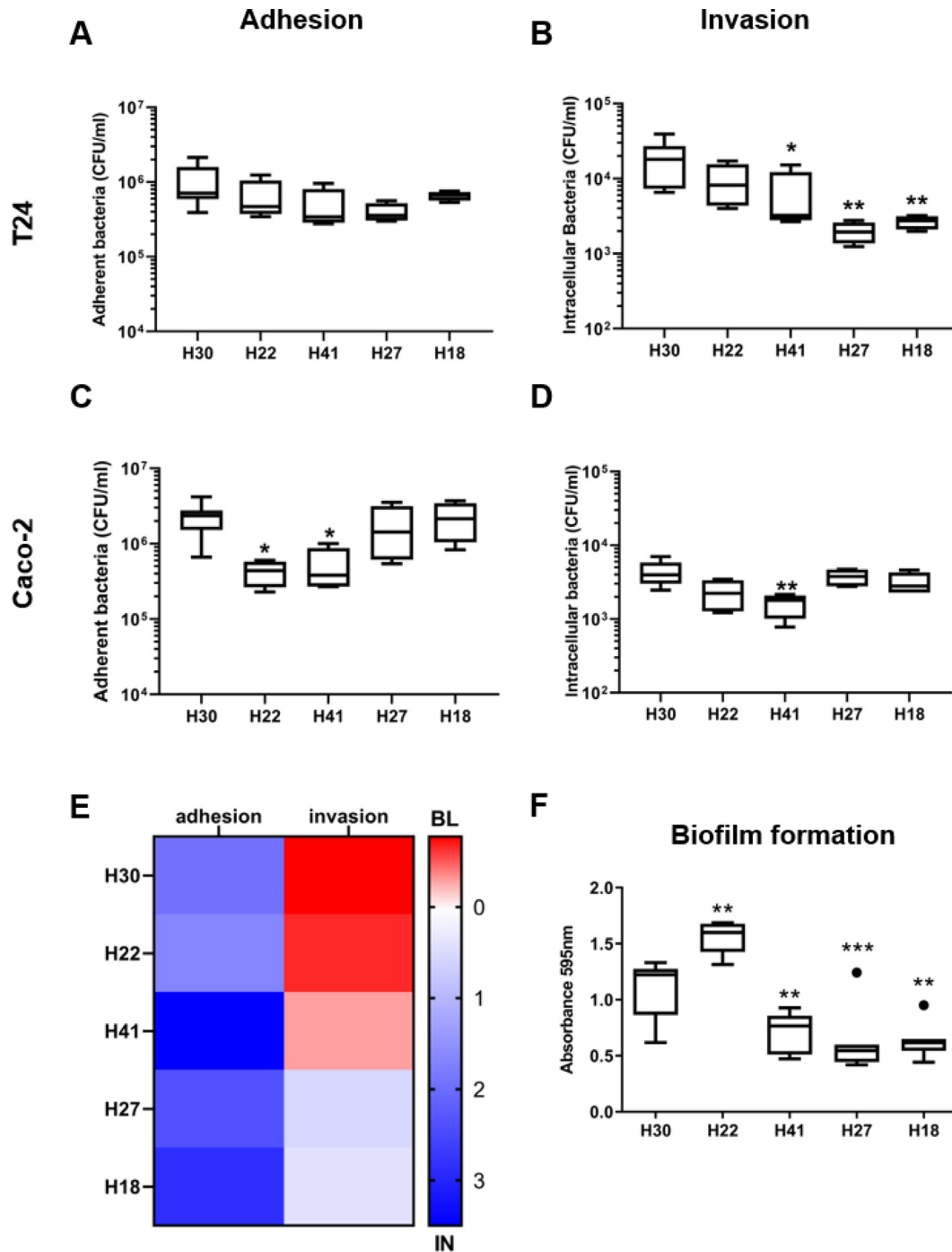
Fig. 6. Natural mutations in FimH impact its cellular function. Adhesion (A&C) and invasion (B&D) of EC958 WT and isogenic *fimH* mutants into human bladder T24 cells (A&B) and colonic Caco-2 cells (C&D). All assays were performed at an MOI of 10. E. Preferential bacterial adhesion and invasion into human intestinal (IN, Caco-2; blue) or bladder (BL, T24; red) epithelial cells by isogenic EC958 *fimH* mutants. Data for each isogenic mutant were centralised against the mean adherent CFU/ml or intracellular CFU/ml from bladder and intestinal cell assays, and then the mean ratio of IN over BL centralised values were calculated and shown as a heatmap for each strain for both adhesion and invasion. F. Biofilm formation by EC958 isogenic *fimH* mutants. Group mean differences were tested by one-way ANOVA followed by Dunn's multiple comparisons tests against H30 with significance set at the 5 % level. *p* values shown as: *: $p < 0.05$; **: $p < 0.01$; and ***: $p < 0.001$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5. Discussion

FimH is the tip adhesin of type 1 fimbriae, which are ubiquitously found in *E. coli*. FimH mediates bacterial attachment to host cells through targeting mannose moieties on host cell surface glycolipids and glycoproteins, contributing to pathogen colonisation and invasion into different host niches. The functional heterogeneity conferred by extensive allelic variation in FimH has been long recognised [29] and has provided new insights into how different FimH variants have evolved to mediate *E. coli* colonisation of various host niches [6,17,30]. With more than 60 wildtype *fimH* alleles reported to date and only two structures available in the Protein Data Bank (FimH2 and FimH27 capturing limited conformational snapshots), we posit there is an important gap in our current

understanding of FimH structure–function in successful *E. coli* lineages. In this respect, the emergence and subsequent global expansion of *E. coli* ST131 MDR clade C [23] prompted us to characterise its defining *fimH*30 allele in comparison to other alleles from this lineage or previously well-characterised alleles from UPEC. Here, we combined MD simulations, using appropriate timescales for statistical sampling of FimH dynamics and non-equilibrium simulations to investigate shear force effects, along with biological assays in a clinically relevant pathogenic background (a reference ST131 multidrug resistant isolate), to directly determine the impact of naturally occurring mutations in FimH on protein allostery and cellular function.

Classically, the allosteric mechanism of FimH has been described as a step-wise process, where FimH on the fimbrial tip is in the associated (T state) conformation under static conditions and undergoes domain separation upon mannose binding or shear force, resulting in increased mannose binding affinity [12,31]. With over 60 PDB files of FimH protein structures available to date [21], different conformations of FimH proteins have been revealed. Recently, Kalas, et al. [15] reported that the apo FimH18 protein in solution samples various conformations including previously solved associated (T) and elongated (R) conformations. The authors then proposed a revised FimH allostery model where the FimH protein pre-exists in both the T and R states, which may follow different pathways to engage with mannose. In our 300 ns nonbiased MD simulations, for all FimH variants that started in the T state bound to mannose, mannose binding alone was not sufficient to induce the domain separation characteristic of the R state. Moreover, when simulations were initiated from the mannose bound R state, none of the FimH variants undergo domain re-association to convert to a T state. This might suggest that the T and R states are distinctive conformations and that there might be a high free energy barrier between the two states. In addition, in our ASMD simulations of the T state, on application of force, the two domains separate to adopt an elongated conformation that readily re-associates back to the T state upon removal of force. Incidentally, the force induced elongated FimH configuration does not correspond to the previously experimentally characterised R state. Taken together our simulation data suggest two distinct scenarios: (*i*) the ASMD simulations did not allow overcoming the energetic barrier between the T and R states and the resulting elongated form may represent an intermediate between the two states; or (*ii*) that FimH can adopt distinct types of R conformations, one arising in a step-wise fashion from the T state upon application of force and another representing a pre-existing relaxed conformation. The latter is also in agreement with experimental data documenting the length of time that *E. coli* pause on mannosylated surfaces, where a double exponential decay in the duration of the pause was observed [32]. These data strongly suggest the co-existence of two distinct FimH conformational states, which engage with mannose through different mechanisms. Collectively, our analysis of natural FimH adhesin variants support the allostery model proposed by Kalas, et al. [15], where FimH exists in two distinctive conformational populations from both the T and R states.

Through MD simulations of clinically relevant FimH variants we demonstrate that naturally occurring mutations in FimH outside the conserved mannose binding pocket can directly affect protein interdomain interactions and influence protein ligand binding stability in different conformational states. In all reported FimH structures solved in R state conformations, residue R166 in the pilin domain is not in direct contact with the lectin domain [21]. In contrast, all reported FimH structures solved in T state conformations maintain interdomain interactions mediated by the R166 residue. We also observed the formation of several hydrogen bonds in all FimH variants, except FimH30 where they were lost upon the R166H mutation. This may facilitate a shift in the equilibrium of

conformational states from T to R by abolishing several key interdomain interactions maintained by R166 in the T state. In consequence, when simulated in the T state, FimH30 maintained the least persistent interactions with mannose in comparison to other tested variants. This is also supported by our structure analysis where the distal end of the pilin domain in FimH30 showed larger disordered movement in the T state compared to other tested variants. In contrast, when simulated in the R state, FimH30 mannose binding was maintained by higher frequency interactions. Together, we propose that the R166H mutation in FimH30 might lead to a conformational equilibrium shift towards R state conformations and a relatively tight binding to mannose. This may in part explain the high mannose binding ability previously reported for FimH30 and related variants with the R166H mutation, despite this mutation localising outside the mannose binding pocket [16]. Also our finding that, in the absence of shear force, the hydrogen network between mannose and residues in the binding pocket is more persistent in R over T state conformations, suggests that FimH30 could serve as a natural FimH variant to reveal another distinctive R state conformational snapshot.

Our chromosomally engineered isogenic *fimH* mutants allowed us to directly interrogate the functional consequences of distinct FimH variants in the most clinically relevant strain background. Therefore, our functional analyses are not subject to strain variation and are more sensitive to functional differences mediated by distinct FimH variants. We showed that H30 outperforms isogenic H22 and H41 strains for adherence to human colonic epithelial cells. A similar trend was previously reported for these dominant ST131 variants for binding to monomanose and trimannnose when overexpressed in an *E. coli* K-12 strain background [16]. Ours findings may also in part explain epidemiological observations reporting ST131 clade C (FimH30) strains as being significantly longer gut colonisers, with an estimated half-life of 13 months, compared to only two to three months half-life by other drug resistant *E. coli* strains [33]. In addition, we found that isogenic FimH variant strains displayed overall higher adhesion levels to colonic cells than to bladder cells and high invasion levels into bladder cells over colonic cells. This suggests that the tissue tropism is not mainly affected by affinity differences between FimH variants, and likely driven by differences in host cell receptors. Indeed, affinity of FimH to different mannosylated molecules can differ by 100-fold [34], while allelic variation in FimH alters the affinity to a specific mannose by only 2–12 fold [35]. Thus, FimH is likely to contribute to human intestinal colonisation, which serves as the main UPEC reservoir, yet favours invasion into bladder cells where it facilitates UTI establishment and likely persistence. This also suggests that higher cell adhesion does not ensure more intracellular bacteria, and that other factors contribute to epithelial cell invasion by UPEC. To our knowledge, this is the first study to report differential tissue tropism by a set of clinically relevant isogenic *fimH* mutants.

In conclusion, we report evidence supporting that the previously solved T and R FimH conformations represent two distinctive populations in the conformational landscape of this protein, which appears more expansive than previously anticipated and that mutations in FimH30 influence the transitions between conformational states, thereby optimising FimH30 adhesin function for ST131 virulence.

6. Materials and methods

6.1. Molecular dynamics simulation methods

FimH was modelled for the FimH2, FimH18, FimH22, FimH27, FimH30, and FimH41 protein variants (Fig. 1B) in the T state con-

formation (PDB ID: 4XOE) and R state conformation (PDB ID: 4XOB) with mannose bound (Fig. S1). In the R state crystal structure (PDB ID: 4XOB), FimH is interacting with the β-strand of FimF (donor strand F or DsF), however DsG of FimG is the native donor peptide to FimH. Therefore, DsF was replaced with the DsG prior to modelling. PropKa [36,37] was run on the crystal structure and all amino acids were determined to be in their canonical protonation state (pH = 7), with the exception of C3 and C161, which form a disulphide bond. Molecular dynamics (MD) simulations were performed using AMBER16 [38] with the AMBER ff14FB force field [39] for the protein, and the Glycam_06j-1 force field [40] for the mannose. Ions were modelled with the parameters developed by Joung and Cheatham [41]. The complexes were solvated in a TIP3P octahedral box such that the minimum distance between the box edge and FimH was 10 Å. NaCl was added to the systems to a final concentration of 0.150 M, while additional Cl^- ions were added to neutralize the charge of each system.

Systems were minimized in a step-wise procedure with minimization of 1) water and counter ions, 2) protein hydrogen atoms, 3) FimH, DsG and mannose, and 4) the entire system. Each step of minimization included 2500 steps of steepest descent and 2500 steps of conjugate gradient minimization, while applying a 50 kcal/(mol·Å^2) restraint on all other atoms. The systems were then heated from 0 to 310 K in 6 steps, with each 10 ps simulation having an increased temperature (i.e., 10, 60, 120, 180, 210, 260, and 310 K). Heating was performed with a 1 fs time step, 10 kcal/(mol·Å^2) restraint on the solute and the Langevin thermostat (γ = 1.0). The systems were then equilibrated over five 20 ps simulations while reducing the constraint on the solute (i.e., 20, 15, 10, 5 and 1 kcal/(mol·Å^2)). Equilibration was performed at 310 K with a 2 fs time step. Unconstrained production simulations of 300 ns were performed in triplicate for each system. Throughout all simulations, the pressure was maintained at 1 bar using isotropic position scaling and the Berendsen barostat (τ_P = 2.0 ps), SHAKE, the periodic boundary condition, a nonbonded cut-off of 8 Å, and long-range electrostatic interactions were treated with the Particle Mesh Ewald (PME) method. Analysis was performed on frames spaced by 0.2 ns using the cpptraj module of AMBER16 [38]. Visualization of the simulations and generation of figures was performed using VMD 1.9.2 [42].

To further investigate the shear force induced conformational change from an initial FimH T state, adaptive bias steered molecular dynamics (ASMD) was used [43,44]. In ASMD the reaction coordinate is separated into stages and the potential of mean force (PMF) is sampled independently in each stage using the Jarzynski equality [45]. Here, a single structure is chosen as the starting point of the subsequent stage based on the simulation in which the associated work is closest to the Jarzynski average. In the simulations, the distance between the mannose ligand (O5) and pilin domain (using the CA of Thr287 from the DsG as a reference) was pulled from 55 to 95 Å with a velocity of 10 Å/ns. The reaction coordinate was separated into 20 stages and 100 simulations were performed for each stage.

6.2. Bacterial strains and culture conditions

E. coli ST131 strain EC958 [26] was used as a reference MDR clinical isolate that naturally encodes the *fimH30* gene. A previously constructed EC958 *ΔfimH* mutant [3] served as the background strain for generating a panel of isogenic EC958 strains each harbouring a different wild-type *fimH* allele (*fimH27, fimH18, fimH22, fimH41* and *fimH30* that was re-introduced as a control; Table S1). Each *fimH* sequence flanked by identical 500-bp upstream and downstream homology regions (from the EC958 genome) [26,46] were synthesized and provided by Epoch Life Science into pBSK plasmids (Table S1). Plasmids carrying distinct

fimH alleles were used as templates for the amplification of linear DNA cassettes that were exchanged into the EC958 *fimH* chromosomal locus by λ-red mutagenesis [47] with modification [26]. All EC958 strains were cultured in LB-Lennox and passaged statically for three rounds of 48 h incubation at 37 °C to maximally switch on type 1 fimbriae production as previously described [27]. Fim-enriched strains were stored at −80 °C and used as inoculating cultures in all assays unless stated otherwise.

6.3. Determination of *fimS* orientation

The orientation of *fimS* was determined by PCR and restriction digestion as described previously [48]. Briefly, a *fimS* containing DNA fragment was PCR amplified from boiled bacterial cell lysate using primers TCGTTTTGCCGGATTATGGG and AGTGAACGGTCCCACCATTAACC. The *fimS* orientation was then determined by Hinf1 restriction digestion of the amplicons followed by separation on a 2 % agarose gel. The percentage of the fragments corresponding to a *fimS* switch-on orientation was estimated by the band intensity ratio of (switch-on/total) using GeneTools (Syngene).

6.4. Confirmation of type 1 fimbriae expression

Semi-quantitative expression of type 1 fimbriae was performed as previously [3]. Briefly, *fim*-enriched bacterial cell lysates were separated by SDS-PAGE and subjected to Western immunoblotting with anti-FimA antibody [3] and anti-GroEL antibody (Invitrogen). Relative band intensities (FimA/GroEL) were normalized against that of WT (defined as 1.00) and were used to compare between blots. Production of functional type 1 fimbriae on the bacterial cell surface was confirmed by yeast cell agglutination [26] with modifications. Bacterial cultures (OD_{600} of 5.0) were 2-fold serially diluted and mixed with an equal volume of 5 % (w/v) yeast cell (*Saccharomyces cerevisiae*) solution in PBS in the presence or absence of 1 % methyl α-D-mannopyranoside (Sigma, M6882). The highest dilution with visible agglutination within 5 min was recorded. Yeast agglutination titers were measured from four independent experiments.

6.5. Mammalian cell culture and in vitro cell infection assays

Human bladder epithelial cells T24 (ATCC HTB-4, in McCoy's 5A medium [modified, Invitrogen]), human colonic epithelial cells Caco-2 (ATCC HTB-37, in Dulbecco's modified Eagle's medium [DMEM]), and T84 (ATCC CCL-248, in DMEM and Ham's F-12 nutrient mixture) were maintained in medium supplemented with 10 % fetal calf serum (Invitrogen). Bacterial adhesion and invasion assays were performed as described previously [3,26]. Briefly, cells grown confluent on a 24-well plate were infected with bacteria at the multiplicity of infection (MOI) of 10 and were incubated at 37 °C, 5 % CO_2 for 1 h, followed by 5 washes with PBS. Cell monolayers were then lysed, and adherent bacteria were enumerated by serial dilution and plating onto LB agar. The intracellular bacterial load was determined by an additional 1 h gentamicin treatment before enumeration as above.

6.6. Bacterial biofilm formation

Biofilm formation was performed as described previously [28]. Fim-enriched bacteria were inoculated into a 96-well round-bottom plate (Greiner) at the OD_{600} of 0.03 in LB and were incubated statically at 37 °C for 18 h. The biofilm was then quantified by staining with 0.1 % (w/v) crystal violet for 30 min at 4 °C, followed by solubilization in ethanol-acetone (80:20 v/v) solution, and the absorbance was measured at 595 nm in a XMark microplate reader (BioRad).

CRediT authorship contribution statement

Jilong Qin: Methodology, Formal analysis, Investigation, Writing – original draft, Visualization. **Katie A. Wilson:** Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Sohinee Sarkar:** Methodology, Formal analysis, Investigation, Visualization, Writing – review & editing. **Begoña Heras:** Investigation, Resources, Visualization, Writing – review & editing. **Megan L. O'Mara:** Conceptualization, Methodology, Formal analysis, Validation, Investigation, Resources, Supervision, Writing – review & editing. **Makrina Totsika:** Conceptualization, Methodology, Formal analysis, Validation, Investigation, Resources, Supervision, Writing – review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Author contributions

J.Q., K.A.W., M.L.M., and M.T. designed research; K.A.W., S.S and M.T, performed research; J.Q., K.A.W., B.H., M.L.M., and M.T. analyzed data; and J.Q., K.A.W., and M.T. wrote the paper. All authors edited the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.08.040.

References

[1] Kaper JB, Nataro JP, Mobley HL. Pathogenic Escherichia coli. Nat Rev Microbiol 2004;2:123–40.
[2] Wright KJ, Seed PC, Hultgren SJ. Development of intracellular bacterial communities of uropathogenic Escherichia coli depends on type 1 pili. Cell Microbiol 2007;9:2230–41.
[3] Sarkar S et al. Intestinal Colonization Traits of Pandemic Multidrug-Resistant Escherichia coli ST131. J Infect Dis 2018;218:979–90.
[4] Spaulding CN et al. Selective depletion of uropathogenic E. coli from the gut by a FimH antagonist. Nature 2017;546:528–32.
[5] Poole NM et al. Role for FimH in extraintestinal pathogenic Escherichia coli invasion and translocation through the intestinal epithelium. Infect Immun 2017;85.
[6] Schwartz DJ et al. Positively selected FimH residues enhance virulence during urinary tract infection by altering FimH conformation. Proc Natl Acad Sci U S A 2013;110:15530–7.
[7] Eris D et al. The conformational variability of FimH: which conformation represents the therapeutic target? ChemBioChem 2016;17:1012–20.
[8] Hung CS et al. Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection. Mol Microbiol 2002;44:903–15.
[9] Zhou G et al. Uroplakin Ia is the urothelial receptor for uropathogenic Escherichia coli: evidence from in vitro FimH binding. J Cell Sci 2001;114:4095–103.
[10] Sauer MM et al. Catch-bond mechanism of the bacterial adhesin FimH. Nat Commun 2016;7:10738.
[11] Aprikian P et al. The bacterial fimbrial tip acts as a mechanical force sensor. Plos Biol 2011;9.
[12] Thomas WE, Trintchina E, Forero M, Vogel V, Sokurenko EV. Bacterial adhesion to target cells enhanced by shear force. Cell 2002;109:913–23.
[13] Thomas WE, Nilsson LM, Forero M, Sokurenko EV, Vogel V. Shear-dependent 'stick-and-roll' adhesion of type 1 fimbriated Escherichia coli. Mol Microbiol 2004;53:1545–57.
[14] Aprikian P et al. Interdomain interaction in the FimH adhesin of Escherichia coli regulates the affinity to mannose. J Biol Chem 2007;282:23437–46.
[15] Kalas V et al. Evolutionary fine-tuning of conformational ensembles in FimH during host-pathogen interactions. Sci Adv 2017;3:e1601944.
[16] Paul S et al. Role of homologous recombination in adaptive diversification of extraintestinal Escherichia coli. J Bacteriol 2013;195:231–42.
[17] Dreux N et al. Point mutations in FimH adhesin of Crohn's disease-associated adherent-invasive Escherichia coli enhance intestinal inflammatory response. PLoS Pathog 2013;9.
[18] Chen SL et al. Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding. Proc Natl Acad Sci U S A 2009;106:22439–44.
[19] Spaulding CN et al. Functional role of the type 1 pilus rod structure in mediating host-pathogen interactions. Elife 2018;7.
[20] Weissman SJ et al. High-resolution two-locus clonal typing of extraintestinal pathogenic Escherichia coli. Appl Environ Microbiol 2012;78:1353–60.
[21] Magala P, Klevit RE, Thomas WE, Sokurenko EV, Stenkamp RE. RMSD analysis of structures of the bacterial protein FimH identifies five conformations of its lectin domain. Proteins 2020;88:593–603.
[22] Choudhury D et al. X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. Science 1999;285:1061–6.
[23] Petty NK et al. Global dissemination of a multidrug resistant Escherichia coli clone. Proc Natl Acad Sci U S A 2014;111:5694–9.
[24] Chen SL et al. Identification of genes subject to positive selection in uropathogenic strains of Escherichia coli: a comparative genomics approach. Proc Natl Acad Sci U S A 2006;103:5977–82.
[25] Tchesnokova V et al. Integrin-like allosteric properties of the catch bond-forming FimH adhesin of Escherichia coli. J Biol Chem 2008;283:7823–33.
[26] Totsika M et al. Insights into a multidrug resistant Escherichia coli pathogen of the globally disseminated ST131 lineage: genome analysis and virulence mechanisms. PLoS ONE 2011;6:e26578.
[27] Sarkar S et al. Comprehensive analysis of type 1 fimbriae regulation in fimB-null strains from the multidrug resistant Escherichia coli ST131 clone. Mol Microbiol 2016;101:1069–87.
[28] Sarkar S, Vagenas D, Schembri MA, Totsika M. Biofilm formation by multidrug resistant Escherichia coli ST131 is dependent on type 1 fimbriae and assay conditions. Pathog Dis 2016;74.
[29] Sokurenko EV, Courtney HS, Abraham SN, Klemm P, Hasty DL. Functional heterogeneity of type 1 fimbriae of Escherichia coli. Infect Immun 1992;60:4709–19.
[30] Pouttu R et al. Amino acid residue Ala-62 in the FimH fimbrial adhesin is critical for the adhesiveness of meningitis-associated Escherichia coli to collagens. Mol Microbiol 1999;31:1747–57.
[31] Rodriguez VB et al. Allosteric coupling in the bacterial adhesive protein FimH. J Biol Chem 2013;288:24128–39.
[32] Thomas W et al. Catch-bond model derived from allostery explains force-activated bacterial adhesion. Biophys J 2006;90:753–64.
[33] Overdevest I, et al., Prolonged colonisation with Escherichia coli O25:ST131 versus other extended-spectrum beta-lactamase-producing E. coli in a long-term care facility with high endemic level of rectal colonisation, the Netherlands, 2013 to 2014. Euro Surveill 21 (2016).
[34] Bouckaert J et al. The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of Escherichia coli pathotypes. Mol Microbiol 2006;61:1556–68.
[35] Sokurenko EV et al. Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin. Proc Natl Acad Sci U S A 1998;95:8922–6.
[36] Sondergaard CR, Olsson MH, Rostkowski M, Jensen JH. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. J Chem Theory Comput 2011;7:2284–95.
[37] Olsson MH, Sondergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. J Chem Theory Comput 2011;7:525–37.
[38] Case RMBDA, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, Giese TJ, et al., AMBER 2016 (University of California, San Francisco, 2016).
[39] Maier JA et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11:3696–713.
[40] Kirschner KN et al. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J Comput Chem 2008;29:622–55.
[41] Joung IS, Cheatham 3rd TE. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. J Phys Chem B 2009;113:13279–90.

[42] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14(33–38):27–38.

[43] Ozer G, Valeev EF, Quirk S, Hernandez R. Adaptive steered molecular dynamics of the long-distance unfolding of neuropeptide Y. J Chem Theory Comput 2010;6:3026–38.

[44] Bureau HR, Merz Jr DR, Hershkovits E, Quirk S, Hernandez R. Constrained unfolding of a helical peptide: implicit versus explicit solvents. PLoS ONE 2015;10:e0127034.

[45] Jarzynski C. Nonequilibrium equality for free energy differences. Phys Rev Lett 1997;78:2690–3.

[46] Forde BM et al. The complete genome sequence of *Escherichia coli* EC958: a high quality reference sequence for the globally disseminated multidrug resistant *E. coli* O25b:H4-ST131 clone. PLoS ONE 2014;9:e104400.

[47] Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 2000;97:6640–5.

[48] Schembri MA, Olsen PB, Klemm P. Orientation-dependent enhancement by H-NS of the activity of the type 1 fimbrial phase switch promoter in *Escherichia coli*. Mol Gen Genet 1998;259:336–44.