Spring 5-15-2020

Conformational Basis and Small Molecule Antagonists of E. coli Adhesion to the Urinary Tract

Vasilios Kalas
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Computational and Molecular Biophysics

Dissertation Examination Committee:
Scott J. Hultgren, Chair
Thomas J. Brett
Gautam Dantas
Jeffrey P. Henderson
Niraj H. Tolia

Conformational Basis and Small Molecule Antagonists of *E. coli* Adhesion to the Urinary Tract

by
Vasilios Kalas

A dissertation presented to the
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2020
St. Louis, Missouri

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

Acknowledgments

Attainment of this Ph.D. would not have been possible without the support of my numerous mentors, friends, and family. First and foremost, I would like to thank Scott Hultgren for accepting me into his lab and providing me the incredible opportunity to work with him. His unwavering support and the immense responsibilities and independence he entrusted me with allowed me to grow and flourish as a research scientist. Scott is an exceptional role model, who taught me that success requires balancing intellect, passion, and perseverance with enthusiasm, humor, and compassion. Further, he taught me the importance of cultivating meaningful relationships and leaning in on them. “Stavros” is a great friend, and I am grateful for everything he’s done for me. I would also like to thank my undergraduate mentor, Wei-Jen Tang, for giving me my first opportunity to conduct structural biology research. Without his commitment to my training, I would not be nearly as confident or successful in my graduate work. In addition, I would like to thank my thesis committee mentors – Tom Brett, Gautam Dantas, Jeff Henderson, and Niraj Tolia – for their guidance and for always challenging me to translate the insights from my rigorous biophysical studies into a compelling biological story.

I am grateful for all the funding sources that supported my graduate research. Specifically, I would like to thank the National Institutes of Health (NIH) Grants R01 DK108840, R01 AI029549, and R01 AI048689 to Scott Hultgren and the MSTP Training Grant T32GM07200, Infectious Disease Training Grant, and Stephen I. Morse Fellowship to me. I would also like to thank the MSTP, the Department of Biochemistry and Molecular Biophysics, and the Department of Molecular Microbiology for supporting me.

Finally, I would like to thank my friends and family for filling my life with so much joy, love, and companionship. I am thankful for the close friendships I developed through medical

school, the MSTP, and the Hultgren Lab. I especially thank all my childhood friends from the Greek community in Chicago with whom I share a strong bond and who continue to foster a warm kinship with me despite my lengthy absence from home during my MD/PhD training here in St. Louis. Most importantly, I want to thank my family from the bottom of my heart for everything they have given me. My parents George and Irene started from humble beginnings in rural Greece and came to the United States with the hopes of a better life. They struggled but they persevered, providing me endless support so that I could realize my dreams and aspirations. Together, they instilled in me the virtues of patience, perseverance, humility, and compassion. My father inspired my love for music, my sense of humor, my pride, and my passion for a career in medicine, while my mother stimulated my ambition in academics, my commitment to family, and my desire to care for others. I can't thank them enough. I thank my sister Tina for always putting up with her big brother and for being such a caring sibling throughout life. I would also like to thank my other parent, my mother-in-law Efi, for reinforcing the same values my parents taught me and for helping me by sharing her experiences as a professional in healthcare. Ultimately, I couldn't have done this without my wife, the other half of my soul, Georgia. Her remarkable strength and unconditional love gave me the resolve to combat the challenges and frequent disappointments in lab. She exudes an effortless and brilliant radiance that warms my heart, motivates me to aim high in my aspirations, and continues to beckon the best version of myself each and every day. I am in constant awe and adoration of the woman that she is, and I am truly blessed to have such a wonderful companion and confidante in life.

Vasilios Kalas

Washington University in St. Louis

May 2020

Dedicated to my parents, George and Irene,
for their unconditional love and support
and for teaching me life's most important lessons.

ABSTRACT OF THE DISSERTATION

Conformational Basis and Small Molecule Antagonists of *E. coli* Adhesion to the Urinary Tract

by

Vasilios Kalas

Doctor of Philosophy in Biology and Biomedical Sciences

Computational and Molecular Biophysics

Washington University in St. Louis, 2020

Professor Scott J. Hultgren, Chair

Urinary tract infections (UTIs) are one of the most prevalent infections, afflicting 15 million women per year in the United States with annual healthcare costs exceeding $2-3 billion. Uropathogenic *Escherichia coli* (UPEC) are the main etiological agent of UTIs and employ numerous virulence factors for host colonization. The most common adhesive mechanism by which UPEC mediate host-pathogen interactions is the chaperone-usher pathway (CUP), which is responsible for the assembly of proteinaceous surface appendages termed pili. Generally, CUP pili function in adherence or invasion of host tissues and in biofilm formation on medical devices and body habitats. CUP pili are highly abundant and diverse among a wide variety of Gram-negative pathogens, with 38 distinct pilus types in *Escherichia* species alone, mediating a considerable range of biological tropisms through adhesins at the distal pilus tip. Typically, these adhesins have a lectin domain, which recognizes a specific carbohydrate receptor, and a pilin domain to anchor the adhesin to the pilus. This thesis specifically examines the structural, dynamic, and allosteric properties of distinct *E. coli* CUP pilus adhesins that govern interactions critical for pilus function at the host-pathogen interface during UTI.

The type 1 pilus adhesin FimH is a critical virulence factor necessary for bacterial attachment to mannosylated receptors on the bladder epithelium during UTI. I determined through molecular and computational biophysics that FimH natively exists in a two-state conformational equilibrium in solution, composed of one low-affinity tense (T) and multiple high-affinity relaxed (R) conformations. I demonstrated that positively selected residues in FimH and ligand binding allosterically modulate this conformational equilibrium and that each of these conformational states engage mannose receptors through distinct binding modes. Mouse models of UTI indicate that FimH has evolved a "moderate" mannose binding affinity through a balanced conformational equilibrium to optimize persistence in the bladder during UTI. Furthermore, I discovered novel small-molecule galactoside antagonists that inhibit the FimH-like adhesin FmlH from binding galactose-containing bladder and kidney epithelial receptors present during chronic UTI. Taken together, this thesis defines the biophysical basis of host receptor recognition and bacterial pathogenesis mediated by FimH and defines the atomic bases of distinct bacterial host tropisms mediated by FimH homologs, which were leveraged to spur the development of antibiotic-sparing, small-molecule glycomimetic antagonists as therapeutics for UTI and other infectious diseases.

Chapter 1: Introduction

Edited from

Adhesive pili from the chaperone-usher family

Vasilios Kalas, Ender Volkan, and Scott J. Hultgren

Escherichia coli: pathotypes and principles of pathogenesis, 2013. Chapter 12: pg 363-386 *ed.* Michael S. Donnenberg. Elsevier/Academic Press, Amsterdam.

and from

Pili and fimbriae of Gram-negative bacteria

Ender Volkan, Vasilios Kalas, and Scott J. Hultgren

Molecular medical microbiology, 2015. Chapter 8: pg 147-162 *ed.* Yi-Wei Tang, Max Sussman, Dongyou Liu, Ian Poxton, and Joseph Schwartzman. Elsevier/Academic Press, Amsterdam.

Abstract

Adhesive pilus fibers mediate the interaction between pathogens and specific host cell-surface ligands that allow a pathogen to establish a foothold in a particular host tissue. The chaperone-usher pathway (CUP) constitutes one of the most prevalent mechanisms among Gram-negative bacteria for the assembly of adhesive pili. Studies at the interface of genetics, biochemistry, and structural biology have detailed the functions of the chaperone and the usher, providing a step-by-step understanding of pilus biogenesis by this sophisticated molecular machine. Furthermore, studies have elucidated the molecular basis of host receptor recognition mediated by adhesins localized at the tips of CUP pili. Together, snapshots in CUP pilus assembly and CUP pilus-mediated host-pathogen interactions have unveiled necessary molecular details for the design and application of novel anti-virulence compounds that promise to prevent and treat acute, chronic, and recurrent bacterial infections in humans.

Introduction

One of the major pathways involved in assembly of adhesive pili in Gram-negative bacteria is the chaperone-usher pathway (CUP). CUP pili are thin, hair-like surface extensions involved in adherence and biofilm formation. One of the first observations of CUP pilus function was likely in 1908 when Guyot reported bacterial-mediated hemagglutination [1]. Since, significant improvements and discoveries have been made in the field of CUP pili revealing unique structural, biophysical, and biochemical phenomena of protein folding and assembly. Studies of CUP pili have exposed their crucial functions and essentiality in bacterial pathogenesis. CUP pili play critical roles in numerous infectious diseases, including type 1 pili in cystitis [2-6], S pili in neonatal meningitis [7, 8], F1 antigen in bubonic plague [9-11], and P pili in pyelonephritis [12-15]. Piliated bacteria adhere to specific receptors on the host cell surface through CUP pilus tip-localized adhesins. These adhesins are typically two-domain proteins comprised of a lectin domain, which recognizes a specific carbohydrate receptor, and a pilin domain to anchor the adhesin to the pilus. As many Gram-negative bacterial adhesins are assembled into polymeric surface appendages, CUP pili provide prevailing models for analysis and understanding of the assembly and function of a wide range of virulence associated proteins. CUP systems, particularly type 1 and P pili from uropathogenic *E. coli*, have been used as prototypes to elucidate how periplasmic protein monomers are assembled into complex structures and arrayed on the extracellular surface of Gram-negative bacteria without requiring cellular energy. In this chapter, I will focus on the structure, assembly mechanism, and receptor recognition mechanism of P pili and type 1 pili in uropathogenic *Escherichia coli* (UPEC), the most common cause of urinary tract infections (UTIs).

Pilus types

CUP pili consist of multiple pilus subunits arranged into long, linear protein polymers. CUP pili can adopt one of two distinct pilus structures depending on the length of the G1 β-strand of the chaperone. F1-G1 long (FGL) chaperones assemble thin, fibrillary, amorphous structures while F1-G1 short (FGS) chaperones assemble thick, helical rods topped by a fibrillar tip [16]. CUP pili also vary according to usher relatedness across six major clades – α, β, γ(1-4), κ, π, and σ – of the 189-membered CUP pilus superfamily [16]. Paradigms for CUP pilus architecture and assembly have been well established with the P pilus and type 1 pilus of UPEC, members of the π and γ1 clades, respectively. These two archetypal pili exhibit a bipartite organization, consisting of a long, helical rod connected to a thin tip fibrillum. The P pilus subunits PapG, PapF, PapE, PapK, PapA, and PapH arrange in order from fibrillar tip to rod base (**Figure 1**). PapG, the adhesin, lies at the distal end of the pilus. The tip adaptor PapF connects PapG to the main tip component PapE, which appears in 5-10 copies and has a width of ~ 2 nm. The adaptor PapK anchors PapE to the main rod component PapA, which appears in >1000 copies and gives rise to a right-handed, helical structure that displays a 6.8 nm width, 2.5 nm pitch, and 3.3 subunits per turn [17-19]. Finally, PapH attaches at the base of the rod and terminates pilus biogenesis [20, 21]. Type 1 pili adopt a similar, yet condensed architecture [22]. The fibrillum consists of single copies of both the adhesin FimH and the tip subunit FimG, while the adaptor FimF links the fibrillum to the major rod component FimA, which appears in ~1000 copies (**Figure 1**). Preliminary studies suggest that FimI functions as the terminator subunit in type 1 pili [23, 24].

Chaperone structure and function

CUP chaperones (~25-30 kDa) are all highly homologous in sequence and structure [25].

They are composed of two complete Ig-like domains (**Figure 2A**). Chaperones function to transiently bind their cognate pilus subunits, shielding their interactive surfaces to facilitate their proper folding and stability in the periplasm [26]. In PapD, the chaperone from the P pilus system, a conserved salt bridge stabilizes the two Ig-like domains to adopt an overall boomerang-like shape. These domains orient to form a cleft that is directly involved in subunit binding. In the absence of subunits, chaperones like PapD and SfaE form dimers and Caf1M forms tetramers as a self-capping mechanism to prevent unfavorable interactions and proteolysis [27-29].

Upon translocation of pilus subunits across the SecYEG translocase to the periplasmic space, they are taken up by their cognate periplasmic chaperones, which use several interactive surfaces to provide stability to the subunits. The active site of the chaperone is comprised of residues R8 and K112, a conserved basic patch in the cleft formed between the two domains. Mutagenesis of these basic residues abolishes the ability of the chaperone to mediate pilus assembly [30, 31]. Crystallography studies carried out on all of the P pili chaperone-subunit complexes [21, 32-36] further confirmed these interactions and the extended conformation by which the C-termini of subunits are anchored at the invariant R8 and K112 residues of the chaperone [31].

Donor strand complementation (DSC)

The absence of the seventh β-strand in the pilus subunit results in a deep groove on its surface that exposes its hydrophobic core. In a process termed donor strand complementation (DSC), the G1 strand of the chaperone is donated *in trans* in a non-canonical, parallel fashion to the exposed hydrophobic groove of the subunit to facilitate its proper folding [32, 37, 38] (**Figure 2B)**. The solvent-exposed set of alternating hydrophobic residues on the chaperone's G1 strand

directly interact with subunit pockets [32]. In the case of PapD, residues L107, I105, and L103 are respectively in register with the subunits' hydrophobic P1-P3 pockets. Mutating residue 105 is particularly detrimental to pilus biogenesis [27]. However, mutagenesis of residue L107 does not alter pilus assembly, which may be due to the plasticity of the P1 pocket [36]. These residues are termed P1-P3 residues based on their corresponding acceptor sites, the subunit P1-P3 pockets (**Figure 2D**). Additionally, N101 forms hydrogen bonds above the shallow P4 subunit pocket. Thus, the interactions involved in DSC facilitate subunit folding and prevent subunit aggregation. Presence of unfolded subunits in the periplasmic space otherwise induces periplasmic stress responses targeting aggregated subunits for degradation by the DegP protease [39, 40].

Donor strand exchange (DSE)

Once the subunit is folded, it remains bound to the chaperone until its incorporation into the pilus. The chaperone is exchanged for the N-terminal extension (Nte) of the incoming subunit via a "zip-in, zip-out" mechanism termed donor strand exchange (DSE). DSE is initiated by the insertion of the Nte P5 residue into the open P5 pocket of the previously assembled subunit (**Figure 2C-E**). The chaperone G1 strand does not occupy the P5 pocket of subunits, which remains easily accessible in a chaperone-subunit complex to the P5 residue of the subunit Nte. Ultimately, insertion of Nte in an antiparallel fashion to the P2-P5 subunit pockets facilitates the removal of the chaperone's parallel-oriented G1 strand. The DSE interaction is more energetically favorable than the DSC interaction, which allows DSE and the stable docking of subunits with each other to occur [33, 41].

Chaperone-usher interactions

PapD-like chaperones have a set of conserved, solvent-exposed, hydrophobic residues (termed Set B: L32, Q34, T53, P54, P55, V56, R68, I93) situated at the N-terminal domain (**Figure 2A**). X-ray crystallography studies demonstrated that the FimC chaperone of the type 1 pilus system interacts with the N-terminal domain of the usher via these Set B residues [42, 43], demonstrating that the Set B patch is a surface that interacts with the usher. Accordingly, point mutations in these residues negatively impact pilus biogenesis [27]. In addition, small molecules called pilicides, designed to block pilus biogenesis [44, 45], were shown to bind Set B residues and prevent the targeting of chaperone-subunit complexes to the usher [44]. These results implicate PapD's conserved hydrophobic patch Set B as the usher-targeting site. Other chaperone residues that may be involved in interactions with other domains of the usher are subjects of investigation.

Another set of highly conserved residues (Set C: L78, P79, D81, R82, E83, S84) is located at the elbow region of PapD, which is not known to interact with any protein partners. The stability of the chaperone depends on the formation of a buried salt bridge composed of D196, E83, and R116, which lies at the interdomain region [27]. E83 is part of the aspartate, arginine, glutamate, and serine (DRES) motif, which is highly conserved in the chaperone superfamily [25]. Situated at the end of the E1-F1 loop, part of the DRES motif packs against the hinge region connecting the two domains [25, 46], where the aspartate and arginine side chains from the DRES motif point out into solution, seemingly suitable for protein-protein interactions (**Figure 2A**). It is plausible that interactions with other proteins (subunits or usher domains) could cause conformational changes that would be transmitted to the interdomain region, possibly disrupting the salt bridge and thus facilitating a reorientation of the domains and causing allosteric conformational changes in the chaperone that play a role in the assembly of the CUP pilus.

Subunit-subunit interactions

There are several major functional types of pilus subunits that may comprise a CUP pilus, including adhesins, adaptors, tip fibrillar subunits, pilus rod subunits, and terminators (**Figure 1)**. Pilins, or pilus subunits, share an outstanding degree of sequence and structural homology with one another, necessary for preserving common mechanisms of subunit-subunit, chaperone-subunit, and usher-chaperone-subunit interactions needed for maturation of a functional pilus. As described, the incomplete nature of the pilin fold ensures either chaperone binding (by DSC) or Nte binding (by DSE) to pilus subunits for their stability, proper folding, and ultimate incorporation into the growing pilus. Thus, the Nte, chaperone G1 strand, and the incomplete Ig-like fold of pilin subunits, particularly the P1-P5 pockets, serve as common recognition motifs for the use of pilins as building blocks in the construction of a CUP pilus. However, structural distinctions across homologous pilins permit their unique positions and functions along the pilus chain. For example, differences in Nte sequence and hydrophobic pocket characteristics dictate subunit ordering. DSE reactions performed by incubating all combinations of Nte peptides (based on the five Nte-containing Pap subunits) with all chaperone-subunit complexes showed a range of reactivities [35, 47]. Reactions that occurred most rapidly were consistently those between cognate groove-Nte partners. In addition to chaperone-subunit and subunit-subunit interactions, subunits also contain specific surfaces that drive chaperone-subunit-usher interactions. Differences in subunit structure and residue side chains dictate selective trafficking to certain periplasmic domains of the usher. Thus, adhesin and terminator subunits, which occupy opposite ends of the pilus, adopt structures much more varied from their pilin counterparts required for their essential functions in host adhesion and pilus assembly termination, respectively.

Terminator structure and function

Studies in the Pap system suggest that pilus biogenesis ends with the incorporation the PapH terminator at the base of the pilus rod. Deletions or mutations of PapH result in the formation of long pili that are shed from the extracellular surface [20]. Overexpression of PapH results in short pili. Further, the structure of PapD-PapH indicates that PapH has an occluded P5 pocket, which prevents the initiation of DSE with further subunits, thereby terminating pilus biogenesis [21]. Analogously, FimI may function as the terminator in type 1 pili [23, 24]. Further studies will elucidate whether incorporation of a P5 pocket-lacking pilin is utilized by other CUP pili for termination.

Adhesin structure and function

A mature CUP pilus typically contains an adhesin at its distal end. Adhesins contain two distinctive domains: a lectin domain, which functions to bind specific receptors with stereochemical specificity and contribute to host and tissue tropisms, and an Ig-like pilin domain, which links the receptor-binding lectin domain to the fibrillar tip [32, 48]. FimH and PapG, adhesins of the prototypical type 1 and P pilus systems, respectively, mediate host-pathogen interactions important in cystitis [5, 49-51] and pyelonephritis [14]. FimH mediates binding to α-D-mannose-containing receptors whereas PapG binds galabiose, a receptor that is present in the globoseries glycolipids found in the kidney. Both lectin domains contain a β-barrel jelly-roll fold that is common to CUP adhesin structures, but they contain different receptor binding pockets. FimH contains a deep, negatively charged pocket at the tip of the lectin domain [52]. In contrast, PapGII (class II of three PapG classes, each of which recognize different globoseries of glycolipids) binds galabiose-containing receptors with a shallow binding pocket formed by three

beta strands and a loop at the side of the lectin domain [53]. CUP adhesins F17-G of ETEC [54] and GafD of UPEC [55] bind N-acetyl-D-glucosamine residues of proteins at a shallow binding site on the side of the lectin domain at a location unrelated to the binding pocket of PapG (**Figure 3**). With the function of adhesion segregated to a distinct protein domain at the pilus tip, various bacterial strains can alter the structure and receptor specificity of the adhesin, thus allowing flexibility and selective advantage in establishing tropism.

Numerous studies suggest that the lectin domain of FimH does not operate independently of the pilin domain but instead interacts intimately with it to influence receptor binding. A crystal structure of the type 1 pilus tip fibrillum [56] indicates a markedly different conformation in FimH, compared to the conformation observed in FimH in the FimC-FimH complex [48] or in the mannose-bound FimC-FimH structure [52]. In the tip structure, the pilin domain interacts with a compacted lectin domain in a stable, noncovalent manner, loosening the mannose-binding pocket at the distal end of the lectin domain and thereby inducing an inactive low-affinity state of the adhesin via a "page-turning" allosteric mechanism [56]. In contrast, FimH, when in complex with FimC and α-D-mannose, adopts an elongated, high affinity state [52], which was later shown to be induced by application of shear force, presumably by destabilization of inter-domain contacts. Indeed, cross-linking mutants along with positively selected residues in the pilin domain suggest that this conformational equilibrium is important for mannose binding and infectivity in murine models of cystitis [56, 57]. Shear force does indeed enhance bacterial adhesion to target cells [58], but whether this mechanism of tensile stress-induced binding enhancement applies to other pili systems or has a definitive role in pathogenesis remains to be seen.

Usher domain function and selectivity

The CUP usher catalyzes the translocation and assembly of the multi-subunit pilus fiber across the outer membrane, while maintaining membrane integrity. The usher is an outer membrane protein comprised of five domains: a 24-stranded beta barrel channel, plug, N-terminal domain (NTD) and C-terminal domains (CTD1, CTD2) [43, 59, 60]. In the apo state of the FimD usher, the plug resides in the lumen of the transmembrane channel, preventing flow of molecules across the outer membrane [61, 62]; in the active form, the plug swings away from the channel, creating an unobstructed opening through which pilins translocate (**Figure 4**). This plug switch induces a conformational change in the β-barrel domain, shifting the shape of the channel pore from ovular (52 Å x 28 Å) to nearly circular (44 Å x 36 Å) [63]. This in turn allows a range of subunit diameters (~20-25 Å) to pass through unobstructed. The plug then docks onto the NTD, as suggested in the crystal structure of the FimD usher [63] and from work in the Pap system [64]. The plug-NTD interaction results in the usher adopting an open state.

NTD and CTDs, which reside in the periplasmic space, serve as the workhorses of assembly within this molecular machine. Functional studies directly implicate the periplasmic domains of the usher in catalysis of pilus formation, as mutations in either NTD or CTD [59, 65] and deletions of the plug [62, 66] abrogate assembly. Measured relative affinities for the interactions of chaperone-subunit complexes with purified NTD, CTD2, and plug domains from the PapC usher suggest that (a) the tight-binding NTD serves as the initial anchoring site for the chaperone-adhesin complex, (b) plug or the NTD-plug complex serves as the docking site for all other chaperone-subunit complexes, and (c) CTD2 likely dissociates chaperone-subunits from NTD and plug, with the exception of the chaperone-terminator complex, which docks on plug or the NTD-plug complex and halts further pilus growth. Analogously to PapC NTD, the FimD NTD is selective; isothermal titration calorimetry (ITC) experiments show that FimD NTD binds the

chaperone when the chaperone is loaded with FimH or FimF but not with FimG or FimA [43]. Furthermore, these structures along with supporting NMR data reveal that the intrinsically disordered 24-residue N-terminal tail of NTD adopts an ordered conformation upon binding. As expected, deletion of the N-terminal tail of NTD completely abolishes pilus assembly *in vivo* [67]. In the ternary complex, the N-terminal tail of NTD contacts the chaperone and the subunit (FimH or FimF), constituting 50-60% of the total contact surface area (1260 $Å^2$). While all components of these two ternary complexes adopt remarkably similar folds, there are subtle yet significant differences in hydrogen bonding at the FimD NTD-subunit interface that likely accounts for subunit selectivity. A similar molecular logic likely holds true for the selectivity of the NTD domains of other CUP ushers, although more work needs to be done to further bolster this argument. While the plug and CTD2 can recognize nearly all chaperone-subunit complexes, the mechanism by which this occurs remains undetermined. Similarly, the basis for the terminator's ability to discriminate between plug and CTD2 remains unknown. More work is needed to elucidate the molecular basis of selectivity by usher domains for chaperone-subunit complexes to better understand the mechanism and subunit ordering of pilus assembly.

Working model of usher-mediated pilus biogenesis

The results and associated interpretations presented above suggest the following working model of CUP pilus biogenesis (**Figure 5**):

(1) The usher adopts a gated conformation in its inactive state, in which its plug domain lies in the center of the kidney-shaped β-barrel. NTD and CTD lie disordered in the periplasm, moving rapidly and randomly.

(2) Once NTD binds the initiator of pilus assembly, the chaperone-adhesin complex, the binding energy induces movement in the NTD linker that thereby transduces conformational changes in the β-barrel domain, relaxing its shape from ovular to circular and preventing plug from re-entering the channel lumen once it has already entered the periplasm. Concurrently, NTD may swing to transfer its bound chaperone-adhesin complex to CTDs. Once NTD has unloaded its cargo, it binds the now periplasmic plug domain to form the NTD-plug complex, which can then bind and dock other incoming chaperone-subunit complexes.

(3) Alternatively or concomitantly, the CTDs may swing over together to the NTD-plug complex to catalyze the dissociation of the chaperone-adhesin complex. A minor conformational species in the FimD-FimC-FimH complex as measured by electron paramagnetic resonance (EPR) [63] suggests the occurrence of an uncharacterized transient binding interaction between CTD2 and the chaperone-adhesin complex that may mediate this transfer. With the chaperone-adhesin complex bound primarily to CTD1, the lectin domain of the adhesin lies in the channel lumen, and the NTD-plug complex is available for the incoming chaperone-tip subunit complex.

(4) Next, the NTD-plug complex recruits the upcoming chaperone-tip subunit complex, while the chaperone-adhesin complex is bound to the CTDs. A unique conformation in the NTD-plug complex brings the tip subunit into the ideal orientation with respect to the adhesin for DSE, allowing for the assembly of the growing pilus fiber. The reaction results in the displacement of the chaperone off of the CTD, allowing the newly incorporated chaperone-subunit complex to dock on the CTDs. The process repeats for all chaperone-subunit complexes in order of the subunit ordering observed in the mature pilus fiber.

(5) Ultimately, the chaperone-terminator complex docks on the NTD-plug complex. Unable to bind CTD2 and unable to undergo DSE due to its lack of a P5 pocket, the chaperone-

terminator complex signals the end of pilus growth. The plug may bind the chaperone-terminator complex with its βC, βB, βE, and βF side and swing into the β-barrel lumen with that same side facing the extracellular space, as seen in the apo state. The mature pilus fiber anchors to the outer membrane in this fashion and is ready to perform its adhesive function.

Overall, this model provides insight into how the usher assembles the CUP pilus. With domain affinities, molecular snapshots of pilus biogenesis steps, and genetic and biochemical evidence, we can now understand how the usher efficiently transfers chaperone-subunit complexes using its periplasmic domains to perform its catalysis of fiber growth. Further work will be required to gain a deeper mechanistic understanding of this molecular machine to provide knowledge of targetable protein interfaces and dynamic processes for next generation small-molecule pilicides that can inhibit these virulent adhesive pili.

Role of CUP pili in infections

UPEC introduced to the urinary tract from the fecal flora is the leading causative agent of urinary tract infections (UTIs), responsible for 85% of community-acquired UTIs [68]. Genetic, biochemical, and imaging studies accompanied by murine models of UTIs revealed that CUP pili are crucial factors for causing this disease. For instance, type 1 pili, encoded by the *fim* operon, are required for bladder infection in a murine model of UTI [2-6] and for biofilm formation in rat kidneys [69]. The adhesin of type 1 pili, FimH, mediates binding to the mannosylated receptors on the surface of bladder urothelial cells for colonization and invasion of the superficial umbrella cells, which line the bladder lumen. Following invasion, UPEC grow inside the cytosol of host cells, giving rise to biofilm-like bacterial aggregates of 10^4-10^5 bacteria, known as intracellular bacterial communities (IBCs) [2, 6, 70]. Type 1 pili are required for the establishment and

maintenance of IBCs, allowing proliferation of UPEC, which upon maturation can detach and spread to neighboring urothelial cells to continue the infection cascade. As IBCs have also been observed in human patients suffering from UTIs [71], it is likely that involvement of type 1 pili and IBCs in human UTI is also common. Furthermore, dysbiosis of the gut microbiota after streptomycin treatment allows UPEC expressing type 1 pili and F17-like pili to colonize the gastrointestinal tract, the source reservoir from which UPEC emerge to infect the urinary tract [Spaulding & Hultgren *Nature* in press].

Another well-studied CUP system that is important in urinary tract infections is P pili, encoded by the *pap* operon. P pilus adhesin PapG was shown to be important in pyelonephritis in primate models of infection [14]. In rat models of infection, P pili were shown to enhance early colonization of the kidney tubular epithelium [69]. Most UPEC strains from human cases of pyelonephritis carry PapGII [72, 73], one of three classes of PapG, which binds to Galα1-4Gal receptor epitopes in the globoseries of glycolipids found in human renal tissue [53].

In addition to P and type 1 pili's importance in UTIs, various other CUP pili from different species are important in bacterial virulence. For instance, S pili, encoded by the *sfa* operon, is involved in meningitis [74, 75], while class 5 ETEC fimbriae are important for the initiation of diarrheal disease [76].

Conventionally, carriage of these CUP pili and other putative virulence factors has been considered the sole determinant of virulence among pathogens. However, recent comparative transcriptomic and phenotypic analyses among human uropathogens indicate that expression of CUP pili such as type 1 pili, and not their carriage, predicts colonization and virulence in the mouse bladder [77]. Additionally, differences in host susceptibility, inflammatory state, and infection history can alter tissue or host habitats to affect recognition by CUP pili, thereby modulating risk

of developing infectious disease, as observed in recurrent UTI [78]. Considering the importance of CUP pili in infection, understanding how these pili are assembled, the mechanisms by which they cause pathogenesis, and their genetic regulation is critical in disease prevention and development of novel therapeutics.

CUP pili as anti-virulence targets

Antibiotics have led to significant improvements of human health over decades, improving quality of life and human longevity. However, antibiotic resistance is escalating [79] and multidrug resistant uropathogens are spreading globally [80]. Accompanied by a lack of a significant effort to develop novel antibiotics and the outstanding risk that antibiotic usage can negatively impact gut microbiota and result in opportunistic infections, the search for anti-virulence therapeutics is gaining attention. Due to the important role of CUP pili in bacterial virulence, they have been important targets for vaccine development studies as well as anti-virulence therapeutics. FimH was shown to be a successful vaccine option in both murine [81] and primate [82] models of infection. Keeping in mind the role of FimH in mediating adhesion with mannosylated uroplakin of uroepithelial cells, alky and phenyl-α-D-mannopyranosides, or mannosides, have been designed to competitively bind with FimH and interfere with adhesion on host cells. These mannosides not only block adhesion but also counteract internalization and *in vitro*/*in vivo* biofilm formation on biotic and abiotic surfaces [83, 84] (**Figure 6A-B**). Mannosides also potentiate the efficacy of existing antibiotics in a murine model of UTI and have great oral bioavailability, increasing hopes that they can be used in drug development [84]. Furthermore, mannosides specifically deplete UPEC residing in the gastrointestinal tract without affecting the resident microbiota, functioning as molecular scalpels that selectively remove virulent *E. coli* [Spaulding & Hultgren *Nature* in

press]. Similar to FimH, PapG was also a target for antivirulence therapeutics design such that p-Methoxy-phenyl derivatives of galabiose inhibit PapG with low micromolar IC_{50} (half-maximal inhibitory concentration) values [85-88].

In addition to targeting the receptor-binding site of the mature pilus, compounds that directly interfere with pilus biogenesis have also been developed. These rationally designed bicyclic 2-pyridone compounds, termed pilicides, inhibit UPEC hemagglutination of erythrocytes and biofilm formation by inhibiting both P and type 1 pilus biogenesis [44]. X-ray crystallography and biochemical studies revealed an interaction of these compounds with PapD [44, 45]. The pilicides were shown to bind with a conserved, hydrophobic, solvent-exposed patch at the N-terminal side of the chaperone (**Figure 6C-D**), the presumed usher-binding site, thus explaining the structural basis of their mechanism of action.

Translated to clinical practice, mannosides, pilicides, and vaccine options can be cost-effective ways to prevent and treat UTIs and other infections that require CUP pili, while reducing antibiotic resistance. Using specific anti-virulence therapeutics should have minimal impact on the composition of host microbiota, reducing the risk of opportunistic or recurrent infections.

Conclusion

Gram-negative bacteria use the chaperone-usher pathway (CUP) to assemble virulent surface appendages called pili. Structural biology combined with genetic and biochemical approaches has elucidated crucial protein-protein interactions made by the dedicated chaperone and ushers to facilitate the ordered assembly of pilins into the final pilus structure on the extracellular surface. This work has generated new insights into protein folding and revealed novel mechanisms of macromolecular assembly. Chaperones stabilize the fold of each subunit in a

chaperone-subunit complex, in which the subunit is held in a high-energy conformation primed to participate in pilus assembly at the usher. Ushers catalyze pilus assembly through interactions with each chaperone-subunit complex, coordinating the release of the chaperone and interactions of subunits with each other as they fold into their final condensed structures and translocate through the outer membrane usher pore. Ultimately, assembled CUP pili display tip-localized adhesins to mediate key host-pathogen interactions required for virulence. These multi-disciplinary approaches have revealed snapshots of a sophisticated protein assembly machinery and mechanisms of adhesin binding, thus spurring the development of therapeutics that suppress infection in animal models. Together, these efforts will lead to a better understanding of CUP pilus-mediated infectious disease, giving rise to potent therapeutics that target acute, chronic, and recurrent infections for prevention and treatment of human disease.

References

1. Guyot, G., *Uber die bakterielle Haemagglutination (Bacterio-Haemagglutination).* Zentralbl. Bakteriol., Parasitenkd., Infektionskr. Hyg., Abt. 1: Orig., 1908. **47**: p. 640-53.
2. Anderson, G.G., et al., *Intracellular bacterial biofilm-like pods in urinary tract infections.* Science, 2003. **301**(5629): p. 105-7.
3. Connell, I., et al., *Type 1 fimbrial expression enhances Escherichia coli virulence for the urinary tract.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9827-32.
4. Hultgren, S.J., et al., *Role of type 1 pili and effects of phase variation on lower urinary tract infections produced by Escherichia coli.* Infect Immun, 1985. **50**(2): p. 370-7.
5. Mulvey, M.A., et al., *Induction and evasion of host defenses by type 1-piliated uropathogenic Escherichia coli.* Science, 1998. **282**(5393): p. 1494-7.
6. Wright, K.J., P.C. Seed, and S.J. Hultgren, *Development of intracellular bacterial communities of uropathogenic Escherichia coli depends on type 1 pili.* Cell Microbiol, 2007. **9**(9): p. 2230-41.
7. Parkkinen, J., et al., *Binding sites in the rat brain for Escherichia coli S fimbriae associated with neonatal meningitis.* J Clin Invest, 1988. **81**(3): p. 860-5.
8. Korhonen, T.K., et al., *Serotypes, hemolysin production, and receptor recognition of Escherichia coli strains associated with neonatal sepsis and meningitis.* Infect Immun, 1985. **48**(2): p. 486-91.
9. Sebbane, F., et al., *The Yersinia pestis caf1M1A1 fimbrial capsule operon promotes transmission by flea bite in a mouse model of bubonic plague.* Infect Immun, 2009. **77**(3): p. 1222-9.
10. Hatkoff, M., et al., *Roles of chaperone/usher pathways of Yersinia pestis in a murine model of plague and adhesion to host cells.* Infect Immun, 2012. **80**(10): p. 3490-500.
11. Felek, S., et al., *Contributions of chaperone/usher systems to cell binding, biofilm formation and Yersinia pestis virulence.* Microbiology, 2011. **157**(Pt 3): p. 805-18.
12. Roberts, J.A., et al., *Prevention of pyelonephritis by immunization with P-fimbriae.* J Urol, 1984. **131**(3): p. 602-7.
13. Roberts, J.A., et al., *Antibody responses and protection from pyelonephritis following vaccination with purified Escherichia coli PapDG protein.* J Urol, 2004. **171**(4): p. 1682-5.
14. Roberts, J.A., et al., *The Gal(alpha 1-4)Gal-specific tip adhesin of Escherichia coli P-fimbriae is needed for pyelonephritis to occur in the normal urinary tract.* Proc Natl Acad Sci U S A, 1994. **91**(25): p. 11889-93.
15. Melican, K., et al., *Uropathogenic Escherichia coli P and Type 1 fimbriae act in synergy in a living host to facilitate renal colonization leading to nephron obstruction.* PLoS Pathog, 2011. **7**(2): p. e1001298.
16. Nuccio, S.P. and A.J. Baumler, *Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek.* Microbiol Mol Biol Rev, 2007. **71**(4): p. 551-75.
17. Jacob-Dubuisson, F., et al., *Initiation of assembly and association of the structural elements of a bacterial pilus depend on two specialized tip proteins.* EMBO J, 1993. **12**(3): p. 837-47.
18. Kuehn, M.J., et al., *P pili in uropathogenic E. coli are composite fibres with distinct fibrillar adhesive tips.* Nature, 1992. **356**(6366): p. 252-5.

19. Striker, R., et al., *Stable fiber-forming and nonfiber-forming chaperone-subunit complexes in pilus biogenesis.* J Biol Chem, 1994. **269**(16): p. 12233-9.
20. Baga, M., M. Norgren, and S. Normark, *Biogenesis of E. coli Pap pili: papH, a minor pilin subunit involved in cell anchoring and length modulation.* Cell, 1987. **49**(2): p. 241-51.
21. Verger, D., et al., *Molecular mechanism of P pilus termination in uropathogenic Escherichia coli.* EMBO Rep, 2006. **7**(12): p. 1228-32.
22. Hahn, E., et al., *Exploring the 3D molecular architecture of Escherichia coli type 1 pili.* J Mol Biol, 2002. **323**(5): p. 845-57.
23. Ignatov, O.V., *The role of FimI protein in the assembly of type 1 pilus from Escherichia coli*, in *Institute of Molecular Biology and Biophysics*. 2009, ETH Zurich.
24. Valenski, M.L., et al., *The Product of the fimI gene is necessary for Escherichia coli type 1 pilus biosynthesis.* J Bacteriol, 2003. **185**(16): p. 5007-11.
25. Hung, D.L., et al., *Molecular basis of two subfamilies of immunoglobulin-like chaperones.* EMBO J, 1996. **15**(15): p. 3792-805.
26. Kuehn, M.J., S. Normark, and S.J. Hultgren, *Immunoglobulin-like PapD chaperone caps and uncaps interactive surfaces of nascently translocated pilus subunits.* Proc Natl Acad Sci U S A, 1991. **88**(23): p. 10586-90.
27. Hung, D.L., et al., *Structural basis of chaperone self-capping in P pilus biogenesis.* Proc Natl Acad Sci U S A, 1999. **96**(14): p. 8178-83.
28. Knight, S.D., et al., *Structure of the S pilus periplasmic chaperone SfaE at 2.2 A resolution.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 6 Pt 2): p. 1016-22.
29. Zavialov, A.V. and S.D. Knight, *A novel self-capping mechanism controls aggregation of periplasmic chaperone Caf1M.* Mol Microbiol, 2007. **64**(1): p. 153-64.
30. Slonim, L.N., et al., *Interactive surface in the PapD chaperone cleft is conserved in pilus chaperone superfamily and essential in subunit recognition and assembly.* EMBO J, 1992. **11**(13): p. 4747-56.
31. Kuehn, M.J., et al., *Structural basis of pilus subunit recognition by the PapD chaperone.* Science, 1993. **262**(5137): p. 1234-41.
32. Sauer, F.G., et al., *Structural basis of chaperone function and pilus biogenesis.* Science, 1999. **285**(5430): p. 1058-61.
33. Sauer, F.G., et al., *Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation.* Cell, 2002. **111**(4): p. 543-51.
34. Verger, D., et al., *Crystal structure of the P pilus rod subunit PapA.* PLoS Pathog, 2007. **3**(5): p. e73.
35. Verger, D., et al., *Structural determinants of polymerization reactivity of the P pilus adaptor subunit PapF.* Structure, 2008. **16**(11): p. 1724-31.
36. Ford, B., et al., *Structure of the PapD-PapGII pilin complex reveals an open and flexible P5 pocket.* J Bacteriol, 2012.
37. Barnhart, M.M., et al., *PapD-like chaperones provide the missing information for folding of pilin proteins.* Proc Natl Acad Sci U S A, 2000. **97**(14): p. 7709-14.
38. Bann, J.G., et al., *Catalysis of protein folding by chaperones in pathogenic bacteria.* Proc Natl Acad Sci U S A, 2004. **101**(50): p. 17389-93.
39. Jones, C.H., et al., *The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems.* EMBO J, 1997. **16**(21): p. 6394-406.
40. Jones, C.H., et al., *Escherichia coli DegP protease cleaves between paired hydrophobic residues in a natural substrate: the PapA pilin.* J Bacteriol, 2002. **184**(20): p. 5762-71.

41. Remaut, H., et al., *Donor-strand exchange in chaperone-assisted pilus assembly proceeds through a concerted beta strand displacement mechanism.* Mol Cell, 2006. **22**(6): p. 831-42.
42. Eidam, O., et al., *Crystal structure of the ternary FimC-FimF(t)-FimD(N) complex indicates conserved pilus chaperone-subunit complex recognition by the usher FimD.* FEBS Lett, 2008. **582**(5): p. 651-5.
43. Nishiyama, M., et al., *Identification and characterization of the chaperone-subunit complex-binding domain from the type 1 pilus assembly platform FimD.* J Mol Biol, 2003. **330**(3): p. 513-25.
44. Pinkner, J.S., et al., *Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria.* Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17897-902.
45. Chorell, E., et al., *Design and synthesis of C-2 substituted thiazolo and dihydrothiazolo ring-fused 2-pyridones: pilicides with increased antivirulence activity.* J Med Chem, 2010. **53**(15): p. 5690-5.
46. Holmgren, A. and C.I. Branden, *Crystal structure of chaperone protein PapD reveals an immunoglobulin fold.* Nature, 1989. **342**(6247): p. 248-51.
47. Rose, R.J., et al., *Unraveling the molecular basis of subunit specificity in P pilus assembly by mass spectrometry.* Proc Natl Acad Sci U S A, 2008. **105**(35): p. 12873-8.
48. Choudhury, D., et al., *X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli.* Science, 1999. **285**(5430): p. 1061-6.
49. Bahrani-Mougeot, F.K., et al., *Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic Escherichia coli virulence determinants in the murine urinary tract.* Mol Microbiol, 2002. **45**(4): p. 1079-93.
50. Martinez, J.J., et al., *Type 1 pilus-mediated bacterial invasion of bladder epithelial cells.* EMBO J, 2000. **19**(12): p. 2803-12.
51. Wu, X.R., T.T. Sun, and J.J. Medina, *In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: relation to urinary tract infections.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9630-5.
52. Hung, C.S., et al., *Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection.* Mol Microbiol, 2002. **44**(4): p. 903-15.
53. Dodson, K.W., et al., *Structural basis of the interaction of the pyelonephritic E. coli adhesin to its human kidney receptor.* Cell, 2001. **105**(6): p. 733-43.
54. Buts, L., et al., *The fimbrial adhesin F17-G of enterotoxigenic Escherichia coli has an immunoglobulin-like lectin domain that binds N-acetylglucosamine.* Mol Microbiol, 2003. **49**(3): p. 705-15.
55. Merckel, M.C., et al., *The structural basis of receptor-binding by Escherichia coli associated with diarrhea and septicemia.* J Mol Biol, 2003. **331**(4): p. 897-905.
56. Le Trong, I., et al., *Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting.* Cell. **141**(4): p. 645-55.
57. Chen, S.L., 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**(52): p. 22439-44.
58. Thomas, W.E., et al., *Bacterial adhesion to target cells enhanced by shear force.* Cell, 2002. **109**(7): p. 913-23.
59. Thanassi, D.G., et al., *Bacterial outer membrane ushers contain distinct targeting and assembly domains for pilus biogenesis.* J Bacteriol, 2002. **184**(22): p. 6260-9.

60. Capitani, G., O. Eidam, and M.G. Grutter, *Evidence for a novel domain of bacterial outer membrane ushers.* Proteins, 2006. **65**(4): p. 816-23.
61. Remaut, H., et al., *Fiber formation across the bacterial outer membrane by the chaperone/usher pathway.* Cell, 2008. **133**(4): p. 640-52.
62. Huang, Y., et al., *Insights into pilus assembly and secretion from the structure and functional characterization of usher PapC.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7403-7.
63. Phan, G., et al., *Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate.* Nature, 2011. **474**(7349): p. 49-53.
64. Volkan, E., et al., *Domain activities of PapC usher reveal the mechanism of action of an Escherichia coli molecular machine.* Proc Natl Acad Sci U S A. **109**(24): p. 9563-8.
65. Henderson, N.S., et al., *Function of the usher N-terminus in catalysing pilus assembly.* Mol Microbiol. **79**(4): p. 954-67.
66. Mapingire, O.S., et al., *Modulating effects of the plug, helix, and N- and C-terminal domains on channel properties of the PapC usher.* J Biol Chem, 2009. **284**(52): p. 36324-33.
67. Nishiyama, M., et al., *Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD.* EMBO J, 2005. **24**(12): p. 2075-86.
68. Ronald, A.R., et al., *Urinary tract infection in adults: research priorities and strategies.* Int J Antimicrob Agents, 2001. **17**(4): p. 343-8.
69. Melican, K., et al., *Uropathogenic Escherichia coli P and Type 1 fimbriae act in synergy in a living host to facilitate renal colonization leading to nephron obstruction.* PLoS Pathog. **7**(2): p. e1001298.
70. Justice, S.S., et al., *Differentiation and developmental pathways of uropathogenic Escherichia coli in urinary tract pathogenesis.* Proc Natl Acad Sci U S A, 2004. **101**(5): p. 1333-8.
71. Rosen, D.A., et al., *Detection of intracellular bacterial communities in human urinary tract infection.* PLoS Med, 2007. **4**(12): p. e329.
72. Marschall, J., et al., *Both host and pathogen factors predispose to Escherichia coli urinary-source bacteremia in hospitalized patients.* Clin Infect Dis. **54**(12): p. 1692-8.
73. Otto, G., et al., *pap genotype and P fimbrial expression in Escherichia coli causing bacteremic and nonbacteremic febrile urinary tract infection.* Clin Infect Dis, 2001. **32**(11): p. 1523-31.
74. Morschhauser, J., et al., *Adhesin regulatory genes within large, unstable DNA regions of pathogenic Escherichia coli: cross-talk between different adhesin gene clusters.* Mol Microbiol, 1994. **11**(3): p. 555-66.
75. Korhonen, T.K., et al., *Escherichia coli fimbriae recognizing sialyl galactosides.* J Bacteriol, 1984. **159**(2): p. 762-6.
76. Chattopadhyay, S., et al., *Adaptive evolution of class 5 fimbrial genes in enterotoxigenic Escherichia coli and its functional consequences.* J Biol Chem. **287**(9): p. 6150-8.
77. Schreiber, H.L.t., et al., *Bacterial virulence phenotypes of Escherichia coli and host susceptibility determine risk for urinary tract infections.* Sci Transl Med, 2017. **9**(382).
78. O'Brien, V.P., et al., *A mucosal imprint left by prior Escherichia coli bladder infection sensitizes to recurrent disease.* Nat Microbiol, 2016. **2**: p. 16196.
79. Boucher, H.W., et al., *Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America.* Clin Infect Dis, 2009. **48**(1): p. 1-12.

80. 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**(10): p. e26578.
81. Langermann, S., et al., *Prevention of mucosal Escherichia coli infection by FimH-adhesin-based systemic vaccination.* Science, 1997. **276**(5312): p. 607-11.
82. Langermann, S., et al., *Vaccination with FimH adhesin protects cynomolgus monkeys from colonization and infection by uropathogenic Escherichia coli.* J Infect Dis, 2000. **181**(2): p. 774-8.
83. Guiton, P.S., et al., *Combinatorial Small-Molecule Therapy Prevents Uropathogenic Escherichia coli Catheter-Associated Urinary Tract Infections in Mice.* Antimicrob Agents Chemother, 2012. **56**(9): p. 4738-45.
84. Cusumano, C.K., et al., *Treatment and prevention of urinary tract infection with orally active FimH inhibitors.* Sci Transl Med, 2011. **3**(109): p. 109ra115.
85. George, S.K., et al., *Chemoenzymatic synthesis of sialylated glycopeptides derived from mucins and T-cell stimulating peptides.* J Am Chem Soc, 2001. **123**(45): p. 11117-25.
86. Hultgren, S.J., et al., *The PapG adhesin of uropathogenic Escherichia coli contains separate regions for receptor binding and for the incorporation into the pilus.* Proc Natl Acad Sci U S A, 1989. **86**(12): p. 4357-61.
87. Lin, L.Y., et al., *Synthetic polymer nanoparticles conjugated with FimH(A) from E. coli pili to emulate the bacterial mode of epithelial internalization.* J Am Chem Soc. **134**(9): p. 3938-41.
88. Ohlsson, J., et al., *Discovery of potent inhibitors of PapG adhesins from uropathogenic Escherichia coli through synthesis and evaluation of galabiose derivatives.* Chembiochem, 2002. **3**(8): p. 772-9.
89. Waksman, G. and S.J. Hultgren, *Structural biology of the chaperone-usher pathway of pilus biogenesis.* Nat Rev Microbiol, 2009. **7**(11): p. 765-74.

Figures

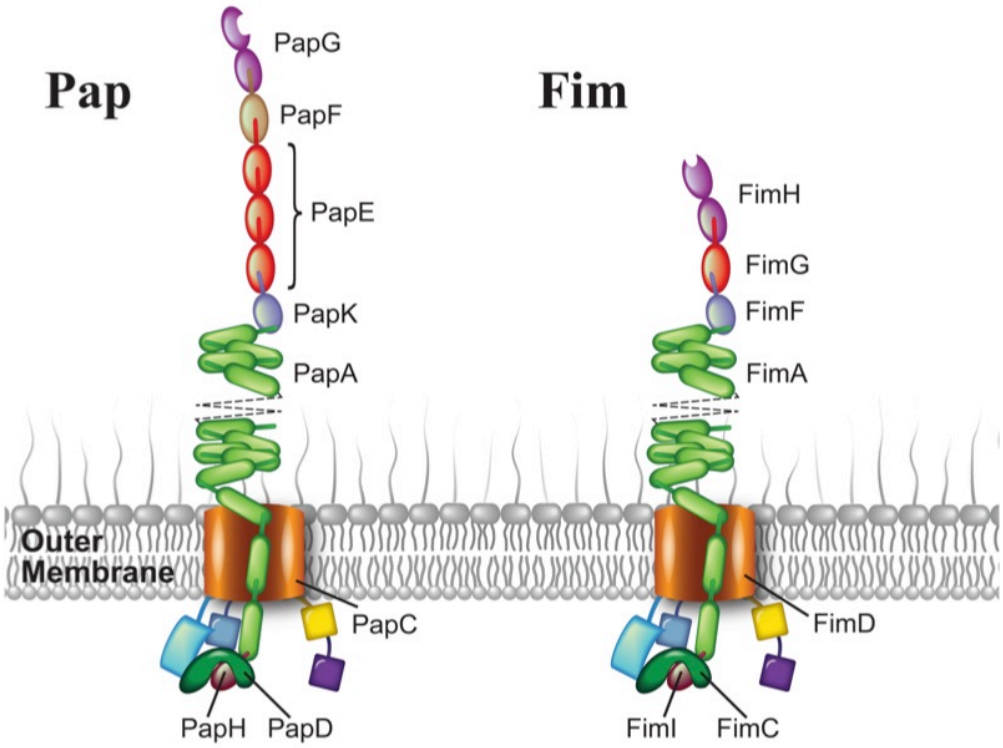


Figure 1. Architecture in P and type 1 pili. A graphic illustrating pili from the Pap (left) and type 1 (right) systems. See text for details.

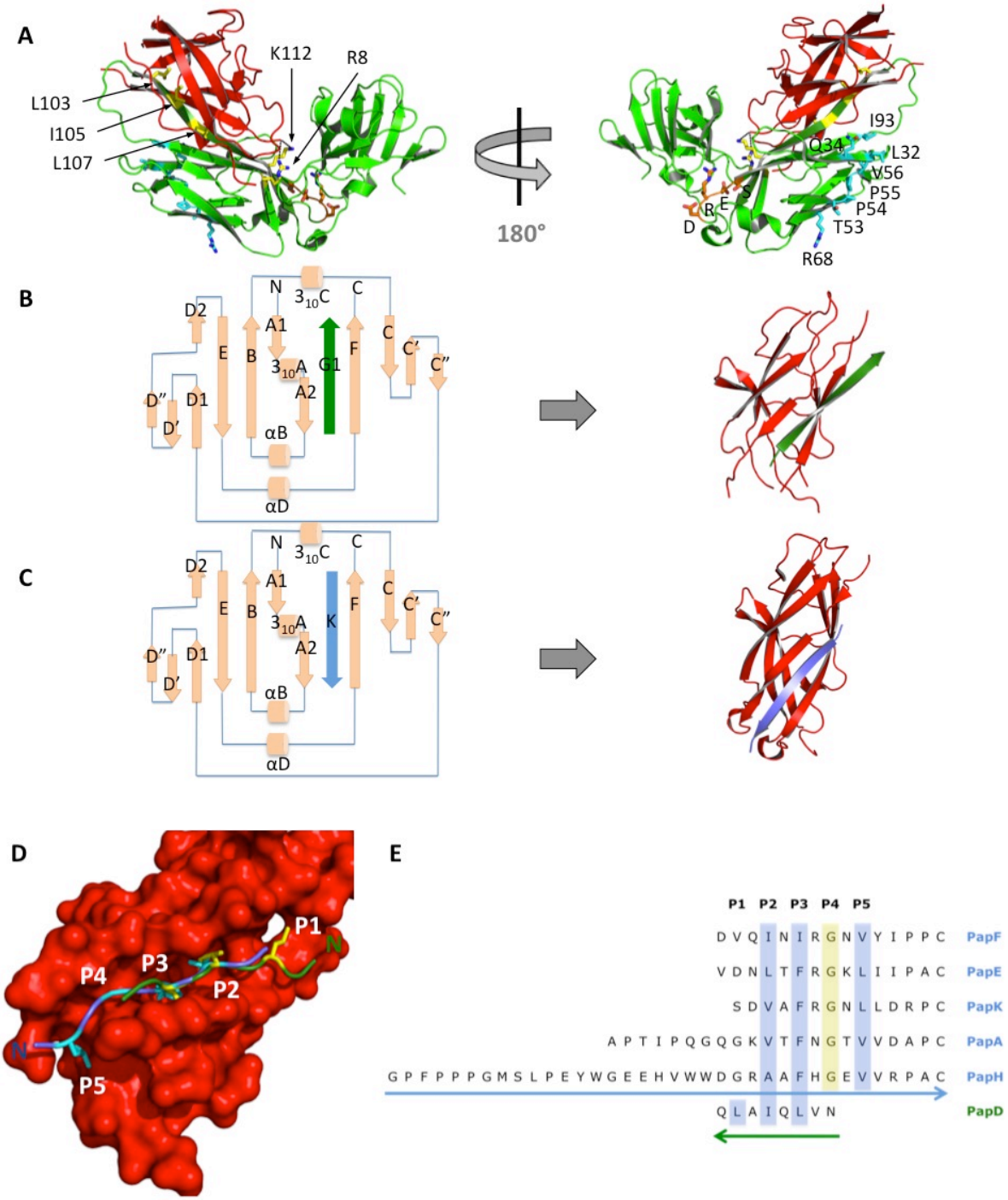


Figure 2. Conserved chaperone residues, donor stand complementation, and donor strand exchange. (A) A ribbon diagram of the PapD-PapE complex. PapE (red) binds the N-terminal domain of PapD (green). The conserved surfaces of PapD are shown in a ball-and-stick

representation and highlighted in yellow, cyan, and orange. Their function is described in the text. **(B)** Donor stand complementation (DSC). The topology diagram on the left shows the secondary structure of donor strand complemented PapE (tan), indicating β-strands as arrows and α-helices as cylinders. PapD donates its G1 strand (green) parallel to the F strand of PapE. The ribbon diagram on the right also depicts DSC, in which the G1 strand of PapD (green) completes the fold of PapE (red). [PDB code: 1N0L]. **(C)** Donor strand exchange (DSE). The topology diagram on the left shows the secondary structure of donor strand exchanged PapE (tan), indicating β-strands as arrows and α-helices as cylinders. PapK donates its N-terminal extension (Nte, blue) antiparallel to the F strand of PapE. The ribbon diagram on the right also depicts DSE, in which the PapK Nte (blue) completes the fold of PapE (red). [PDB code: 1N12]. **(D)** The P1-P5 pockets of PapE. The surface representation of PapE (red) emphasizes the pockets that allow the PapK Nte (blue) and the PapD G1 strand (green) to mediate DSE and DSC, respectively. PapD residues L103, I105, and L107 (yellow, ball-and-stick representation) project into the P1-P3 pockets and are correspondingly referred to as P1-P3 residues. PapK Nte residues (cyan, ball-and-stick representation) project into the P2-P5 pockets and are correspondingly referred to as P2-P5 residues. The N-terminal ends of the Nte and G1 strand are labeled in the corresponding color. Note the shallow nature of the P4 pocket, which only accommodates a glycine residue in the Nte sequence of Pap subunits, as shown in (E). **(E)** Sequence alignment of the PapD G1 strand and Nte's of all Pap subunits (except PapG, which lacks an Nte). Pap subunits have conserved P2-P5 residues at their Nte. P2, P3, and P5 residues are hydrophobic (blue), while the P4 residue is strictly glycine (yellow) to prevent steric clashes with the shallow P4 binding groove. PapD has three hydrophobic residues that correspond to the P1-P3 residues. The arrows indicate the N- to C-terminal direction for the sequences of Pap subunits (blue) and PapD (green).

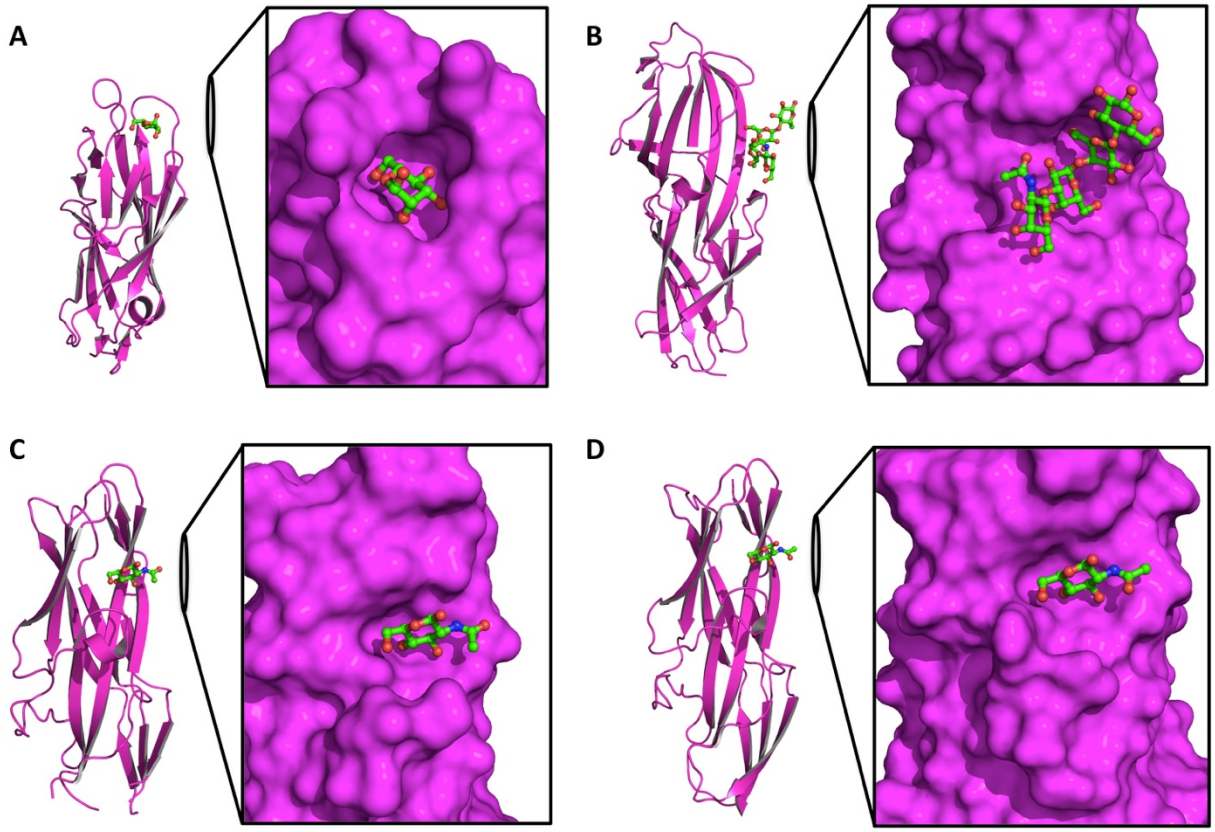


Figure 3. Co-crystal structures of CUP adhesins with their receptors. (A) FimH in complex with D-mannose [PDB code:1KLF] **(B)** PapG in complex with galabiose [PDB code: 1J8R] **(C)** F17-G in complex with N-acetyl-D-glucosamine [PDB code: 1O9W] **(D)** GafD in complex with N-acetyl-D-glucosamine [PDB code: 1OIO] In every complex, the adhesin (magenta) is shown as a ribbon representation (left) and a magnified surface representation (right), with the bound ligand depicted as a ball-and-stick representation (green). The structures demonstrate that the binding pocket is in different locations of the adhesin fold and has different morphology, which dictates receptor specificity.

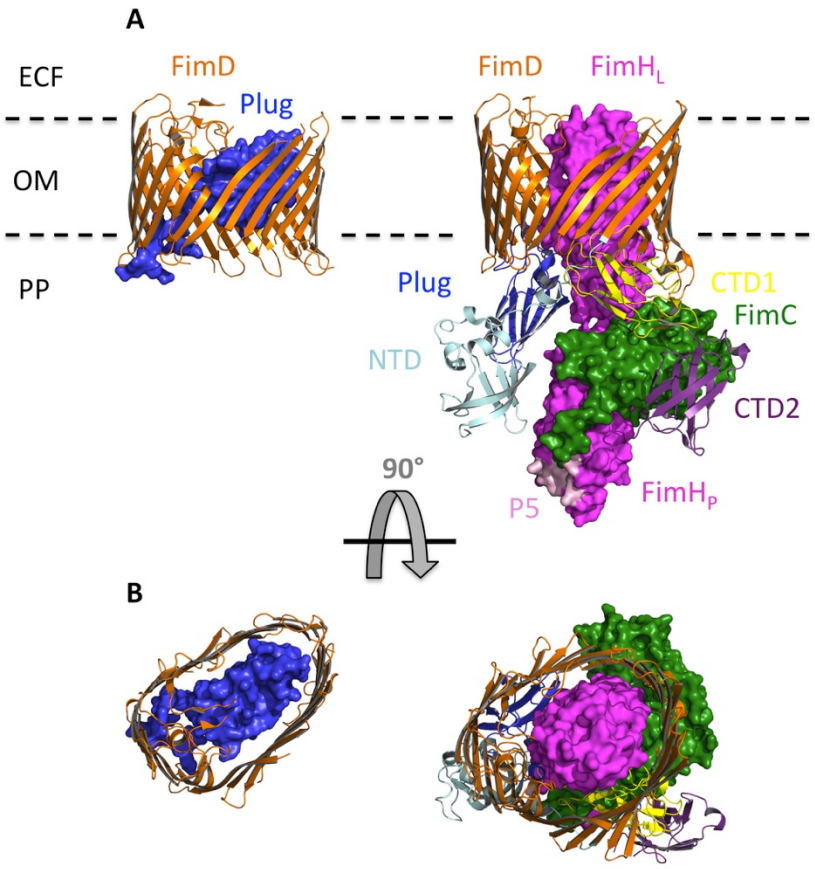


Figure 4. Structures of the gated and open FimD usher. **(A)** In the gated apo state, the plug (blue, surface representation) resides in the lumen of the usher pore (orange, ribbon diagram), preventing flow of molecules across the OM and maintaining membrane integrity. The remaining usher domains are not depicted since a FimD truncate was used to solve the apo structure [PDB code: 3OHN] **(B)** In the FimD-FimC-FimH complex, the plug swings into the periplasm and binds NTD (light blue), while FimC (green) docks at CTD1 (yellow) and CTD2 (purple) with its bound FimH (magenta). The lectin domain of FimH now resides in the lumen of the usher. Note the unoccupied P5 pocket (light pink) in the FimH pilin domain, which is targeted by the Nte of the incoming subunit FimG for DSE [PDB code: 3RFZ]. Rotation of the usher by 90°, such that the

usher is seen from the top, shows a kidney-shaped pore in the plug-gated usher but a nearly circular pore in the secreting usher. ECF, extracellular fluid; OM, outer membrane; PP, periplasm.

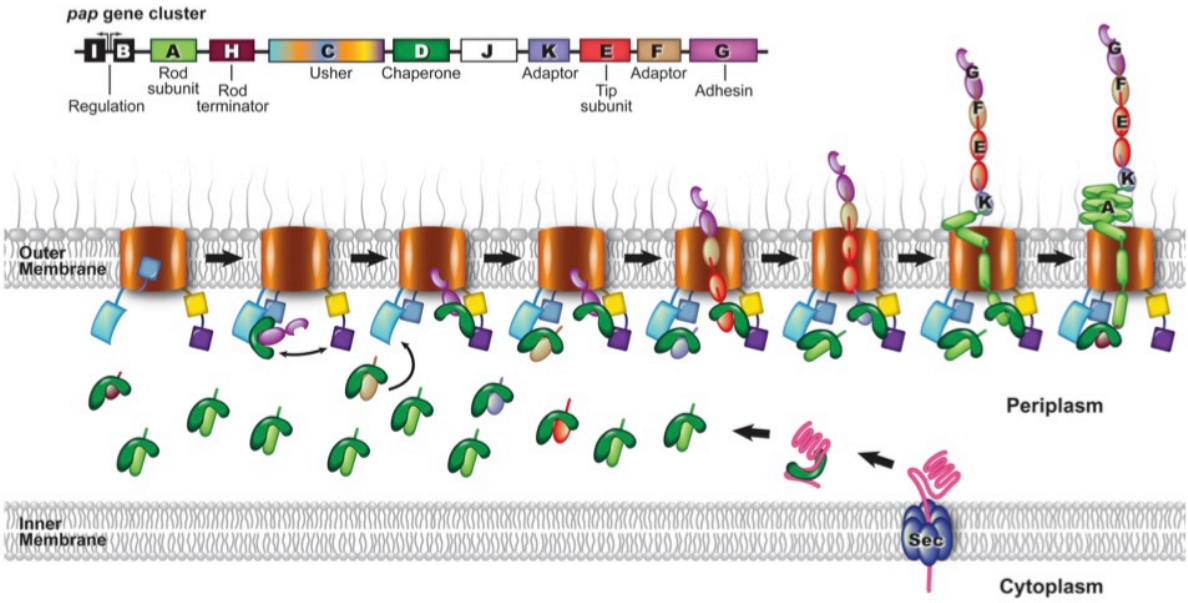


Figure 5. Model of pilus assembly. Upon translocation of subunits into the periplasm via the SEC machinery, chaperone proteins bind the subunits to maintain their proper fold and stability. At the outer membrane, the usher lies in a gated state, with the plug domain (blue) residing in the lumen of the usher pore (orange). Once the chaperone-adhesin complex binds the usher NTD (cyan), the plug domain swings away from the pore and docks onto the NTD, thus preparing the usher for pilus secretion. Soon thereafter, CTD1 (yellow) and CTD2 (purple) may swing over to bind and carry the chaperone-adhesin complex, freeing the NTD-plug complex in the process. The freed NTD-plug complex recruits the incoming chaperone-subunit complex, orienting it properly so that the Nte can participate in DSE with the CTD-docked chaperone-adhesin complex and the newly exchanged chaperone-subunit complex can transfer to the CTDs. The process repeats until the chaperone-terminator complex, bound at the NTD-plug complex, cannot transfer to the CTDs or engage in DSE, signaling the end of pilus assembly.

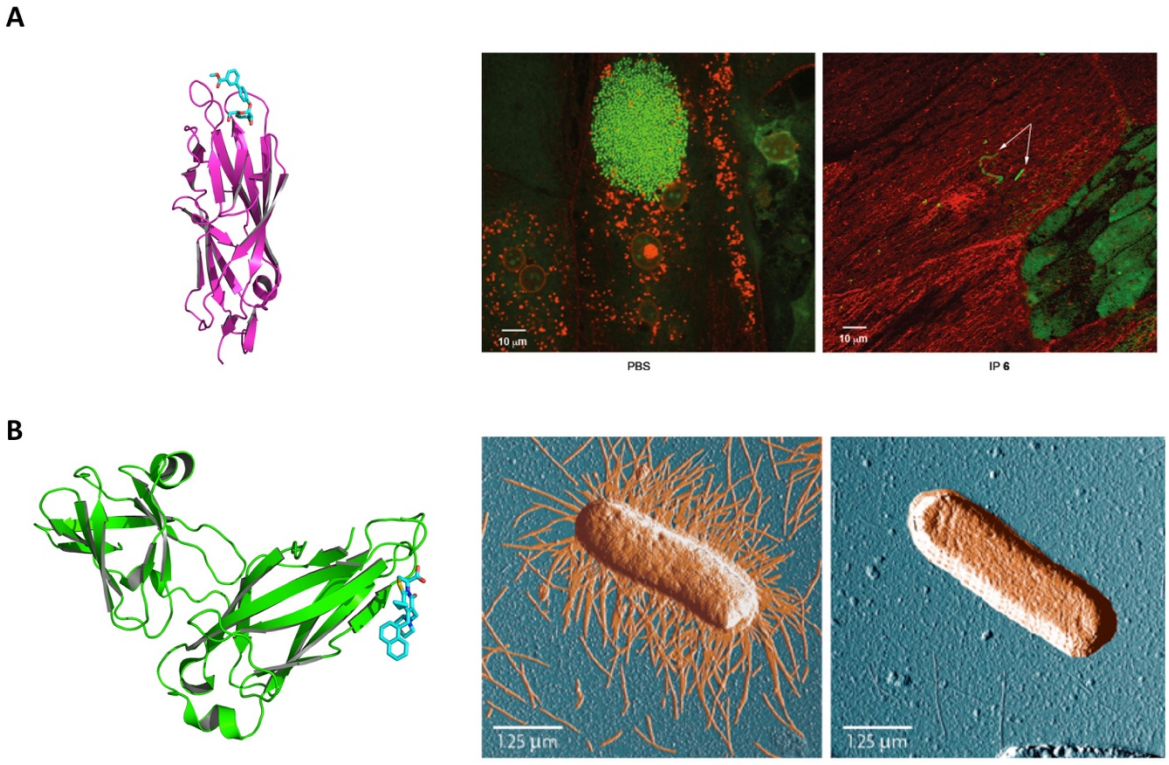


Figure 6. CUP pili as antibacterial targets. **(A)** Crystal structure of a mannoside (cyan) in complex with the FimH lectin domain (magenta) [PDB code: 3MCY]. **(B)** Mannoside prevents IBC formation and can treat established infections [84]. Confocal microscopy of mice bladders depicts an intracellular bacterial community (IBC, green) in the left panel and prevention of IBC formation with mannoside treatment in the right panel (white arrows indicate luminal bacteria). **(C)** Crystal structure of a pilicide (cyan) bound to PapD (green). The pilicide binds to the F1, C1 and D1" strands, a region thought to be the usher-targeting site of the protein, where many of the Set B residues coincide [PDB code: 2J7L]. **(D)** Pilicide inhibits type 1 pilus and P pilus assembly. Atomic force microscopy images show a piliated, untreated bacterium (left panel) and a naked,

pilicide-treated bacterium (right panel), indicating that pilicides suppress pilus biogenesis. Figure parts (B) and (D) are reproduced, with permission, from [84] and [89].

Chapter 2: 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

Science Advances. 2017 Feb 10;3(2):e1601944. PMCID: PMC5302871

Abstract

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 decoy 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 adhesive virulence mechanisms figure prominently in the colonization of the bladder by uropathogenic *Escherichia coli* (UPEC), which account for approximately 85% of all urinary tract infections (UTIs) [6,7]. Mouse models have shown that adhesive hair-like surface appendages called type 1 pili allow UPEC to colonize the bladder epithelium during UTIs by binding to mannosylated receptors on the urothelial surface through the tip-localized adhesin FimH [8]. Clinical observations and mouse models of UTI have shown that FimH mediates the invasion of UPEC into bladder superficial umbrella cells via endocytosis [9-12]. Escape from the endocytic vesicle allows UPEC to replicate within the cytosol to form intracellular bacterial communities (IBCs), which is one mechanism used by UPEC to subvert neutrophil attack, thus facilitating survival and dissemination during UTI [9-12].

Consistent with its vital role in bacterial colonization of the bladder, FimH has evolved within human uropathogenic strains of *E. coli* by positive selection [13]. FimH is composed of the following: (i) an N-terminal lectin domain ($FimH_{LD}$) that binds mannose via a pocket formed by three loops, (ii) a C-terminal pilin domain ($FimH_{PD}$) 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 $FimH_{PD}$ exists in one conformation but $FimH_{LD}$ 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 $FimH_{PD}$ 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 $FimH_{LD}$ and $FimH_{PD}$ do not interact with each other and $FimH_{LD}$ 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 $FimH_{PD}$ [21-23]. Crystal structures show that FimH adopts a compact conformation when bound to FimG, wherein $FimH_{LD}$ and $FimH_{PD}$ interact closely with one another and $FimH_{LD}$ 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 $FimH_{PD}$ can allosterically diminish the ability of $FimH_{LD}$ to bind mannose through interactions with the base of $FimH_{LD}$. Conversely, mannose binding to $FimH_{LD}$ appears to promote $FimH_{LD}$ conformations that no longer interact with $FimH_{PD}$. 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 FimCFFGH 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 mannose-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 $FimG_{Nte}H$ A27V/V163A reveals a novel **R** conformation, wherein the orientations of the $FimH_{LD}$ and $FimH_{PD}$ form a 90° bend (Fig. 2D). This observed conformation represents a substantial structural deviation from the elongated orientation of $FimH_{LD}$ and $FimH_{PD}$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 $FimF_{Nte}H$ bound to heptyl mannoside, which shows $FimH_{LD}$ adopting a ~45° bend angle and a different orientation with respect to $FimH_{PD}$ (fig. S2) [26]. Structural alignment of $FimH_{LD}$'s of the **bent** and **elongated R** conformations indicates that the $FimH_{LD}$'s of the two **R** structures are nearly identical [root mean square deviation (RMSD) = 0.6 Å], displaying the same binding loop and side-chain orientations that are required for high-affinity interactions with mannose (Fig. 2F) [16]. Furthermore, the bent conformation exhibits a new interface between $FimH_{LD}$ and $FimH_{PD}$, composed of a dislodged insertion loop and an extended, solvent-exposed linker (Fig. 2E). No specific hydrogen bonding, electrostatics, or hydrophobic interactions are observed at this new interface, which would allow for a high degree of conformational flexibility in the relative orientations between $FimH_{LD}$ and $FimH_{PD}$ about the extended linker. Instead, multiple contacts formed between symmetry partners within the crystal lattice are coordinated by Ca^{2+}, which likely allowed for crystallographic trapping of the elusive **bent R** conformation (fig. S1, C and D).

Comparison of the **bent R** conformation to the **T** conformation indicates that the orientation of $FimH_{LD}$ and $FimH_{PD}$ in the **bent R** conformation lies on a different bend and rotation axis than their orientation in the **T** conformation, hinting at considerable conformational freedom and possible transition pathways between **T** and **bent/elongated R** states. Together, these two FimH conformations (**T** and **bent R**) provide the first atomic-level description of a functionally diverse structural ensemble adopted by FimH in a native tip-like setting in the absence of ligand.

Positively selected residues and ligand binding modulate conformational ensembles of $FimG_{Nte}H$ variants

Structural ensembles of FimH in solution were probed by small-angle x-ray scattering (SAXS) of FimH variants in $FimG_{Nte}H$ 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 (χ^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 (χ = 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 $FimG_{Nte}H$ complex, each variant adopted very different relative conformation(s) (χ^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.

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 $FimG_{Nte}H$ 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 (χ = 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 $FimG_{Nte}H$ 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 $FimG_{Nte}H$ 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 $FimG_{Nte}H$ 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 $FimG_{Nte}H$ and 4Z269-bound $FimG_{Nte}H$ 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 $FimG_{Nte}H$ 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 $FimG_{Nte}H$ complexes were each similarly best explained by two curves that have mean CCS values of 2464 to 2472 $Å^2$ and 2606 to 2636 $Å^2$, which we respectively label ***A*** and ***B*** (Fig. 4B and table S2). In contrast, the A27V/V163A $FimG_{Nte}H$ profile was best explained by two highly overlapping curves with mean CCS values of 2480 and 2565 $Å^2$, 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 $FimG_{Nte}H$ 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 FimG$_{Nte}$H 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 FimG$_{Nte}$H 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, FimG$_{Nte}$H WT, Q133K, and A62S in the gas phase take on an equilibrium of two distinct conformations with partially overlapping CCS distributions, whereas FimG$_{Nte}$H 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 FimG$_{Nte}$H 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**, **T**, 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.

FimG$_{Nte}$H samples expansive conformational phase space composed of restrained T state and dynamic R state

To gauge the dynamic behavior of FimH in the FimG$_{Nte}$H WT tip-like setting, we performed unrestrained molecular dynamics (MD) simulations in four replicates, each starting from the **T**, **bent R**, and **elongated R** conformations in the presence or absence of ligands. MD trajectories initiated from the **T** state revealed very little structural fluctuation (RMSD) over time, vibrating around a fixed protein shape (R_g distribution), whereas those initiated from the **bent**

R state carved a vast conformational landscape with various overall protein shapes (Fig. 5A and movies S1 and S2). To best compare and conceptualize the sampled conformational landscapes, we parameterized $FimG_{Nte}H$ conformation within a phase space composed of three measurable angular dimensions: (i) bend, (ii) twist, and (iii) orientation of $FimH_{LD}$ relative to $FimH_{PD}$ (fig. S5A). This analysis showed the breadth of conformations sampled in the **R** state compared to the restrained 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 accessible but restrained by the architecture of the $FimH_{LD}$ and $FimH_{PD}$ interface, which functions as a "socket." We also observed considerable 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.

Dynamics and binding mechanisms of conformational populations in $FimG_{Nte}H$ WT

Mannose was then included in simulations of $FimG_{Nte}H$ 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 occured 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 anomeric 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(β1–4)-GlcNAc(β1–4)-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 $FimH_{LD}$ 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 $FimG_{Nte}H$ 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 $FimH_{LD}$ 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 mannoside 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 facet 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 facet cells that line the undisrupted 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 mannosylated 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 preexisting 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 $FimH_{LD}$ 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 mannose-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 mannose 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 mannose-binding pocket mutant Q133K, which cannot bind mannose, 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]. $FimG_{Nte}H$ complexes were assembled by a spontaneous in vitro DSE reaction, in which $FimG_{Nte}$ 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 hours. $FimG_{Nte}$ displaces FimC in this reaction, and resultant $FimG_{Nte}H$ variant complexes were

purified away from excess $FimG_{Nte}$ 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 $FimG_{Nte}H$ 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 $FimG_{Nte}H$ variant complexes for at least 1 hour at 4°C before biophysical analysis at molar ratios indicated below.

Differential scanning fluorimetry

$FimG_{Nte}H$ 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 [$y = A_2 + (A_1 - A_2)/(1 + \exp((x - x_o)/dx))$, where x_o 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 Sigma except crude THP, which was isolated by ammonium sulfate precipitation of urine donated by healthy volunteers. Coated wells were then incubated with 200 µl of blocking buffer [phosphate-buffered saline (PBS) + 2% bovine serum albumin (BSA)] for 2 hours at 23°C, followed by incubation with 100 µl of $FimG_{Nte}H$ variants diluted in blocking buffer to 1 µg/ml for 1 hour at 37°C. After washing three times with PBS + 0.05% Tween 20, 100 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 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 H_2SO_4, and absorbance was measured at 450 nm.

Small-angle x-ray scattering

Before sample submission, FimCH and $FimG_{Nte}H$ 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. $FimG_{Nte}H$:4Z269 complexes were prepared in the same manner after $FimG_{Nte}H$ 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 (D_{max}) 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 multistate 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

$FimG_{Nte}H$ 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) (>99.99% 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. $FimG_{Nte}H$:4Z269 complexes were prepared in the same manner after $FimG_{Nte}H$ 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 $FimG_{Nte}$, 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]. $FimG_{Nte}H$ 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 $FimG_{Nte}H$ complex at low CE but steadily dissociated with increasing CE (fig. S3A). These spectra also revealed that $FimG_{Nte}H$ 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 $FimG_{Nte}H$ 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 $FimG_{Nte}H$ A62S grew under numerous polyethylene glycol (PEG) conditions at 20°C by hanging drop vapor diffusion in 96-well plates. The diffracted crystals of $FimG_{Nte}H$ 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 rectangular prism morphology and were picked directly from the 96-well screen. They were transferred 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 $FimG_{Nte}H$ 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 cadmium 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 marmux 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) $FimH_{PD}$ (from PDB ID 1KLF) and $FimH_{LD}$ in the **T** state (from PDB ID 3JWN) as separate domains and (ii) $FimH_{PD}$ (from PDB ID 1KLF) and $FimH_{LD}$ in the **R** state (from PDB ID 1KLF) as separate domains. The best solutions were identified as a **T** conformation for $FimG_{Nte}H$ A62S and a **bent R** conformation (90° between $FimH_{LD}$ and $FimH_{PD}$) for $FimG_{Nte}H$ A27V/V163A. Several rounds of refinements were performed in phenix.refine to improve the final models.

Molecular dynamics

Crystal structures of $FimG_{Nte}H$ published in this study along with models of $FimG_{Nte}H$ based on previous crystal structures (G_{Nte} docked on H from CH structure) in the absence or presence of ligands were used to initiate MD simulations. Structural models of $FimG_{Nte}H$ 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** (5JQI), **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 coordinates of the following atoms were tracked and assigned the corresponding labels: Leu225 and CD2 (pip1 for "point in pilin 1"), Leu193 and CD1 (pip2), Ala247 and N (pip3), Leu193 and CA (pip4), Leu129 and CD2 (pil1 for "point in lectin 1"), Val20 and CG1 (pil2), and Val105 and N (pil3). The points {pip1, pip2} and {pip2, pil1} define vectors along the longest dimension of $FimH_{PD}$ and $FimH_{LD}$, respectively. The points {pip2, pip3, pip4} and {pil1, pil2, pil3} define transverse planes that respectively bisect $FimH_{PD}$ and $FimH_{LD}$ perpendicular to the axis of the longest dimension. Bend angle is defined as the {pip1, pip2, pil1} angle. Twist angle is defined as the {pip1, pip2, *x*} angle, where *x* is defined as the point on the {pil1, pil2, pil3} plane at a fixed radial distance from pil1 that is closest to the {pip2, pip3, pip4} plane. Orientation

angle is defined as the {pip3, pip2, *y*} angle, where *y* is the point on the {pip2, pip3, pip4} plane closest to *x*. These angles are schematically represented in fig. S5A.

Mouse infections

For the uncomplicated UTI model, 7- to 8-week-old female C3H/HeN mice were obtained from Envigo. Mice were anesthetized and inoculated via transurethral catheterization with 50 μl of bacterial suspension (~1×10^7 to 2×10^7 CFU in total) in PBS. At times indicated, mice were sacrificed, and bladders were aseptically removed and processed for CFU determination [37]. For the CAUTI model, a 5-mm piece of silicon tubing (RenaSil 0.635-mm outer diameter) was inserted transurethrally into the bladders of 6-week-old female C57BL/6 mice, as previously described [37]. Twenty-four hours after implantation, the mice were transurethrally infected with ~2×10^7 CFU of UTI89 expressing either the WT or A27V/V163A allele of FimH with the *fim* operon locked in the phase on state [13]. Mice were sacrificed 24 or 72 hpi, and bacteria colonizing the bladder and implant were plated for quantification.

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% CO_2. 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.

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.

References

1. Croxen M. A., Finlay B. B., Molecular mechanisms of *Escherichia coli* pathogenicity. Nat. Rev. Microbiol. 8, 26–38 (2010).
2. Rasko D. A., Sperandio V., Anti-virulence strategies to combat bacteria-mediated disease. Nat. Rev. Drug Discov. 9, 117–128 (2010).
3. Persat A., Nadell C. D., Kim M. K., Ingremeau F., Siryaporn A., Drescher K., Wingreen N. S., Bassler B. L., Gitai Z., Stone H. A., The mechanical world of bacteria. Cell 161, 988–997 (2015).
4. Kline K. A., Fälker S., Dahlberg S., Normark S., Henriques-Normark B., Bacterial adhesins in host-microbe interactions. Cell Host Microbe 5, 580–592 (2009).
5. Flemming H.-C., Wingender J., Szewzyk U., Steinberg P., Rice S. A., Kjelleberg S., Biofilms: An emergent form of bacterial life. Nat. Rev. Microbiol. 14, 563–575 (2016).
6. Ronald A. R., Nicolle L. E., Stamm E., Krieger J., Warren J., Schaeffer A., Naber K. G., Hooton T. M., Johnson J., Chambers S., Andriole V., Urinary tract infection in adults: Research priorities and strategies. Int. J. Antimicrob. Agents 17, 343–348 (2001).
7. Waksman G., Hultgren S. J., Structural biology of the chaperone-usher pathway of pilus biogenesis. Nat. Rev. Microbiol. 7, 765–774 (2009).
8. Mulvey M. A., Lopez-Boado Y. S., Wilson C. L., Roth R., Parks W. C., Heuser J., Hultgren S. J.,Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. Science 282, 1494–1497 (1998).
9. Anderson G. G., Palermo J. J., Schilling J. D., Roth R., Heuser J., Hultgren S. J., Intracellular bacterial biofilm-like pods in urinary tract infections. Science 301, 105–107 (2003).
10. Justice S. S., Hung C., Theriot J. A., Fletcher D. A., Anderson G. G., Footer M. J., Hultgren S. J.,Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. Proc. Natl. Acad. Sci. U.S.A. 101, 1333–1338 (2004).
11. Rosen D. A., Hooton T. M., Stamm W. E., Humphrey P. A., Hultgren S. J., Detection of intracellular bacterial communities in human urinary tract infection. PLOS Med. 4, e329 (2007).
12. Robino L., Scavone P., Araujo L., Algorta G., Zunino P., Pírez M. C., Vignoli R., Intracellular bacteria in the pathogenesis of *Escherichia coli* urinary tract infection in children. Clin. Infect. Dis. 59, e158–e164 (2014).
13. Chen S. L., Hung C. S., Pinkner J. S., Walker J. N., Cusumano C. K., Li Z., Bouckaert J., Gordon J. I., Hultgren S. J., 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. 106, 22439–22444 (2009).
14. Schwartz D. J., Kalas V., Pinkner J. S., Chen S. L., Spaulding C. N., Dodson K. W., Hultgren S. J., Positively selected FimH residues enhance virulence during urinary tract infection by altering FimH conformation. Proc. Natl. Acad. Sci. U.S.A. 110, 15530–15537 (2013).
15. Choudhury D., Thompson A., Stojanoff V., Langermann S., Pinkner J., Hultgren S. J., Knight S. D., X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. Science285, 1061–1066 (1999).
16. Hung C.-S., Bouckaert J., Hung D., Pinkner J., Widberg C., DeFusco A., Auguste C. G., Strouse R., Langermann S., Waksman G., Hultgren S. J., Structural basis of tropism

of *Escherichia coli* to the bladder during urinary tract infection. Mol. Microbiol. 44, 903–915 (2002).
17. Le Trong I., Aprikian P., Kidd B. A., Forero-Shelton M., Tchesnokova V., Rajagopal P., Rodriguez V., Interlandi G., Klevit R., Vogel V., Stenkamp R. E., Sokurenko E. V., Thomas W. E., Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like β sheet twisting. Cell 141, 645–655 (2010).
18. Phan G., Remaut H., Wang T., Allen W. J., Pirker K. F., Lebedev A., Henderson N. S., Geibel S., Volkan E., Yan J., Kunze M. B. A., Pinkner J. S., Ford B., Kay C. W. M., Li H., Hultgren S. J., Thanassi D. G., Waksman G., Crystal structure of the FimD usher bound to its cognate FimC–FimH substrate. Nature474, 49–53 (2011).
19. Geibel S., Procko E., Hultgren S. J., Baker D., Waksman G., Structural and energetic basis of folded-protein transport by the FimD usher. Nature 496, 243–246 (2013).
20. Sauer F. G., Fütterer K., Pinkner J. S., Dodson K. W., Hultgren S. J., Waksman G., Structural basis of chaperone function and pilus biogenesis. Science 285, 1058–1061 (1999).
21. Jones C. H., Pinkner J. S., Roth R., Heuser J., Nicholes A. V., Abraham S. N., Hultgren S. J., FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc. Natl. Acad. Sci. U.S.A. 92, 2081–2085 (1995).
22. Sauer F. G., Pinkner J. S., Waksman G., Hultgren S. J., Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. Cell 111, 543–551 (2002).
23. Remaut H., Rose R. J., Hannan T. J., Hultgren S. J., Radford S. E., Ashcroft A. E., Waksman G., Donor-strand exchange in chaperone-assisted pilus assembly proceeds through a concerted β strand displacement mechanism. Mol. Cell 22, 831–842 (2006).
24. Yakovenko O., Sharma S., Forero M., Tchesnokova V., Aprikian P., Kidd B., Mach A., Vogel V., Sokurenko E., Thomas W. E., FimH forms catch bonds that are enhanced by mechanical force due to allosteric regulation. J. Biol. Chem. 283, 11596–11605 (2008).
25. Chen S. L., Hung C.-S., Xu J., Reigstad C. S., Magrini V., Sabo A., Blasiar D., Bieri T., Meyer R. R., Ozersky P., Armstrong J. R., Fulton R. S., Latreille J. P., Spieth J., Hooton T. M., Mardis E. R., Hultgren S. J., Gordon J. I., Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: A comparative genomics approach. Proc. Natl. Acad. Sci. U.S.A. 103, 5977–5982 (2006).
26. Sauer M. M., Jakob R. P., Eras J., Baday S., Eriş D., Navarra G., Bernèche S., Ernst B., Maier T., Glockshuber R., Catch-bond mechanism of the bacterial adhesin FimH. Nat. Commun. 7, 10738 (2016).
27. Cusumano C. K., Pinkner J. S., Han Z., Greene S. E., Ford B. A., Crowley J. R., Henderson J. P., Janetka J. W., Hultgren S. J., Treatment and prevention of urinary tract infection with orally active FimH inhibitors. Sci. Transl. Med. 3, 109ra115 (2011).
28. Schneidman-Duhovny D., Hammel M., Sali A., FoXS: A web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res. 38, W540–W544 (2010).
29. Bouckaert J., Mackenzie J., De Paz J. L., Chipwaza B., Choudhury D., Zavialov A., Mannerstedt K., Anderson J., Piérard D., Wyns L., Seeberger P. H., Oscarson S., De Greve H., Knight S. D., The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of *Escherichia coli* pathotypes. Mol. Microbiol. 61, 1556–1568 (2006).
30. Wellens A., Garofalo C., Nguyen H., Van Gerven N., Slättegård R., Hernalsteens J.-P., Wyns L., Oscarson S., De Greve H., Hultgren S., Bouckaert J., Intervening with urinary

tract infections using anti-adhesives based on the crystal structure of the FimH–oligomannose-3 complex. PLOS ONE 3, e2040 (2008).
31. Wellens A., Lahmann M., Touaibia M., Vaucher J., Oscarson S., Roy R., Remaut H., Bouckaert J., The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. Biochemistry 51, 4790–4799 (2012).
32. Rodriguez V. B., Kidd B. A., Interlandi G., Tchesnokova V., Sokurenko E. V., Thomas W. E., Allosteric coupling in the bacterial adhesive protein FimH. J. Biol. Chem. 288, 24128–24139 (2013).
33. Fogh J., Cultivation, characterization, and identification of human tumor cells with emphasis on kidney, testis, and bladder tumors. Natl. Cancer Inst. Monogr. 5–9 (1978).
34. Guiton P. S., Hung C. S., Hancock L. E., Caparon M. G., Hultgren S. J., Enterococcal biofilm formation and virulence in an optimized murine model of foreign body-associated urinary tract infections. Infect. Immun. 78, 4166–4175 (2010).
35. Murphy C. N., Mortensen M. S., Krogfelt K. A., Clegg S., Role of *Klebsiella pneumoniae* type 1 and type 3 fimbriae in colonizing silicone tubes implanted into the bladders of mice as a model of catheter-associated urinary tract infections. Infect. Immun. 81, 3009–3017 (2013).
36. Flores-Mireles A. L., Pinkner J. S., Caparon M. G., Hultgren S. J., EbpA vaccine antibodies block binding of *Enterococcus faecalis* to fibrinogen to prevent catheter-associated bladder infection in mice. Sci. Transl. Med. 6, 254ra127 (2014).
37. Conover M. S., Flores-Mireles A. L., Hibbing M. E., Dodson K., Hultgren S. J., Establishment and characterization of UTI and CAUTI in a mouse model. J. Vis. Exp., e52892 (2015).
38. Raffi H. S., Bates J. M., Flournoy D. J., Kumar S., Tamm-Horsfall protein facilitates catheter associated urinary tract infection. BMC Res. Notes 5, 532 (2012).
39. Monod J., Changeux J. P., Jacob F., Allosteric proteins and cellular control systems. J. Mol. Biol. 6, 306–329 (1963).
40. Koshland D. E. Jr, Némethy G., Filmer D., Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5, 365–385 (1966).
41. Perutz M. F., Stereochemistry of cooperative effects in haemoglobin. Nature 228, 726–739 (1970).
42. Cui Q., Karplus M., Allostery and cooperativity revisited. Protein Sci. 17, 1295–1307 (2008).
43. Motlagh H. N., Wrabl J. O., Li J., Hilser V. J., The ensemble nature of allostery. Nature 508, 331–339 (2014).
44. Thomas W. E., Trintchina E., Forero M., Vogel V., Sokurenko E. V., Bacterial adhesion to target cells enhanced by shear force. Cell 109, 913–923 (2002).
45. Thomas W. E., Nilsson L. M., Forero M., Sokurenko E. V., Vogel V., Shear-dependent 'stick-and-roll' adhesion of type 1 fimbriated *Escherichia coli*. Mol. Microbiol. 53, 1545–1557 (2004).
46. Yakovenko O., Tchesnokova V., Sokurenko E. V., Thomas W. E., Inactive conformation enhances binding function in physiological conditions. Proc. Natl. Acad. Sci. U.S.A. 112, 9884–9889 (2015).
47. Greene S. E., Hibbing M. E., Janetka J., Chen S. L., Hultgren S. J., Human urine decreases function and expression of type 1 pili in uropathogenic *Escherichia coli*. MBio 6, e00820 (2015).

48. Dyer K. N., Hammel M., Rambo R. P., Tsutakawa S. E., Rodic I., Classen S., Tainer J. A., Hura G. L.,High-throughput SAXS for the characterization of biomolecules in solution: A practical approach. Methods Mol. Biol. 1091, 245–258 (2014).
49. Petoukhov M. V., Franke D., Shkumatov A. V., Tria G., Kikhney A. G., Gajda M., Gorba C., Mertens H. D. T., Konarev P. V., Svergun D. I., New developments in the *ATSAS* program package for small-angle scattering data analysis. J. Appl. Crystallogr. 45, 342–350 (2012).
50. Hura G. L., Budworth H., Dyer K. N., Rambo R. P., Hammel M., McMurray C. T., Tainer J. A., Comprehensive macromolecular conformations mapped by quantitative SAXS analyses. Nat. Methods 10, 453–454 (2013).
51. Svergun D. I., Petoukhov M. V., Koch M. H., Determination of domain structure of proteins from X-ray solution scattering. Biophys. J. 80, 2946–2953 (2001).
52. Baker D., Krukowski A. E., Agard D. A., Uniqueness and the ab initio phase problem in macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 49, 186–192 (1993).
53. Michaelevski I., Kirshenbaum N., Sharon M., T-wave ion mobility-mass spectrometry: Basic experimental procedures for protein complex analysis. J. Vis. Exp., 1985 (2010).
54. Battye T. G. G., Kontogiannis L., Johnson O., Powell H. R., Leslie A. G., *iMOSFLM*: A new graphical interface for diffraction-image processing with *MOSFLM*. Acta Crystallogr. D Biol. Crystallogr. 67, 271–281 (2011).
55. Winn M. D., Ballard C. C., Cowtan K. D., Dodson E. J., Emsley P., Evans P. R., Keegan R. M., Krissinel E. B., Leslie A. G. W., McCoy A., McNicholas S. J., Murshudov G. N., Pannu N. S., Potterton E. A., Powell H. R., Read R. J., Vagin A., Wilson K. S., Overview of the *CCP*4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242 (2011).
56. Adams P. D., Afonine P. V., Bunkóczi G., Chen V. B., Davis I. W., Echols N., Headd J. J., Hung L.-W., Kapral G. J., Grosse-Kunstleve R. W., McCoy A. J., Moriarty N. W., Oeffner R., Read R. J., Richardson D. C., Richardson J. S., Terwilliger T. C., Zwart P. H., *PHENIX*: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221 (2010).
57. Berendsen H. J. C., van der Spoel D., van Drunen R., GROMACS: A message-passing parallel molecular-dynamics implementation. Comput. Phys. Commun. 91, 43–56 (1995).

Figures

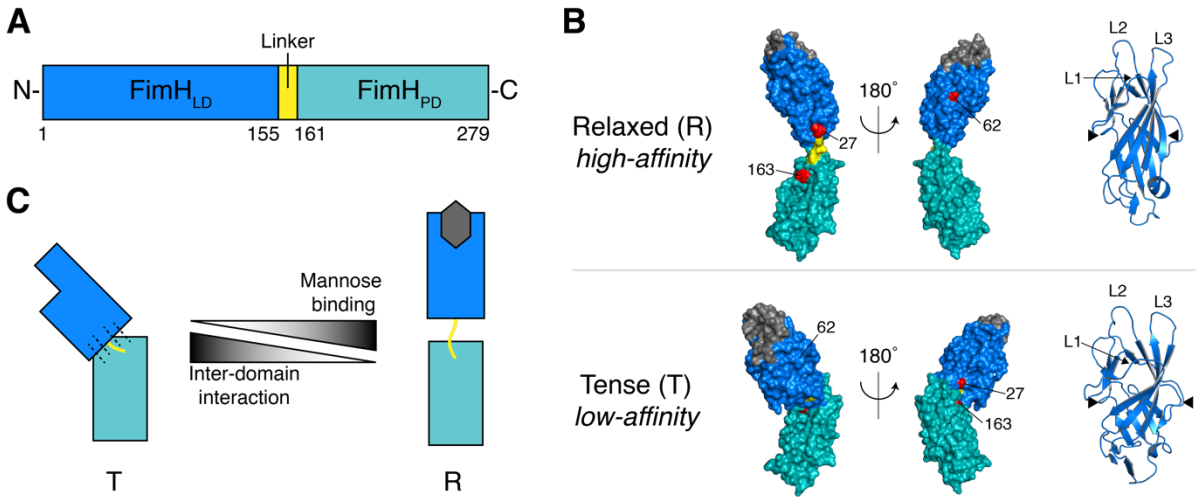


Figure 1. Structure-function relationship in the type 1 adhesin FimH. (**A**) Schematic representation of FimH sequence. The lectin domain ($FimH_{LD}$, residues 1 to 155) is colored blue, the linker is colored yellow (residues 156 to 160), and the pilin domain ($FimH_{PD}$, residues 161 to 279) is colored teal. (**B**) Comparison of previously identified conformations of FimH. FimH bound to FimC in a FimCH complex exists in a high-affinity conformer, or **R** state, with an elongated orientation between $FimH_{LD}$ and $FimH_{PD}$, a narrowly packed β-sandwich fold in $FimH_{LD}$ (as highlighted between the two black triangles), and packed mannose-binding loops (labeled as L1, L2, and L3). FimH in a tip assembly (FimCFFGH complex) adopts a low-affinity conformer, or **T** state, with a compacted orientation between $FimH_{LD}$ and $FimH_{PD}$, a widened β-sandwich fold in $FimH_{LD}$ (as highlighted between the two black triangles), and displaced binding loops (particularly L1). Positively selected residues are indicated as red spheres, and the mannose-binding pocket is shaded gray. (**C**) Schematic representation of the negatively coupled allosteric relationship between mannose and the interface between FimH domains, whereby increases in mannose binding disfavor contacts between $FimH_{LD}$ and $FimH_{PD}$ and vice versa.

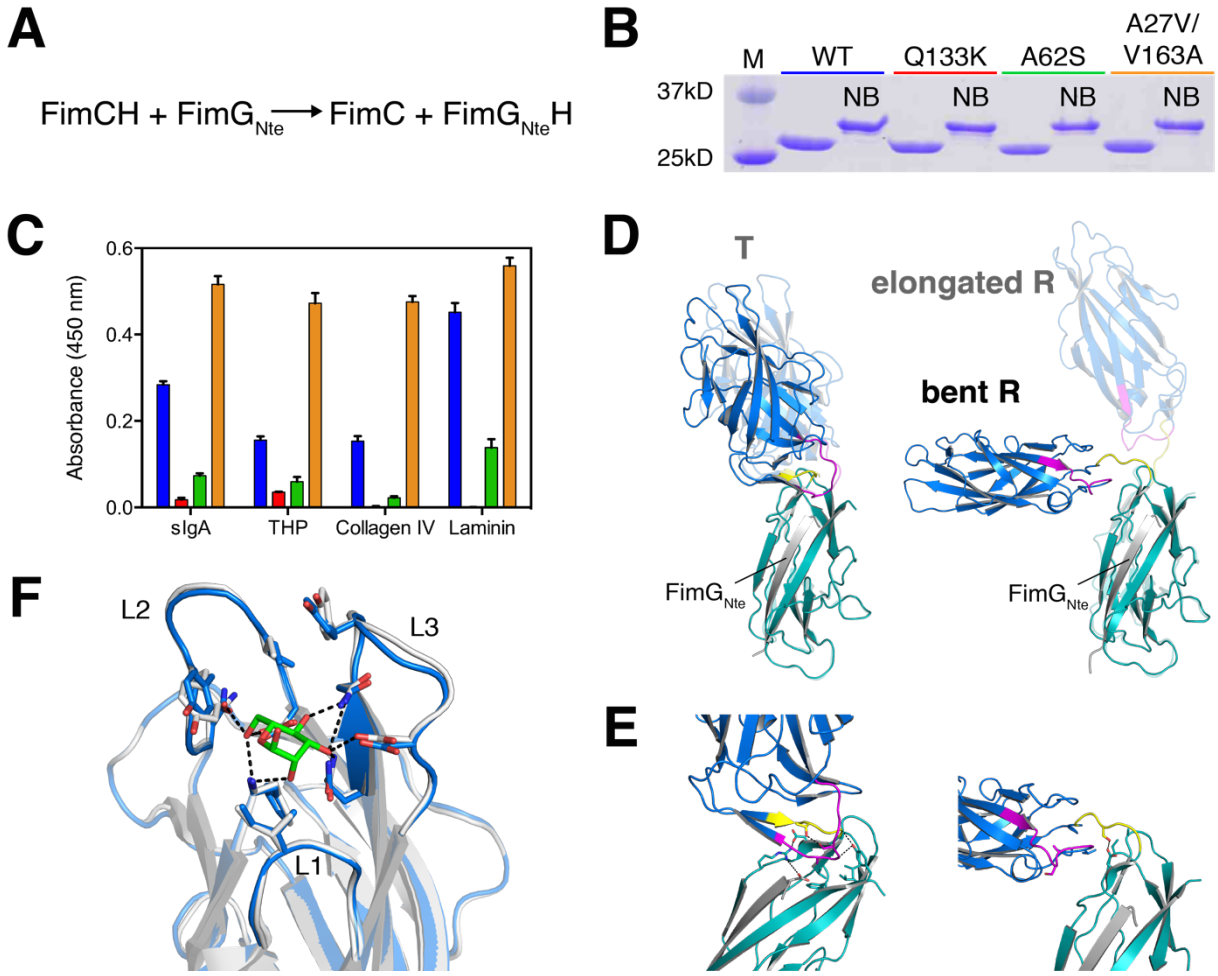


Figure 2. Crystal structures and mannose binding of FimH variants in a tip-like setting. (**A**) Reaction scheme of in vitro DSE reaction to produce tip-like FimG$_{Nte}$H complexes. (**B**) Representative SDS–polyacrylamide gel electrophoresis of purified FimG$_{Nte}$H variants either boiled (not labeled) or not boiled (NB). (**C**) Enzyme-linked immunosorbent assay (ELISA) measuring binding of FimG$_{Nte}$H variant complexes to surface-coated glycoproteins, which include secretory IgA (sIgA), Tamm-Horsfall protein (THP), collagen IV, and laminin. (**D**) Crystal structures of FimG$_{Nte}$H A62S [Protein Data Bank (PDB) ID 5JQI, left] and FimG$_{Nte}$H A27V/V163A (PDB ID 5JR4, right) depicted as ribbons. These structures are overlaid on previously solved crystal structures of FimH in a FimCFFGH complex (3JWN) and FimCH complex (1KLF), respectively. Conformations are labeled accordingly. FimH$_{LD}$, linker, and FimH$_{PD}$ are colored as in

(A), the insertion loop (residues 109 to 124) is colored purple, and $FimG_{Nte}$ is colored gray. (**E**) $FimH_{LD}$-$FimH_{PD}$ interface in $FimG_{Nte}H$ A62S (left) and $FimG_{Nte}H$ A27V/V163A (right). Contacts between residues are indicated as black dotted lines. (**F**) Structural alignment of $FimG_{Nte}H$ A27V/V163A (colored blue) to $FimH_{LD}$ of mannose-bound FimCH (colored white). Residue side chains and mannose in green are depicted as sticks. Contacts between mannose and FimH are indicated as black dotted lines.

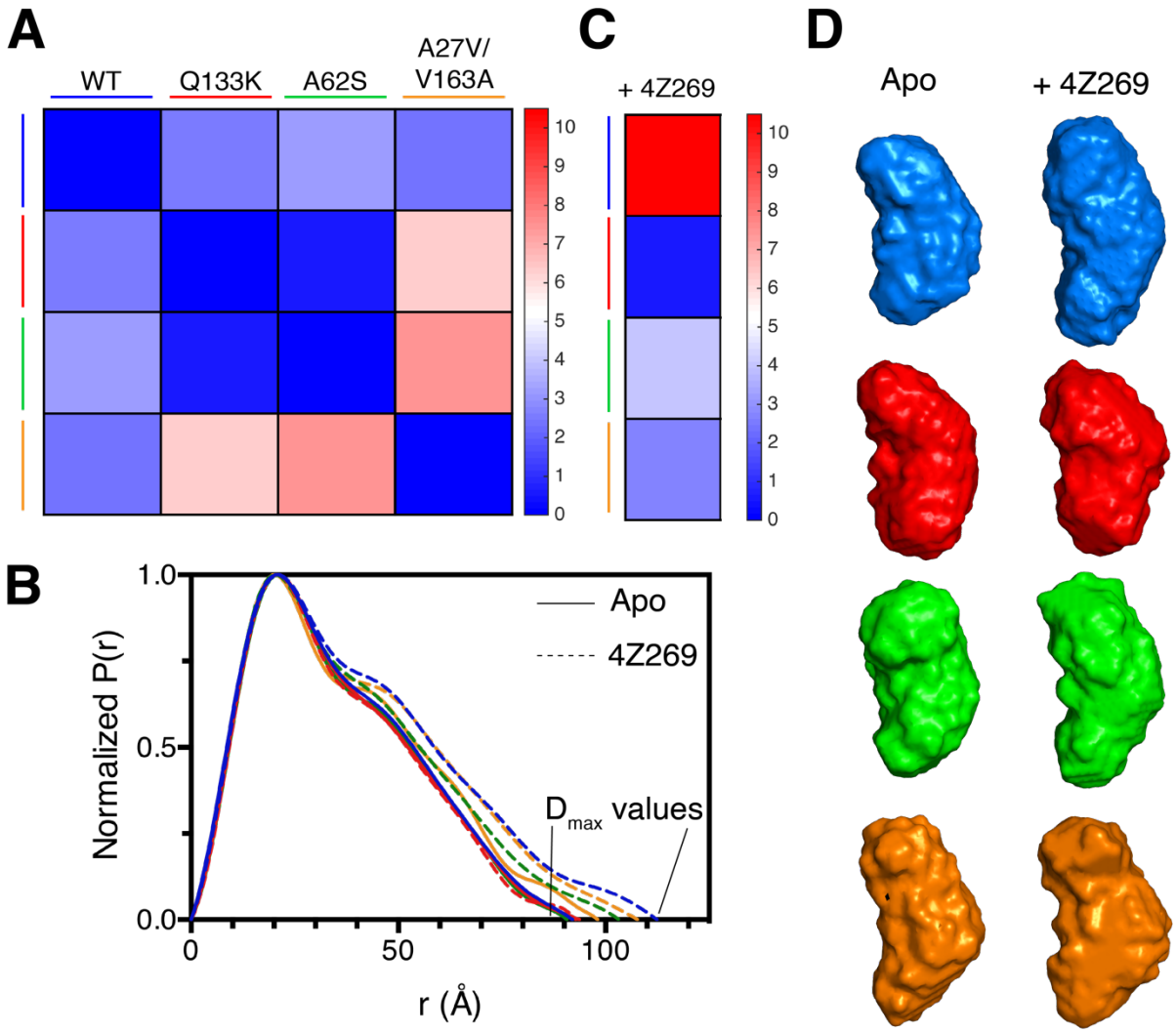


Figure 3. Conformational ensembles of apo and ligated FimH variants in solution. (**A**) Structural comparison heat map of SAXS profiles indicates varying degrees of conformational similarity among $FimG_{Nte}H$ variants, as measured by χ^2, ranging from high (blue) to low (red) similarity. (**B**) Normalized pair distance distributions of $FimG_{Nte}H$ variants in the absence or presence of 4Z269. $FimG_{Nte}H$ variants are represented as solid lines in the absence of 4Z269 or dotted lines in the presence of 4Z269 and color-coded, as indicated by the colored lines in (A). (**C**) Structural comparison heat map indicates varying degrees of conformational similarity of each $FimG_{Nte}H$ variant in the absence or presence of mannoside 4Z269 at a 2× molar ratio. Color-coded

as in (A). (**D**) Averaged ab initio models of $FimG_{Nte}H$ variants in the absence or presence of 4Z269 are color-coded, as previously indicated.

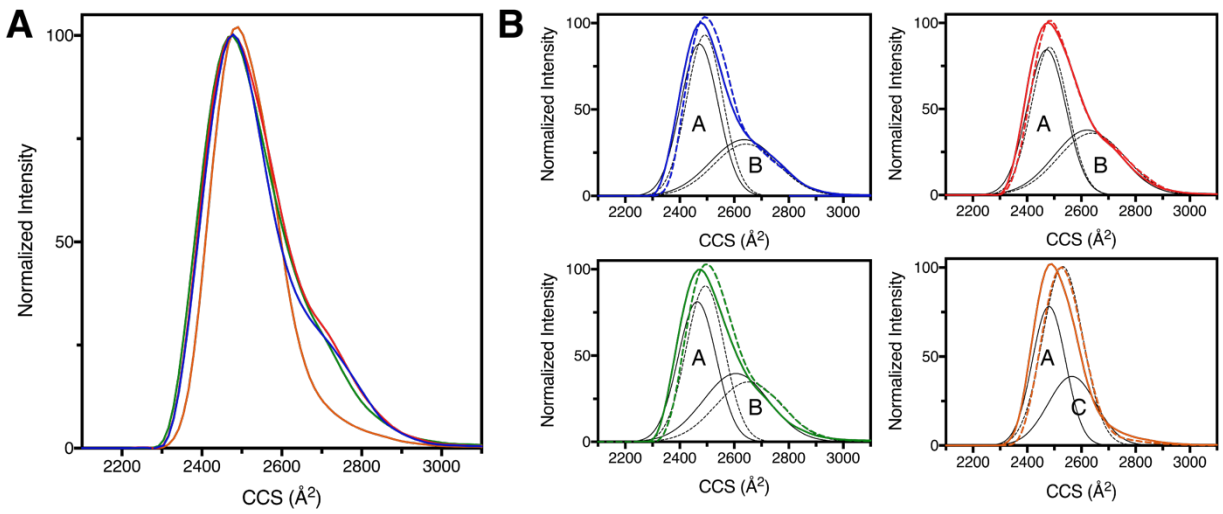


Figure 4. Conformational distributions of free and 4Z269-bound FimG$_{Nte}$H variants isolated in the gas phase, as revealed by IMMS. (**A**) CCS distributions of intact FimG$_{Nte}$H variant complexes measured by IMMS. (**B**) Comparison of CCS distributions of free (solid line) and 4Z269-bound (dotted line) FimG$_{Nte}$H variants. The solid and dotted black lines represent fitted Gaussian distributions to apo and ligated FimG$_{Nte}$H, 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 FimG$_{Nte}$H Q133K spectral peak from the sample that was treated with 4Z269.

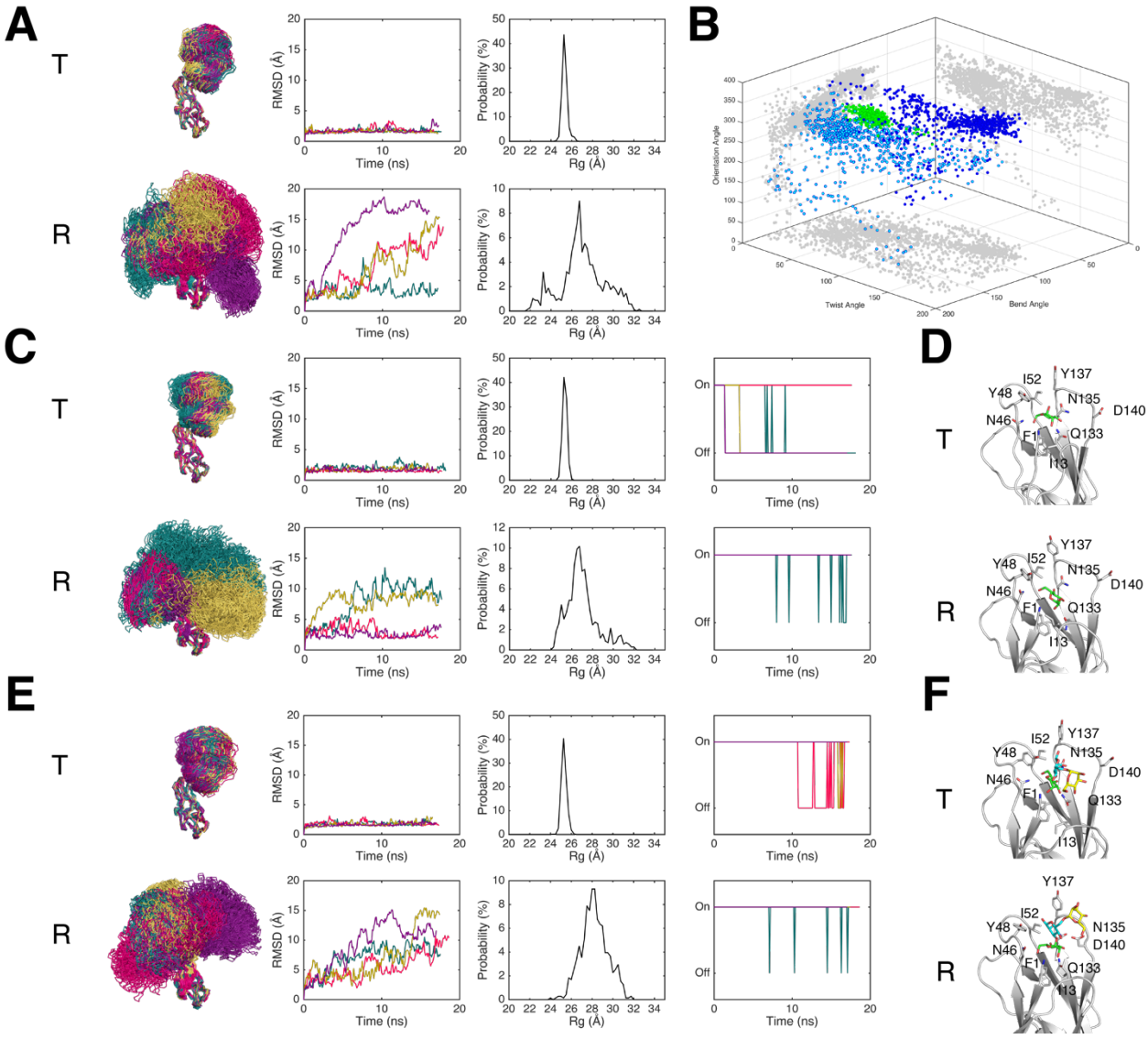


Figure 5. Dynamics and binding mechanisms of conformational populations in FimG$_{Nte}$H WT. (**A**) Structures revealed by MD simulations of FimG$_{Nte}$H WT in a **T** conformation (top) or **bent R** conformation (bottom) with corresponding measures of structural fluctuation over time (RMSD) and distributions of sampled protein shapes (R_g). Different colors correspond to four independent simulation replicates. (**B**) Three-dimensional conformational phase space of FimG$_{Nte}$H as defined by bend, twist, and orientation angles for simulations initiated from the **T** (green), **bent R** (blue), or **elongated R** (cyan) conformation. Shadows are cast on the grid panels and colored in gray. (**C**) Structures revealed by MD simulations of FimG$_{Nte}$H WT in

a **T** conformation (top) or **bent R** conformation (bottom) in the presence of mannose, with corresponding measures of structural fluctuation over time (RMSD), distributions of sampled protein shapes (R_g), and mannose binding. "On" and "Off" measure whether the center of mass of mannose is within or outside 10 Å of the carbonyl of residue F1 in the binding pocket. (**D**) Representative binding modes of mannose for **T** (top) and **bent R** (bottom) after 5 ns. Mannose is depicted as sticks, whereas FimH is shown as a ribbon representation. (**E**) Structures revealed by MD simulations of $FimG_{Nte}H$ WT in a **T** (top) or **bent R** (bottom) conformation in the presence of oligomannose-3 (Man(α1–3)-[Man(α1–6)]-Man), with corresponding measures of structural fluctuation over time (RMSD), distributions of sampled protein shapes (R_g), and mannose binding. "On" and "Off" measure whether the center of mass of oligomannose-3 is within or outside 20 Å of the carbonyl of residue F1 in the binding pocket. (**F**) Representative binding modes of oligomannose-3 for **T** (top) and **bent R** (bottom) after 8.5 ns. Man(α1–3)-[Man(α1–6)]-Man is depicted as sticks and colored green, yellow, and cyan, respectively.

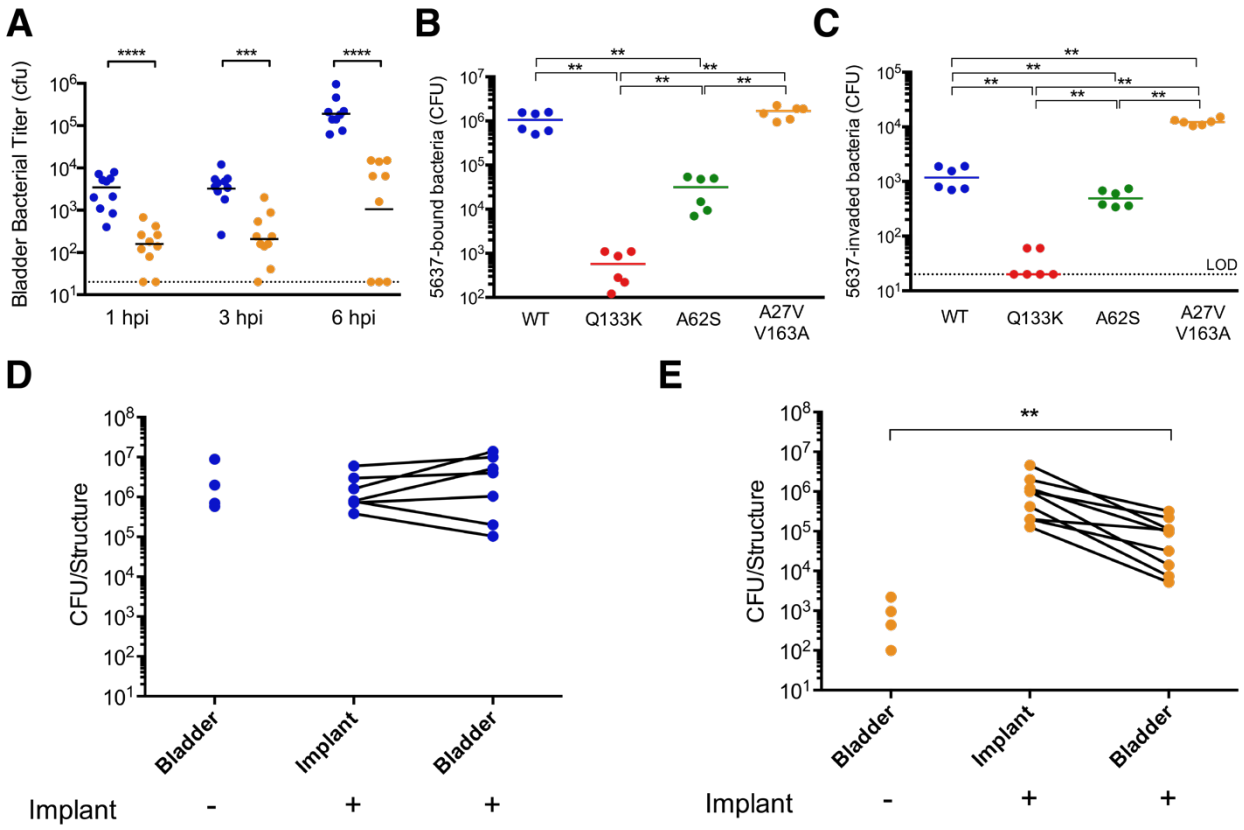


Figure 6. Role of FimH conformation in bladder colonization during UTI. (**A**) Bacterial titers of mouse bladders infected with UTI89 harboring either FimH WT (blue) or A27V/V163A (orange) at an inoculum of 10^7colony-forming units (CFU) measured at 1, 3, and 6 hpi. (**B**) Total bacterial titers of 5637 bladder epithelial cells (no gentamicin treatment) infected with UTI89 harboring WT (blue), Q133K (red), A62S (green), or A27V/V163A (orange) FimH at an inoculum of 10^7 CFU. (**C**) Invaded bacterial titers of 5637 bladder epithelial cells (treated with gentamicin) infected with UTI89 harboring WT (blue), Q133K (red), A62S (green), or A27V/V163A (orange) FimH at an inoculum of 10^7 CFU. LOD, limit of detection. (**D** and **E**) Bacterial titers of C57BL/6 mouse bladders without catheterization or bladders and implants infected with UTI89 harboring either FimH WT (blue) or A27V/V163A (orange) at an inoculum of 10^7 CFU 24 hours after catheterization. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, two-tailed Mann-Whitney *U* test.

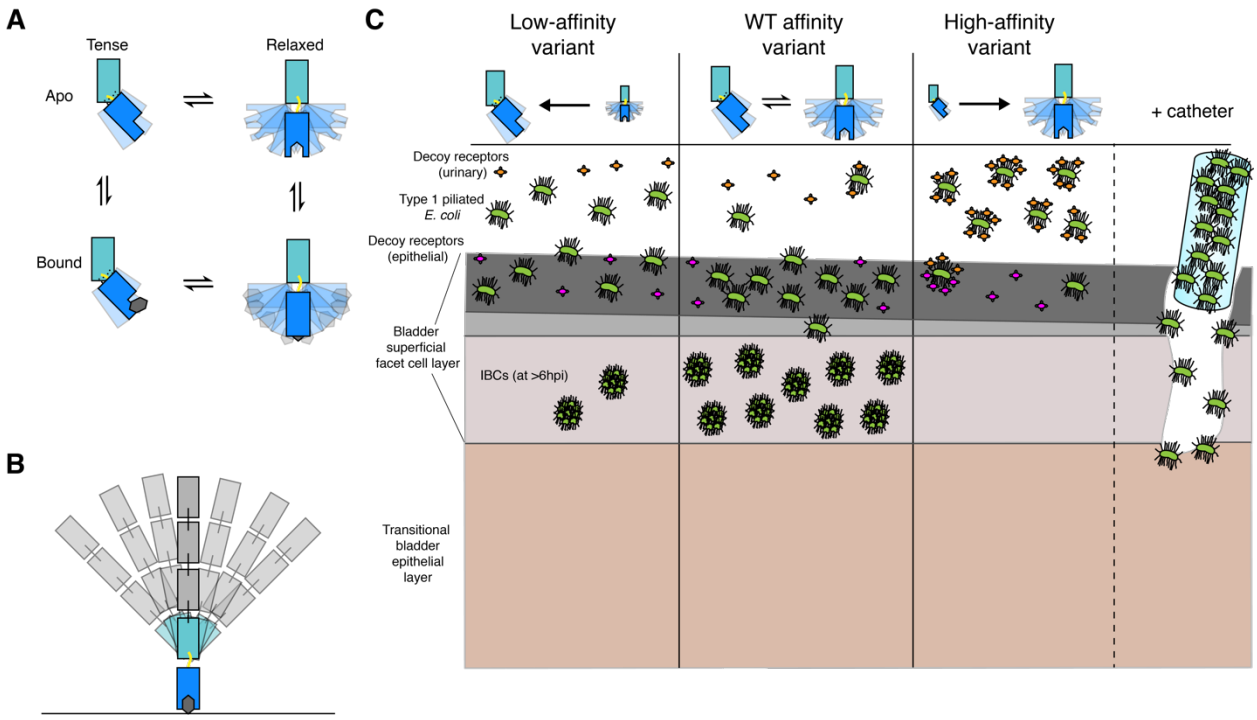


Figure 7. Proposed model of FimH conformational ensembles, mannose binding, and virulence in UTI. (**A**) Two-state conformational landscape of FimH. FimH at the pilus tip natively adopts an equilibrium of a single, dynamically restrained, low-affinity **T** state and multiple, highly dynamic, high-affinity **R** states with various bends, twists, and orientations. Positively selected residues can shift this preexisting conformational equilibrium and thereby influence mannose-binding affinity. The **T** and **R** states can bind mannose. Mannose in a tilted orientation rapidly enters into the widened and shallow binding pocket of the **T** state. Mannosylated ligands in a bound **T** state can then rotate in a high-affinity orientation and allosterically trigger structural perturbations that disrupt $FimH_{LD}$ and $FimH_{PD}$interactions and facilitate conversion to the bound **R** state. In addition, mannose in a horizontal orientation can less rapidly engage the **R** state but does so very tightly through hydrogen bond interactions with several binding loop residues. Positive selection, in modulating a native conformational equilibrium, likely alters flux through these two distinct binding mechanisms. (**B**) Schematic model of the FimH molecular tether. The

bends, twists, and orientations between $FimH_{LD}$ and $FimH_{PD}$ adopted in bound **R** states argue for a model in which the pilus tip can bend and rotate at the site of the FimH linker with an immobilized, bound $FimH_{LD}$. This physical tethering in theory increases the biophysical and functional adaptability of the pilus and thereby allows bacteria to remain attached to the bladder epithelium. (**C**) Pathogenesis outcomes depend on the preexisting equilibrium and affinity of FimH, whereby moderate affinity is ideal for successful colonization of the bladder epithelium and formation of IBCs. Catheterization allows the high-affinity variant A27V/V163A to partially circumvent the colonization resistance property observed in the intact, unperturbed bladder habitat.

Supplementary Figures

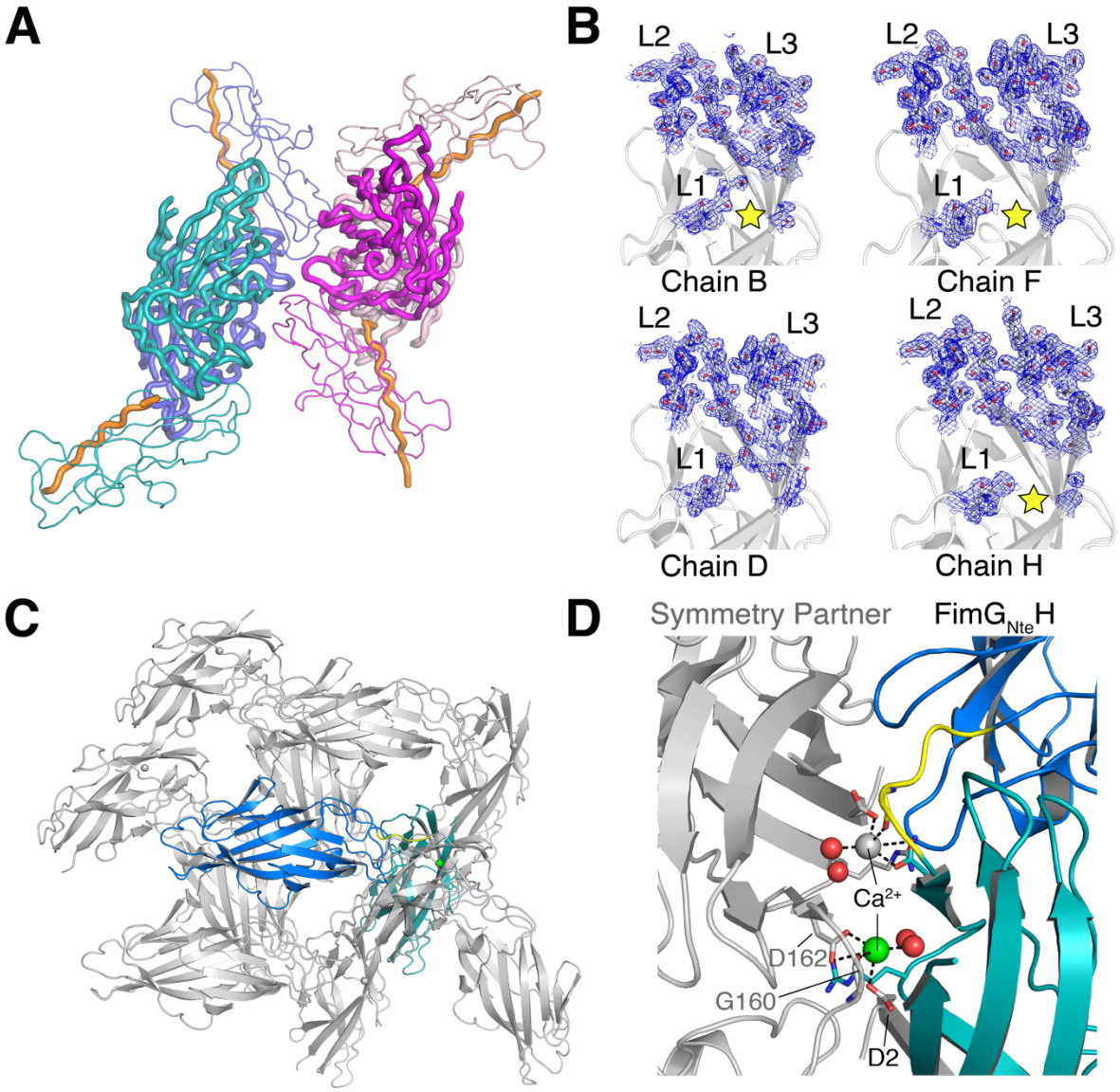


Figure S1. Structural analysis of solved FimG$_{Nte}$H complex crystal structures. (**A**) FimG$_{Nte}$H A62S (PDB ID 5JQI) crystallizes with four subunits in the unit cell. FimH$_{LD}$ is represented as tubes, FimH$_{PD}$ is represented as thin ribbons, and FimG$_{Nte}$ is represented as orange tubes. (**B**) 2Fo-Fc maps of FimG$_{Nte}$H A62S contoured at 1σ reveal well-resolved electron density for L2 and L3 but moderate to weak electron density for L1 in chains B, F, and H, likely indicative of a high

degree of conformational flexibility in L1 in the crystal. (**C**) Crystal packing of $FimG_{Nte}H$ A27V/V163A (PDB ID 5JR4) with symmetry partners colored in gray. (**D**) Close up of the packing interface reveals a calcium ion (green sphere) coordinated by D2 of $FimG_{Nte}$, waters (red spheres), and residues G159 and D162 from a neighboring FimH symmetry partner. Calcium acetate was absolutely required for formation of $FimG_{Nte}H$ A27V/V163A crystals.

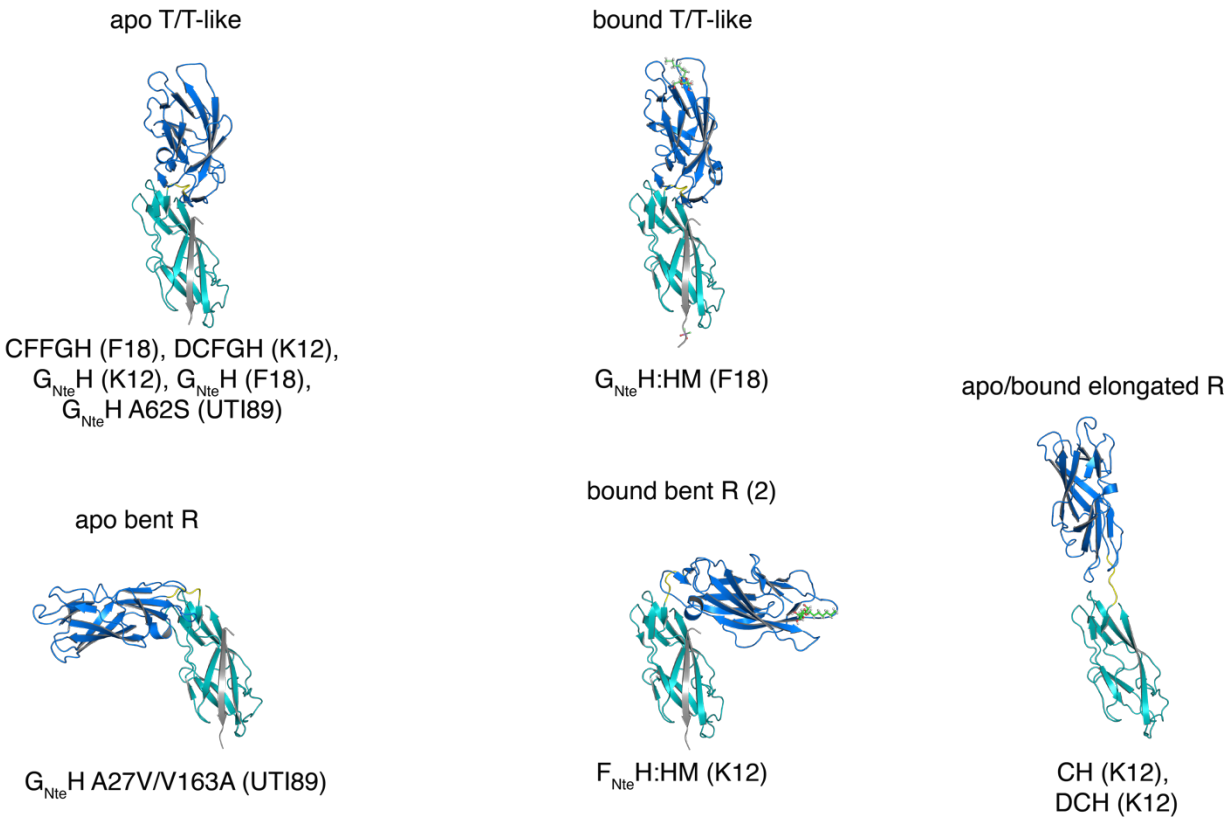


Figure S2. Structural comparison of all known FimH conformations. To date, four truly distinct conformational shapes (**T**, **bent R**, a distinct **bent R**, and **elongated R**) of FimH have been elucidated using X-ray crystallography under various conditions and with different binding partners, as indicated below each structural model, and with different FimH sequences, as indicated in parentheses. These four distinct conformational shapes reflect only two conformational states, given the structural homology and mannose binding propensity of the two unique folds of $FimH_{LD}$ identified within these crystal structures.

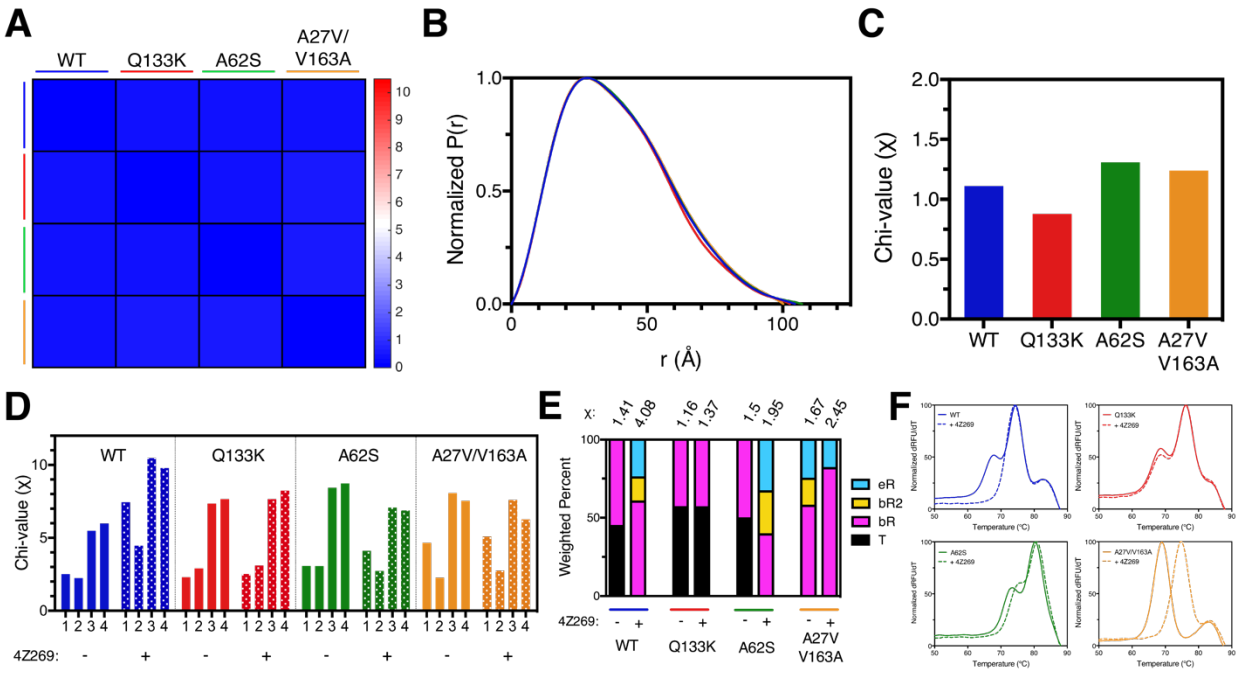


Figure S3. Solution analysis of FimCH and FimG$_{Nte}$H variants. (**A**) Structural comparison heat map of SAXS profiles indicates high degree of conformational similarity among FimCH variants as measured by χ^2, ranging from high (blue) to low (red) similarity. (**B**) Normalized pair distance distributions of FimCH, color-coded as indicated by the colored lines in (**A**). (**C**) Rigid body modeling of the FimCH crystal structure (1KLF) to the scattering profiles of FimCH variants indicates excellent agreement between crystallographic and solution conformations for all variants. (**D**) Rigid body modeling of four structural models (1: 5JQI="**T**"; 2: 5JR4= "**bent R**"; 3: derived from 4XOB = "**bent R2**"; 4: derived from 1KLF = "**elongated R**") to scattering profiles of FimG$_{Nte}$H variants in the absence or presence of 4Z269 results in varying goodness of fit. (**E**) Best-fitting multi-state models of FimG$_{Nte}$H scattering data. The fitted conformations are colored black for **T**, purple for **bent R**, yellow for a distinct, previously identified **bent R**, and cyan for **elongated R**. The relative abundances of these conformations are depicted as weighted percentages, with the corresponding goodness of fit shown above each bar. (**F**) Derivative melting curves obtained by

differential scanning fluorimetry (DSF) reveal distinct numbers of and positions in melting transitions for $FimG_{Nte}H$ variants in the absence or presence of 4Z269.

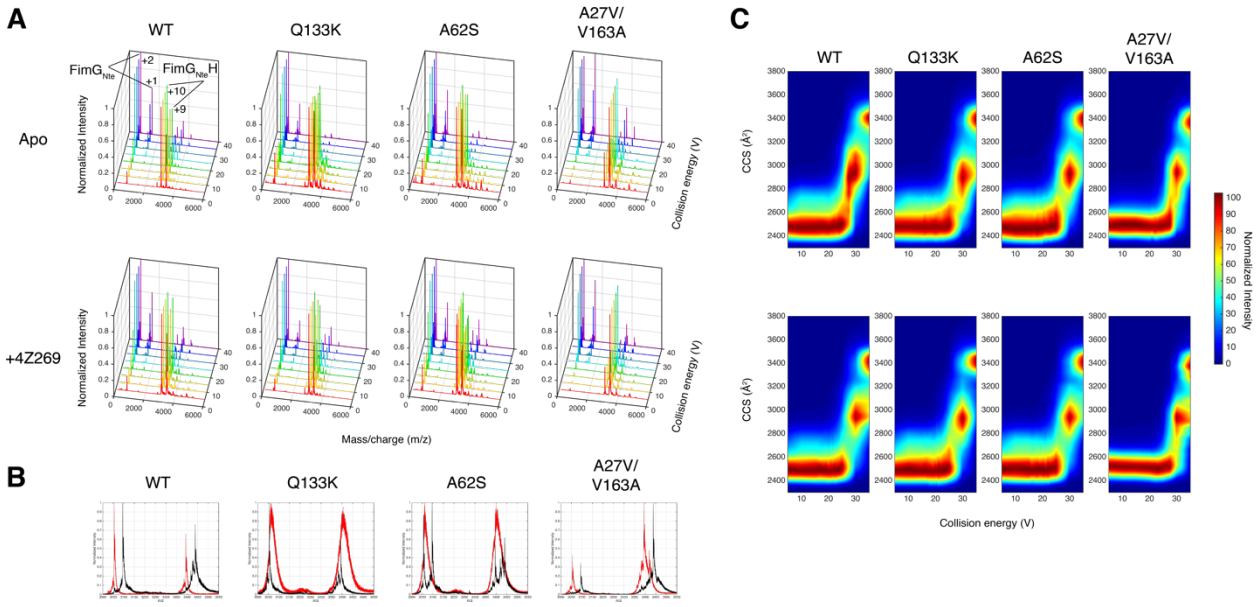


Figure S4. Native and ion mobility mass spectra of $FimG_{Nte}H$ variants in the absence or presence of 4Z269. (**A**) Native mass spectra of $FimG_{Nte}H$ variants in the absence (top) or presence (bottom) of 4Z269 at various collision energies are shown, with an increase in free $FimG_{Nte}$ and decrease in intact $FimG_{Nte}H$ complex as CE increases. (**B**) Comparison of native mass spectra of $FimG_{Nte}H$ variants in the absence (red) or presence (black) of 4Z269. (**C**) CCS distributions of $FimG_{Nte}H$ variants in the absence (top) or presence (bottom) of 4Z269 over various CE.

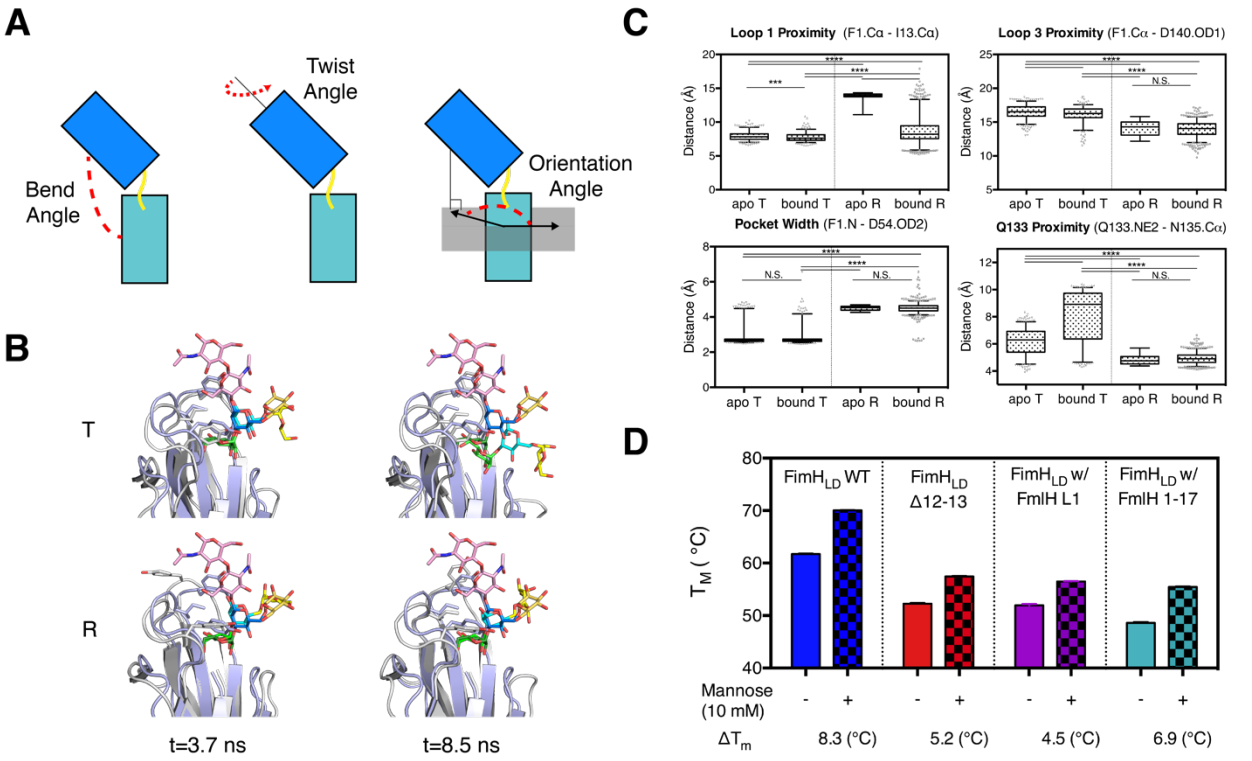


Figure S5. Conformational dynamics and binding mechanisms of $FimG_{Nte}H$ WT. (**A**) Schematic representation of the angles measured to define conformational phase space of FimH. (**B**) Structural alignment of oligomannose-3-chitobiose (Man(α1-3)-[Man(α1-6)]-Man(β1- 4)-GlcNAc(β1-4)-GlcNAc) to the "horizontal" and "tilted" mannose orientations observed in the bound **T** and **R** states ("tilted" orientation shown at t=8.5 ns for **T** state in upper right quadrant; "horizontal" orientation shown in all other quadrants). Comparison to the crystal structure of $FimH_{LD}$ bound to this high-mannose epitope suggests a mechanism by which mannose may slide and rotate and press against the "tyrosine gate" (Y48, Y137) in a high-affinity orientation. Man(α1-3)-[Man(α1-6)]-Man is depicted as sticks and colored green, yellow, and cyan, respectively, while the chitobiose component (GlcNac(β1-4)- GlcNac) is represented as pink sticks. (**C**) Comparison of the distributions of interatomic distances (indicated in graph title) among apo **T** (n=406), bound **T** (n=293), apo **R** (n=10), and bound **R** (n=687) states from simulations in which mannose was present. Movement of loop 1 toward the binding pocket is 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 can remain stably associated with **R**. Loop 3 is consistently more proximal to the pocket in the **R** state over the **T** state, contributing to high-affinity interactions. In addition, pocket width (defined as the distance between the two residues F1 and D54 which coordinate the C6 hydroxyl group of mannose) is larger by 2 Å in the **R** state and thus more capable of forming stronger hydrogen bonds with mannose compared to the **T** state. Furthermore, residue Q133 is generally further away from the binding pocket in the **T** state than in the **R** state due to structural difference in **T** and **R** pockets, but when Q133 does move closer to the center of the pocket in the **T** state, it tends to disfavor mannose binding. Thus, the Q133K mutation abolishes mannose binding through steric hindrance of mannose entry in either the **T** or **R** pockets. Statistical significance was determined by the non-parametric Mann-Whitney U test, where *** indicates $p<0.005$ and **** indicates $p<0.0001$. (**D**) Melting temperatures observed by DSF for $FimH_{LD}$ variants in the presence or absence of 10 mM mannose.

Table S1. Data collection and refinement statistics

	$FimG_{Nte}H$ A62S (5JQI)	$FimG_{Nte}H$ A27V/V163A (5JR4)
Data collection		
Space group	P1	C2
Cell dimensions□□		
a, *b*, *c* (Å)	56.4, 75.6, 78.4	122.3, 33.0, 72.5
□□*α, β, γ* (°)	68.4, 69.0, 77.4	90.0, 120.6, 90.0
Resolution (Å)	49.8-1.96 (2.03-1.96) *	62.4-2.60 (2.73-2.60) *
R_{merge} (%)[a]	14.1 (70.6)	16.4 (48.2)
R_{pim} (%)[b]	2.4 (41.8)	3.8 (21.2)
$I / \sigma I$	8.5 (1.6)	8.1 (3.3)
Completeness (%)	83.4 (90.4)	99.6 (95.1)
Multiplicity	3.9 (3.8)	6.4 (6.0)
$CC_{1/2}$	0.99 (0.85)	1.0 (0.68)
Total / Unique reflections	297,496 / 75,411(40,054 / 10,548)	51,190 / 7,964 (6,606 / 1,106)
Refinement		
R_{work}[c]/ R_{free}[d]	20.1 / 24.7	21.4 / 26.5
No. atoms		
Protein	8,522	2,138
Ligand/ion	24	7
Water	890	24
B-factors		
Protein	23.0	31.9
Ligand/ion	27.9	34.8
Water	27.3	30.9
R.m.s. deviations		
Bond lengths (Å)	0.004	0.012
Bond angles (°)	0.79	1.03
Ramachandran plot		
favored (%)	98.0	95.0
allowed (%)	1.8	4.6
outliers (%)	0.2	0.4
Clashscore	2.30	9.17

[a]$R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl)-\langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where the sum *i* is over all separate measurements of the unique reflection *hkl*.

[b]$R_{pim} = \Sigma_{hkl} [1/(n-1)]^{1/2} \Sigma_i |I_i(hkl)-\langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$

[b]$R_{work} = \Sigma_{hkl} ||F_{obs}| - |F_{calc}|| / \Sigma_{hkl} |F_{obs}|$

[c]$\boldsymbol{R}_{free}$, calculated the same as for $\mathbf{R}_{work}$ but on the 5% data randomly excluded from the refinement calculation.

[*]the outer resolution shell. Values in parentheses indicate the highest resolution shell

Table S2. Fitted parameter values and analysis of Gaussian peaks on CCS distributions

	Amp1	Mean 1	SD1	Amp2	Mean2	SD2	R^2	Peak A (% AUC)	Peak B (% AUC)	Peak C (% AUC)
WT	87.81 (2.27)	2472 (0.69)	69.0 (0.87)	32.54 (0.82)	2636 (8.1)	123 (4.50)	0.9985	60	40	---
WT + 269	93.04 (3.68)	2492 (0.88)	66.43 (1.22)	30.01 (1.42)	2643 (13.4)	117.9 (7.08)	0.9972	64	36	---
Q133K	84.44 (4.84)	2472 (1.21)	71.11 (1.72)	37.75 (2.05)	2622 (14.3)	122.1 (6.97)	0.997	57	43	---
Q133K + 269	85.84 (2.47)	2483 (0.63)	67.56 (0.97)	35.92 (0.94)	2639 (7.95)	124.2 (4.24)	0.9985	57	43	---
A62S	81.16 (4.80)	2464 (1.07)	69.91 (1.72)	39.64 (2.16)	2606 (13.1)	120 (6.05)	0.9975	57	43	---
A62S + 269	90.28 (3.27)	2494 (0.94)	70.09 (1.14)	34.82 (1.26)	2653 (10.7)	121.2 (5.66)	0.9979	60	40	---
AV	78.14 (18.3)	2480 (2.51)	61.77 (3.48)	38.86 (11.8)	2565 (32.2)	88.06 (9.93)	0.9969	59	---	41
AV + 269[a]	100.5 (0.43)	2531 (0.37)	74.25 (0.37)	---	---	---	0.9968		100	

[a] All CCS distributions were fit by either one Gaussian or a sum of two Gaussian peaks. The best fits are presented in this table. Two Gaussians explained the data better than one Gaussian, increasing R^2 by 5% and improving the residual plots in all cases except for A27V/V163A + 4Z269, in which one Gaussian was sufficient.

Chapter 3: Structure-based discovery of glycomimetic FmlH ligands as inhibitors of bacterial adhesion during urinary tract infection

Vasilios Kalas, Michael E. Hibbing, Amarendar Reddy Maddirala, Ryan Chugani, Jerome S. Pinkner, Laurel K. Mydock-McGrane, Matt S. Conover, James W. Janetka, and Scott J. Hultgren

PNAS. 2018 Mar 20;115(12):E2819-E2828. PMCID: PMC5866590

Abstract

Treatment of bacterial infections is becoming a serious clinical challenge due to the global dissemination of multidrug antibiotic resistance, necessitating the search for alternative treatments to disarm the virulence mechanisms underlying these infections. Uropathogenic Escherichia coli (UPEC) employs multiple chaperone–usher pathway pili tipped with adhesins with diverse receptor specificities to colonize various host tissues and habitats. For example, UPEC F9 pili specifically bind galactose or N-acetylgalactosamine epitopes on the kidney and inflamed bladder. Using X-ray structure-guided methods, virtual screening, and multiplex ELISA arrays, we rationally designed aryl galactosides and N-acetylgalactosaminosides that inhibit the F9 pilus adhesin FmlH. The lead compound, 29β-NAc, is a biphenyl N-acetyl-β-galactosaminoside with a K_i of ∼90 nM, representing a major advancement in potency relative to the characteristically weak nature of most carbohydrate–lectin interactions. 29β-NAc binds tightly to FmlH by engaging the residues Y46 through edge-to-face π-stacking with its A-phenyl ring, R142 in a salt-bridge interaction with its carboxylate group, and K132 through water-mediated hydrogen bonding with its N-acetyl group. Administration of 29β-NAc in a mouse urinary tract infection (UTI) model significantly reduced bladder and kidney bacterial burdens, and coadministration of 29β-NAc and mannoside 4Z269, which targets the type 1 pilus adhesin FimH, resulted in greater elimination of bacteria from the urinary tract than either compound alone. Moreover, FmlH specifically binds healthy human kidney tissue in a 29β-NAc–inhibitable manner, suggesting a key role for F9 pili in human kidney colonization. Thus, these glycoside antagonists of FmlH represent a rational antivirulence strategy for UPEC-mediated UTI treatment.

Introduction

Urinary tract infections (UTIs) are one of the most prevalent infections, afflicting 15 million women per year in the United States alone, with annual healthcare costs exceeding $2 to $3 billion [1-3]. Nearly 50% of women will experience at least one UTI in their lifetime. Despite appropriate and often successful clearance of bacteriuria by antibiotic treatment, 20% to 30% of women will experience a recurrence within 6 mo of the initial acute UTI [1, 4]. Kidney infection, or pyelonephritis, represents a severe manifestation of UTI, with ∼250,000 cases and 100,000 hospitalizations per year in the United States [5]. Acute pyelonephritis requires hospital admission and i.v. antibiotics to thwart the long-term sequelae of kidney failure and renal scarring, and, together with bacteremia, results in a mortality rate of 10% to 20% [6–8]. With the global dissemination and increase of antibiotic resistance, treatment of UTI is becoming a serious clinical challenge [9]. Antibiotic susceptibility tests indicate that many uropathogens are resistant to traditional first-line antibiotics like trimethoprim-sulfamethoxazole (TMP-SMZ) and even to last-line antibiotics such as ciprofloxacin and colistin [10-15]. The diminishing efficacy of antibiotic therapies toward UTIs and other infectious diseases demands alternative antibiotic-sparing approaches to combat bacterial pathogens. Recently, promising efforts have been made to target the virulence mechanisms that cause bacterial infection. These studies have provided much-needed therapeutic alternatives, which simultaneously reduce the burden of antibiotic resistance and minimize disruption of gastrointestinal microbial communities that are beneficial to human health [16].

Uropathogenic Escherichia coli (UPEC) is the main etiological agent of UTIs, accounting for greater than 80% of community-acquired UTIs [17, 18]. Comparative genomic studies have revealed that UPEC strains are remarkably diverse, such that only 60% of the genome is shared

among all strains [19]. As a consequence, UTI risk and outcome are determined by complex interactions between host susceptibility and diverse bacterial urovirulence potentials, which can be driven by differences in the expression and regulation of conserved functions. The ability of UPEC to colonize various habitats, such as the gut, kidney, and bladder, depends in large part on the repertoire of adhesins encoded in their genome. The most common mechanism for adhesion utilized by UPEC is mediated through the chaperone–usher pathway (CUP), which generates extracellular fibers termed pili that can confer bacterial adhesion to host and environmental surfaces, facilitate invasion into host tissues, and promote interaction with other bacteria to form biofilms [20]. Phylogenetic analysis of Escherichia genomes and plasmids predicts at least 38 distinct CUP pilus types, with single organisms capable of maintaining as many as 16 distinct CUP operons [21]. Many of these CUP pilus operons contain two-domain, tip-localized adhesins, each of which likely recognize specific ligands or receptors to mediate colonization of a host and/or environmental niche. For example, the type 1 pilus adhesin FimH binds mannosylated glycoproteins on the surface of the bladder epithelium, which is crucial for the establishment of cystitis [22, 23]. The structural basis of mannose (Man) recognition by the N-terminal–receptor binding domain, or lectin domain (LD), of FimH has been leveraged to rationally develop high-affinity aryl mannosides [24-32]. In mouse models of UTI, we have previously demonstrated that orally bioavailable mannosides that tightly bind FimH can prevent acute UTI, treat chronic UTI, and potentiate the efficacy of existing antibiotic treatments like TMP-SMZ, even against antibiotic-resistant E. coli strains [28]. Thus, use of mannosides that target the adhesin FimH represents the first successful application of an antivirulence strategy in the treatment of UTI.

A homolog of the type 1 pilus, the F9 pilus, is one of the most common CUP pili in the E. coli pan genome and an important urovirulence factor employed by UPEC for the maintenance of

UTI [21, 33]. Our recent work has demonstrated that UPEC up-regulates the expression of F9 pili in response to bladder inflammation and epithelial remodeling induced upon UPEC infection [34]. These pili display the FimH-like adhesin FmlH, which is capable of binding terminal galactose (Gal), N-acetylgalactosamine (GalNAc), or Thomsen-Friedenreich antigen (TF) [Gal(β1-3)GalNAc(α)]. FmlH was shown to bind TF within naïve or infected kidneys and to Thomsen nouvelle antigen (Tn) (GalNAc) within the inflamed bladder epithelium during chronic, unresolved UTI. Deletion of FmlH in the urosepsis isolate CFT073 resulted in a competitive defect in the ability of this strain to maintain murine UTI in C3H/HeN female mice. Furthermore, vaccination with the LD of FmlH ($FmlH_{LD}$) as the challenge antigen significantly protected mice from developing UTI. Thus, we have shown that FmlH serves a key role in the UPEC pathogenesis cascade and represents a promising target for antivirulence therapies for UTI in both the bladder and kidney habitats.

Herein, we describe the discovery and structure-based optimization of high-affinity aryl galactoside and N-acetylgalactosaminoside FmlH ligands that potently inhibit the function of FmlH. Treatment with these FmlH antagonists significantly reduced bacterial burdens in the kidneys and bladders of infected mice, thereby demonstrating promising translational value in the treatment of UTI in humans. The results of these studies, together with our previous work on FimH mannosides, further support the mechanistic and therapeutic value of antivirulence strategies that leverage structure-function relationships of diverse bacterial adhesins for the rational design of high-affinity glycosides for the treatment of UTI and other bacterial infections.

Results

O-nitrophenyl β-Galactoside Identified as Early Lead Inhibitor of the F9 Pilus Adhesin FmlH

We revealed in a previous communication that FmlH binds surface glycan receptors containing terminal Gal, GalNAc, or TF residues [34]. Given the role of FmlH in UTI pathogenesis, we aimed to develop high-affinity galactoside antagonists of FmlH through an X-ray structure-guided medicinal chemistry approach. This strategy entailed (i) screening a select library of galactosides through multiplex ELISA arrays for initial lead compound identification; (ii) an iterative process of cocrystal structure determination, virtual screening, structure-based ligand design, and in vitro biochemical characterization; and (iii) evaluation of the top lead compound in a mouse model of UTI (Fig. 1A). Toward these goals, we first investigated whether Gal, GalNAc, and TF could be adapted to function as soluble, competitive inhibitors of FmlH. To that end, an ELISA-based competition assay was developed to detect binding of $FmlH_{LD}$ to surface-immobilized desialylated bovine submaxillary mucin (ds-BSM) in the presence or absence of soluble compounds (Fig. 1A). As expected, Gal, GalNAc, and TF were each capable of inhibiting $FmlH_{LD}$ at a concentration of 1 mM, with GalNAc exerting greater inhibitory potency than TF or Gal. However, neither Man nor glucose (Glc) had any detectable effect on the ability of $FmlH_{LD}$ to bind ds-BSM (Fig. 1B). Lactose (Lac), or Gal(β1-4)Glc, was also incapable of inhibiting $FmlH_{LD}$, demonstrating the high selectivity in which $FmlH_{LD}$ engages Gal-containing glycans (Fig. 1B). O-nitrophenyl β-galactoside (ONPG) and isopropyl β-thiogalactoside (IPTG) were also tested for inhibition in this exploratory phase of our search for FmlH inhibitors. While IPTG exerted minor inhibitory activity at 100 μM, ONPG was found to block $FmlH_{LD}$ from interacting with ds-BSM more effectively than Gal, GalNAc, or TF (Fig. 1B). The strong inhibitory potency of ONPG suggested that β-galactosides could potentially be rationally designed with higher affinity by specifically targeting residues within and surrounding the sugar binding pocket of FmlH.

Therefore, X-ray crystallography was implemented to elucidate the 3D structures of both apo and ligand-bound $FmlH_{LD}$ (SI Appendix, Table S1). First, a crystal structure of apo $FmlH_{LD}$ was solved at 1.6 Å resolution by molecular replacement using $FimH_{LD}$ [Protein Data Bank (PDB) ID 3MCY] as the search model. Within this structure, two copies of $FmlH_{LD}$ are found in the asymmetric unit, each of which adopts a canonical β-sandwich fold, with three distinct binding loops (loop 1: residues 10 to 15; loop 2: residues 44 to 53; and loop 3: residues 132 to 142) that form a wide, shallow, solvent-exposed binding pocket (Fig. 1 C and D). Within the binding pocket of both copies resides a sulfate ion, which interacts with residues implicated in Gal binding (Fig. 1D). Cocrystal structures of $FmlH_{LD}$ bound to TF and of $FmlH_{LD}$ bound to ONPG were also solved to 2.1 Å and 1.8 Å, respectively. Structural overlay of the apo and ligated crystal structures yields root-mean-square deviation (RMSD) values that fall within 0.6 Å, suggesting that $FmlH_{LD}$ generally adopts the same active or functional conformational state in the absence or presence of ligand (Fig. 1C). This functional conformational state most likely corresponds to a high-affinity conformation of FmlH, as the $FmlH_{LD}$ structures exhibit a higher degree of structural homology to the high-affinity conformation of FimH (RMSD values of 0.8 to 0.9 Å) than to the low-affinity conformation of FimH (RMSD values of 1.7 to 1.9 Å) [34–38].

The cocrystal structure of $FmlH_{LD}$-TF reveals two copies of $FmlH_{LD}$-TF in the unit cell, in which each TF adopts a distinct ligand conformation (Fig. 1D). In both copies, the terminal Gal in TF occupies the cleft of the binding pocket through direct polar interactions with residues F1, D53, K132, and N140. In contrast, the orientation of the GalNAc in TF differs significantly between the two copies of FmlH. In chain A, the GalNAc sugar points toward loop 3, with the carbonyl group of GalNAc forming a hydrogen bond (H-bond) with the guanidinium group of R142. In chain B, however, the GalNAc packs against and forms a H-bond with the hydroxyl group of Y46.

Accordingly, the differences in the orientation of bound ligand across the two copies are accompanied by slight differences in orientation of the side chains of the interacting residues Y46 and R142. The multiple binding modes observed for a single ligand suggests that the wide, shallow nature of the Gal binding pocket in FmlH would enable galactosides to possibly bind FmlH with diverse interactions and conformations.

The $FmlH_{LD}$-ONPG cocrystal structure also shows two copies of $FmlH_{LD}$ in the unit cell, in which a sulfate ion occupies the binding pocket of chain A while ONPG occupies the binding pocket of chain B (Fig. 1D). As expected, the Gal component of ONPG resides in the cleft of the binding pocket, while the solvent-exposed nitrophenyl group mediates a polar or salt-bridge interaction with R142 through an intricate network of H-bonds with water molecules. Furthermore, the positioning of the Gal component of ONPG aligns with that of the Gal residue of TF (Fig. 1C). Moreover, the conformation of the FmlH binding pocket observed in this $FmlH_{LD}$-ONPG cocrystal structure resembles the binding pocket conformation in the $FmlH_{LD}$-TF cocrystal structure, reflecting a high-affinity binding orientation that can be targeted for drug discovery. These results and observations strongly suggested that the $FmlH_{LD}$-ONPG cocrystal structure represents an appropriate structural candidate for use in virtual screening to aid in the design of galactoside compounds specific for FmlH.

Virtual Screen Identifies and Informs the Design of FmlH-Targeting Galactosides

An exhaustive virtual screen was performed using AutoDock Vina to computationally dock ~1,800 known galactosides in the binding pocket of $FmlH_{LD}$ (from an $FmlH_{LD}$-ONPG cocrystal structure; PDB ID 6AOY), generating a ranked list of top binding poses and associated docking scores for each galactoside (SI Appendix, Fig. S1A). Top hits from the virtual screen were filtered according to group efficiency values and then visually inspected to aid and inform structure-guided

drug design. In all cases, the Gal component of top-scoring galactosides bound to the cleft of the binding pocket, as expected. In addition, most of the high-scoring hits also interacted with specific hot-spot residues near the Gal binding pocket, which we sought to leverage for compound optimization. These hot-spot residues included (i) residue Y46, which caps the top of the binding pocket and can contribute hydrophobic interactions; (ii) residue K132, which lies at the bottom of the sugar binding pocket and can engage polar groups linked to the Gal sugar; and (iii) residue R142, which points toward an empty, solvent-exposed cleft near the binding pocket and can contribute electrostatic interactions (SI Appendix, Fig. S1B). These visual insights were then considered in our rational design strategy for FmlH-targeting galactoside antagonists.

Design and Synthesis of FmlH-Targeting Galactoside Antagonists

To increase FmlH binding affinity and explore structure-activity relationships (SARs), we constructed a large library of galactoside analogs (Fig. 2). Based on the docking results, we predicted that β-Gal isomers would be preferred over α-Gal and that ortho positioning of functional groups on a phenyl scaffold would best facilitate interactions with specific sites within the binding pocket, namely hot-spot residues Y46 and R142. Accordingly, we synthesized and evaluated small sets of phenyl galactosides with ortho-substituted functional groups (2 to 6; Fig. 2A). We also either purchased or synthesized several other phenyl galactosides, which contained meta or para substituents on the aglycone ring (7 to 11; Fig. 2A), and other aryl and heterocyclic galactosides (12 to 22; Fig. 2 B and C). This allowed us to derive meaningful SARs for informing further design and optimization of improved galactosides. In addition, we tested natural-product galactosides isolated from cranberries and other natural sources (23 to 27; Fig. 2D). The promising activity of the simple galactoside ONPG (4β) in the initial screen, coupled with the hot-spot residues identified in virtual screening, prompted us to expand our FmlH-ligand design strategy with a

compound series containing biphenyl aglycones (28 to 32; Fig. 2E), such as 29β-NAc, the N-acetyl-β-galactosaminoside with an m-carboxylic acid on the B-ring designed to directly interact with the hot-spot residue R142 (SI Appendix, Fig. S1 B and C). To confirm the predicted preference for the β-Gal isomers, we also synthesized and tested many corresponding α-Gal isomers. Compounds were synthesized by using one of two general synthetic glycosylation methods involving either a reaction between Gal pentaacetate and phenols promoted by boron trifluoride or a Koenigs–Knorr-type reaction of galactosyl halide with aryl alcohols (SI Appendix, Fig. S2).

Biochemical Characterization of FmlH Antagonists

Selected top-hit glycosides and a few low-scoring analogs from the virtual screen, as well as synthetic galactosides, were tested in the ELISA-based competition assay for their ability to inhibit binding of $FmlH_{LD}$ to ds-BSM. Direct comparison of inhibitory potency among galactosides led to delineation of basic SARs (Fig. 3 A–C and SI Appendix, Table S2). When tested at 100 μM, the phenyl β-galactoside 1β (beta isomer of 1; Fig. 2A) exhibited significantly higher binding inhibition (77%) than Gal (8.1%), indicating that the phenyl group significantly enhances binding to $FmlH_{LD}$ (Fig. 3A). Various ortho substituents on the phenyl ring additionally conferred substantial improvements in inhibitory potency, as observed with 2β (87%), 3β (95%), 4β (ONPG; 93%), 5β (97%), and 6β (90%). In contrast, the meta-methoxy groups in compound 7β (76%) did not enhance binding strength compared with 1β. Further, para-substituted functional groups displayed variable inhibitory potencies relative to 1β, with enhancements observed in 8β (86%) and 9β (86%), with no significant effect observed in 11β (78%) or 11β-thio (72%), and with a reduction observed in 10β (65%). Thus, we deduced that the ortho-substituted phenyl β-galactosides generally outperformed other simple phenyl galactosides.

Complex heterocyclic galactosides, such as coumarins 12β (85%) and 14β (89%), which differ only by a methyl group, displayed significant inhibitory potencies against $FmlH_{LD}$, while the related galactoside 13β (50%) displayed reduced inhibitory activity, likely because of its fluoro substituents (Fig. 3A and SI Appendix, Table S2). Resorufin galactoside 15β (80%) also showed similar potency compared with the phenyl β-galactoside 1β. These combined results suggest that the substituents of 12β are responsible for augmenting affinity relative to 1β. In contrast, indoles 16β (22%) and 17β (41%) performed poorly as inhibitors of $FmlH_{LD}$. Naphthyl galactosides 18β (46%) and 19β (79%), in addition to isoquinoline 21β (15%), showed no improvement in activity relative to 1β. However, quinoline 20β (95%) displayed significantly higher inhibition than 1β and 18β. This advocates that the electron pair-donating nitrogen atom in 20β is making a specific interaction with FmlH. This observation is consistent with the pattern of SARs, indicating that the ortho position is key to enhancing inhibitory potency against $FmlH_{LD}$.

We also evaluated naturally occurring galactosides derived from cranberries and other natural sources in this screen (Fig. 3A and SI Appendix, Table S2). These compounds included anthocyanidin (pelargonidin, 23β; cyanidin, 24β; peonidin, 25β) and flavonol (quercetin, 26β; myricetin, 27β) β-galactosides. Generally, these compounds exhibited moderate to weak inhibition of $FmlH_{LD}$ binding, with little enhancement in inhibition relative to Gal (8.1%). The only significant binders were 24β (29%) and 26β (14%). Comparison of the anthocyanidin family indicates that the 3′ or meta-substituted hydroxyl group on the B-ring of 24β is critical for its specific interaction with FmlH. Absence of this meta substituent in 23β (0.7%) or methylation of the hydroxyl group in 25β (3.6%) abrogates potency, suggesting that the hydroxyl group of 24β might participate in a H-bond to a specific residue in the $FmlH_{LD}$ binding pocket. Additional inhibitory screens performed with cranberry-derived compounds and fractions at 1 mM confirmed

the specificity and necessity of the Gal sugar for inhibiting the binding pocket of FmlH (SI Appendix, Fig. S3 A and B).

Interestingly, the tested GalNAc-derived compounds possessed significantly higher inhibitory potency compared with their matched-pair Gal-derived counterparts, as exemplified with 4β-NAc (87%) relative to 4β (31%) when tested for inhibition at 10 μM (Fig. 3B and SI Appendix, Table S2). These results taught us that the N-acetyl group, together with other functional groups, contributes to binding by targeting distinct components of the binding pocket of FmlH. In contrast, the galactosides with α-linkages (28α-30α) or disaccharides with aglycone moieties (33 to 35) were generally poor inhibitors of FmlH, except for 11α-NAc (82%) (Fig. 3A and SI Appendix, Table S2).

Consistent with the above-mentioned SARs, the ortho biphenyl galactoside 28β (91%) was more potent than the meta 31β (57%) or para 32β (30%) analogs (Fig. 3A and SI Appendix, Table S2). Next, we installed a carboxylate group at the meta position on the biphenyl B-ring (29β), intended to target the pocket formed by N140 and R142, and found that 29β exhibited greater inhibition (99%) compared with 28β when tested at 100 μM. This pronounced difference in activity was further highlighted when these compounds were tested for inhibition at 10 μM and 1 μM (Fig. 3 B and C and SI Appendix, Table S2). Importantly, 30β (87%), the methyl ester of 29β, tested at 100 μM resulted in a reduction in binding, suggesting that the negative charge of the carboxylic acid likely mediates a critical electrostatic interaction with R142 of $FmlH_{LD}$. Lastly, we synthesized the GalNAc version of 29β to increase its binding affinity and found that 29β-NAc (93%) had significant improvement in activity over 29β (75%) when tested at 10 μM. Final evaluation of the highest performing galactosides in the ELISA-based competition assay at concentrations of 10 μM and 1 μM allowed for a clearer ranking of compounds, where 29β-NAc

clearly stood out as the most potent (Fig. 3 B and C and SI Appendix, Table S2).

Determination of FmlH–Galactoside Binding Affinities

Bio-layer interferometry (BLI) was pursued to quantitate the binding affinity of the most promising FmlH antagonists. First, biotinylated serine-linked TF (Ser-TF) immobilized on streptavidin pins was incubated with varied titrations of $FmlH_{LD}$ in solution, and steady-state analysis of binding responses revealed a dissociation constant (K_d) of 15.0 ± 0.8 μM (Fig. 3D). Next, immobilized Ser-TF was incubated in solutions comprising a fixed concentration of $FmlH_{LD}$ but varying concentrations of galactosides to determine their inhibitory or dissociation constant (K_i or K_d) values (Fig. 3E). The BLI-based affinity determinations correlated well with the relative binding strengths measured in the ELISA-based competition assay (Fig. 3 A–C and SI Appendix, Table S2). The two lead compounds, 29β-NAc and 29β, bound tightly to $FmlH_{LD}$, with respective K_i values of ∼90 nM and 2.1 μM, which represent a ∼7,800-fold and ∼330-fold enhancement in binding affinity relative to Gal. Another promising compound, 4β-NAc, bound $FmlH_{LD}$ with a K_i value of 2.3 μM. In summary, a combinatorial approach based on virtual screening and structure-guided ligand design led to the discovery of small–molecular weight monomeric glycosides derived from Gal and GalNAc that function as effective antagonists of FmlH. Optimization of early hits to high-affinity o-biphenyl Gal and GalNAc antagonists was realized via ortho substitution on phenyl aglycones to facilitate interactions that significantly enhanced binding to FmlH.

Structural Basis of Galactoside Inhibition of FmlH

To elucidate the molecular basis for galactoside inhibition of FmlH, cocrystal structures of $FmlH_{LD}$ bound to 4β, 5β, 20β, and 29β-NAc were determined (Fig. 4 A and B). These galactosides share a common aglycone motif consisting of a phenyl ring with an ortho-substituted functional

group. As predicted from computational studies, the sugar portion of all these galactosides resides within the cleft of the binding pocket. The phenyl groups directly attached to the sugar portion of all four compounds lie along the same 3D plane. In this nearly identical conformation, the phenyl ring is oriented perpendicularly to the side chain of residue Y46, revealing edge-to-face π-stacking, which likely contributes to the affinity enhancement observed for all β-galactosides. For 4β, 5β, and 20β, the ortho substituents point toward R142 but are too distant (>7 Å) for direct interaction and, instead, form H-bonds with water molecules that, in turn, interact with residues K132 and R142 (Fig. 4A). Thus, we deduced that the marked affinity enhancement observed for 4β, 5β, and 20β is due to a combination of (i) indirect interactions between the ortho substituent and residues K132 and R142 formed by an intricate network of water-mediated H-bonds and (ii) edge-to-face π-stacking between the phenyl ring and residue Y46.

In contrast to simple phenyl galactosides, the biphenyl scaffold of 29β-NAc presents the carboxylic acid to engage in a direct charge–charge interaction with the guanidinium side chain of R142 (Fig. 4B). The lower potency of the methyl ester derivative 30β is further evidence that the charge–charge interaction likely drives the observed affinity enhancement (Fig. 4C). The improved affinity of 29β-NAc relative to 29β is also due to additional interactions mediated by the N-acetyl group in H-bonding to a water molecule captured by the biphenyl aglycone and the side chain of residue K132 (Fig. 4 B and C). Altogether, analysis of all X-ray crystal structures of ligand-bound FmlH offers two general mechanisms for the significant enhancement in binding affinity of galactosides relative to Gal: edge-to-face π-stacking with Y46 and polar or electrostatic charge–charge interactions with K132 and R142.

FmlH Antagonist Effectively Treats Murine UTI in Vivo and Prevents Binding to Human Kidney Tissue

We previously reported that FmlH binds to naïve kidney and inflamed bladder tissue and plays a critical role in chronic UTI, as abrogation of its function through genetic deletion or vaccination results in significant attenuation in the ability of UPEC to cause chronic UTI [34]. Thus, we hypothesized that galactosides that inhibit the function of FmlH would have efficacy in the treatment and/or prevention of UTI. To assess therapeutic potential, the lead compound 29β-NAc was evaluated for its ability to reduce bacterial burdens in the urinary tracts of C3H/HeN mice during chronic UTI. We previously defined chronic cystitis in C3H/HeN mice as urine titers of $>10^4$ CFU/mL lasting at least 2 to 4 wk, as well as bladder inflammation and edema at euthanasia [39]. Further, C3H/HeN mice are genetically predisposed to vesicoureteral reflux (retrograde flow of urine from the bladder to the kidneys), which can lead to bacterial colonization of the kidneys, renal abscess formation, scarring, and atrophy [40]. Accordingly, we observed high levels of kidney colonization by CFT073 in control (vehicle-treated) animals. When delivered intravesically, 29β-NAc significantly reduced bacterial burdens in both the bladder and the kidneys of these mice (Fig. 5 A and B). For comparison, mannoside 4Z269, which inhibits the type 1 pilus adhesin FimH, also significantly reduces titers of CFT073 from the bladders and kidneys of infected mice relative to vehicle control (Fig. 5 A and B). When administered together, 29β-NAc and 4Z269 eradicated bacteria from the kidney in nearly all mice while also reducing bacterial titers in the bladder, suggesting that FimH mannosides and FmlH galactosides may function synergistically to target distinct bacterial adhesins or communities within the kidney habitat (Fig. 5 A and B).

To show relevance to human UTI, we assessed FmlH and FmlH-targeting galactosides through immunofluorescence analysis of $FmlH_{LD}$ binding to human kidney and bladder biopsied tissue determined to be nonmalignant. While $FmlH_{LD}$ does not appear to bind to healthy human

bladder tissue, $FmlH_{LD}$ does bind to healthy human kidney tissue, particularly in regions resembling the collecting ducts and distal tubules of the kidney (Fig. 5C and SI Appendix, Figs. S4 and S5). As a negative control, the binding null mutant $FmlH_{LD}$ K132Q, which lacks the ability to bind ds-BSM in vitro (SI Appendix, Fig. S6), was incapable of binding kidney tissue, suggesting that $FmlH_{LD}$ specifically recognizes receptors naturally present in human kidney tissue (Fig. 5C). These observations are consistent with the previously reported binding phenotypes in mice, in which FmlH can bind naïve mouse kidney tissue, but not naïve mouse bladder tissue, and can bind to receptors in inflamed bladder tissue [34]. Moreover, incubation of 29β-NAc with $FmlH_{LD}$ prevented binding to human kidney tissue, suggesting that these results may translate to humans. Importantly, these collective data provide substantial evidence that aryl glycoside–based FmlH antagonists derived from β-Gal or β-GalNAc can serve as an effective therapy for persistent UTIs, including pyelonephritis, for which there is an enormous unmet medical need.

Discussion

UPEC is the causative agent of most UTIs, a common and very costly disease in women, children, and the elderly that is becoming increasingly resistant to antibiotic treatment. By leveraging our expertise in UPEC pathogenesis and structure-based drug design, we developed small-molecule Gal-based FmlH antagonists that show in vivo efficacy in the treatment of chronic UTI in mouse models. Virtual screening combined with rational design led to the identification of several naturally occurring cranberry and synthetic galactosides, the most potent of which binds FmlH with nanomolar affinity. X-ray crystallography revealed that potent galactosides achieve significant enhancements in binding affinity through interactions on opposite sides of the wide Gal binding pocket of FmlH. Appropriately substituted aryl groups, like those found in 4β/4β-NAc,

5β, and 20β, are seen to mediate edge-to-face π-stacking interactions with Y46 of $FmlH_{LD}$. Further, the optimized biphenyl aglycone of compound 29β-NAc contains an ideally positioned carboxyl group to mediate electrostatic interactions with R142 in addition to π-stacking interactions with Y46. Evaluation of the lead candidate 29β-NAc in a mouse model of chronic UTI demonstrated significant reductions of bacterial burdens in the mouse kidney and bladder. Combination dosing with mannoside and galactoside resulted in near complete clearance of bacteria from the kidney and significant elimination of bacteria from the bladder. Furthermore, FmlH was shown to bind specifically to human kidney tissue, which could be inhibited by 29β-NAc. Additionally, FmlH has been shown to be up-regulated in urine samples directly isolated from human patients with UTI compared with expression during in vitro growth in media or normal urine [41], suggesting an important role for FmlH in human UTI. Thus, FmlH-targeting galactosides represent a rational antivirulence modality for the effective treatment of UPEC-mediated UTI.

Our rational strategy to discover receptor-mimicking galactosides targeting FmlH was similar to the strategy we followed for the development of FimH mannosides. However, the design of the galactoside and N-acetylgalactosaminoside antagonists of FmlH was met with distinct challenges. The most striking difference between FmlH and FimH is the binding affinity for their respective ligands: FimH binds soluble Man with a moderate binding affinity of ∼5 to 10 μM, and FmlH binds soluble Gal with a weak binding affinity of ∼700 μM [34, 42, 43]. The weak binding affinity of FmlH, which is quite common for most carbohydrate–lectin interactions, rendered the development of high-affinity galactosides much more challenging. This disparity in affinity is a direct consequence of the substantial variance in the shape of the binding pocket. FimH binds Man with high affinity because of the deep, narrow pocket formed by loops 1, 2, and 3, in which loops 2 and 3 mediate specific polar interactions directly to Man and a water molecule and loop 1 serves

as an affinity clamp to stymie dissociation of Man [35]. In contrast, loop 1 in FmlH is more distant from loops 2 and 3 than it is in FimH and does not contribute to binding, which results in a widened, solvent-exposed pocket for weak Gal binding (Fig. 1D). In addition, the differences in binding pocket architecture dictate the sterically allowed linkage types. FimH has space at the tip of the LD between its parallel tyrosine gate (residues Y48 and Y137) to accept α-linked moieties, of which biaryl groups confer drastic enhancements in affinity through strong parallel face-to-face π-stacking interactions. In contrast, FmlH is capped at the very tip of the pocket with Y46, which biases specificity toward β-linked moieties, of which biaryl groups confer moderate enhancements in affinity through significant edge-to-face π-stacking interactions. Having accounted for these variations, our structure-guided medicinal chemistry approach, coupled with our in vivo work, has clearly demonstrated the future translational impact of galactosides as treatments for UTI.

It is noteworthy that our collective search for high-affinity antagonists of FmlH and FimH has led to discovery of biphenyl moieties as the preferred aglycone groups for high-affinity galactosides and mannosides, respectively. Pocket geometry dictates the type of biphenyl scaffold that is optimal. Thus, the best FimH-targeting mannosides contain para biphenyls in the alpha stereochemistry, while the best FmlH-targeting galactosides contain ortho biphenyls in the beta orientation. However, in both cases, H-bonding donors or acceptors on the B-ring result in significant enhancement in binding affinity through specific interactions outside the sugar binding pocket. Intriguingly, the inhibitory potency conferred by the meta carboxyl on the B-ring of 29β-NAc is also appreciated in the significant inhibitory role of the meta-substituted group on the B-ring in cranberry compounds 24β and 26β, which suggests a common pharmacophore between our optimal synthetic compound and natural-product compounds in targeting FmlH. This study provides evidence that specific glycosidic compounds in cranberry can specifically bind and

inhibit a bacterial adhesin. Furthermore, our work exposes a trend indicating that π-stacking of aromatic aglycones with binding pocket residues in the adhesin is essential in mimicking glycoprotein receptors and for developing tight-binding ligands in each lectin. Mimicking carbohydrates with small molecules is a long-sought-after goal in medicinal chemistry and chemical biology [44-46], and we believe that these results add significantly to this understanding and goal. This information can now be utilized not only in the future optimization of lead compound 29β-NAc as a treatment for UTIs, but also in the rational design of numerous other lectin antagonists for the development of small-molecule glycoside-based drugs aimed at treating infections mediated by E. coli or other microbes [44].

The rapid increase and spread of antibiotic resistance, including multidrug-resistant forms of bacteria, has rendered many antibacterial therapies ineffective and threatens to undermine the biomedical strides made to promote human health and longevity [9]. Selection pressures imposed by antibiotics on bacterial pathogens have promoted their proliferation, especially through overuse of antibiotics within the farming industry and inappropriate use or misuse among patients [47-50]. Recent reports indicate that patients are now succumbing to bacterial strains which possess broad-spectrum resistance to all last-resort antibiotics, which many fear signals that antibiotic resistance will pave the way for the "next pandemic" [15]. Antivirulence strategies that aim to reduce the pathogenicity of bacterial pathogens promise to provide the same therapeutic efficacy as antibiotics without introducing selective pressures that would promote widespread dissemination of resistance [16]. Multiple antivirulence efforts will be required to combat the multiple mechanisms by which diverse bacterial pathogens colonize the host, which can include, for example, the targeting of CUP pilus adhesins or the biogenesis machinery responsible for the assembly of CUP pili [51]. As highlighted in this work, UPEC employs an armament of diverse

CUP pili to colonize and persist within changing local environments encountered during UTI pathogenesis, which suggests that targeting more than one CUP adhesin may indeed be a more effective strategy for combating UTIs. Herein, we have highlighted the overwhelming value of applying a deep mechanistic understanding of structure-function-virulence relationships of bacterial adhesins to the rational design of high-affinity carbohydrate glycomimetics for the treatment of UTI. This demonstration serves as a general model for the rational approach necessary to target virulence factors and disrupt their role in bacterial infections.

Materials and Methods

Ethics Statement

All animal experiments were conducted according to the National Institutes of Health (NIH) guidelines for housing and care of laboratory animals and performed in accordance with institutional regulations after pertinent review and approval by the Institutional Animal Care and Use Committee at Washington University School of Medicine (protocol 20150226). Deidentified human tissue was obtained from the Tissue Procurement Core at Washington University School of Medicine.

Protein Expression and Purification

FmlH residues 1 to 160 from UPEC strain UTI89 with a C-terminal six-histidine tag (i.e., $FmlH_{LD}$) were cloned into the IPTG-inducible plasmid pTrc99A. This construct was then transformed into and expressed in E. coli strain C600. Periplasms were isolated as previously described and dialyzed four times against PBS plus 250 mM NaCl [34]. $FmlH_{LD}$ was purified from this periplasmic fraction by cobalt affinity chromatography through elution with 150 mM imidazole. $FmlH_{LD}$ was buffer exchanged into 10 mM Hepes [4-(2-hydroxyethyl)-1-

piperazineethanesulfonic acid] (pH 7.5) and 50 mM NaCl, concentrated to 6 mg/mL, and stored stably at 4 °C for use in biochemical and biophysical assays.

In Silico Docking and Virtual Screening

Structure-based virtual screening through in silico docking was performed with AutoDock Vina [52]. Existing Gal-based derivatives were identified through the ZINC12 database [53]. Their 3D structures were extracted from the downloaded mol2 file as pdb coordinates and converted to pdbqt format using Open Babel [54]. The crystal structure of apo FmlH was converted to its topology file using AutoDock Tools. The grid box was centered at the Gal binding pocket of FmlH, and its dimensions ($26 \times 26 \times 26$ Å^3) were chosen to accommodate bulky compounds and multiple potential binding modes at or near the binding pocket. The exhaustiveness of the search was set to a value of 15. The top binding modes and scores within this grid space were generated by AutoDock Vina. Custom in-house scripts in Bash and MATLAB were used to link these binding scores with compound properties such as molecular weight. Top binding modes were visualized in PyMOL.

Virtual screening of this library, which comprised galactosides ranging from 150 to 900 Da in molecular mass, yielded a mean docking score of 6.3 kcal/mol (1 kcal = 4.18 kJ), with a standard deviation of 0.73 kcal/mol and a range of 4 to 9 kcal/mol (SI Appendix, Fig. S1A). To prioritize hits, we abstained from directly comparing raw binding scores, as large, lipophilic molecules tend to have artificially high predicted binding interactions due to their contribution to hydrophobic interactions as calculated by the empirical scoring function of AutoDock Vina [52]. Instead, the results of the virtual screen were evaluated per group efficiency (GE), which, in this context, measures the contribution of the aglycone group within each galactoside (indicated as X in the following equation) to the docking score (DS) with respect to the number of heavy atoms

(HA) present in the aglycone group [GE = ($DS_X - DS_{Gal}$)/($HA_X - HA_{Gal}$)]. Top hits were defined as galactosides with a GE value greater than 1.25 times the SD (σ = 0.0016 kcal/mol per HA) above the library mean (μ = 0.0011 kcal/mol per HA), which constituted the top ~10% of highest scoring galactosides (SI Appendix, Fig. S1A).

Synthesis of Galactosides and N-acetyl Aminogalactosides

Galactosides and N-acetyl aminogalactosides were synthesized by standard glycosylation chemistry, including boron trifluoride-mediated glycosidation and the Koenigs–Knorr reaction, respectively (SI Appendix, Fig. S2). In method A, boron trifloride-promoted glycosylation of phenols with Gal pentaacetate yielded corresponding acetylated aryl galactosides, which were treated with sodium methoxide in methanol to provide the corresponding aryl galactosides (1β to 3β, 5β to 9β, 18β to 19β, and 28β to 32β; 2α to 3α, 18α to 19α, and 28α to 32α). In method B, final GalNAc and Gal analogs (20β-NAc, 21β, 28β-NAc, and 29β-NAc) were synthesized from galactosyl halide and aryl alcohols via a Koenigs–Knorr-type reaction, which yielded aryl galactosides that were then deacetylated by treatment with methylamine in ethanol.

ELISA

Immulon 4HBX 96-well plates were coated overnight with 1 μg of bovine submaxillary mucin (Sigma). Coated wells were then treated with 100 μL of Arthrobacter ureafaciens sialidase (10 mU/mL) diluted in PBS for 1 h at 37 °C. Thereafter, wells were incubated with 200 μL of blocking buffer (PBS plus 1% BSA) for 2 h at 23 °C, followed by incubation with 100 μL of biotinylated $FmlH_{LD}$ diluted in blocking buffer to 20 μg/mL in the presence or absence of galactoside compounds for 1 h at 23 °C. After washing three times with PBS plus 0.05% TWEEN-20, 100 μL of streptavidin-HRP conjugate (BD Biosciences; 1:2,000 dilution in blocking buffer) was added to each well for 1 h at 23 °C. After a final round of washing, plates were developed

with 100 μL of tetramethylbenzidine (BD Biosciences) substrate and quenched within 1 to 2 min with 50 μL of 1 M H_2SO_4, and absorbance was measured at 450 nm. This assay was used to determine percent inhibition values and inhibitory constant (IC_{50}) values where indicated.

BLI

Streptavidin pins were first dipped in a baseline in PBS (pH 7.4) for 120 s, followed by loading of 5 to 10 μg/mL biotinylated Ser-TF (Toronto Research Chemicals) in PBS for 300 s, quenching by 10 μg/mL biocytin in PBS for 240 s, and another baseline step in PBS for 120 s. Thereafter, pins were dipped in PBS for 120 s and transferred to protein samples (varying concentration of $FmlH_{LD}$ or fixed concentration of $FmlH_{LD}$ with varying concentration of galactoside compounds) for association for 300 to 600 s. Equilibrium binding response values were used to determine the affinity of interaction between $FmlH_{LD}$ and immobilized Ser-TF under a 1:1 binding model or between $FmlH_{LD}$ and galactosides in solution under a competitive one-site binding model.

Protein Crystallization and Structure Determination

Crystals of apo $FmlH_{LD}$ in 10 mM Hepes (pH 7.5) and 50 mM NaCl were grown by mixing 2 μL of protein (6 mg/mL) with 2 μL of mother liquor [0.2 M ammonium sulfate, 0.1 M NaCl, 0.1 M Mes [2-(N-morpholino)ethanesulfonic acid] (pH 5.6), and 28% PEG 3350] and equilibrated against 1 mL of mother liquor in the reservoir. Cocrystals of $FmlH_{LD}$ bound to TF or galactosides 4β (in space group P 2 21 21), 5β, and 20β were grown by mixing 2 μL of protein (6 mg/mL) in the presence of 5 mM compound with 2 μL of mother liquor [0.2 M ammonium sulfate, 0.1 M NaCl, 0.1 M Mes (pH 5.6), and 32% PEG 3350] and equilibrated against 1 mL of mother liquor in the reservoir. These crystals were transferred into cryoprotectant [0.2 M ammonium sulfate, 0.1 M NaCl, 0.1 M Mes (pH 5.6), 35% PEG 4000, and 10% glycerol] and then flash frozen in liquid

nitrogen. Cocrystals of $FmlH_{LD}$ bound to the galactoside 29β-NAc were grown by mixing 2 μL of protein complex (10 mg/mL $FmlH_{LD}$ with a 1.2:1 molar ratio of 29β-NAc to $FmlH_{LD}$) with 2 μL of mother liquor (0.7 M LiSO4 and 2% PEG8000) and equilibrated against 1 mL of mother liquor in the reservoir. These crystals were transferred into cryoprotectant (1 M $LiSO_4$, 10% PEG8000, and 25% glycerol). Diffraction data for TF, 4β (in space group C 1 2 1), and 29β-NAc structures were collected at 100 K at an in-house facility equipped with a rotating anode Rigaku MicroMax 007 generator, a Rayonix Marmux X-ray source, and a Mar345 image plate detector. Diffraction data for apo, 4β, 5β, and 20β structures were collected at 100 K at the ALS Beamline 4.2.2. Data were indexed and integrated in iMosflm [55], XDS [56], or HKL2000 and scaled by Scala [57]. The phase problem was solved by molecular replacement using Phaser-MR in PHENIX [58] with $FimH_{LD}$ from PDB ID 3MCY. Several rounds of refinements were performed in PHENIX to improve the final models.

Mouse Infections

Seven- to 8-wk-old female C3H/HeN mice were obtained from Envigo. Mice were anesthetized and inoculated via transurethral catheterization with 50 μL of CFT073 bacterial suspension (∼1×10^8 to 2×10^8 CFU in total) in PBS. Mice experiencing high titers of bacteriuria (>10^4 CFU/mL) and edematous and inflamed bladders when killed after 2 wk, or chronic cystitis [39], were then transurethrally inoculated either with 50 mg/kg compound or vehicle control (10% DMSO). Mice were killed 6 h posttreatment, and bacteria colonizing the bladder or kidney were plated for quantification.

Immunofluorescence

Frozen, deidentified human bladder and kidney sections were obtained from the Tissue Procurement Core and stored stably at −80 °C. These tissue section slides were removed from the

freezer and allowed to thaw at room temperature for 10 to 20 min. After applying a hydrophobic barrier pen around the tissue, slides were rehydrated in 200 µL buffer (5% BSA and 0.2% Triton X-100 in PBS) for 10 min. Buffer was gently aspirated and slides were blocked for 1 h at room temperature with 200 µL of buffer. Thereafter, buffer was gently aspirated and slides were incubated with 200 µL of sample overnight at 4 °C. Samples diluted in buffer included 50 µg/mL $FmlH_{LD}$ wild-type (WT), 50 µg/mL $FmlH_{LD}$ K132Q, and 50 µg/mL $FmlH_{LD}$ WT incubated with 100 µM 29β-NAc. Samples were gently aspirated and slides were washed three times in buffer for 5 min each. Next, slides were incubated with our mouse anti-FmlH polyclonal antibody (1:500 dilution in buffer) for 1 h at room temperature. Slides were washed again three times in buffer and then incubated in the dark with donkey anti-mouse IgG Alexa Fluor 594 and Wheat Germ Agglutinin Alexa Fluor 633 (each 1:500 dilution in buffer) for 1 h at room temperature. Slides were washed once with buffer and then incubated in the dark with DAPI (1:1,000 dilution in buffer) for 5 min at room temperature. After washing twice with buffer, coverslips were mounted using 80 µL of mounting media. Slides were loaded onto a Zeiss LSM 880 Confocal Laser Scanning Microscope (Carl Zeiss, Inc.) equipped with a diode 405 to 430 laser, a HeNe 543 laser, and a HeNe 633 laser. Images were acquired with a 20×, 0.8 numerical aperture Zeiss Plan Apochromat objective using ZEN 2 imaging software.

Statistics

Mouse data are compiled from two (4Z269 plus 29β-NAc) or three (all other treatments) independent experiments, with four or five mice per group per experiment. These data were analyzed using the uncorrected two-tailed Mann–Whitney U test in GraphPad Prisim v.5. ELISA data are reported as box-and-whisker plots indicating the mean, 2.5th, 25th, 75th, and 97.5th percentiles of at least two independent experiments, with three technical replicates per experiment.

Acknowledgments

We thank members of the S.J.H. laboratory for helpful suggestions; Rick Stegeman at Washington University and Jay Nix at ALS Beamline 4.2.2 for technical assistance in X-ray data collection; and Wandy Beatty at Washington University for assistance and expertise in confocal microscopy. We thank Ocean Spray for their helpful advice and the Alvin J. Siteman Cancer Center at Washington University School of Medicine, the Barnes-Jewish Hospital, and the Institute of Clinical and Translational Sciences (ICTS) at Washington University in St. Louis, for the use of the Tissue Procurement Core, which provided human urinary tract tissue. The Alvin J. Siteman Cancer Center is supported, in part, by National Cancer Institute Cancer Center Support Grant P30 CA091842. The ICTS is funded by NIH National Center for Advancing Translational Sciences Clinical and Translational Science Award Program Grant UL1 TR002345. J.W.J. and S.J.H. were supported by NIH National Institute of Diabetes and Digestive and Kidney Diseases Grant R01 DK108840. V.K. was supported by Medical Scientist Training Program Grant T32GM07200.

Author contributions: V.K., M.E.H., A.R.M., J.S.P., J.W.J., and S.J.H. designed research; V.K. and J.W.J. designed compounds; V.K., M.E.H., A.R.M., R.C., J.S.P., and L.K.M.-M. performed research; A.R.M., R.C., and L.K.M.-M. synthesized compounds; V.K., M.S.C., and J.S.P. purified proteins; V.K. and J.S.P. performed virtual screening, X-ray crystallography, ELISA, bio-layer interferometry, and immunofluorescence experiments; M.E.H. performed animal experiments; V.K., A.R.M., R.C., J.S.P., L.K.M.-M., and M.S.C. contributed new reagents/analytic tools; V.K., M.E.H., A.R.M., J.S.P., J.W.J., and S.J.H. analyzed data; V.K., M.E.H., J.S.P., J.W.J., and S.J.H. interpreted all data; and V.K., M.E.H., A.R.M., J.W.J., and S.J.H. wrote the paper.

Conflict of interest statement: J.W.J. and S.J.H. are inventors on US patent US8937167 B2, which covers the use of mannoside-based FimH ligand antagonists for the treatment of disease. J.W.J., M.E.H., and S.J.H. have ownership interests in Fimbrion Therapeutics and may benefit if the company is successful in marketing mannosides.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID codes 6AOW, 6AOX, 6AOY, 6ARM, 6ARN, 6ARO, and 6AS8).

References

1. Foxman B. Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. Dis Mon. 2003;49:53–70.
2. Griebling TL. Urologic diseases in America project: Trends in resource use for urinary tract infections in women. J Urol. 2005;173:1281–1287.
3. Gupta K, Hooton TM, Roberts PL, Stamm WE. Patient-initiated treatment of uncomplicated recurrent urinary tract infections in young women. Ann Intern Med. 2001;135:9–16.
4. Foxman B. The epidemiology of urinary tract infection. Nat Rev Urol. 2010;7:653–660.
5. Ramakrishnan K, Scheid DC. Diagnosis and management of acute pyelonephritis in adults. Am Fam Physician. 2005;71:933–942.
6. Pertel PE, Haverstock D. Risk factors for a poor outcome after therapy for acute pyelonephritis. BJU Int. 2006;98:141–147.
7. Efstathiou SP, et al. Acute pyelonephritis in adults: Prediction of mortality and failure of treatment. Arch Intern Med. 2003;163:1206–1212.
8. Roberts FJ, Geere IW, Coldman A. A three-year study of positive blood cultures, with emphasis on prognosis. Rev Infect Dis. 1991;13:34–46.
9. WHO 2014. Antimicrobial resistance: Global report on surveillance 2014 (WHO, Geneva)
10. Guneysel O, Onur O, Erdede M, Denizbasi A. Trimethoprim/sulfamethoxazole resistance in urinary tract infections. J Emerg Med. 2009;36:338–341.
11. Raz R, et al. Israeli Urinary Tract Infection Group Empiric use of trimethoprim-sulfamethoxazole (TMP-SMX) in the treatment of women with uncomplicated urinary tract infections, in a geographical area with a high prevalence of TMP-SMX-resistant uropathogens. Clin Infect Dis. 2002;34:1165–1169.
12. Aypak C, Altunsoy A, Düzgün N. Empiric antibiotic therapy in acute uncomplicated urinary tract infections and fluoroquinolone resistance: A prospective observational study. Ann Clin Microbiol Antimicrob. 2009;8:27.
13. Kallen AJ, Welch HG, Sirovich BE. Current antibiotic therapy for isolated urinary tract infections in women. Arch Intern Med. 2006;166:635–639.
14. Karlowsky JA, Hoban DJ, Decorby MR, Laing NM, Zhanel GG. Fluoroquinolone-resistant urinary isolates of *Escherichia coli* from outpatients are frequently multidrug resistant: Results from the North American Urinary Tract Infection Collaborative Alliance-Quinolone Resistance study. Antimicrob Agents Chemother. 2006;50:2251–2254.
15. McGann P, et al. *Escherichia coli* harboring mcr-1 and blaCTX-M on a novel IncF plasmid: First report of mcr-1 in the United States. Antimicrob Agents Chemother. 2016;60:4420–4421.
16. Rasko DA, Sperandio V. Anti-virulence strategies to combat bacteria-mediated disease. Nat Rev Drug Discov. 2010;9:117–128.
17. Ronald A. The etiology of urinary tract infection: Traditional and emerging pathogens. Dis Mon. 2003;49:71–82.
18. Ronald AR, et al. Urinary tract infection in adults: Research priorities and strategies. Int J Antimicrob Agents. 2001;17:343–348.
19. Schreiber HL, 4th, et al. Bacterial virulence phenotypes of *Escherichia coli* and host susceptibility determine risk for urinary tract infections. Sci Transl Med. 2017;9:eaaf1283.

20. Waksman G, Hultgren SJ. Structural biology of the chaperone-usher pathway of pilus biogenesis. Nat Rev Microbiol. 2009;7:765–774.
21. Wurpel DJ, Beatson SA, Totsika M, Petty NK, Schembri MA. Chaperone-usher fimbriae of *Escherichia coli*. PLoS One. 2013;8:e52835.
22. Jones CH, et al. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc Natl Acad Sci USA. 1995;92:2081–2085.
23. Mulvey MA, et al. Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. Science. 1998;282:1494–1497.
24. Mydock-McGrane LK, Hannan TJ, Janetka JW. Rational design strategies for FimH antagonists: New drugs on the horizon for urinary tract infection and Crohn's disease. Expert Opin Drug Discov. 2017;12:711–731.
25. Jarvis C, et al. Antivirulence isoquinolone mannosides: Optimization of the biaryl aglycone for FimH lectin binding affinity and efficacy in the treatment of chronic UTI. ChemMedChem. 2016;11:367–373.
26. Mydock-McGrane L, et al. Antivirulence C-mannosides as antibiotic-sparing, oral therapeutics for urinary tract infections. J Med Chem. 2016;59:9390–9408.
27. Han Z, et al. Structure-based drug design and optimization of mannoside bacterial FimH antagonists. J Med Chem. 2010;53:4779–4792.
28. Cusumano CK, et al. Treatment and prevention of urinary tract infection with orally active FimH inhibitors. Sci Transl Med. 2011;3:109ra115.
29. Grabosch C, Hartmann M, Schmidt-Lassen J, Lindhorst TK. Squaric acid monoamide mannosides as ligands for the bacterial lectin FimH: Covalent inhibition or not? ChemBioChem. 2011;12:1066–1074.
30. Kleeb S, et al. FimH antagonists: Bioisosteres to improve the in vitro and in vivo PK/PD profile. J Med Chem. 2015;58:2221–2239.
31. Chalopin T, et al. Second generation of thiazolylmannosides, FimH antagonists for *E. coli*-induced Crohn's disease. Org Biomol Chem. 2016;14:3913–3925.
32. Mydock-McGrane LK, Cusumano ZT, Janetka JW. Mannose-derived FimH antagonists: A promising anti-virulence therapeutic strategy for urinary tract infections and Crohn's disease. Expert Opin Ther Pat. 2016;26:175–197.
33. DJ, et al. F9 fimbriae of uropathogenic *Escherichia coli* are expressed at low temperature and recognise Galβ1-3GlcNAc-containing glycans. PLoS One. 2014;9:e93177.
34. Conover MS, et al. Inflammation-induced adhesin-receptor interaction provides a fitness advantage to uropathogenic *E. coli* during chronic infection. Cell Host Microbe. 2016;20:482–492.
35. Kalas V, et al. Evolutionary fine-tuning of conformational ensembles in FimH during host-pathogen interactions. Sci Adv. 2017;3:e1601944.
36. Le Trong I, et al. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. Cell. 2010;141:645–655.
37. Sauer MM, et al. Catch-bond mechanism of the bacterial adhesin FimH. Nat Commun. 2016;7:10738.
38. Mayer K, et al. Urinary tract infection: Which conformation of the bacterial lectin FimH is therapeutically relevant? J Med Chem. 2017;60:5646–5662
39. Hannan TJ, Mysorekar IU, Hung CS, Isaacson-Schmid ML, Hultgren SJ. Early severe inflammatory responses to uropathogenic *E. coli* predispose to chronic and recurrent urinary tract infection. PLoS Pathog. 2010;6:e1001042.

40. Murawski IJ, et al. The C3H/HeJ inbred mouse is a model of vesico-ureteric reflux with a susceptibility locus on chromosome 12. Kidney Int. 2010;78:269–278.
41. Subashchandrabose S, et al. Host-specific induction of *Escherichia coli* fitness genes during human urinary tract infection. Proc Natl Acad Sci USA. 2014;111:18327–18332.
42. Bouckaert J, et al. The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of *Escherichia coli* pathotypes. Mol Microbiol. 2006;61:1556–1568.
43. Hung CS, et al. Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. Mol Microbiol. 2002;44:903–915.
44. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. Nat Rev Drug Discov. 2009;8:661–677.
45. Hsu CH, et al. The dependence of carbohydrate-aromatic interaction strengths on the structure of the carbohydrate. J Am Chem Soc. 2016;138:7636–7648.
46. Hudson KL, et al. Carbohydrate-aromatic interactions in proteins. J Am Chem Soc. 2015;137:15152–15160.
47. Ter Kuile BH, Kraupner N, Brul S. The risk of low concentrations of antibiotics in agriculture for resistance in human health care. FEMS Microbiol Lett. 2016;363:fnw210.
48. Llor C, Bjerrum L. Antimicrobial resistance: Risk associated with antibiotic overuse and initiatives to reduce the problem. Ther Adv Drug Saf. 2014;5:229–241.
49. Kardas P, Devine S, Golembesky A, Roberts C. A systematic review and meta-analysis of misuse of antibiotic therapies in the community. Int J Antimicrob Agents. 2005;26:106–113.
50. Chang Q, Wang W, Regev-Yochay G, Lipsitch M, Hanage WP. Antibiotics in agriculture and the risk to human health: How worried should we be? Evol Appl. 2015;8:240–247.
51. Pinkner JS, et al. Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria. Proc Natl Acad Sci USA. 2006;103:17897–17902.
52. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455–461.
53. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A free tool to discover chemistry for biology. J Chem Inf Model. 2012;52:1757–1768.
54. O'Boyle NM, et al. Open Babel: An open chemical toolbox. J Cheminform. 2011;3:33.
55. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr. 2011;67:271–281.
56. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010;66:125–132.
57. Winn MD, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235–242.
58. Adams PD, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66:213–221.

Figures

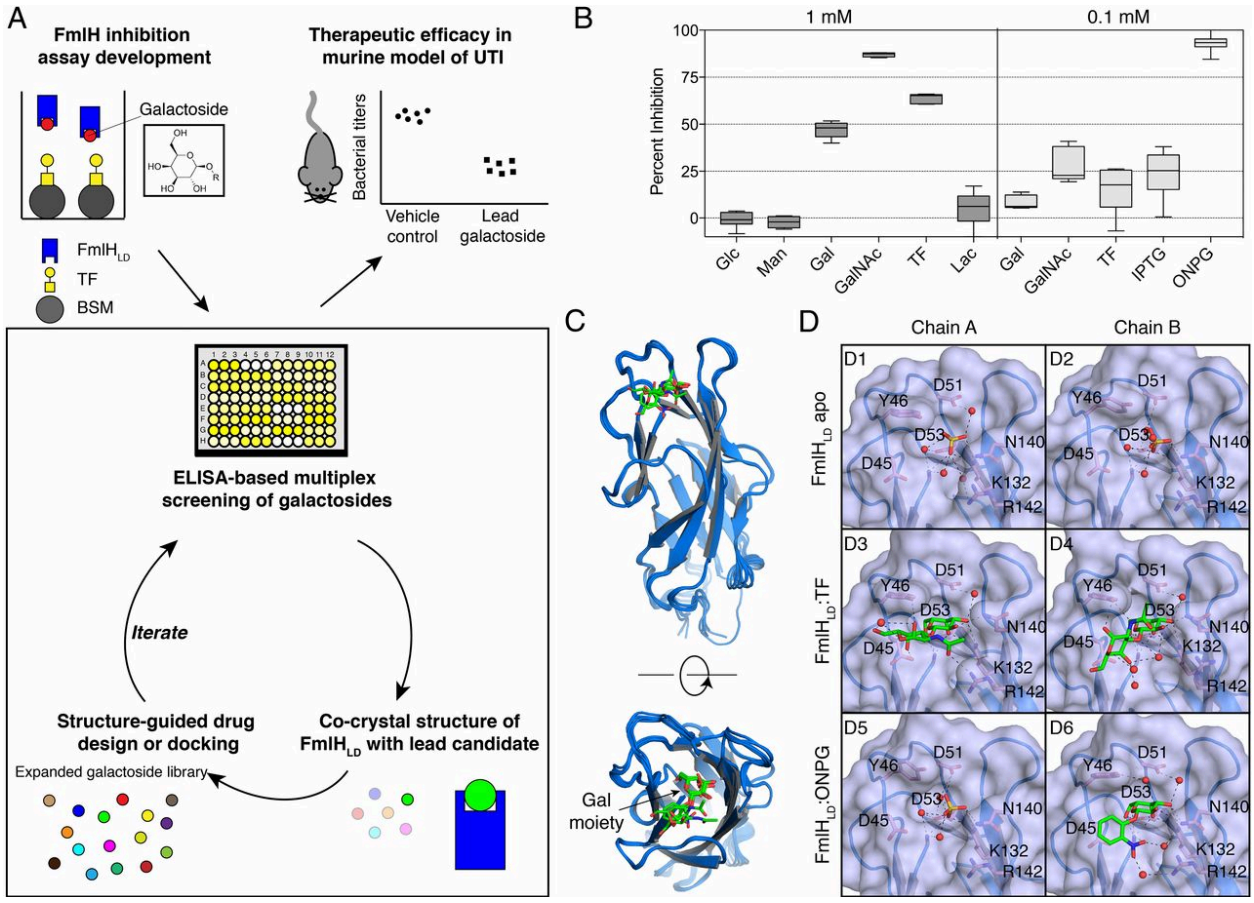


Figure 1. Biochemical and structural characterization of early galactoside antagonists of FmlH. (**A**) Strategy for structure-guided drug design and evaluation of FmlH-targeting galactosides. A select library of galactosides were initially assessed in an ELISA-based competition assay for inhibition of FmlH binding to sialidase-treated BSM, with BSM indicated by gray circles, TF residues indicated by the yellow square-circle conjugates, biotinylated $FmlH_{LD}$ by blue rectangles, and galactosides shown as colored circles. Cocrystal structures of $FmlH_{LD}$ bound to a lead compound facilitated virtual screening and structure-guided drug design for biochemical evaluation of an expanded galactoside library. The top lead compound would then be tested as a treatment in a mouse model of UTI. (**B**) ELISA-based competition assay performed in triplicate in

the absence or presence of 1 mM or 0.1 mM compounds with at least two biological replicates. Data are reported as the mean percent inhibition, with the box indicating the 25th to 75th percentiles and the whiskers indicating the 2.5th and 97.5th percentiles. (**C**) Structural alignment of $FmlH_{LD}$ from an apo $FmlH_{LD}$ crystal structure (PDB ID 6AOW), a $FmlH_{LD}$-TF cocrystal structure (PDB ID 6AOX), and a $FmlH_{LD}$-ONPG cocrystal structure (PDB ID 6AOY). (**D**) Crystal structures of sulfate ions or ligands bound in the $FmlH_{LD}$ binding pocket, with H-bonding (black dashed lines) indicated between sulfate ions (yellow sticks), ligands (green sticks), water molecules (red spheres), or side chains (pink sticks).

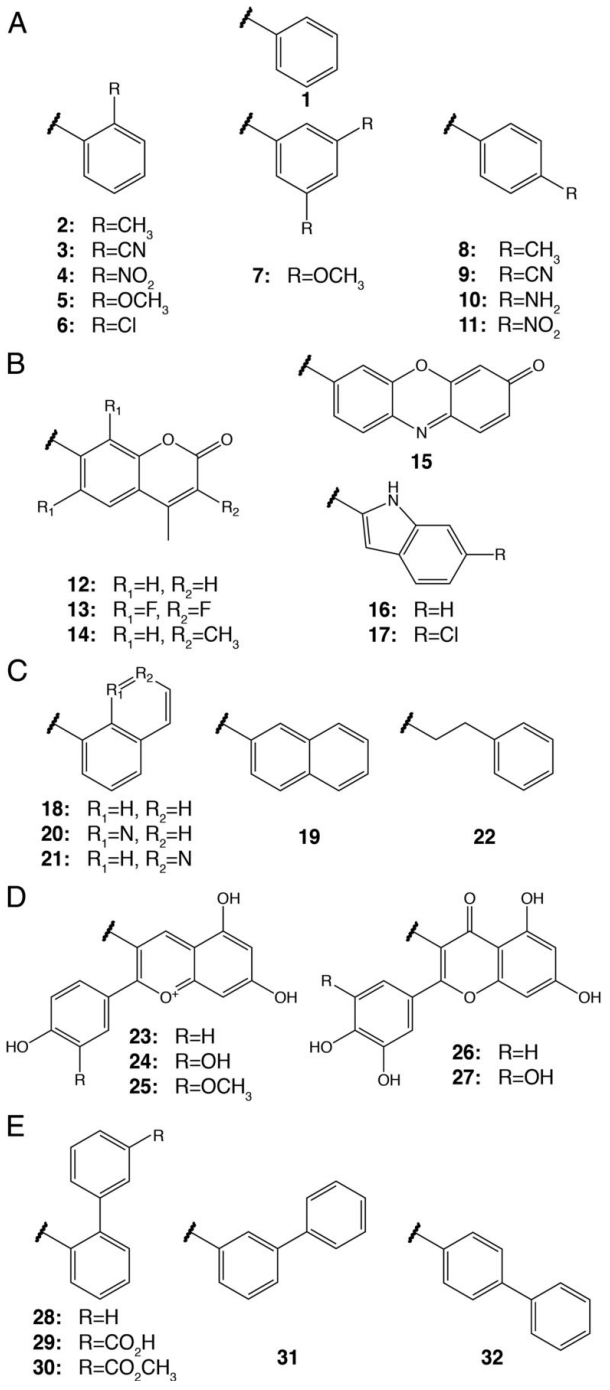
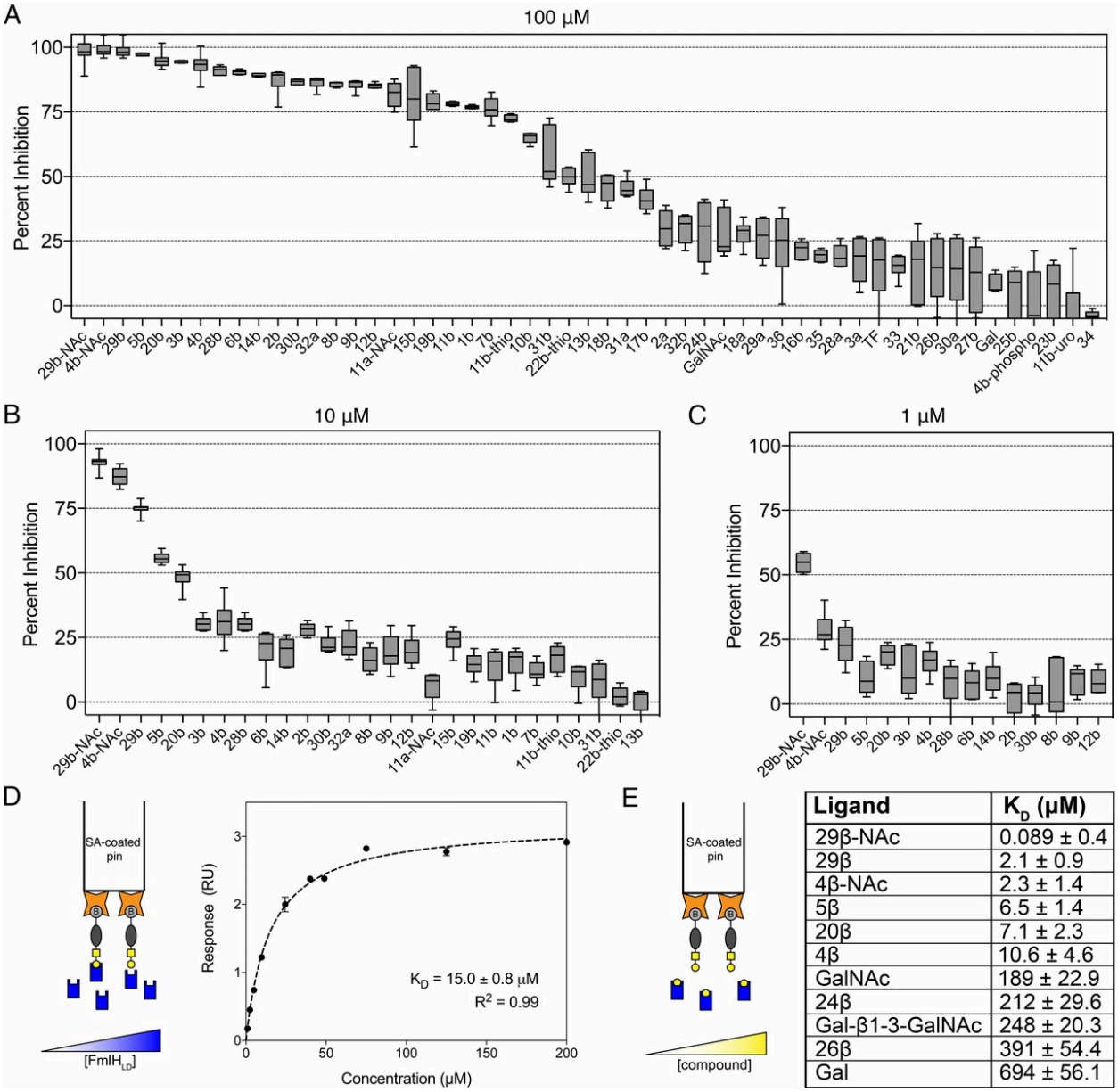


Figure 2. Grouped organization of galactosides evaluated for $FmlH_{LD}$ inhibition. The major groups include the phenyl (**A**), heterocyclic (**B**), napthyl/quinoline/phenylethyl (**C**), natural product (**D**), and biphenyl (**E**) series.



Ligand	K_D (µM)
29β-NAc	0.089 ± 0.4
29β	2.1 ± 0.9
4β-NAc	2.3 ± 1.4
5β	6.5 ± 1.4
20β	7.1 ± 2.3
4β	10.6 ± 4.6
GalNAc	189 ± 22.9
24β	212 ± 29.6
Gal-β1-3-GalNAc	248 ± 20.3
26β	391 ± 54.4
Gal	694 ± 56.1

Figure 3. In vitro screening and affinity determination of galactosides against FmlH$_{LD}$. (**A–C**) ELISA-based competition assay performed in triplicate in the absence or presence of (**A**) 100 µM, (**B**) 10 µM, and (**C**) 1 µM compounds with at least two biological replicates. Data are reported as the mean percent inhibition, with the box indicating the 25th to 75th percentiles and the whiskers indicating the 2.5th and 97.5th percentiles. a, α; b, β. (**D**, *Left*) Schematic of conventional BLI experiment, in which pins coated with streptavidin (orange stars) are loaded with biotinylated Ser-

TF (gray ovals and yellow square-circle conjugates) and dipped into solutions of varying concentrations of $FmlH_{LD}$ (blue rectangles). (*Right*) Equilibrium analysis of soluble $FmlH_{LD}$ binding to immobilized Ser-TF according to a 1:1 binding model. (**E**, *Left*) Schematic of competitive BLI experiment, in which streptavidin-coated pins are dipped into a solution composed of a fixed concentration of $FmlH_{LD}$ in the presence of varying concentrations of galactoside (yellow circles). (*Right*) Equilibrium constants of soluble galactoside-mediated inhibition of $FmlH_{LD}$ in binding immobilized Ser-TF in accord with $R^2 > 0.85$.

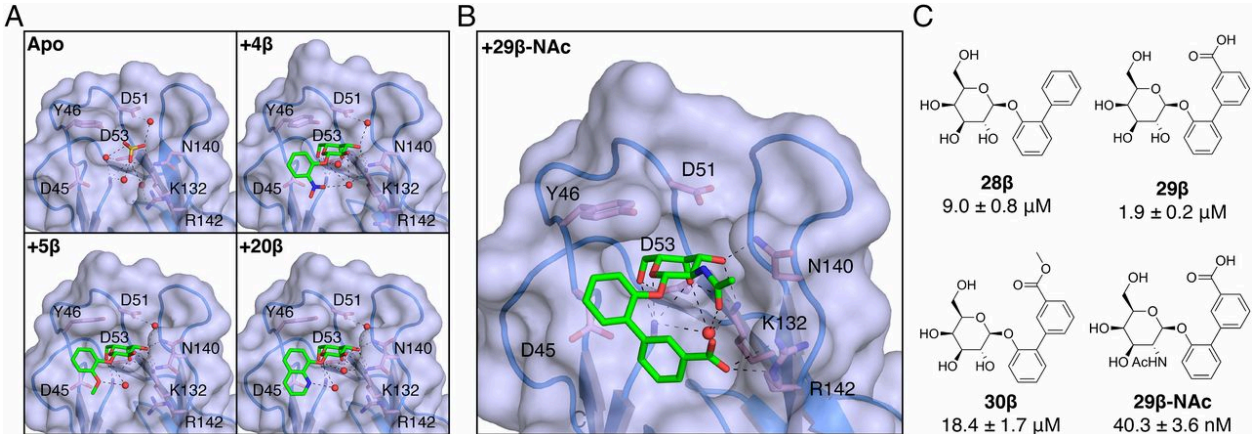


Figure 4. Structural basis of galactoside inhibition of $FmlH_{LD}$. (**A**) Crystal structures of sulfate ions or galactosides bound in the $FmlH_{LD}$ binding pocket, with H-bonding (black dashed lines) indicated between sulfate ions (yellow sticks), ligands (green sticks), water molecules (red spheres), or side chains (pink sticks). Crystal structures shown here include an apo $FmlH_{LD}$ crystal structure (PDB ID 6AOW), a $FmlH_{LD}$-4β cocrystal structure (PDB ID 6ARM), a $FmlH_{LD}$-5β cocrystal structure (PDB ID 6ARN), and a $FmlH_{LD}$-20β cocrystal structure (PDB ID 6ARO). (**B**) Cocrystal structure of 29β-NAc bound to $FmlH_{LD}$ (PDB ID 6AS8). (**C**) SARs for 29β-NAc and related compounds, with their corresponding IC_{50} values derived from the ELISA-based competition assay. IC_{50} values are reported for six replicates as the mean with SEM.

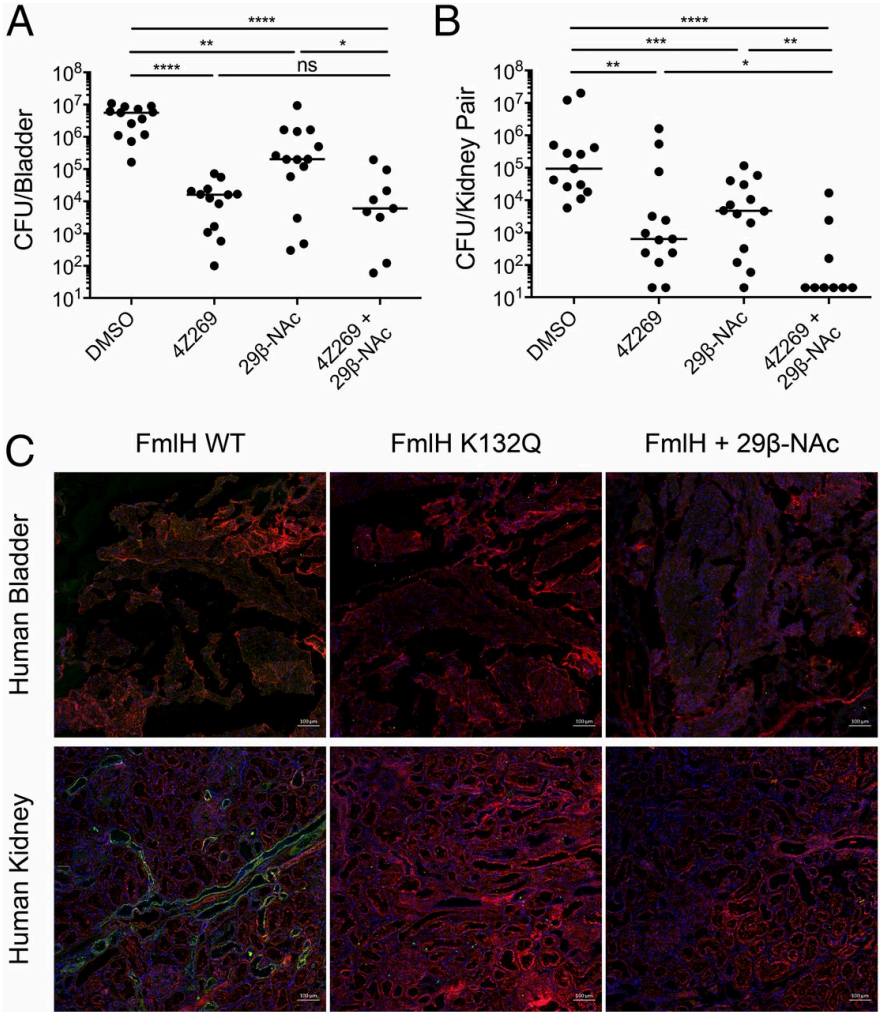


Figure 5. Evaluation of galactosides for treatment of UTI and relevance in humans. (**A** and **B**) Bacterial titers in bladders (**A**) or kidneys (**B**) from C3H/HeN mice experiencing chronic cystitis transurethrally inoculated with 10% DMSO (three replicates, n = 13), or 50 mg/kg of 4Z269 (three replicates, n = 13), of 29β-NAc (three replicates, n = 14), or of both 4Z269 and 29β-NAc (two replicates, n = 9). Bars indicate median values. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001; ns, not significant; two-tailed Mann–Whitney U test. (**C**) Immunofluorescence analysis of $FmlH_{LD}$ WT, $FmlH_{LD}$ K132Q, or $FmlH_{LD}$ WT in the presence of 29β-NAc binding to human bladder or human kidney tissue. Green corresponds to FmlH, red corresponds to Wheat Germ

Agglutinin, and blue corresponds to DAPI. Each image is representative of nine total images (three imaged areas of three tissue slices). (Scale bars: 100 μm.)

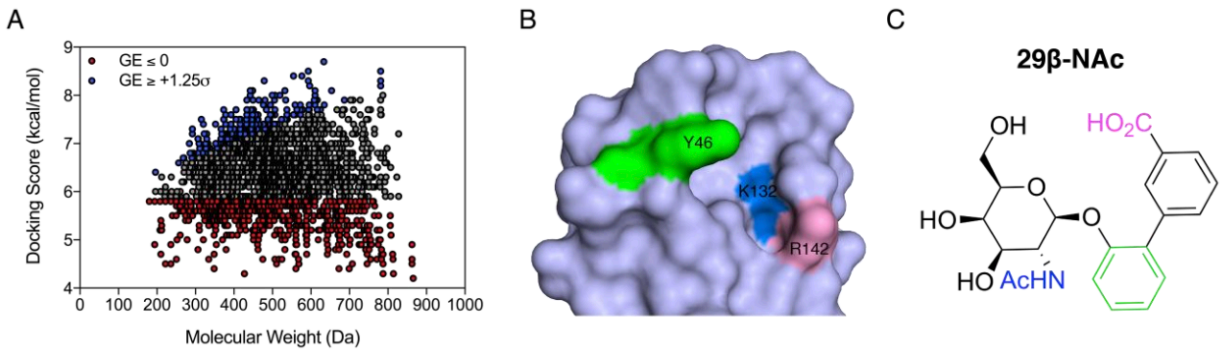


Figure S1. Virtual screen and structure-guided design of galactosides targeting $FmlH_{LD}$. (**A**) Summary of virtual screen of galactosides against $FmlH_{LD}$, in which the docking score of the top predicted binding mode is plotted against the molecular weight for each galactoside. Compounds with GE values 1.25σ above the mean are colored blue while compounds with GE values below 0 are colored red. (**B**) Surface representation of $FmlH_{LD}$ with hot spot residues Y46, K132, and R142 colored green, blue, and pink, respectively. Hot spot residues were identified as common targets observed in the binding modes of the top compounds from virtual screening. (**C**) Chemical structure of compound 29β-NAc, with the carboxylic acid in pink designed to interact with R142, the phenyl ring in green designed to interact with Y46, and the N-Acetyl group in blue designed to interact with K132.

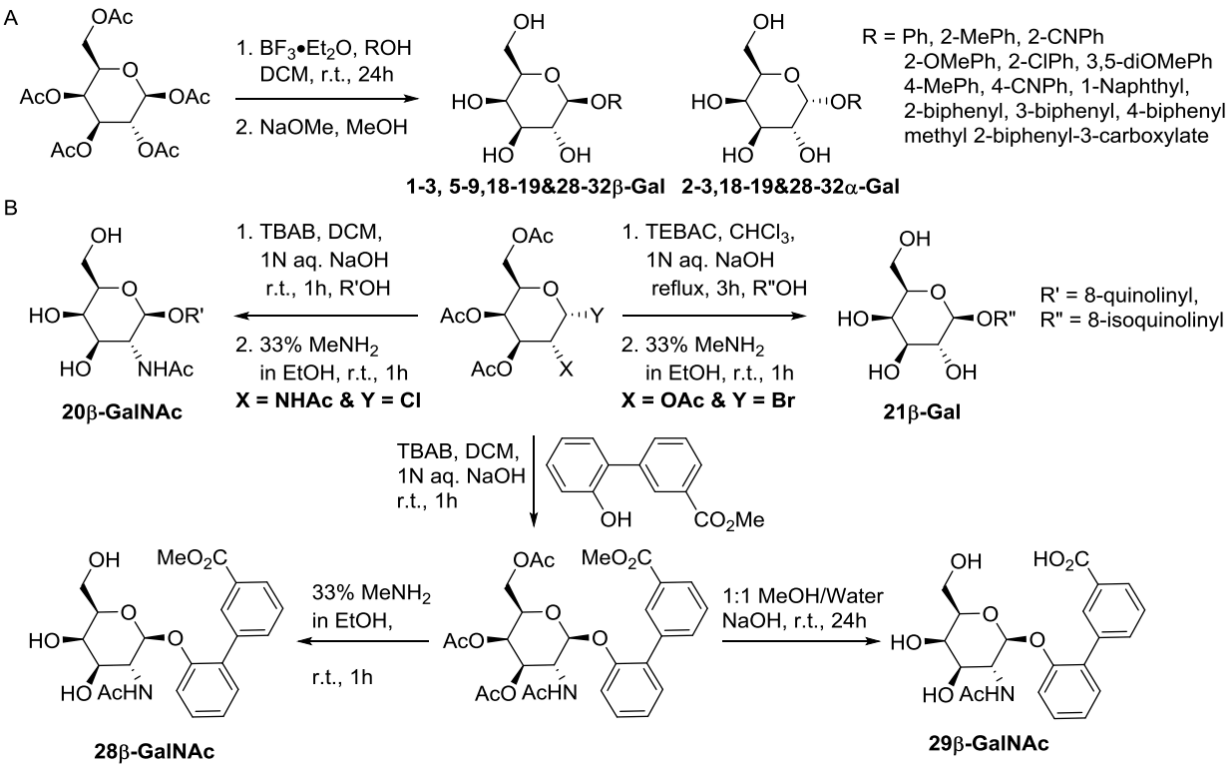


Figure S2. Synthesis of galactoside-based FmlH ligands. (**A**) Boron trifluoride promoted glycosidation reaction of protected galactose followed by deprotection to yield galactosides. (**B**) Koenigs-Knorr type reaction followed by deprotection for the substitution of a galactosyl halide with an alcohol to yield galactoside and galactosaminosides.

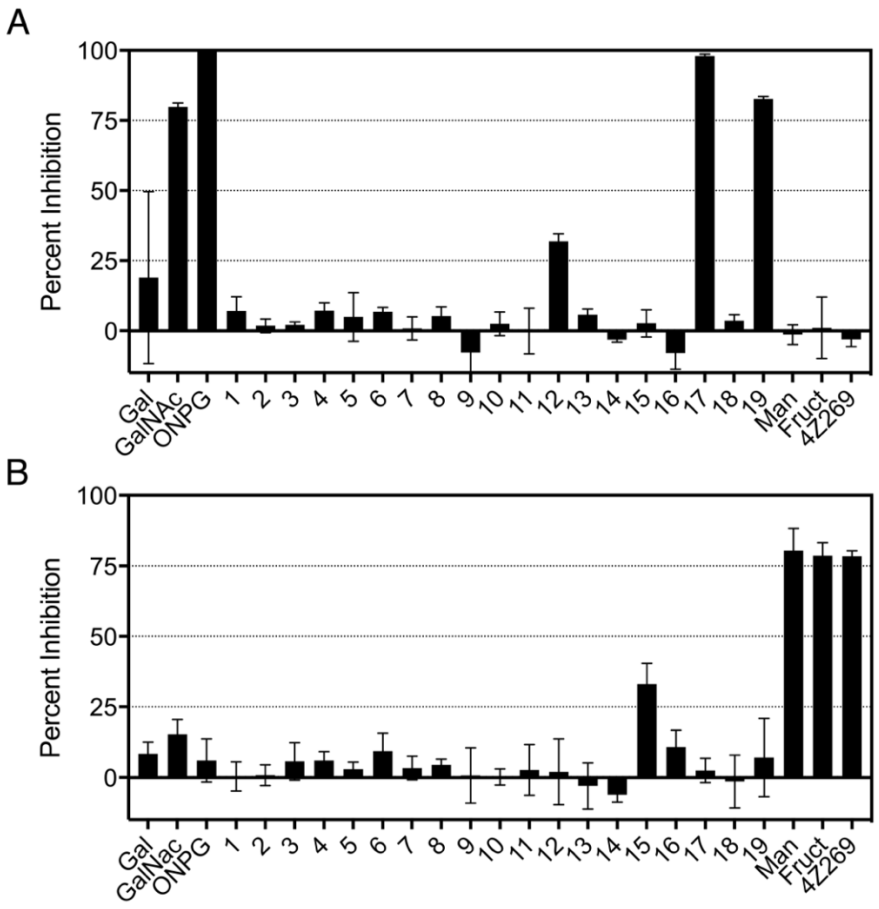


Figure S3. Cranberry-derived galactosides can inhibit FmlH. (**A**) Cranberry-derived compounds and fractions were tested at 1 mM in the ELISA-based competition assay for inhibition of $FmlH_{LD}$. (**B**) As a control, cranberry-derived compounds and fractions were tested at 1 mM in the ELISA-based competition assay for inhibition of $FimH_{LD}$. The identifies of the compounds indicated above are as follows: 1 (quinic acid), 2 (gallic acid), 3 (p-coumaric acid), 4 (2,4-dihydrobenzoic acid), 5 (protocatechuic acid), 6 (ferulic acid), 7 (vanillic acid), 8 (catechin), 9 (epicatechin), 10 (quecetin), 11 (quercitrin), 12 (quercetin galactoside; 26β) 13 (myricetin), 14 (myricetrin), 15 (cranberry fraction 1 – oligosaccharide), 16 (cranberry fraction 2 –anthocyanins/flavonols), 17 (cranberry fraction 3 mixed-sized proanthocyanidins), 18 (cyanidin arabinoside), 19 (cyanidin galactoside; 24β), Fruct (fructose).

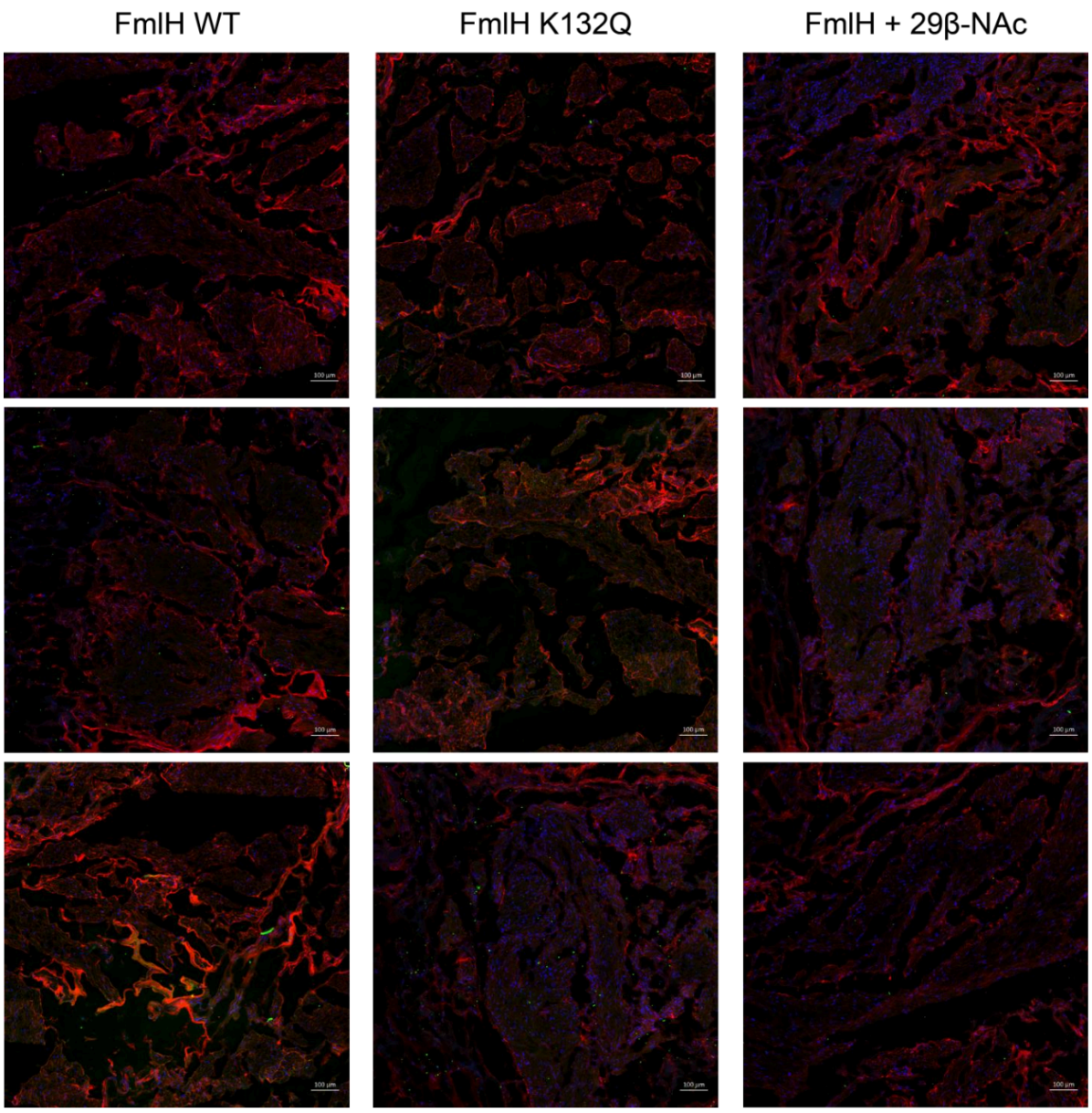


Figure S4. Immunofluorescence analysis of $FmlH_{LD}$ WT, $FmlH_{LD}$ K132Q, or $FmlH_{LD}$ WT in the presence of 29β-NAc binding to human bladder tissue. Green corresponds to FmlH, red corresponds to Wheat Germ Agglutinin, and blue corresponds to DAPI.

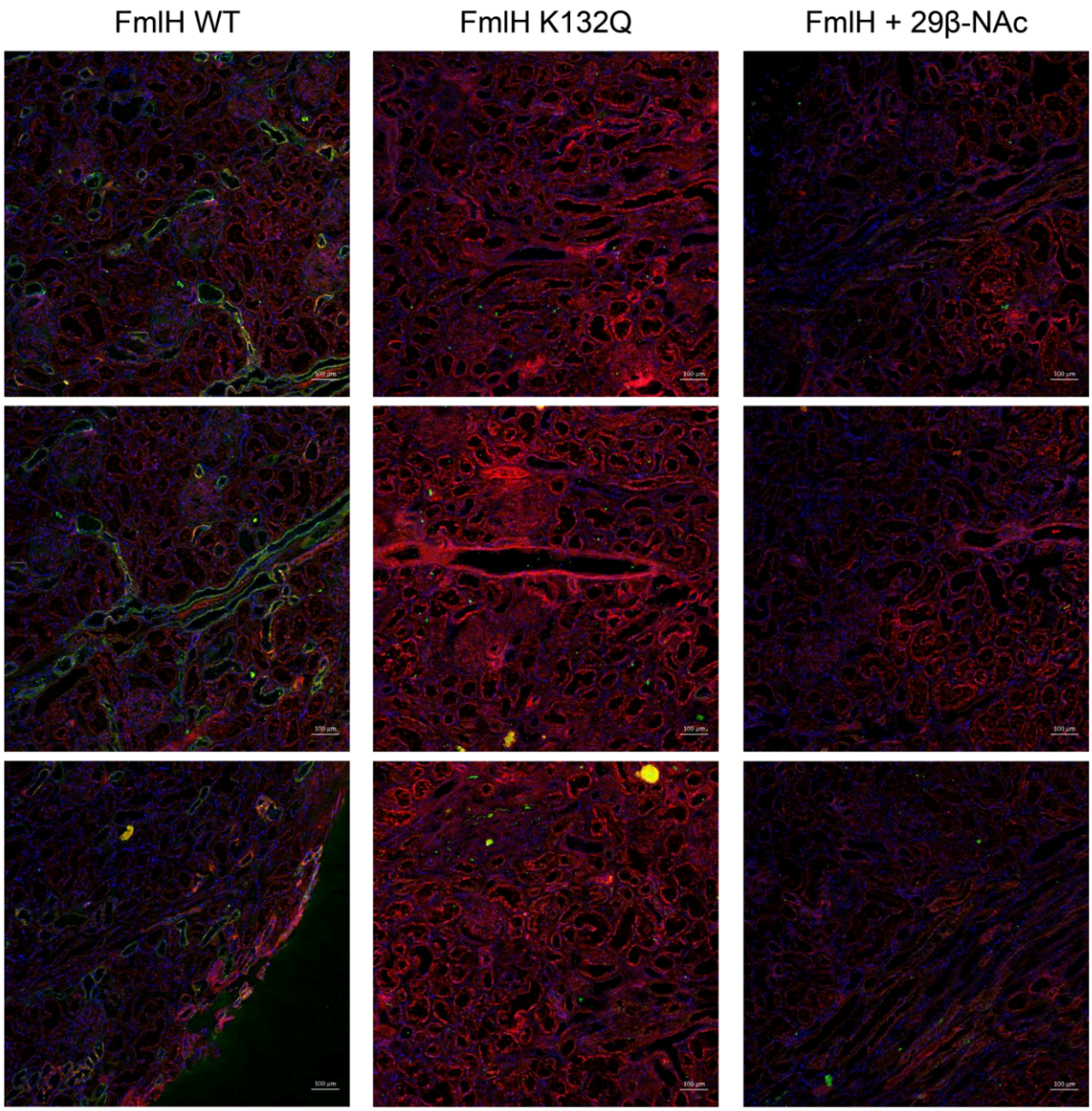


Figure S5. Immunofluorescence analysis of $FmlH_{LD}$ WT, $FmlH_{LD}$ K132Q, or $FmlH_{LD}$ WT in the presence of 29β-NAc binding to human kidney tissue. Green corresponds to FmlH, red corresponds to Wheat Germ Agglutinin, and blue corresponds to DAPI.

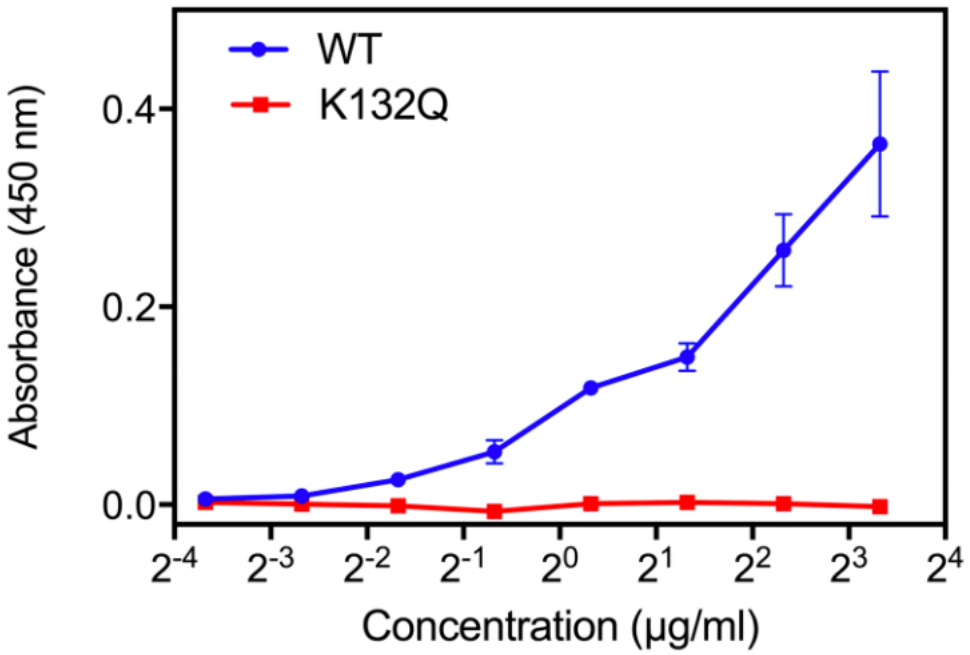


Figure S6. Mutagenesis of FmlH binding pocket abrogates function. Varying concentrations of $FmlH_{LD}$ WT and K132Q were tested for binding to sialidase-treated BSM by ELISA.

Table S1. Data collection and refinement statistics

	apo $FmlH_{LD}$ (6AOW)	$FmlH_{LD}$:TF (6AOX)	$FmlH_{LD}$:ONPG (6AOY)	$FmlH_{LD}$:**4β** (6ARM)	$FmlH_{LD}$:**5β** (6ARN)	$FmlH_{LD}$:**20β** (6ARO)	$FmlH_{LD}$:**29β-Nac** (6AS8)
Data collection							
Space group	C 2	C 2 2 21	C 2	P 2 21 21	P 21 21 2	P 21 21 2	P 2 21 21
Cell dimensions							
a, *b*, *c* (Å)	65.5, 78.3, 58.5	67.4, 78.1, 105.5	66.2, 78.3, 58.5	50.0, 51.3, 114.6	51.3, 116.1, 50.3	51.3, 116.3, 50.6	51.0, 51.5, 117.5
α, *β*, *γ* (°)	90.0, 97.0, 90.0	90.0, 90.0, 90.0	90.0, 97.5, 90.0	90.0, 90.0, 90.0	90.0, 90.0 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50.0-1.6 (1.63-1.60)	52.8-2.1 (2.21-2.10)	40.0-1.8 (1.90-1.80)	57.3-1.50 (1.53-1.50)	51.3-1.25 (1.27-1.25)	58.1-1.15 (1.17-1.15)	50.0-2.10 (2.14-2.10)
R_{merge} (%)[a]	9.3 (139)	27.2 (189)	8.1 (67.8)	6.8 (137)	6.5 (192)	11.4 (86.9)	14.6 (51.3)
R_{pim} (%)[b]	5.7 (87.6)	8.0 (55.7)	3.2 (32.2)	2.8 (68.7)	2.7 (102)	4.8 (63.7)	4.1 (26.0)
$I / \sigma I$	8.7 (0.9)	8.0 (1.6)	15.6 (2.8)	16.0 (0.8)	15.0 (0.7)	9.2 (1.0)	16.1 (1.5)
Completeness (%)	98.5 (99.8)	99.6 (99.9)	99.6 (98.5)	96.4 (78.3)	99.2 (91.5)	93.6 (58.5)	87.7 (44.2)
Multiplicity	3.5 (3.4)	12.0 (12.0)	6.9 (5.1)	6.2 (4.1)	6.5 (4.2)	6.0 (2.6)	10.7 (2.5)
$CC_{1/2}$	1.0 (0.38)	0.99 (0.39)	1.0 (0.79)	1.0 (0.28)	1.0 (0.31)	0.99 (0.46)	0.99 (0.82)
Total / Unique reflections	134,257/38,073(6,454/1,891)	199,439/16,562(28,511/2,379)	188,555/27,380(19,891/3,397)	288,591/46,203(7,374/1,811)	544,119/83,215(15,705/3,708)	610,631/101,126(7,812/3,048)	176,770/16,455(932/412)
Refinement							
R_{work}[c]/ R_{free}[d]	19.5 / 23.3	22.5 / 25.9	18.5 / 22.0	20.1 / 24.0	20.8 / 22.4	17.6 / 19.0	20.1 / 24.0
No. atoms							
Protein	2307	2247	2345	2309	2321	2431	2322
Ligand/ion	15	52	26	47	45	49	60
Water	262	140	263	326	350	513	126
B-factors							
Protein	24.9	35.6	26.9	23.2	19.7	14.5	30.2
Ligand/ion	40.6	31.5	40.1	25.2	19.4	14.7	32.3
Water	35.5	38.2	32.8	30.8	29.5	30.5	32.6
R.m.s. deviations							
Bond lengths (Å)	0.009	0.004	0.008	0.008	0.006	0.013	0.011
Bond angles (°)	1.19	0.98	1.11	1.19	1.10	1.54	0.81
Ramachandran plot							
favored (%)	97.0	98.0	96.0	94.0	97.0	97.0	96.0
allowed (%)	3.0	2.0	3.0	5.7	2.3	2.7	3.7
outliers (%)	0	0	1.0	0.3	0.7	0.3	0.3
Clashscore	2.6	1.6	3.9	9.1	6.1	3.5	0.4

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl) \rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where the sum *i* is over all separate measurements of the unique reflection *hkl*.

[b] $R_{pim} = \Sigma_{hkl} [1/(n-1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl) \rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$

[b] $R_{work} = \Sigma_{hkl} ||F_{obs}| - |F_{calc}|| / \Sigma_{hkl} |F_{obs}|$

[c] R_{free}, calculated the same as for R_{work} but on the 5% data randomly excluded from the refinement calculation.

Values in parentheses indicate the highest resolution shell

Table S2. Galactoside inhibition of FmlH

	PI_{100}	PI_{10}	PI_1
29β-NAc	99.5 ± 2.2	93.0 ± 1.5	54.7 ± 2.8
4β-NAc	99.4 ± 1.3	87.2 ± 1.7	28.6 ± 4.5
29β	99.0 ± 1.3	75.1 ± 1.0	22.9 ± 5.1
5β	97.2 ± 0.4	55.7 ± 1.6	10.0 ± 4.3
20β	95.1 ± 1.1	48.1 ± 2.0	19.2 ± 2.8
3β	94.5 ± 0.3	30.4 ± 1.8	12.1 ± 6.3
4β (ONPG)	93.0 ± 1.4	31.1 ± 3.3	16.5 ± 3.8
28β	91.1 ± 1.2	31.0 ± 2.1	8.1 ± 6.0
6β	90.5 ± 0.6	20.7 ± 5.5	8.0 ± 4.0
14β	89.2 ± 0.4	19.8 ± 3.8	10.2 ± 4.3
2β	87.3 ± 3.7	28.2 ± 1.8	1.8 ± 5.7
30β	86.6 ± 0.8	22.4 ± 2.6	3.7 ± 3.5
32α	86.5 ± 1.7	22.6 ± 3.9	
8β	85.8 ± 0.7	16.5 ± 3.3	4.6 ± 7.9
9β	85.7 ± 1.6	19.2 ± 4.8	9.5 ± 3.7
12β	85.3 ± 0.7	19.7 ± 4.2	8.7 ± 3.1
11α-NAc	82.0 ± 2.7	6.3 ± 3.7	
15β	80.4 ± 6.6	24.0 ± 3.2	
19β	78.8 ± 2.1	14.6 ± 3.1	
11β	78.1 ± 0.6	13.7 ± 5.3	
1β	76.9 ± 0.4	15.5 ± 4.2	
7β	76.4 ± 2.4	11.7 ± 2.7	
11β-thio	72.5 ± 0.9	17.0 ± 3.7	
10β	65.1 ± 1.4	9.8 ± 3.9	
31β	56.9 ± 6.2	7.8 ± 5.5	
22β-thio	49.8 ± 2.5	2.3 ± 2.5	
13β	49.7 ± 4.5	0.8 ± 3.1	
18β	45.9 ± 3.7		
31α	45.6 ± 2.5		
17β	41.2 ± 3.4		
2α	30.0 ± 5.4		
32β	30.0 ± 3.9		
24β	28.8 ± 6.5		
GalNAc	28.5 ± 5.2		
18α	28.1 ± 3.4		
29α	26.2 ± 5.9		
36β-thio (IPTG)	23.7 ± 5.3		
16β	21.7 ± 2.4		
35	19.4 ± 1.6		
28α	19.2 ± 3.0		
3α	17.8 ± 6.3		
TF	15.5 ± 6.7		
33	15.3 ± 3.1		
21β	15.2 ± 7.1		
26β	14.4 ± 6.8		
30α	13.1 ± 7.6		
27β	10.7 ± 7.2		
Gal	8.1 ± 2.6		
25β	3.6 ± 6.2		
4β-phospho	1.3 ± 4.7		
23β	0.7 ± 10.3		
11β-uro	-2.4 ± 5.2		
34	-3.6 ± 1.0		

Percent inhibition (PI) values are reported as the mean with standard error of the mean for galactosides tested at 100 μM (PI_{100}), 10 μM (PI_{10}), and 1 μM (PI_1) in the ELISA-based competition assay. Compounds indicated above not shown in Figure S1 include 33 (p-nitrophenyl Gal-β1-3-GalNAc), 34 (p-nitrophenyl Gal-β1-3-GlcNAc), 35 (p-nitrophenyl Gal-β1-6-Gal), and 36β-thio (IPTG). The "thio" designation indicates a sulfur linkage between the sugar and the aglycone group, the "phospho" designation indicates a phosphate group attached the C6-hydroxyl group on the sugar, and the "uro" designation indicates galacturonide as the sugar.

Chapter 4: Conclusions and Future Directions

The global emergence and dissemination of antibiotic resistance among Gram-negative pathogens has crippled the ability of current last-resort therapies to effectively treat bacterial infections, including urinary tract infections (UTIs). This has become a looming worldwide crisis that has prompted the development of novel treatment and prevention strategies for infections, including the targeting of bacterial virulence mechanisms. Extracellular fibers termed chaperone-usher pathway (CUP) pili are critical virulence factors in a wide variety of pathogenic Gram-negative bacteria, functioning in adherence and invasion of host tissues and in biofilm formation on medical devices and body habitats. These fibers are each assembled by an outer membrane β-barrel usher from multiple pilus subunits properly folded by a periplasmic chaperone. There are more than a hundred different types of CUP pili, which mediate a considerable range of biological tropisms, frequently, through presentation of two-domain adhesins at the distal pilus tip. These adhesins have a lectin domain, which often recognizes a specific carbohydrate receptor, and a pilin domain to anchor the adhesin to the pilus. The type 1 pilus adhesin FimH, for example, binds mannosylated receptors on the bladder epithelium, mediates invasion of superficial facet cells, and promotes formation of biofilm-like collections termed intracellular bacterial communities (IBCs), all critical events in the establishment of urinary tract infection (UTI) by uropathogenic *E. coli* (UPEC). Studies that uncovered allelic variation in FimH among clinical UPEC isolates, which drastically impacts FimH function and virulence, have challenged our understanding of the FimH structure-function relationship, necessitating new approaches to understand the mechanisms of adhesin function and the pathogenic correlates of allelic diversity. Furthermore, UPEC encode several CUP adhesins, including FmlH and SfaH, which presumably confer distinct tropisms but remain poorly characterized. The work presented in this thesis uncovers the structural basis of UPEC adhesion in the urinary tract mediated by distinct CUP adhesins. I have discovered that

positively selected residues in FimH alter a pre-existing equilibrium of two conformational states to fine-tune mannose binding strength and host-pathogen interactions in order to optimize persistence in the bladder during UTI. Further, I defined the structural basis of the FimH-like adhesin FmlH in binding to galactose-containing receptors exposed in kidneys and inflamed bladders during chronic cystitis. I then leveraged these structural insights to develop novel small-molecule galactosides that function as high-affinity competitive antagonists of FmlH with therapeutic efficacy in mouse models of UTI. Together, the work presented in this thesis provides fundamental, biophysical underpinnings of host-pathogen interactions in UTI and provides novel treatment modalities that may help reduce the economic and health burdens imposed by this common and increasingly antibiotic-resistant infectious disease.

Discussion

Conformational ensembles of FimH alleles

The type 1 pilus adhesin FimH binds mannosylated receptors on the bladder epithelium, mediates invasion of superficial facet cells, and promotes formation of tight-knit biofilm-like communities termed intracellular bacterial communities (IBCs), all critical to the establishment of UTIs by UPEC [1-7]. FimH has a mannose binding lectin domain ($FimH_{LD}$) and a pilin domain ($FimH_{PD}$), which is joined via DSE to the FimG pilin [8]. Structural snapshots of pilus assembly [9-12] have revealed two distinct conformations in FimH, which I refer to as a low-affinity **T** state and high-affinity **R** state. Functional assays have hinted that mannose binds tightly to the **R** state but weakly, if at all, to the **T** state [9, 13], raising the question as to how FimH in a tip, which is observed by crystallography to adopt the **T** state, manages to bind mannose in solution. While the mannose binding pocket of FimH is invariant in all clinical UPEC strains, natural variation occurs outside of the binding pocket with residues at positions 27, 62, and 163 under positive selection [14]. In particular, the A27V/V163A double mutation confers increased mannose binding *in vitro* but results in a 10,000-fold reduction in mouse bladder colonization 24 hours post infection (hpi) and an inability to form IBCs. Thus, the A27V/V163A double mutation revealed either the necessity of an unknown function of FimH or the requirement for the regulation of mannose binding for *in vivo* fitness. The observation that positions which lie far away from the mannose binding pocket can, by some allosteric mechanism, impact FimH-dependent functions in UPEC pathogenesis is profound, but was not well understood.

These questions prompted the investigation of the structural and evolutionary basis for conformational allostery in FimH at the pilus tip. FimH alleles purified within a $FimG_{Nte}H$ complex, which functioned as a tip-like setting, were interrogated by a battery of molecular

biophysics techniques, including X-ray crystallography, small-angle X-ray scattering, ion mobility mass spectrometry, and molecular dynamics. These methods revealed that FimH in a tip-like setting natively adopts a pre-existing two-state conformational equilibrium in solution comprised of a single low-affinity **T** conformation and multiple, dynamic high-affinity **R** conformations. Positively selected residues can influence the relative proportions of these conformations. With respect to FimH from UPEC strain UTI89 (defined as WT), A62S skews the equilibrium toward the low-affinity **T** state while A27V/V163A pushes the equilibrium toward the high-affinity **R** state. In addition, high affinity mannose analogs termed mannosides can shift the conformational ensembles of FimH toward the **R** state. Yet despite their high-affinity for the **R** state, mannosides can engage both the **T** and **R** states through distinct binding modes, in which mannoside can enter and dock onto the low-affinity, solvent-exposed binding pocket of the **T** state in a "vertical" orientation or tightly interact with the compressed binding loops of the **R** state in a "horizontal" orientation. Together, the conformational landscape and ligation states indicate FimH exists in a thermodynamic cycle whereby binding can possibly proceed through induced fit (apo **T** → bound **T** → bound **R**) or conformational selection (apo **R** → bound **R**) or a mixture of both. The contribution of these binding mechanisms for each FimH allele requires further exploration through detailed kinetic studies capable of resolving each of these apo and ligated FimH conformational states. Furthermore, the conformational flexibility and 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 but further work is required to demonstrate that entropic freedom in the bound **R** state influences bacterial persistence at the host-pathogen interface to function as a molecular tether.

Our work also revealed that FimH alleles skewed toward the **T** state (A62S) or toward the **R** state (A27V/V163A) exhibit defects in pathogenesis in a mouse model of UTI, suggesting that an equilibrium between **T** and **R** states is optimal for bladder colonization. The mechanism by which the bladder habitat selects against the high-affinity **R** state as early as 1 hour post infection remains unknown and requires further investigation. Some possible explanations for this pathogenesis defect include a slow rate of association of the **R** state under conditions of flow, an inhibitory property mediated by intact superficial facet cells of the bladder epithelium, or soluble decoy receptors in the urine that preferentially antagonize variants that favor the **R** state. Together, our data suggest that the **T** state may serve to temporarily mask the strong affinity that FimH has for mannose to avoid restrictive factors or properties native to the bladder habitat to initiate productive binding when engaging the host epithelium.

Overall, in-depth analysis of sequence variation in a virulence factor expressed by clinical *E. coli* isolates critical for UTI led to the elucidation of fundamental principles underlying host-pathogen interactions. The work presented herein exposes the biophysical basis for evolutionary pressures that selected specific residues in FimH for fitness in the urinary tract and the structural basis for mannose binding by FimH. The conformational phase space of FimH in solution and population shifts spurred by positively selected residues further provides a framework for understanding the structural basis of allosteric coupling between inter-domain interactions and mannose binding. Finally, mannosides demonstrate considerable promise in the treatment of UTI, especially given their ability to target both the low-affinity **T** and high-affinity **R** conformations of FimH, both of which contribute to host-pathogen interactions in the urinary tract.

Galactoside inhibition of FmlH during chronic cystitis

In addition to type 1 pili, UPEC express the Fml/F9/Yde pilus tipped with the FmlH adhesin to bind the inflamed bladder epithelium that exists during chronic cystitis [15]. Chronic cystitis is characterized by long-term UPEC colonization of a chronically inflamed bladder epithelium that is hyperplastic and devoid of terminally differentiated superficial facet cells, likely due to inflammation induced mucosal wounding, increased urothelial turnover and unchecked bacterial replication [16]. FmlH specifically binds to Tn antigen which appears as part of the remodeled glycan profile of the mouse bladder epithelium during chronic cystitis. Consistent with this, FmlH provides a fitness advantage during chronic cystitis but not in acute cystitis due to its ability to bind to inflamed but not to naïve bladder tissue.

Furthermore, FmlH binds to TF epitopes found in naïve kidney tissue, thus promoting UPEC colonization of the kidneys [15]. As such, FmlH may play a critical role in the development of pyelonephritis, which contributes to high rates of hospitalization, severe kidney morbidities, and mortality [17]. Interestingly, FmlH is upregulated in urines directly isolated from patients with UTI compared to expression during *in vitro* growth in media or normal urine, suggesting a host condition-specific induction of Fml pili [18]. Thus, FmlH represents a promising target for the development of anti-adhesive compounds for use in the treatment of UTI and in particular pyelonephritis.

Structure-guided drug design directed the development of high-affinity galactoside inhibitors of FmlH. Initial biochemical screens led to the identification of o-nitrophenyl β-galactoside (ONPG) as an early lead FmlH inhibitor. A FmlH-ONPG co-crystal structure then facilitated *in silico* virtual screening of all known galactose analogs against FmlH, providing structural insights for the rational design of novel galactosides specific for FmlH. ELISA-based

screening and iterative rounds of structure-guided drug design led to the discovery of numerous phenyl and biphenyl galactoside antagonists of FmlH, including our lead compound, a biphenyl N-acetyl-β-galactosamine **29β-NAc**, which exhibits a K_D of ~90 nM, a ~7,800-fold enhancement in binding affinity relative to D-galactose. The design strategy for **29β-NAc** accounted for all the unique structural features of the binding pocket of FmlH, including: (i) the wide and solvent-exposed nature of the binding pocket; (ii) the considerable distance of Loop 1 from Loop 2 and Loop 3; (iii) the capping of the binding pocket by residue Y46; (iv) the positively charged side chain of residue R142 on the edge of the binding pocket; and (v) the capacity to accommodate the N-acetyl group of N-acetyl galactosamine. Necessarily, FmlH galactoside design differed in approach to that of FimH mannoside design due to differences in receptor specificity, stereoisomerism at the anomeric carbon, binding affinity, binding pocket architecture, and accessory side chains in the vicinity of the binding pocket (including orientation of tyrosine residues). Yet, similarities in the properties of these compounds, including the biphenyl scaffold in the aglycone group and hydrogen bond interactions mediated through functional groups on the B ring, may be a reflection of universal features of adhesin-carbohydrate interactions and a common mechanism by which aromatic aglycones mimic those interactions.

Our *in vivo* work clearly demonstrates the translational impact of galactosides as treatments for UTI. Transurethral delivery of **29β-NAc** resulted in significant reduction of bacterial titers in the bladder and especially the kidney. Furthermore, a combinatorial approach in which both **29β-NAc** and mannoside 4Z269 were administered resulted in synergistic efficacy in ridding the kidney of UPEC, suggesting that FimH and FmlH occupy niches within the urinary tract simultaneously during chronic UTI. Future studies aim to further this progress in galactoside development through optimization of oral bioavailability, pharmacokinetics, and metabolic stability, which will provide

lead candidates that can be tested individually or together with mannosides for treatment or prevention of human cystitis and pyelonephritis in the clinic.

Future directions

FimH conformations in biofilm formation and maturation

Type 1 pili and FimH play a critical role in the formation and maturation of biofilms on abiotic surfaces [19]. Deletion of type 1 pili results in an attenuated ability of *E. coli* to mediate microcolony formation. Further, mannose or α-methyl-D-mannose can inhibit biofilm formation on diverse abiotic substrates, including polyvinyl chloride, polycarbonate, polystyrene, and borosilicate glass. A fascinating corollary exists between the propensity of *E. coli* to form biofilms and the mannose binding affinity and/or conformational state of FimH [14]. Yet, the mechanism by which FimH adsorbs to a surface and promotes biofilm formation among *E. coli* remains elusive. I hypothesize that the high-affinity **R** state is dually critical for biofilm formation: (i) for abiotic surface attachment with a hydrophobic surface of FimH exposed in the **R** state during biofilm initiation and (ii) for high-affinity mannose binding during biofilm expansion/maturation.

To address this hypothesis, a two-pronged approach can be pursued. First, biofilm and hemagglutination screening can be performed on a FimH mutant library generated by error-prone PCR to identify mutation(s) that preserve the structure of FimH in the **R** state but prevents mannose binding. In addition, molecular dynamics simulations of FimH adsorption to a graphene sheet can reveal specific residues implicated in surface interactions that could then spur site-directed mutagenesis. Together, these studies could decouple surface adhesion from biofilm maturation and reveal the atomic basis for FimH-mediated biofilm formation, providing key molecular details for understanding and combatting biofilms that form on catheters and lead to serious nosocomial

infections.

In vivo colonization resistance to high-affinity R state

UTI89 engineered to express A27V/V163A FimH, which predominantly occupies the high-affinity **R** state, lacks the ability to effectively colonize the urinary tract and establish UTI. I hypothesize that the high mannose binding affinity of A27V/V163A is directly responsible for its colonization defect. To uncover the mechanism which mediates this colonization resistance, we first examined whether Tamm-Horsfall protein (THP) is responsible for the pathogenesis defect of A27V/V163A. In a preliminary study, THP knock-out mice (of the 129S1 background) acquired from Dr. Satish Kumar were inoculated with 10^7 CFU/ml of UTI89 expressing either WT FimH or A27V/V163A FimH. Bacterial titer analysis indicates that A27V/V163A remains attenuated relative to WT FimH in this mouse model, suggesting that THP does not mediate the pathogenesis defect of A27V/V163A (**Fig. 1**). Further work will explore whether other decoy receptors on the surface of superficial facet cells or in urine, such as secretory IgA or soluble oligomannose, may be involved in mediating this pathogenesis defect. Mouse and human urine filtration experiments and outer membrane preparations of superficial facet cells derived from *in vitro* bladder spheroid models will provide the cleanest approaches to pulling down or identifying receptors that may be implicated in the colonization resistance property of the bladder habitat.

SfaH structure-function relationship and biological tropism of S pili

E. coli expressing S pili have been implicated in the onset or maintenance of neonatal meningitis and UTI [20-23]. However, the role of S pili in infectious disease has not be clearly delineated. Moreover, the mechanism underlying the sialic acid specificity of S pili is not clearly understood. Several studies implicate different components of S pili with adhesive functions, with debates regarding the true adhesin [24, 25]. Some experimental evidence suggests that SfaG, a

pilus tip adapter, is responsible for recognition of sialic acid receptors while bioinformatics suggests that the SfaH adhesin would exert the function of sialic acid binding. Through a panel of glycan arrays, I have shown that SfaH in a SfaCH complex does not bind sialic acid residues or any other tested glycan structure (**Fig. 2A-D**). Yet, SfaH does possess some adhesive ability, as SfaH in a SfaCH complex strongly binds secretory IgA and laminin in a protease-sensitive manner, suggesting that SfaH may play a role in binding uncharacterized protein tertiary structures or motifs (**Fig. 3**). A crystal structure of $SfaH_{LD}$ indicates that it adopts a similar fold to FimH (**Fig. 4A**), resembling $FimH_{LD}$ in the **T** conformation (RMSD=1.7 Å) more than $FimH_{LD}$ in the **R** conformation (RMSD=3.2 Å) (**Fig. 4B-C**). However, the side chain packing in the canonical binding pocket appears to restrict space necessary for carbohydrates to interact with SfaH, likely due to steric occlusion mediated by residues S14 and E148 (**Fig. 4D**). Thus, while SfaH closely resembles an adhesin homologous to FimH and FmlH, SfaH lacks the ability to bind glycans due to the geometry of residues within the canonical adhesin binding pocket that preclude the entrance of carbohydrate moieties. More work is required to demonstrate that SfaG is the true sialic acid binder within S pili through purification of SfaG and screening in glycan arrays and glycoprotein binding assays. Thereafter, the binding functions of SfaG and SfaH should be evaluated using immunohistochemistry methods against tissues at implicated host sites, such as meninges, choroid plexus, kidneys, and bladder, and of different host ages, including neonates and adults, to determine their role in pathogenesis.

Conformational ensembles in CUP adhesins as a universal feature of bacterial adhesion

CUP pili typically present tip-localized two-domain adhesins to mediate attachment to a particular host or tissue. These two-domain adhesins are comprised of a lectin domain, which often recognizes a specific carbohydrate receptor, and a pilin domain, which anchors the adhesin to the

distal end of the pilus. FimH has been demonstrated to adopt diverse structures within a two-state conformational landscape, a property that may represent a universal feature of two-domain CUP adhesins. Ion mobility mass spectrometry analysis of FmlG_{Nte}H and SfaG_{Nte}H complexes reveal that these two homologs of FimH do in fact adopt a two-state conformational equilibrium similar in distribution to FimG_{Nte}H, lending support to the supposition that conformational phase space is a natural property of two-domain CUP adhesins (**Fig. 5**). Further studies will explore whether these two-state conformational equilibria in FmlH and SfaH impact ligand binding affinity and pathogenesis, as is observed in FimH. Future studies will also aim to investigate FimH homologs and related two-domain CUP adhesins among Gram-negative bacteria (**Fig. 6**) to characterize diverse conformational landscapes and receptor specificities to understand mechanisms underlying *E. coli* adhesion and tissue and host tropisms.

Closing remarks

Studies rooted in structural biology have uncovered fundamental physicochemical principles of the chaperone-usher pathway. Protein crystallography, in particular, has provided views of key macromolecular interactions at the host-pathogen interface at the atomic level. However, these static structures only serve as signposts for the dynamic processes that underlie CUP pilus function. This thesis builds on previous structural studies with a focus on a biophysical understanding of the allosteric and dynamic molecular details of CUP systems using a broad range of techniques including bacterial genetics, bioinformatics, computational biology, biochemistry, biophysics, structural biology, and cell biology. Understanding how protein-protein interactions and ligand binding can regulate a dynamic conformational equilibrium of the receptor binding domain of bacterial adhesins is changing the established view of UTI pathogenesis. Elucidating

structure-function relationships in distinct adhesins has helped build a molecular platform for interrogation and identification of host selective pressures that have acted to diversify ligand specificity and has contributed to the design of novel anti-virulence compounds for the treatment of UTI. Together, this work has unraveled the dynamics, allostery, and atomic determinants that govern CUP function and specificity, which is critical for understanding virulence in Gram-negative pathogens and is spawning new ways of thinking about drug development for the treatment of infectious diseases.

References

1. Eto, D.S., et al., *Integrin-mediated host cell invasion by type 1-piliated uropathogenic Escherichia coli.* PLoS Pathog, 2007. **3**(7): p. e100.
2. Wu, X.R., T.T. Sun, and J.J. Medina, *In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: relation to urinary tract infections.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9630-5.
3. 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**(Pt 22): p. 4095-103.
4. Mulvey, M.A., et al., *Induction and evasion of host defenses by type 1-piliated uropathogenic Escherichia coli.* Science, 1998. **282**(5393): p. 1494-7.
5. Martinez, J.J., et al., *Type 1 pilus-mediated bacterial invasion of bladder epithelial cells.* EMBO J, 2000. **19**(12): p. 2803-12.
6. Justice, S.S., et al., *Differentiation and developmental pathways of uropathogenic Escherichia coli in urinary tract pathogenesis.* Proc Natl Acad Sci U S A, 2004. **101**(5): p. 1333-8.
7. Anderson, G.G., et al., *Intracellular bacterial biofilm-like pods in urinary tract infections.* Science, 2003. **301**(5629): p. 105-7.
8. Jones, C.H., et al., *FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae.* Proc Natl Acad Sci U S A, 1995. **92**(6): p. 2081-5.
9. Le Trong, I., et al., *Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting.* Cell, 2010. **141**(4): p. 645-55.
10. Phan, G., et al., *Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate.* Nature, 2011. **474**(7349): p. 49-53.
11. Geibel, S., et al., *Structural and energetic basis of folded-protein transport by the FimD usher.* Nature, 2013. **496**(7444): p. 243-6.
12. Choudhury, D., et al., *X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli.* Science, 1999. **285**(5430): p. 1061-6.
13. Aprikian, P., et al., *The bacterial fimbrial tip acts as a mechanical force sensor.* PLoS Biol, 2011. **9**(5): p. e1000617.
14. Chen, S.L., 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**(52): p. 22439-44.
15. Conover, M.S., et al., *Inflammation-Induced Adhesin-Receptor Interaction Provides a Fitness Advantage to Uropathogenic E. coli during Chronic Infection.* Cell Host Microbe, 2016. **20**(4): p. 482-492.
16. Hannan, T.J., et al., *Early severe inflammatory responses to uropathogenic E. coli predispose to chronic and recurrent urinary tract infection.* PLoS Pathog, 2010. **6**(8): p. e1001042.
17. Ramakrishnan, K. and D.C. Scheid, *Diagnosis and management of acute pyelonephritis in adults.* Am Fam Physician, 2005. **71**(5): p. 933-42.
18. Subashchandrabose, S., et al., *Host-specific induction of Escherichia coli fitness genes during human urinary tract infection.* Proc Natl Acad Sci U S A, 2014. **111**(51): p. 18327-32.

19. Pratt, L.A. and R. Kolter, *Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili.* Mol Microbiol, 1998. **30**(2): p. 285-93.
20. Parkkinen, J., et al., *Identification of the O-linked sialyloligosaccharides of glycophorin A as the erythrocyte receptors for S-fimbriated Escherichia coli.* Infect Immun, 1986. **54**(1): p. 37-42.
21. Parkkinen, J., et al., *Binding sites in the rat brain for Escherichia coli S fimbriae associated with neonatal meningitis.* J Clin Invest, 1988. **81**(3): p. 860-5.
22. Ott, M., et al., *Analysis of the genetic determinants coding for the S-fimbrial adhesin (sfa) in different Escherichia coli strains causing meningitis or urinary tract infections.* Infect Immun, 1986. **54**(3): p. 646-53.
23. Korhonen, T.K., et al., *Binding of Escherichia coli S fimbriae to human kidney epithelium.* Infect Immun, 1986. **54**(2): p. 322-7.
24. Schmoll, T., et al., *Analysis of genes coding for the sialic acid-binding adhesin and two other minor fimbrial subunits of the S-fimbrial adhesin determinant of Escherichia coli.* Mol Microbiol, 1989. **3**(12): p. 1735-44.
25. Prasadarao, N.V., et al., *Adhesion of S-fimbriated Escherichia coli to brain glycolipids mediated by sfaA gene-encoded protein of S-fimbriae.* J Biol Chem, 1993. **268**(14): p. 10356-63.

Figures

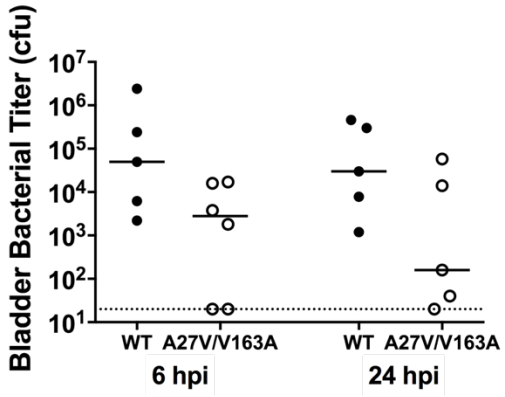


Figure 1. Role of THP in mediating colonization resistance to FimH variants. Bacterial titers in the bladders of THP knockout mice infected with 10^7 CFU/ml UTI89 expressing either WT FimH or A27V/V163A FimH at 6 or 24 hpi.

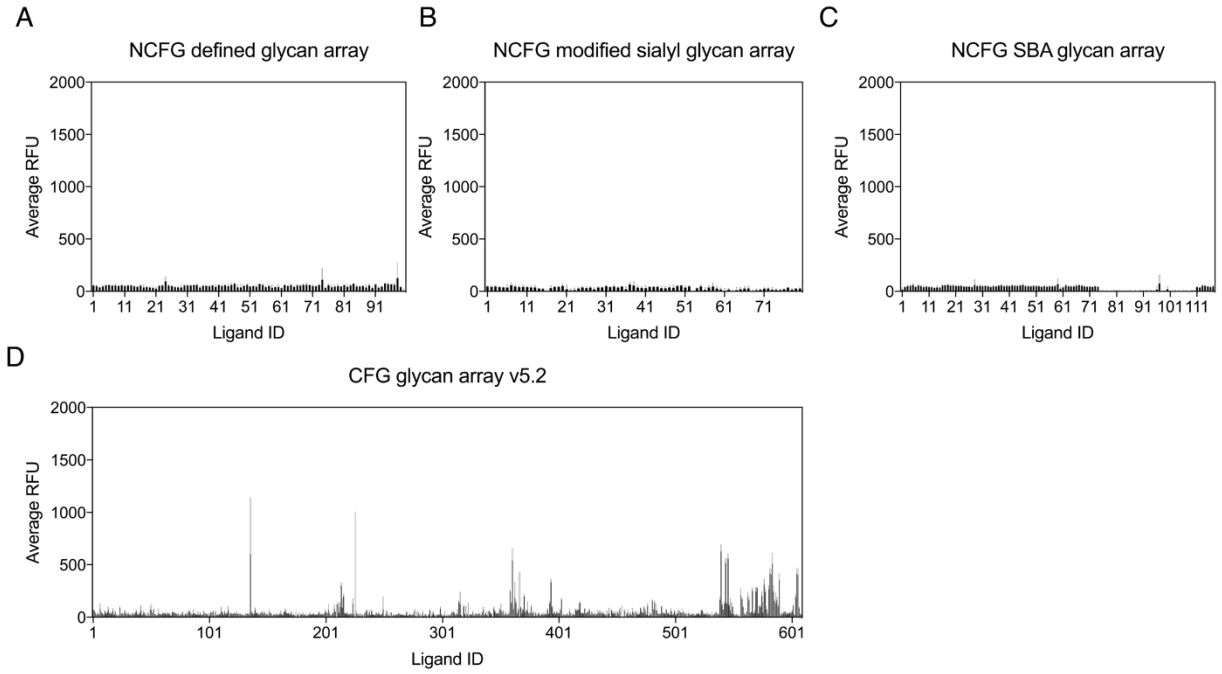


Figure 2. Glycan array screens for determination of SfaH specificity. Biotinylated SfaCH complex (100-200 μg) was incubated with various glycan arrays, which include (**A**) the National Center for Functional Glycomics (NCFG) defined glycan array, (**B**) the NCFG modified sialyl glycan array, (**C**) the NCFG SBA glycan array, and (**D**) the Center for Functional glycomics (CFG) glycan array v5.2. Bound SfaCH was then quantitated after incubation of streptavidin labeled with Alexa Fluor 488. A positive glycan binding result is typically associated with an RFU value greater than 10,000.

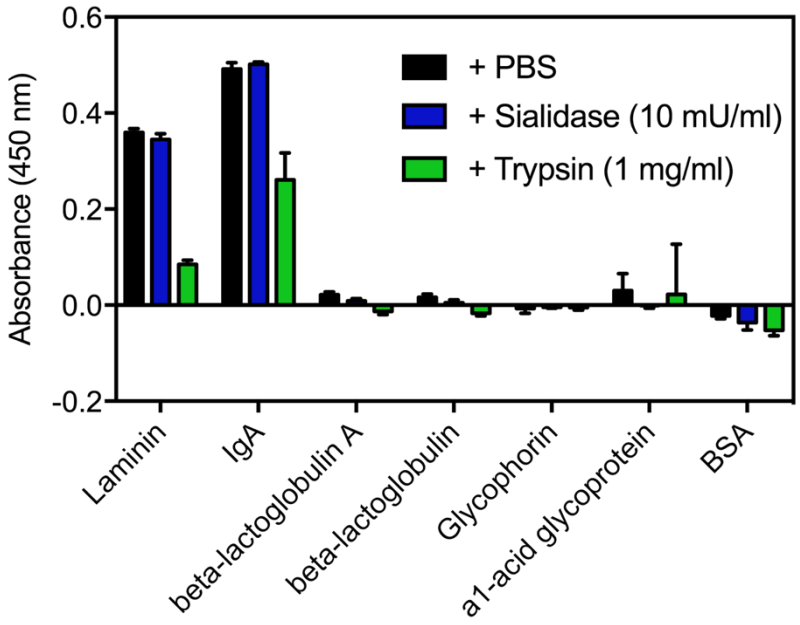


Figure 3. Glycoprotein binding by SfaH. Glycoproteins (1 μg) coated on an ELISA plate, as indicated on the x-axis, were pre-treated either with PBS, sialidase, or trypsin for 1 hour at 37 ˚C. Thereafter, biotinylated SfaCH complex (5 μg) was incubated with glycoproteins for 4 hours at 4 ˚C, and binding was detected by incubation with streptavidin-HRP conjugate for 1 hour at 4 ˚C.

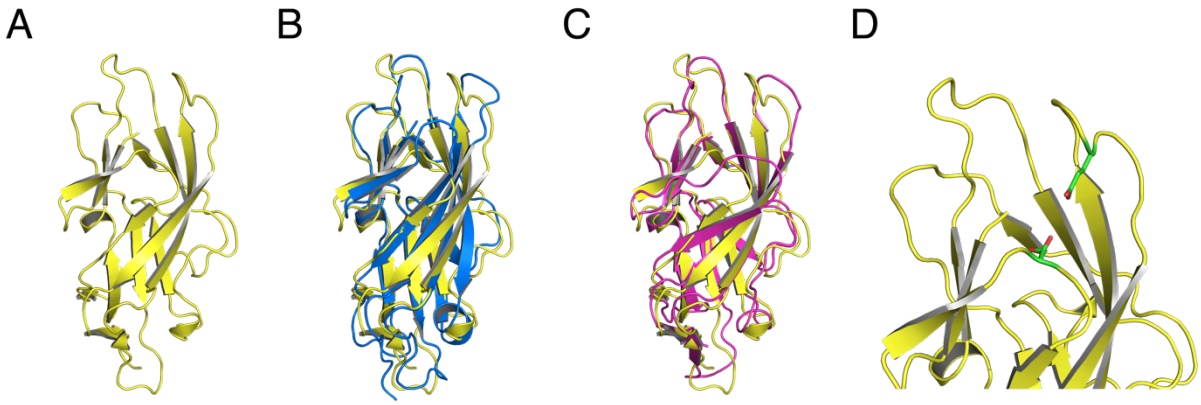


Figure 4. Structural analysis of SfaH. (**A**) Crystal structure of $SfaH_{LD}$ solved to a resolution of 1.6 Å (R_{work}=16.1%; R_{free}=19.4%). (**B**) Structural overlay of $SfaH_{LD}$ (yellow) and $FimH_{LD}$ in the **R** conformation (marine). (**C**) Structural overlay of $SfaH_{LD}$ (yellow) and $FimH_{LD}$ in the **T** conformation (magenta). (**D**) Binding pocket of $SfaH_{LD}$ indicating the proximity between and steric occlusion mediated by residues S14 and E148 (green sticks).

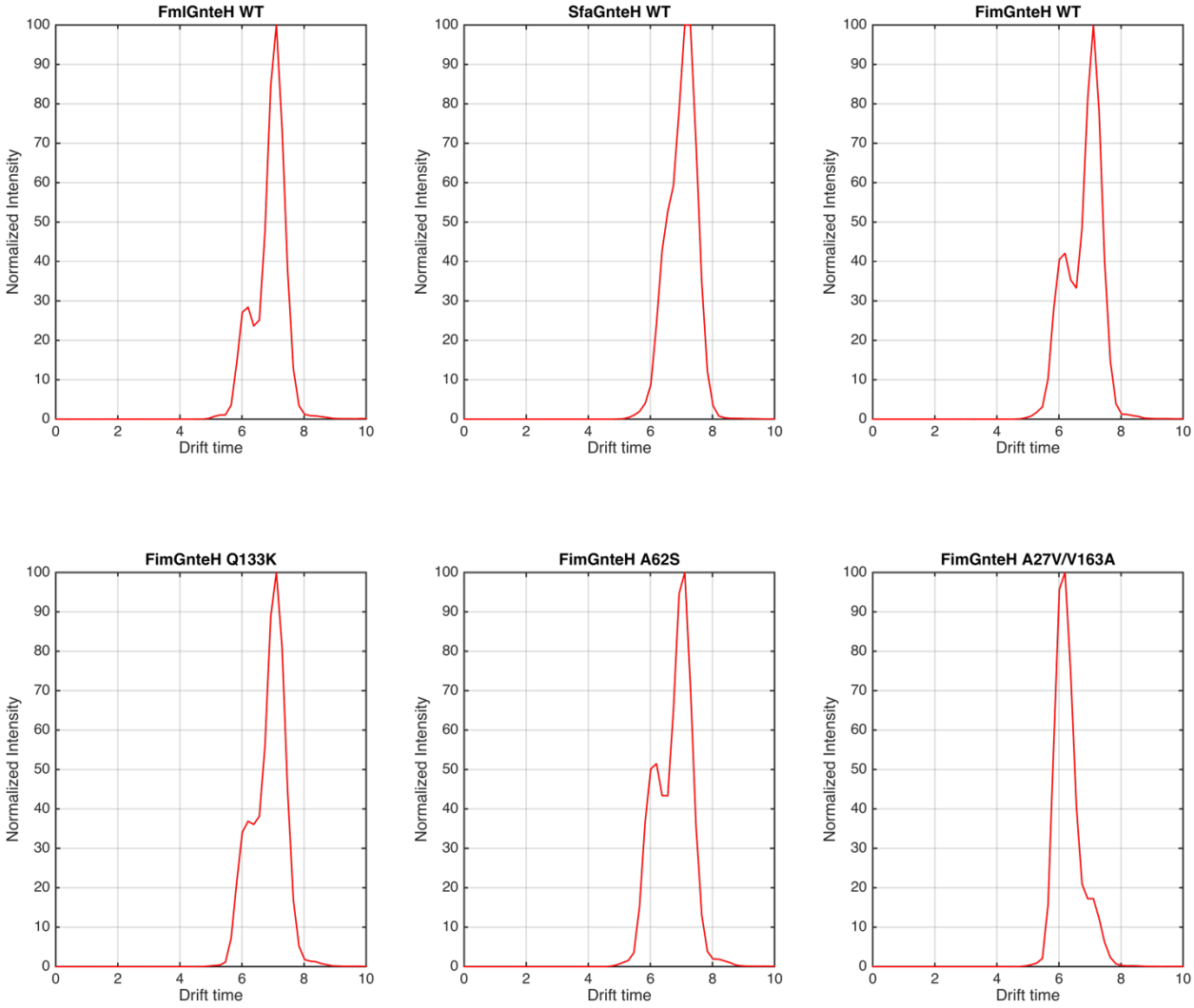


Figure 5. Conformational ensembles of fimbrial adhesins. Ion mobility mass spectrometry reveals drift time distributions of adhesins in tip-like settings ($G_{Nte}H$ complexes), including Fml$G_{Nte}H$ WT, Sfa$G_{Nte}H$ WT, Fim$G_{Nte}H$ WT, Fim$G_{Nte}H$ Q133K, Fim$G_{Nte}H$ A62S, and Fim$G_{Nte}H$ A27V/V163A. These distributions indicate two distinct conformations centered around t=6 ms and t=7 ms. Fim$G_{Nte}H$ A27V/V163A, as expected, predominantly adopts a single conformation characterized by a drift time of 6 ms while all other tested adhesins display an equilibrium of two conformations that favors the conformation corresponding to a drift time of 7 ms.

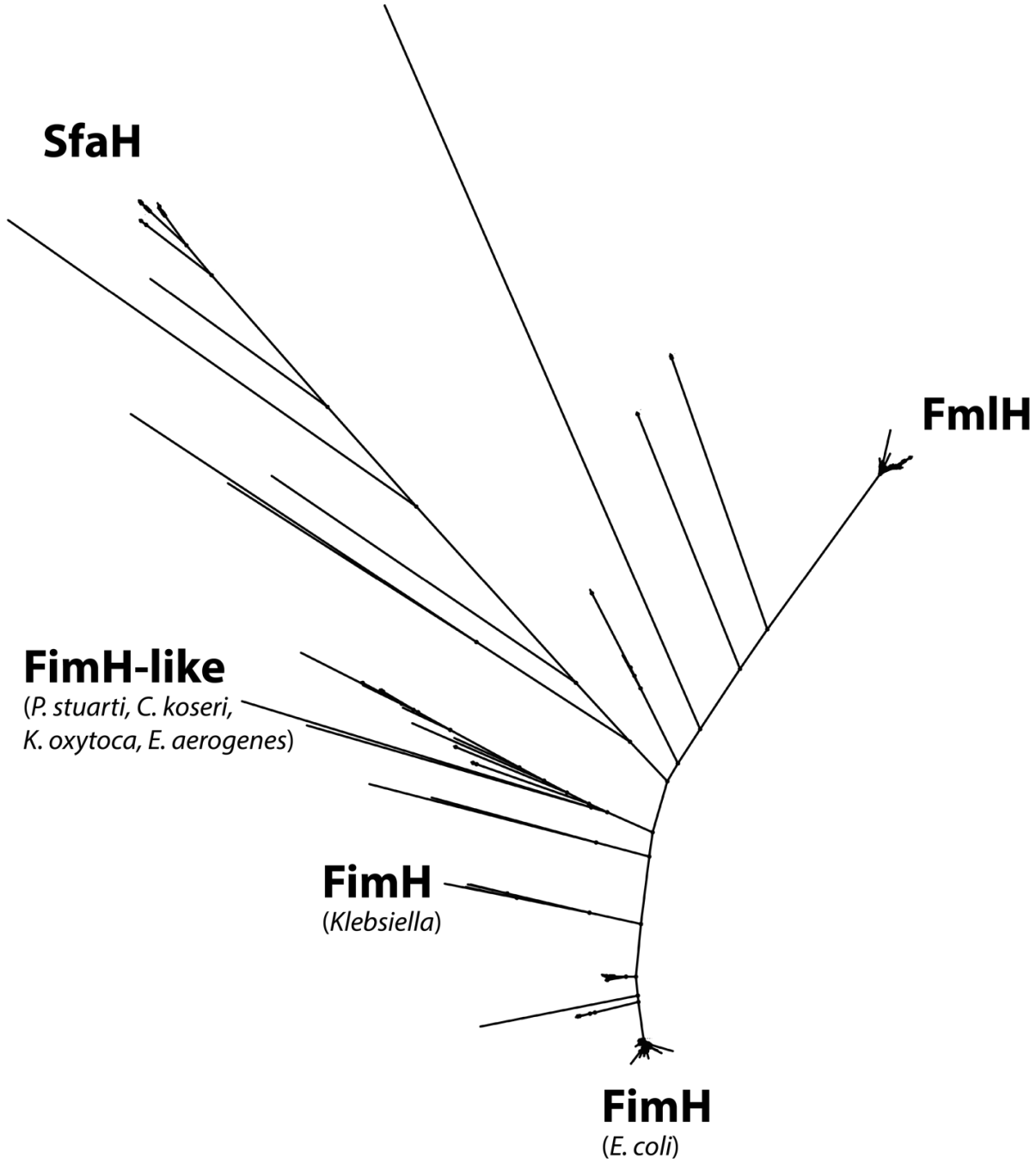


Figure 6. Diversification of CUP adhesins. Maximum likelihood phenogram depicting the evolutionary relationship among FimH, FmlH, and SfaH homologs in Gram-negative bacteria. Distinct groupings of branches are annotated according to relatedness to the three aforementioned adhesins.