Positive selection in the two-domain type 1 pilus adhesin FimH enhances *Escherichia coli* fitness in urinary tract infection (UTI). We report a comprehensive atomic-level view of FimH in two-state conformational ensembles in solution, composed of one low-affinity tense (T) and multiple high-affinity relaxed (R) conformations. Positively selected residues allosterically modulate the equilibrium between these two conformational states, each of which engages mannose through distinct binding orientations. A FimH variant that only adopts the R state is severely attenuated early in a mouse model of uncomplicated UTI but is proficient at colonizing catheterized bladders in vivo or bladder transitional-like epithelial cells in vitro. Thus, the bladder habitat has barrier(s) to R state–mediated colonization possibly conferred by the terminally differentiated bladder epithelium and/or decay receptors in urine. Together, our studies reveal the conformational landscape in solution, binding mechanisms, and adhesive strength of an allosteric two-domain adhesin that evolved “moderate” affinity to optimize persistence in the bladder during UTI.

**INTRODUCTION**

Bacterial pathogens adhere to and establish footholds in favorable habitats within their host. Adhesion often confers fitness advantages to bacterial pathogens by promoting access to essential nutrients, resistance to displacement by fluid flow, or evasion of immune responses (1–3). Bacterial pathogens have evolved numerous ways in which to achieve adhesion either specifically or nonspecifically through extracellular structures that include capsules, extracellular polymeric substances, pili/fimbriae, flagella, and other diverse adhesins (1, 4, 5). Many of these adhesiveness mechanisms are conformationally flexible and can undergo changes in response to environmental stimuli. For example, the type 1 pilus adhesin FimH is a multifunctional adhesin that noncovalently joins FimH to the pilus tip, and (iii) a five–amino acid linker that connects the two domains (Fig. 1A). It has been shown that three positively selected residues (residues 27, 62, and 163), which lie outside the mannose-binding pocket, modulate conformational changes in FimH (Fig. 1B) (14). Crystal structures of FimH in various stages of pilus assembly have revealed that FimH forms a single conformation but FimHLD has at least two conformational states that show distinct affinities for mannose (Fig. 1B) (15–19). Before pilus assembly, the incomplete immunoglobulin (Ig)–like fold of FimHNR is stabilized via a donor strand complementation interaction with the chaperone FimC in the periplasm (15, 20). Crystal structures have revealed that FimH adopts an elongated conformation when bound to FimC, wherein FimHLD and FimHNR do not interact with each other and FimHLD is in a high-affinity mannose-binding state. Thereafter, the FimC–FimH (FimCH) complex initiates pilus assembly at the outer membrane, where the FimD usher catalyzes donor strand exchange (DSE), in which the N-terminal extension (Nte) of FimG displaces FimC to complete the Ig-like fold of FimHNR (21–23). Crystal structures show that FimH adopts a compact conformation when bound to FimG, wherein FimHLD and FimHNR interact closely with one another and FimHLD adopts a low-affinity mannose-binding state. To succinctly describe the correlation between all the structural and functional properties of FimH in its two known conformations (Fig. 1B), we adapted the nomenclature used in the Monod-Wyman-Changeux model of allostery to herein refer to the high-affinity conformation as the relaxed (R) state and the low-affinity conformation as the tense (T) state. On the basis of these structures, an allosteric model that links the conformation of FimH to its mannose-binding function has been proposed (Fig. 1C) (17, 24). This model posits that FimHNR can allosterically diminish the ability of FimHLD to bind mannose through interactions with the base of FimHLD. Conversely, mannose binding to FimHLD appears to promote FimHLD conformations that no longer interact with FimHNR. However, the structural and evolutionary basis for conformational allostery in FimH at the pilus tip remains incompletely understood. Here, we characterize the conformational ensembles, dynamics, and binding mechanisms of FimH variants in a tip-like setting in solution. We discovered a mechanism whereby positively
selected residues allosterically modulate a two-state conformational equilibrium in FimH to adapt a “moderate” mannose-binding affinity for optimal colonization of the bladder during UTI.

RESULTS

FimG_{Nte}H functions as a type 1 pilus tip-like setting

Biophysical experiments to directly probe conformations of FimH within an in vitro reconstituted pilus tip have remained challenging because of the inherent tendency of pilus subunits, or pilins, to self-polymerize and introduce heterogeneity in solution (14). Thus, we created a “minimal” tip-like FimH complex through a spontaneous in vitro DSE reaction between FimH and a peptide corresponding to the Nte of FimG (FimG_{Nte} residues 1 to 15) (Fig. 2A). This strategy was based on previous work that demonstrated how Nte peptides stabilize pilin domains (22). We investigated the conformational and functional effects of allelic variation in positively selected residues A27, A62, and V163 (13, 14) on FimH within the FimG_{Nte}H complex. This was done using the FimH sequence of the UPEC strain UTI89 (25) as the wild-type (WT) background. We produced FimG_{Nte}H complexes with WT and variant FimH sequences (A62S, A27V/V163A) and the Q133K variant, a binding pocket mutation that abrogates adhesion (13). As with other DSE interactions (14), all purified FimG_{Nte}H complexes exhibited high stability because all FimH variants remained bound to FimG_{Nte} in the SDS sample buffer at room temperature (~31 kDa) but dissociated when boiled (~29 kDa) (Fig. 2B). Moreover, these variant FimG_{Nte}H complexes displayed the same differential binding affinities toward human mannosylated glycoproteins as those affinities previously determined by equilibrium binding analysis of FimCGH variants (Fig. 2C). A62S shows low affinity, WT displays intermediate or moderate affinity, and A27V/V163A exhibits high affinity for mannose. However, each of the binding-competent FimH variants displayed equal affinity for mannose when in a FimCH complex (14). Thus, we hypothesize that the observed adhesive differences among FimG_{Nte}H complexes are the result of shifts in conformational equilibria of FimH in a tip-like setting caused by sequence variation. In summary, the FimG_{Nte}H complex reliably functions as a minimal tip-like setting and appears well suited for the examination of FimH conformation by molecular biophysics.

Crystal structures of ligand-free FimG_{Nte}H variants reveal diverse conformations in a tip-like setting

Apo x-ray crystal structures of FimG_{Nte}H A62S and FimG_{Nte}H A27V/V163A were solved at 1.96 and 2.6 Å resolutions, respectively, to investigate whether conformational changes determine the differences in mannose affinity of FimH in a tip-like setting. Both structures were solved by molecular replacement (MR) using FimH_{PD} and either the low-affinity or high-affinity FimH_{LD} from previously solved crystal structures as search models (table S1). The FimG_{Nte}H A62S crystal structure reveals four FimG_{Nte}H copies in the unit cell (fig. S1A), each adopting a T conformation (Fig. 2D). This T conformation is nearly equivalent to the T conformations of FimH (from E. coli strain F18) observed in (i) a FimCFGH crystal structure (17) and (ii) the apo and heptyl mannoside–bound FimG_{Nte}H crystal structures (fig. S2) (26). In these structures, FimH_{LD} adopts the same overall compacted shape, widened β-sandwich fold, and displaced binding loops that confer weak affinity for mannose (fig. S2). However, structural alignment of FimH_{PD} in the FimG_{Nte}H A62S complex to FimH_{PD} from the WT complexes in the T conformation exposes a ~7 Å rigid-body tilt of FimH_{LD}, suggesting a small vibrational degree of conformational flexibility in the T state. Despite this tilt, the T conformer of FimG_{Nte}H A62S forms the same FimH_{PD}–FimH_{LD} interface, in which hydrophobic interactions and interdomain hydrogen bonds bury the linker and an “insertion loop” (residues 109 to 124) (17) within a local hydrophobic core (Fig. 2E). Moreover, three of the four FimG_{Nte}H A62S copies lack visible electron density in regions of mannoside-binding loop 1 (residues 11 to 16), particularly at residues 13 and 14, suggesting that loop 1 is highly dynamic within the crystal (fig. S1B).
In contrast, the crystal structure of FimGNteH A27V/V163A reveals a novel R conformation, wherein the orientations of the FimHLD and FimHPD form a 90° bend (Fig. 2D). This observed conformation represents a substantial structural deviation from the elongated orientation of FimHLD and FimHPD observed in the R conformation of FimH bound to the FimC chaperone; thus, we refer to this crystal structure as the bent R conformation and the previously identified chaperoned state as the elongated R conformation (15). This bent R conformation is also distinct from a recently solved crystal structure of FimFNteH bound to the pental mannoside, which shows FimHLD adopting a ~45° bend angle and a different orientation with respect to FimHPD (fig. S2) (26). Structural alignment of FimHLD of mannose-bound FimCH (colored white) to FimHLD of mannose-bound FimCH (colored blue) are indicated as black dotted lines.

Positively selected residues and ligand binding modulate conformational ensembles of FimGNteH variants

Structural ensembles of FimH in solution were probed by small-angle x-ray scattering (SAXS) of FimH variants in FimGNteH or FimCH complexes. In a FimCH complex, all binding-competent FimH variants have the same affinity for mannose (14). They also appeared to adopt the same high-affinity elongated R conformation based on the structural comparison heat map ($\chi^2 = 0.4$ to 0.6) (fig. S3, A and B) and rigid-body modeling of the FimCH crystal structure to the SAXS profiles for each of the FimCH variants ($\chi = 0.9$ to 1.3) (fig. S3C). In contrast, each of these FimH variants had a drastically different affinity for mannose when in complex with FimG (Fig. 2C) (14). Correspondingly,
in the FimGNeH complex, each variant adopted very different relative conformation(s) ($\chi^2 = 0.6$ to 7.4), with the low-affinity A62S and high-affinity A27V/V163A variants being most dissimilar from one another (Fig. 3A). Further, radius of gyration ($R_g$) and maximal intramolecular distance ($D_{max}$) measurements indicated that A27V/V163A FimH is larger in overall shape by ~2 Å and more extended by ~5 Å than all other tested variants (Fig. 3B). Thus, we propose that positively selected residues allosterically influence the conformation(s) of FimH in pilus tips in solution, thus modulating mannose-binding affinity.

Binding of the FimGNeH variants to a high-affinity mannose analog, termed mannoside 4Z269 (27), caused various degrees of conformational perturbation, with the WT variant undergoing the greatest change, as indicated by the structural comparison map (Fig. 3C). Upon binding 4Z269, all binding-competent FimGNeH variants underwent a rightward shift in the pair distance distribution function [$\rho(r)$] with corresponding increases in $D_{max}$ and $R_g$, suggesting extension of the shape of FimH (Fig. 3B). These changes were less pronounced for the A27V/V163A variant, which preexists in a high-affinity state. These effects were not observed in the control, binding null mutant Q133K. Ab initio models of the average conformation in solution, as represented by low-resolution molecular envelopes, revealed distinctive changes in overall protein shape among variants and between apo and 4Z269-bound states (Fig. 3D).

To resolve the solution structure(s) of FimH variants in a tip-like setting in the absence or presence of 4Z269, we evaluated the agreement between SAXS profiles and crystal structures of FimGNeH in four distinct conformations [T, bent R, a previously identified distinct bent R (26), and elongated R] either individually or in combination with one another. Rigid-body modeling of individual structures (28) against these scattering profiles resulted in moderate-to-poor goodness of fit ($\chi^2 = 2$ to 10) (fig. S3D). In contrast, multistate rigid-body modeling determined specific weighted combinations of multiple conformational models that fit the data better than any individual model. This analysis indicated that in the absence of 4Z269, the A27V/V163A variant exists entirely in the R state, comprising a mix of elongated and bent shapes, whereas WT, Q133K, and A62S adopt various proportions of T and bent R states (fig. S3E). Addition of 4Z269 shifted the equilibrium away from the T conformation and toward a mix of bent and elongated R conformations for all variants except Q133K. However, multistate rigid-body modeling of SAXS profiles for FimGNeH WT and A27V/V163A variants in the presence of 4Z269 resulted in moderate goodness of fit, which indicates that the four abovementioned static conformations do not completely represent the solution ensemble and implies additional solution conformations that have not been determined by crystallography. Together, the improvements in goodness of fit from multistate modeling strongly intimate a conformational equilibrium and suggest that the crystal structures used in rigid-body modeling likely represent conformational snapshots and may not necessarily represent “preferred” states or capture the expanse or diversity of the conformational space of FimH in solution. In agreement with the multistate modeling, differential scanning fluorimetry revealed a two-state unfolding behavior for apo FimGNeH WT, Q133K, and A62S, with shifts toward one-state unfolding for A62S and WT variants in the presence of 4Z269 (fig. S3F). Single-state unfolding was observed for FimGNeH A27V/V163A in the presence and absence of 4Z269. Together, these
data strongly suggest that the identity of positively selected residues and mannose binding allosterically influence a preexisting equilibrium of distinct conformations in FimH in solution.

Further, ion mobility–mass spectrometry (IMMS) was used as a direct method to investigate conformational distributions of FimH due to the ability of IMMS to resolve protein collision cross sections (CCSs). All FimGNteH and 4Z269-bound FimGNteH complexes remained intact when they were electrosprayed into the mass spectrometer, facilitating downstream analysis of the tip-like setting by IMMS (fig. S4, A and B). The CCS distributions at low collision energy (CE) of the WT, Q133K, and A62S variant FimGNteH complexes revealed a major peak and shoulder, whereas only one predominant peak with positive skew was observed for the A27V/V163A variant (Fig. 4A). Comparison of these distributions indicated that the A27V/V163A distribution is shifted to the right (or “right-shifted”) and displays a greater mean CCS than do the other variant distributions, consistent with its larger average shape (by \( R_g \) and \( D_{max} \)) in solution. Given their non-normal shape, these distributions were modeled as a sum of two Gaussian curves. The WT, Q133K, and A62S variant FimGNteH complexes were each similarly best explained by two curves that have mean CCS values of 2464 to 2472 Å² and 2606 to 2636 Å², which we respectively label \( A \) and \( B \) (Fig. 4B and table S2). In contrast, the A27V/V163A FimGNteH profile was best explained by two highly overlapping curves with mean CCS values of 2480 and 2565 Å², which we respectively label \( A \) and \( C \) given their positions relative to the curves assigned for the other three FimH variants. Binding of 4Z269 caused a rightward shift in the overall CCS distributions of WT, A62S, and A27V/V163A (Fig. 4B). 4Z269-bound WT and A62S FimGNteH displayed CCS distributions with right-shifted \( A \) and right-shifted \( B \), indicating that the two major structural species can engage 4Z269. In contrast, 4Z269-bound A27V/V163A FimGNteH displayed a normal CCS distribution adequately explained by one well-fitting curve because modeling with two Gaussian curves indicated that right-shifted \( A \) and \( C \) overlap too extensively for an unambiguous assignment.

As expected, no FimGNteH Q133K:4Z269 complex was observed (fig. S4B). In all cases, CCS distributions steeply increased with elevation of CE, indicative of protein unfolding and confirming that the CCS distributions analyzed above represent folded protein conformations (fig. S4C). In summary, FimGNteH WT, Q133K, and A62S in the gas phase take on an equilibrium of two distinct conformations with partially overlapping CCS distributions, whereas FimGNteH A27V/V163A adopts an equilibrium of two very similar conformations with highly overlapping CCS distributions. Thus, we have resolved two to three different structural conformations among FimGNteH variants in the gas phase. However, the width of the measured CCS distributions and mixture of structural populations likely mask multiple underlying conformational substates that dynamically interconvert within this conformational landscape. The conformational states represented by peaks \( A \), \( B \), and \( C \) may relatively correspond to bent \( R \), \( R \), and elongated \( R \) states, respectively, but a detailed examination of the presumed conformational dynamics of the \( T \) and \( R \) states is necessary to assess the validity of these structural assignments.

**FimGNteH samples expansive conformational phase space composed of restrained \( T \) state and dynamic \( R \) state**

To gauge the dynamic behavior of FimH in the FimGNteH WT tip-like setting, we performed unrestrained molecular dynamics (MD) simulations in four replicates, each starting from the \( T \), bent \( R \), and elongated \( R \) states, respectively, but a detailed examination of the presumed conformational dynamics of the \( T \) and \( R \) states is necessary to assess the validity of these structural assignments.

**Fig. 4. Conformational distributions of free and 4Z269-bound FimGNteH variants isolated in the gas phase, as revealed by IMMS. (A) CCS distributions of intact FimGNteH variant complexes measured by IMMS. (B) Comparison of CCS distributions of free (solid line) and 4Z269-bound (dotted line) FimGNteH variants. The solid and dotted black lines represent fitted Gaussian distributions to apo and ligated FimGNteH, respectively. Fitted Gaussian distributions are labeled by letters, given their mean CCS values. Note that Q133K cannot bind mannose and that the dotted lines for this variant represent CCS distributions and Gaussian fits to an independently measured apo FimGNteH Q133K spectral peak from the sample that was treated with 4Z269.**
of FimHLD relative to FimHPD (fig. S5A). This analysis showed the breadth of conformations sampled in the R state compared to the re- 
strained phase space of the T state simulations (Fig. 5B). Motions among the R conformations resembled movement about a ball-and- 
socket joint, in which a wide range of bends and rotations are acces-
sible but restrained by the architecture of the FimHLD and FimHPD 
interface, which functions as a “socket.” We also observed consider-
able overlap in phase spaces from simulations initiated from the bent 
and elongated R states (Fig. 5B and movies S2 and S3), but no exchange between T and R states under these in silico conditions or in preliminary
trials at 200 to 600 ns, suggesting a high-energy barrier to transition between T and R states that occurs on a much longer time scale (microseconds to milliseconds). Thus, the bent and elongated R conformations represent distinct snapshots of a highly dynamic conformational state, whereas the T conformation is a snapshot of a conformationally restrained structural state, which argues for enthalpy-entropy compensation in the equilibrium between T and R states.

Mannose was then included in simulations of FimGn6H WT to probe its effect on conformational dynamics and mechanisms of ligand binding. The presence of mannose did not significantly alter the simulated dynamic behaviors of the T and R states (Fig. 5C and movies S4 to S6). Strikingly, the R states remained bound to mannose even during the bending, twisting, and rotation that occurred between the two domains throughout the duration of the simulations (Fig. 5C and movies S5 and S6). As expected, the R states oriented their binding pocket loops to clasp mannose tightly with rare, short-lived dissociation events, a signature of high-affinity binding (Fig. 5, C and D). Mannose sat in the R pocket rigidly in a “horizontal” orientation, held in place by hydrogen bonds to residues in loop 2 (N46 and D54) and loop 3 (Q133, N135, and D140) and through packing against I13, I52, and the “tyrosine gate” (defined as binding pocket residues Y48 and Y137), as previously observed (16). In contrast, T states weakly interacted with mannose, characterized by continual motion of mannose within the open T pocket and multiple, long-lasting dissociation events (Fig. 5C and movie S4). When bound to the T state, the mannose ring occasionally sampled the horizontal orientation, as observed in the pocket of the bound R state, but was primarily observed rotated ~45° relative to its horizontal orientation (Fig. 5D). In the T state pocket, the hydroxyl group off the achiral carbon (C6) of mannose was bound to the N terminus of FimH and the carboxyl side chain of FimH residue D54. These two interactions were also observed with mannose bound to FimH in the R state. In contrast to what was observed in the R state, the axial hydroxyl group off the mannose anemic carbon (C1) faced away from the rest of the binding pocket. This “tilted” mannose orientation represents a unique bound conformation and provides the first mechanistic insights into how the T state may contribute to host-pathogen interactions. FimH engages complex mannose-containing glycans on glycoproteins expressed on host epithelial cells (29, 30). Thus, to have biological significance, the tilted mannose orientation must allow stereochemical space for FimH to engage the mannose-containing glycan without steric clashes with the extended glycan chain. Simulations indicated that oligomannose-3 (Man(α1–3)-[Man(α1–6)]-Man) approached and bound the T state pocket in either the tilted or horizontal orientation with about equal frequency, revealing the sterically unhindered manner in which the T state can engage oligomannose receptors (Fig. 5, E and F, and movie S7). As expected, oligomannose-3 bound to the R state only sampled the horizontal orientation (Fig. 5, E and F, and movie S8). Docking of oligomannose-3-chitobiose (Man(α1–3)-[Man(α1–6)]-Man(B1–4)-GlcNac(B1–4)-GlcNac) to distinct time points within the T state simulations indicated that the 1–3 branched terminal mannose of the oligomannose epitope is capable of freely rotating by ~45° within the binding pocket to allow for intercalation of the Man-GlcNac moiety within the sterically accessible, dynamic tyrosine gate (fig. S5B). Previous crystallographic studies and isothermal titration calorimetry experiments indicate that the tyrosine gate displays a high degree of conformational dynamics in the presence of hydrophobic functional groups, such as those within the aglycon moiety of mannosides (31). Thus, the proposed ligand entry and rotation pathways may represent a stepwise mechanism through which induced-fit binding proceeds, whereby interactions in the T state trigger structural perturbations within FimH$_{TD}$ that culminate in T-R conformational changes.

Comparison of the distributions of binding loop and residue positions among apo T, mannose-bound T, apo R, and mannose-bound R states for WT FimGn6H revealed specific motions that are correlated with mannose binding (fig. S5C). Among these changes, movement of loop 1 toward the binding pocket was most strikingly associated with mannose binding in the R state but marginally so in the T state. Despite the occasional displacement of loop 1 away from the R pocket, mannose often remained stably associated with R, suggesting that loop 1 positioning is not the only determinant of mannose binding. Consistent with this observation, binding assays performed on loop 1 deletion mutants and chimeras in WT FimH$_{TD}$ showed a significant reduction, but not total abrogation, in mannose binding, demonstrating a role for loop 1 as an affinity clamp in mannose recognition (fig. S5D). These findings agree with the weakened mannose affinity observed in mutations upstream of loop 1 that prevent β-hairpin formation necessary to bring loop 1 in close proximity to the binding pocket and observed in a recent cocrystal structure showing loop 1 in close proximity to heptyl mannose in a bound T state due to stabilization from crystal packing (26, 32). In all, these computational studies establish a vast conformational phase space in the absence of ligand characterized by a conformationally restrained T state and multiple, shape-shifting R state conformations, and they elucidate structural and dynamic insights into two distinct mechanisms of mannose recognition by the low-affinity T and high-affinity R conformations.

Positive selection in FimH promotes moderate mannose-binding affinity to facilitate bladder colonization

Positively selected residues, which were identified through sequence analysis of FimH alleles enriched in urinary *E. coli* isolates, significantly affect adhesive function and bacterial fitness in the urinary tract (13). Despite A27V/V163A existing entirely in the high-affinity R state, previous work paradoxically demonstrated that UTI89 engineered to encode FimH A27V/V163A in place of FimH WT was severely attenuated and unable to form IBCs in the C3H/HeN mouse model of UTI at 6 to 24 hours post-infection (hpi) (13, 14). We investigated the kinetics of this virulence defect during the acute stages of bladder colonization. UTI89 expressing A27V/V163A FimH exhibited attenuated colonization as early as 1 hpi, which suggests a defect in the ability of A27V/V163A to bind and/or invade superficial facial cells (Fig. 6A). Yet, UTI89 expressing A27V/V163A FimH bound and invaded 5637 bladder cells in vitro more efficiently than UTI89 expressing WT FimH (Fig. 6, B and C), suggesting that the high-affinity variant does not lack the capacity to bind bladder tissue per se. 5637 bladder cells, a cancer cell line, exhibit more similarities with undifferentiated transitional bladder epithelial cells than with the terminally differentiated superficial facial cells that line the undisturbed bladder lumen (33). Thus, attenuation in the mouse model may reflect colonization resistance properties of the bladder habitat that are specifically selective against A27V/V163A FimH over WT FimH. To address this hypothesis, we changed the bladder habitat by inserting a catheter implant and subsequently tested these FimH variants in a model of catheter-associated UTI (CAUTI). In this model, a 5-mm piece of silicon tubing is implanted into the C57BL/6 mouse bladder, which mechanically disrupts regions of the bladder epithelium, exposing the underlying transitional epithelium and inducing inflammation. It has previously been shown that this catheterization increases the efficiency of bladder colonization by otherwise attenuated species of bacteria (34–37). In the absence of an implant, again, a colonization
defect was observed in C57BL/6 mice at 24 hpi (Fig. 6, D and E). Within this CAUTI model, in the presence of an implant, UTI89 with either WT or A27V/V163A FimH robustly colonized both the implant and the implanted mouse bladder tissue at 1 day after infection, although this rescue was partial because WT still outperformed A27V/V163A by 10-fold in bacterial titers in the bladder (Fig. 6, D and E). How the implant facilitates colonization may be multifactorial. First, disruption of the terminally differentiated epithelium may allow bacteria expressing the A27V/V163A variant to bypass the intact superficial facet cell layer and colonize deeper epithelial layers. Second, the catheter provides another surface to which the bacteria can bind and form a “staging ground” for dissemination to the bladder tissue. This is particularly possible because implanted catheters become coated by host proteins, including THP (36, 38), which we have shown is tightly bound by A27V/V163A (Fig. 2C). In addition, A27V/V163A can directly interact with implanted catheters, and in vitro assays indicate that it forms biofilms on abiotic surfaces to a higher degree than does WT (13). In contrast, without the catheter, the tighter binding to soluble THP, or other soluble host proteins or mannose-containing oligosaccharides, by A27V/V163A may be particularly detrimental to colonization because it prevents binding to urinary tract surfaces. Thus, in the naive bladder, soluble host proteins and oligosaccharides may act as decoy receptors, whereas in the catheterized bladder, soluble host proteins may aid in binding of bacteria to the implant. Together, these studies demonstrate that the preexisting conformational ensemble and corresponding mannose affinity of FimH determine the outcome of infection. Further, the results suggest that positive selection balances the conformational equilibrium of FimH in solution between R and T states to tailor moderate affinity to UPEC for the urothelium, thus facilitating adherence to and persistence in the bladder.

**DISCUSSION**

Our interdisciplinary studies combining evolutionary biology and computational and molecular biophysics allowed us to directly interrogate, at the atomic and structural level, the multiple conformations that FimH adopts in solution in a tip-like setting within a two-state, T-R conformational landscape (Fig. 7A). Positively selected residues influence the preexisting conformational equilibrium of the two-domain FimH by shifting the relative occupancies of a dynamically restrained T state conformation and multiple R state conformations that sample a great expanse of bends, twists, and orientations through ball-and-socket joint-like motions. As a result of these population shifts, natural sequence variation alters the apparent affinity of FimH toward mannansylated ligands. Thus, the conformational phase space of FimH in solution and population shifts spurred by positively selected residues further provide a framework for understanding the structural basis of allosteric coupling between interdomain interactions and mannose binding. Classically, the mechanism underlying protein allostery has been conceptualized as a deterministic process, in which information is transmitted through structural perturbations from one site of a protein to another in a sequential or concerted manner (39–41). However,
recent synergy in the fields of protein biophysics and protein evolution has given rise to the emerging perspective that protein allostery is a natural and statistical consequence of shifts in the relative populations and/or dynamics of preexisting conformational ensembles (42, 43), which is supported by this work. Furthermore, by influencing the pre-existing equilibrium of FimH, positively selected residues alter the very pathway through which binding occurs: If a particular FimH variant favors the T state, it will favor an induced-fit mechanism of binding; if a FimH variant favors the R state, it will use conformational selection as its mechanism of binding.

Our structural, biophysical, and computational work indicate that the low-affinity T and high-affinity R conformations can bind mannose through distinct binding orientations and pathways. Consistent with previous cocrystal structures, simulations of the R state indicate that mannose and oligomannose-3 bind in a high-affinity horizontal orientation coordinated through hydrogen bond formation with residue F1 and residues in loops 2 and 3 and clamped by residue I13 in loop 1. Our simulations and a recent crystal structure of the T state bound to heptyl mannoside (26) also indicate that the T state may bind mannose in this horizontal orientation, particularly when loop 1 is proximal to the binding pocket. In addition to this binding mode, mannose also approaches and binds the T state pocket in a tilted orientation, specifically to residue F1 of the N terminus and D54 of loop 2, in a low-affinity interaction. Mannose glycans may slide and rotate and interact with the conformationally dynamic tyrosine gate of FimH. This novel ligand entry and rotation pathway may represent a stepwise mechanism through which induced-fit binding proceeds, whereby interactions in the T state trigger structural perturbations within FimH that culminate in T-R conformational changes. Furthermore, tight binding of mannose through multiple different bends and orientations of the R state, as indicated in our simulations, provides the basis for a physical model, here termed molecular tethering, by which bacteria can remain bound to their mannosylated receptors on the bladder surface (Fig. 7B). Entropic freedom of multiple bound conformations within the high-affinity R state would theoretically enhance the lifetime of mannose binding relative to a single high-affinity R state. Moreover, the enhanced conformational flexibility and the number of viable bound R conformations in FimH may increase the biophysical adaptability of...
type 1 pili while tethered to a surface in part to allow attached bacteria to resist urine flow while bound to the bladder epithelium (44, 45). Molecular tethering likely represents a universal feature of macromolecular interactions involving Gram-negative adhesins or any other two-domain, allosterically regulated protein containing a flexible linker.

Coupled with genetic and in vivo pathogenesis work, we have demonstrated how evolutionary pressures through positive selection in FimH allosterically shape the conformational dynamics and phase space of FimH to maintain a balanced conformational equilibrium between R and T states to allow UPEC to colonize the urinary tract (Fig. 7C). Surprisingly, a FimH variant that only adopts the high-affinity R state is severely attenuated early in a mouse model of uncomplicated UTI but is proficient at colonizing catheterized bladders in vivo or bladder transitional-like epithelial cells in vitro. Given how early the pathogenesis defect of A27V/V163A is observed relative to WT (within 1 hpi), some preset aspect(s) of the bladder habitat likely select against the high-affinity R conformation(s). One possibility may relate to the slower kinetics of association of the R state under conditions of flow, as part of a catch-bond mechanism (44–46), which, in the case of A27V/V163A, would translate to less frequent interactions with the bladder epithelium and increased clearance of bacteria from the bladder during periods of urination. A second possibility is that some inherent property of the bladder epithelium serves as a restrictive factor against the R state. We suggest that the superficial facet cells may mediate this restriction, because the A27V/V163A variant is capable of binding 5637 bladder cells in vitro and can mediate the colonization of the catheter-implanted bladder habitat. A third possibility is that soluble mannosylated glycoproteins, such as THP, and mannos-containing oligosaccharides (47) may serve as decoy receptors that reduce the ability of A27V/V163A to colonize the bladder epithelium because it binds more tightly than WT to THP. These same host proteins that serve as decoys in host defense may facilitate bacterial colonization by coating the catheter (36, 38). More work is required to dissect the individual and combinatorial contributions that the abovementioned possibilities may exert in preventing the high-affinity R state in mediating successful bladder colonization. Together, our data suggest that the T state may serve to temporarily mask the strong affinity that FimH has for mannoside to avoid restrictive factors or properties native to the bladder habitat to initiate productive binding when engaging the host epithelium.

Fundamentally, this encompassing study of solution protein dynamics, structure, and function exemplifies the importance of defining the native conformational ensembles of a protein in solution and its population shifts in the presence of ligand for a complete model of allosteric regulation. Understanding mechanistic and structural aspects of adhesin allostery, conformation, and function is critical in efforts to further develop antibiotic-sparing small molecules and vaccines for the treatment of acute and recurrent infections caused by UPEC and other pathogens.

**Materials and Methods**

**Ethics Statement**

All animal experiments were conducted according to the National Institutes of Health (NIH) *Guide for the Care and Use of Laboratory Animals* and performed in accordance with institutional regulations after pertinent review and approval by the Animal Studies Committee at Washington University School of Medicine (protocol number 20150226). Human urine collection was performed with informed consent and approved under Institutional Review Board protocol 201207143.

**FimH Variants**

In silico analysis of FimH sequences among 287 *E. coli* strains previously revealed three specific residues (positions 27, 62, and 163) evolving under positive selection. Variations at these three positions in the FimH sequence from UPEC strain UTI89 were examined in this study. These include the following: (i) the WT variant (A27/A62/V163), in which these three amino acids are together observed in *E. coli* isolates from infected urine (19 of 254 strains) but not in isolates from healthy feces (0 of 33 strains); (ii) the A62S variant (A27/S62/V163), in which these three amino acids are together observed in high abundance in isolates from both infected urine (169 of 254 strains) and healthy feces (28 of 33 strains); and (iii) the engineered A27V/V163A variant (V27/A62/A163), in which these three amino acids have not been observed in tandem and have been shown to negatively affect pathogenesis in vivo. The mannoside-binding pocket mutant Q133K, which cannot bind mannoside, was also incorporated in this study as a negative control. These variants were generated by site-directed mutagenesis of the WT FimH allele from UTI89, as described in a previous report (13).

**Protein expression and purification**

FimCH variant complexes were purified from periplasm preparations, as previously described (14). FimG_{Nle}H complexes were assembled by a spontaneous in vitro DSE reaction, in which FimG_{Nle} peptide (FimG residues 1 to 15; EZBiolab) was mixed in ~10× molar excess with FimCH variant complexes in 15 mM MES (pH 5.6) and 50 mM NaCl and incubated at 37°C for 16 h. FimG_{Nle} displaces FimC in this reaction, and resultant FimG_{Nle}H variant complexes were purified away from excess FimG_{Nle} peptide and free FimC using a SOURCE 15S column (GE) in 15 mM MES (pH 5.6) with a gradient of 0 to 400 mM NaCl. Pooled fractions containing FimG_{Nle}H variant complexes were dialyzed against 15 mM MES (pH 5.6) and 50 mM NaCl, concentrated to 1 to 5 mg/ml, and stored stably at 4°C for use in biophysical assays.

**Mannoside Incubations**

Mannoside compound 4Z269 (*para*-biphenyl-2-methyl-3′-methyl amide mannoside) (27) was incubated with FimG_{Nle}H variant complexes for at least 1 hour at 4°C before biophysical analysis at molar ratios indicated below.

**Differential Scanning Fluorimetry**

FimG_{Nle}H variants (10 μM) in the absence or presence of 4Z269 (10 μM) were combined with 5× SYPRO Orange (Sigma; 1:1000 dilution of 5000× stock) in 50 μl of reaction mixture buffered in 15 mM MES (pH 5.6), 50 mM NaCl, and 0.4% dimethyl sulfoxide. Binding equilibria were established by allowing the reaction mixtures to incubate at 23°C for 30 min. These reaction mixtures were then placed in 96-well clear-bottom polymerase chain reaction plates and subjected to a melt curve from 20° to 90°C in 0.5°C increments of 15 s, each followed by a fluorescence read of the “HEX” channel in a Bio-Rad CFX96 thermocycler. Melt curves were fitted to the Boltzmann equation: 

\[ y = A_2 + (A_1 - A_2)/(1 + \exp((x - x_c)/dx)) \]

where \( x_c \) is the \( T_m \) to determine the melting temperature (\( T_m \)).

**Enzyme-Linked Immunosorbent Assay**

Immulon 4HBX 96-well plates were coated overnight with 1 μg of human glycoproteins reported to be ligands of FimH (secretory IgA, laminin, collagen IV, and THP). All glycoproteins were ordered from...
Within 1 min with 50 μl of polyclonal rabbit antibody raised against FimH residues 1 to 165 (from E. coli strain J96) with a C-terminal six-histidine tag (Sigma Genesis; 1:5000 dilution in PBS + 2% BSA) was added to each well for 1 hour at 37°C. After washing three times with PBS + 0.05% Tween 20, 100 μl of polyclonal goat anti-rabbit antibody conjugated to horseradish peroxidase (KPL; 1:5000 dilution in PBS + 2% BSA) was added to each well for 1 hour at 37°C. After washing three times with PBS + 0.05% Tween 20, 100 μl of polyclonal goat anti-rabbit antibody conjugated to horseradish peroxidase (KPL; 1:5000 dilution in PBS + 2% BSA) was added to each well for 1 hour at 37°C. After a final round of washing, plates were developed with 100 μl of tetramethylbenzidine substrate (BD Biosciences) and quenched within 1 min with 50 μl of 1 M H₂SO₄, and absorbance was measured at 450 nm.

Small-angle x-ray scattering

Before sample submission, FimCH and FimGn thêm variant complexes were buffer-exchanged three to five times in Spin-X UF 5K concentrators (Corning) against freshly prepared buffer [15 mM MES (pH 5.6) and 50 mM NaCl] to a final concentration of 5 to 8 mg/ml. FimGn thêm:4Z269 complexes were prepared in the same manner after FimGn thêm variants were incubated with 4Z269 at a 1:2 molar ratio. Samples were diluted to 1 to 5 mg/ml using the final filtrate to ensure identical buffer conditions between buffer and sample. Samples were then shipped to the SIBYLS beamline at the Advanced Light Source for data collection and basic data processing, including scattering integration and buffer subtraction, as detailed extensively in the high-throughput mail-in SAXS protocol (48). Integrated scattering profiles were then inspected and analyzed using the ATSAS program suite (49). At least 9 to 12 scattering profiles representing multiple concentrations and exposures were merged in PRIMUS to generate an average scattering profile for each sample tested. Pair distance distributions [p(r)] and maximal intramolecular distances (Dmax) were determined by GNOM. Structural comparison maps were calculated, as previously described, using the SIBYLS SAXS Similarity online tool (http://sibyls.als.lbl.gov/saxs_similarity/) (50). Fits of crystal structures to merged scattering profiles were evaluated by the FoXs and MultiFoXs servers in the case of single-state and multisite rigid-body modeling, respectively. For each sample, 10 ab initio models were generated through simulated annealing by GASBOR (51) and averaged by DAMAVER (52). The resultant low-resolution molecular envelope of each sample was represented in PyMOL by setting the van der Waals distance (vdw) to 5.0 Å and the solvent radius (solvent_radius) to 4.5 Å.

Native electrospray ionization and IMMS

FimGn thêm variant complexes in 15 mM MES (pH 5.6) and 50 mM NaCl were diluted to 10 to 20 μM protein concentration and underwent multiple rounds of buffer exchange through dilution and concentration into 100 mM ammonium acetate (pH 6.5) (>99.9% trace metals basis) such that the final salt concentration was reduced to lower than 10 to 20 μM for clean native electrospray ionization (nESI) spectra. FimGn thêm:4Z269 complexes were prepared in the same manner after FimGn thêm variants were incubated with 4Z269 at a 1:1 molar ratio. Samples were loaded at 5 to 10 μl in custom-made electrospray capillaries and injected into a hybrid ion mobility quadrupole time-of-flight mass spectrometer (SYNAPT G2 High Definition Mass Spectrometry, Waters). The instrument was operated under gentle ESI conditions (capillary voltage, 1.5 to 2.2 kV; sampling cone, 2 V; extraction cone, 1 V; source temperature, 36°C). CE was varied to observe dissociation events between FimGn thêm FimH and 4Z269. The pressure of the vacuum/backing region was 5.1 to 5.6 mbar. For the ion mobility measurements, the helium gas flow to the collision cell was 70 ml/min, the ion-mobility spectrometry (IMS) gas flow was 35 ml/min, the IMS wave velocity was 500 m/s, and the IMS wave height was 20 V. Nitrogen was used as the mobility carrier gas. The instrument was externally calibrated up to 8000 mass/charge ratio (m/z) with the clusters produced by ESI of NaI solution (100 mg/ml). The peak picking and data processing were achieved by using MassLynx (version 4.1) and DriftScope software (Waters). The CCSs for protein ions were converted using previously published calibration protocols and databases (53). FimGn thêm variants in the presence or absence of 4Z269 were aerosolized by nESI, and native mass spectra were acquired under a range of CEs (5 to 40 V). Each spectrum was acquired every 1 s for 2 min and gated at m/z values of 3054 to 3056 (apo) and 3094 to 3096 (+4Z269) for the +10 charge state. These spectra indicated that all FimH variants entered the gas phase as an intact FimGn thêm complex at low CE but steadily dissociated with increasing CE (fig. S3A). These spectra also revealed that FimGn thêm variants stably associated with 4Z269 in accordance with the relative binding strengths of the FimH alleles: 4Z269 occupies ~100% of A27V/V163A, ~90% of WT, ~50% of A62S, and ~0% of Q133K (fig. S3B). The +10 charge state corresponding to intact FimGn thêm variants was then subjected to ion mobility analysis at various CEs. CCS distributions were modeled as a sum of two Gaussian curves because modeling by one Gaussian curve resulted in significantly poorer fits in nearly all cases (table S2).

Protein crystallization and structure determination

Crystals of FimGn thêm A62S grew under numerous polyethylene glycol (PEG) conditions at 20°C by hanging drop vapor diffusion in 96-well plates. The diffracting crystals of FimGn thêm A62S were grown within the PEG II crystallization screen (Qiagen) by mixing 100 nl of protein (7.5 mg/ml) with 100 nl of mother liquor [0.2 M calcium acetate, 0.1 M Hepes (pH 7.5), and 10% PEG 8000] and equilibrated against 75 μl of mother liquor in the reservoir. These crystals took on a thick-plate rectangular prism morphology and were picked directly from the 96-well screen. They were dissolved into a cryoprotectant [0.2 M calcium acetate, 0.1 M MES (pH 5.6), 25% PEG 8000, and 20% glycerol] and then flash-frozen in a cryostream. Crystals of FimGn thêm A27V/V163A were grown at 20°C by hanging drop vapor diffusion by mixing 1 μl of protein (5 mg/ml) with 1 μl of mother liquor [0.2 M calcium acetate, 0.1 M Hepes (pH 7.5), and 5% PEG 8000] and 0.5 μl of 0.1 M sodium chloride and equilibrated against 1 ml of mother liquor in wells of a 24-well plate. These crystals took on a thin-plate morphology. They were transferred into a cryoprotectant [0.2 M calcium acetate, 0.1 M Hepes (pH 7.5), 25% PEG 8000, and 20% glycerol] and then flash-frozen in a cryostream. Diffraction data were collected at 100 K at an in-house facility equipped with a rotating anode (Rigaku MicroMax-007 generator), a Rayonix marx mnx x-ray source, and a Mar345 image plate detector. Data were indexed and integrated in iMosflm (54) and scaled by Scala (55). The phase problem was solved by MR using Phaser-MR in Phenix (56) with two distinct search ensembles: (i) FimH₁D (from PDB ID 1KLF) and FimH₁D in the T state (from PDB ID 3JWN) as separate domains and (ii) FimH₁D (from PDB ID 1KLF) and FimH₁D in the R state (from PDB ID 1KLF) as separate domains. The best solutions were identified...
as a T conformation for FimG_{NeH} A62S and a bent R conformation (90° between FimH_{LD} and FimH_{PD}) for FimG_{NeH} A27V/V163A. Several rounds of refinements were performed in phenix.refine to improve the final models.

**Molecular dynamics**

Crystal structures of FimG_{NeH} published in this study along with models of FimG_{NeH} based on previous crystal structures (G_{NeH} docked on H from CH structure) in the absence or presence of ligands were used to initiate MD simulations. Structural models of FimG_{NeH} variants were created with the in silico mutagenesis wizard in PyMOL after the mutant side-chain rotamer with the fewest steric clashes was selected. Structural models were then prepared within the CHARMM36 force field and the TIP3P water model by the MD software package GROMACS-4.6.7 (57). The system was first solvated in a cubic box, with sides equal to 130 Å in the presence of 50 mM NaCl. The structure was then energy-minimized to reduce steric clashes, and the system was finally equilibrated at constant temperature (37°C) and pressure (1 atm). Conventional unrestrained MD simulations were then performed with steps of 2 fs for a total time of ~15 to 20 ns for each run, collecting four replicates for simulations on the T (5JQ), bent R (5JR4), and elongated R (based on 1KLF) conformations for WT FimH. Longer simulations were also performed at 100 to 200 ns for three replicates for all FimH variants in these conformations with similar results. Coordinates were saved every 0.1 ns. Commands in GROMACS and graphical output in MATLAB were used to analyze simulation trajectories.

**Conformational phase space**

Conformational phase space calculations were performed with custom in-house scripts in PyMOL and MATLAB after alignment of the pilin domain to a reference structure (FimH_{PD} from FimCH). The in-house scripts in PyMOL and MATLAB after alignment of the pilin domain to a reference structure (FimH_{PD} from FimCH). The in-house scripts in PyMOL and MATLAB after alignment of the pilin domain to a reference structure (FimH_{PD} from FimCH). The in-house scripts in PyMOL and MATLAB after alignment of the pilin domain to a reference structure (FimH_{PD} from FimCH). The in-house scripts in PyMOL and MATLAB after alignment of the pilin domain to a reference structure (FimH_{PD} from FimCH). The in-house scripts in PyMOL and MATLAB after alignment of the pilin domain to a reference structure (FimH_{PD} from FimCH). The in-house scripts in PyMOL and MATLAB after alignment of the pilin domain to a reference structure (FimH_{PD} from FimCH).

**Bladder epithelial cell studies**

Human bladder epithelial cells, designated 5637 (ATCC HTB-9) cells, were obtained from the American Type Culture Collection and maintained in RPMI 1640 supplemented with heat-inactivated 10% (v/v) fetal bovine serum at 37°C in the presence of 5% CO₂. Confluent, serum-starved 5637 cells in 24-well plates were infected with UTI89 strains at a multiplicity of infection of 10. After 30 min, culture media were replaced either by fresh culture media or by media with gentamicin sulfate (120 µg/ml) (Sigma-Aldrich) to kill extracellular bacteria. Cells were further incubated for 1 hour, washed rigorously, solubilized with 1% Triton X-100, quenched with PBS, and plated for bacterial CFU quantification.

**REFERENCES AND NOTES**


Acknowledgments: We thank members of the S.J.H. laboratory, C. Frieden, and J. Cooper for helpful suggestions. We also thank R. Stegeman for technical assistance in x-ray data collection, Z. Han for synthesizing and providing 4Z269, and G. Hura and J. Tainer at the Advanced Light Source for SAXS data collection and processing. Funding: S.J.H. was supported by NIH National Institute of Allergy and Infectious Diseases (NIAID) grants R01 AI029549, R01 AI048689, and U01 AI095542 and National Institute of Diabetes and Digestive and Kidney Diseases grant R01 DK051406. V.K. was supported by the Medical Scientist Training Program through NIH training grant T32 GM07200. T.J.H. was supported by NIH NIAID grant U01 AI095542. M.L.G. was supported by NIH National Institute for General Medical Sciences (NIGMS) grant P41GM103422. R.V.P. was funded by NIH NIGMS grant P41GM103422. Author contributions: V.K., N.H.T., M.L.G., R.V.P., J.J., and S.J.H. designed the research plan. V.K. and J.S.P. purified the proteins. V.K. performed x-ray crystallography, SAXS, ELISA, and in vitro bladder binding and invasion experiments. V.K. and H.Z. performed the IMMS experiments. V.K. and A.S.H. performed the simulations. T.J.H. and M.E.H. performed the animal experiments. V.K., T.J.H., M.E.H., and S.J.H. analyzed and interpreted all data. V.K., K.W.D., and S.J.H. wrote the manuscript. Competing interests: S.J.H., J.J., and T.J.H. have ownership interest in Fimbrion Therapeutics Inc. and may benefit if the company is successful in marketing the mannosides that are related to this research. S.J.H. and J.J. are cofounders, stockholders, and board members of Fimbrion Therapeutics Inc., and T.J.H. is an employee of the same organization with stock options. J.J. has three patents related to the compounds described in the work: U.S. Patent US8937167, “Compounds and methods for treating bacterial infections,” J.J., S.J.H., Z. Han, J. Pinkner, C. Cusumano (Washington University School of Medicine), 20 January 2015; PCT Application WO2012109263, “Mannoside compounds and methods of use thereof,” J.J., S.J.H., Z. Han, J. Pinkner, C. Cusumano (Washington University School of Medicine), 16 August 2012; and U.S. Application US20120309701, “Compounds and methods useful for treating bacterial infections,” J.J., S.J.H., Z. Han, J. Pinkner, C. Cusumano (Washington University School of Medicine), 6 December 2012. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. The atomic coordinates and structure factors have been deposited in the PDB under accession codes 5JQI and 5JR4.

Submitted 16 August 2016
Accepted 25 January 2017
Published 10 February 2017

Evolutionary fine-tuning of conformational ensembles in FimH during host-pathogen interactions
Vasilios Kalas, Jerome S. Pinkner, Thomas J. Hannan, Michael E. Hibbing, Karen W. Dodson, Alex S. Holehouse, Hao Zhang, Niraj H. Tolia, Michael L. Gross, Rohit V. Pappu, James Janetka and Scott J. Hultgren

Sci Adv 3 (2), e1601944.
DOI: 10.1126/sciadv.1601944