THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 13, pp. 8409–8419, March 27, 2015

Structural and Functional Insight into the Carbohydrate Receptor Binding of F4 Fimbriae-producing Enterotoxigenic *Escherichia coli*[*]

Received for publication, October 14, 2014, and in revised form, January 22, 2015 Published, JBC Papers in Press, January 28, 2015, DOI 10.1074/jbc.M114.618595

Kristof Moonens[‡§1], Imke Van den Broeck[‡§], Maia De Kerpel[‡§], Francine Deboeck[¶], Hanne Raymaekers[‡§], Han Remaut[‡§2], and Henri De Greve[‡§3]

From the [‡]Structural and Molecular Microbiology, VIB Structural Biology Research Center, 1050 Brussels, the [§]Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, and the [¶]Viral Genetics Laboratory, Vrije Universiteit Brussel, 1050 Brussels, Belgium

Background: F4 fimbriae produced by enterotoxigenic *Escherichia coli* mediate attachment to eukaryotic host receptors.
Results: The structure of lactose bound to the F4 fimbrial adhesin $FaeG_{ad}$ was elucidated.
Conclusion: Lactose interacts at a subdomain grafted on the $FaeG_{ad}$ core domain.
Significance: The co-complex structure explains the finely tuned receptor specificity of $F4_{ad}$ fimbriae; additionally, the carbohydrate binding site differs among FaeG variants.

Enterotoxigenic *Escherichia coli* (ETEC) strains are important causes of intestinal disease in humans and lead to severe production losses in animal farming. A range of fimbrial adhesins in ETEC strains determines host and tissue tropism. ETEC strains expressing F4 fimbriae are associated with neonatal and post-weaning diarrhea in piglets. Three naturally occurring variants of F4 fimbriae ($F4_{ab}$, $F4_{ac}$, and $F4_{ad}$) exist that differ in the primary sequence of their major adhesive subunit FaeG, and each features a related yet distinct receptor binding profile. Here the x-ray structure of $FaeG_{ad}$ bound to lactose provides the first structural insight into the receptor specificity and mode of binding by the poly-adhesive F4 fimbriae. A small D′-D″-α1-α2 subdomain grafted on the immunoglobulin-like core of FaeG hosts the carbohydrate binding site. Two short amino acid stretches Phe^{150}–Glu^{152} and Val^{166}–Glu^{170} of $FaeG_{ad}$ bind the terminal galactose in the lactosyl unit and provide affinity and specificity to the interaction. A hemagglutination-based assay with *E. coli* expressing mutant $F4_{ad}$ fimbriae confirmed the elucidated co-complex structure. Interestingly, the crucial D′-α1 loop that borders the $FaeG_{ad}$ binding site adopts a different conformation in the two other FaeG variants and hints at a heterogeneous binding pocket among the FaeG serotypes.

Enterotoxigenic *Escherichia coli* (ETEC)[4] are important causes of intestinal disease in humans and animals (1, 2). Infections with ETEC strains affect hundreds of millions of persons worldwide annually, mainly travelers and young children in developing countries (3). ETEC that are taken up by the consumption of contaminated water or food colonize the small intestine where they secrete heat-labile and/or heat-stable enterotoxins, thereby causing diarrhea (4). As an initial and crucial step in pathogenesis, ETEC express fimbrial colonization factors that steer species-specific adherence and colonization of the intestinal tract. Human ETEC possess over 20 distinct colonization factors (5). In porcine ETEC, F4 fimbriae (previously termed K88 fimbriae) form the most common colonization factor associated with intestinal infection of newborn, suckling, and newly weaned piglets, leading to significant death and morbidity (6).

Most fimbriae belong to a class of cell surface organelles that are assembled by the conserved chaperone/usher pathway, named after two essential components: a periplasmic chaperone protein and outer membrane pilus assembly platform, termed usher (7). The chaperone stabilizes and delivers pilus subunits to the usher, where subunits are added to the base of a non-covalent polymer by a fold complementation mechanism (8, 9). F4 fimbriae are encoded by the *fae* operon, which comprises genes coding for the regulatory proteins FaeA and FaeB (10), the distal tip protein FaeC (11), the usher FaeD (12), the periplasmic chaperone FaeE (13), the minor fimbrial shaft subunits FaeF and FaeH (14), the major subunit FaeG (15), and the minor subunits FaeI and FaeJ, whose roles remained obscure until now. FaeG forms the major structural component of the F4 fimbriae and also incorporates the adhesive properties of the fibers. Therefore, contrary to the most well studied chaperone/usher pili, where receptor binding is determined by a single two-domain adhesin incorporated at the fimbrial tip, binding characteristics of F4 fimbriae are part of the major subunit and are displayed throughout the length of the fiber (15, 16). Three naturally occurring serological variants of F4 fimbriae ($F4_{ab}$, $F4_{ac}$, and $F4_{ad}$) each feature a related but yet different binding

* This research was granted by Fonds voor Wetenschappelijk Onderzoek (FWO) Project G030411N.
The atomic coordinates and structure factors (codes 4WE2 and 4WEI) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] A doctoral fellow of the FWO – Vlaanderen.
[2] Supported by a VIB Young principal investigator project grant and the Odysseus program of the FWO-Vlaanderen.
[3] To whom correspondence should be addressed: Structural Biology Research Center, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium. Tel.: 32-2-629-18-44; Fax: 32-2-629-19-63; E-mail: hdegreve@vub.ac.be.
[4] The abbreviations used are: ETEC, enterotoxigenic *E. coli*; IGLad, intestinal neutral glycosphingolipid; Cer, ceramide; TDA, two-domain tip adhesin(s).

and hemagglutination profile (17, 18). Differences in binding specificity were attributed to each of the three F4 fimbriae variants ($F4_{ab}$, $F4_{ac}$, and $F4_{ad}$), and could be interchanged by replacement of a residue stretch in the FaeG protein (19). Based on the receptor profiles in the brush borders, six different pig phenotypes can be distinguished that differ in susceptibility to specific FaeG variants (20). An extensive literature search describes studies seeking to unravel the F4 receptor specificities. Both specific piglet intestinal glycoproteins and glycolipids have been identified as F4 receptors, and a three-receptor model was proposed, which was by later observations adapted and changed to include five putative receptors (21–24). A subset of brush borders glycoproteins ranging in ranging in molecular mass from 45 to 70 kDa was demonstrated to interact with all three types of F4 fimbriae (25). Two intestinal mucin-type sialoglycoproteins (IMTGP-1 and IMTGP-2) were identified as receptors for $F4_{ab}$ and $F4_{ac}$ fimbriae (22, 26, 27), and later a β-linked galactose was shown to constitute an important component in their recognition by $F4_{ac}$ fimbriae (28). Intestinal transferrin (GP74) was shown to be a F4ab-specific receptor (29), with exoglycosidase treatment pointing toward the pivotal role of GlcNAc residues in the core of the *N*-glycan chain on GP74. A $F4_{ad}$-specific receptor was identified as an intestinal neutral glycosphingolipid (IGLad) (24). Grange *et al.* (24) demonstrated terminal β-linked galactose to be an essential component of IGLad, and furthermore their results strongly indicated neolactotetraosylceramide (Galβ1-4GlcNAcβ1-3 Galβ1-4Glcβ1Cer) was the IGLad receptor. The important role for galactose and/or *N*-acetyl galactosamine residue in the F4 receptor structure was highlighted in different studies; most often these residues were present at the non-reducing end in β-linkage (24, 28, 30–36). In other studies, the interaction of all F4 variants with different glycosphingolipids was demonstrated, such as lactosylceramide, gangliotriaosylceramide, gangliotetraosylceramide, globotriaosylceramide, lactotetraosylceramide, and lactotetraosylceramide (24, 30, 32, 37). A more recent study investigated the glycosphingolipid recognition by the different F4 variants and found that the $F4_{ab}$ and $F4_{ac}$ variants showed more similarities in their glycosphingolipid recognition patterns when compared with the $F4_{ad}$ variant (30). For example, $F4_{ab}$ and $F4_{ac}$ fimbriae interacted with both sulfatide and galactosylceramide, whereas $F4_{ad}$ fimbriae did not (30). In contrast, $F4_{ad}$ fimbriae bound to gangliotriaosyl- and gangliotretraosylceramide, unlike the ab or ac variants.

Previously, we determined the crystal structure of monomeric N-terminally deleted, donor strand complemented FaeG ($FaeG_{ntd/dsc}$) of variant ad (38). Equivalent constructs for the ab, ac, and ad variants are employed in this study in co-crystallization trials with the above described carbohydrate receptors, resulting in the first structural insight into ligand binding by the poly-adhesive F4 fimbriae. The co-complex structure between $FaeG_{ad}$ and lactose allowed us to explain the finely tuned receptor specificity of $F4_{ad}$ fimbriae. We also obtained the apo structure of $FaeG_{ntd/dsc}$ variant ab, and together these results hinted at the non-conserved nature of the lactose binding pocket on the FaeG surface as the crucial D′-α1 loop significantly differs among the FaeG variants.

EXPERIMENTAL PROCEDURES

Expression and Purification of the $FaeG_{ntd/dsc}$ Variants—The different constructs of N-terminally deleted, donor strand complemented FaeG ($FaeG_{ntd/dsc}$) were expressed and purified as described previously (38). In short, *E. coli* C43 (DE3) cells containing a plasmid encoding the *$faeG_{ntd/dsc}$* construct were grown in LB medium while shaking at 37 °C and were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside for overnight expression at 37 °C upon reaching an A_{600} of 1. The cytoplasmic content was collected by breaking the cells using a Constant Systems cell cracker at 20,000 p.s.i. at 4 °C and removal of cell debris by centrifugation at 18,000 × *g* for 15 min. Initially, the cytoplasmic extract was subjected to a nickel affinity purification step, and $FaeG_{ntd/dsc}$ was eluted by increasing the concentration of imidazole. $FaeG_{ntd/dsc}$-containing fractions were pooled and dialyzed against 20 mM formic acid, pH 3.8, after which many contaminants readily precipitated, and these were removed by centrifugation. $FaeG_{ntd/dsc}$ was further purified using cation exchange chromatography (HiTrap SP FF, GE Healthcare Life Sciences) at pH 3.8 and eluted by increasing the concentration of NaCl to 1 M. Afterward, depending on the purity, $FaeG_{ntd/dsc}$-containing fractions were subjected to anion exchange chromatography (HiTrap Q FF, GE Healthcare Life Sciences) using 20 mM HEPES, pH 8, to equilibrate the column and load the protein, and using 20 mM HEPES, pH 8.0, 1 M NaCl for elution. As a final step, protein solutions were subjected to gel filtration chromatography (Sephacryl S-100 HiPrep 26/60 column, GE Healthcare Life Sciences). The runs were performed in 20 mM Tris, pH 8.0, 150 mM NaCl.

Crystallization, Data Collection, Structure Determination, and Refinement—Apo $FaeG_{ntd/dsc}$ variant ab (23 mg/ml) and the complex between $FaeG_{ntd/dsc}$ variant ad and lactose (20 mg/ml of protein, 25 mM lactose) were crystallized against a solution containing 18 mM Na/K-phosphate and 18% PEG 3350, respectively, and in condition B12 of the Morpheus crystallization screen (Molecular Dimensions; 100 mM Trizma (Tris base)/Bicine, pH 8.5, 0.03 M of each halogen (sodium fluoride, sodium bromide, sodium iodide), 37.5% v/v MPD_P1K_P3350 (2-methyl-2,4-pentanediol, PEG1000, PEG3350). The crystals were flash-cooled to 100 K in their crystallization solution supplemented with 15% glycerol for data collection. Single wavelength data were collected at 0.98 Å wavelength, and data were processed with XDS and XSCALE (39) and Truncate from the CCP4 suite (40). Data were phased by molecular replacement with Phaser (40) using coordinates of the previously determined $FaeG_{ntd/dsc}$ variant ad structure (Protein Data Bank (PDB) code 3GGH) (38). Additional electron density was observed for a lactose ligand. Galactose and glucose monosaccharides were obtained from the HIC-Up database (41), positioned into the electron density map, and restrained by the monomer library of Refmac. The models were further improved using the graphics program COOT (42), and maximum likelihood was refined using Refmac5.5 (40) against the native dataset. Crystal parameters and data processing statistics for all structures are summarized in Table 1.

Design of Mutant F4 Fimbriae—Full-length *$faeG_{ad}$* was amplified using primers K88-64/K88-65 that contain the attB1

and attB2 sites from total genomic DNA of the $F4_{ad}$-positive *E. coli* isolate C1360-79 and subsequently introduced into the Gateway entry vector pDONR221 (Invitrogen), yielding pENT39. The site-specific mutations in $faeG_{ad}$ were introduced by overlap PCR (primers are given in Table 2; the first PCR reactions were performed with inward primers SeqLA1 and SeqLB and mutation-specific outward primers, and the second overlap PCR reaction was done with primers K88-152 and K88-153) and subsequent reintroduction of the PCR fragments in the unique EcoRI and Eco47III sites of the pENT39 using the In-fusion Advantage PCR cloning kit (Clontech). Sequences of all mutants were confirmed by nucleotide sequencing.

The $F4_{ab}$ gene cluster comprising all genes (*faeCDEFGHI*) except the regulatory genes *faeA* and *faeB* and the minor subunit *faeJ* was amplified using primers K88-156 and K88-157 (containing the attB1 and attB2 sites) from total genomic DNA of the $F4_{ab}$-positive *E. coli* C1289-78 isolate and introduced into the pGV5558 plasmid, which is a derivative of the pTrc99A vector (43) containing the pDONR221 Gateway cassette in its unique SmaI site. Transformed *E. coli* DH5α were assayed for expression of the F4 fimbriae by colony blotting and agglutination with polyclonal antibodies directed against the FaeG major subunit. To delete the $faeG_{ab}$ gene, the plasmid pHD935 carrying the $F4_{ab}$ encoding gene cluster was digested with the restriction enzyme Eco47III, only present in the *faeF* and *faeG* genes. To delete the *faeG* gene and to restore the *faeF* gene, a PCR fragment obtained with primers K88-164 and K88-165 and containing the *faeF* missing part as well as a unique PmeI site was introduced with In-Fusion cloning (Clontech) in the Eco47III-digested pHD935. The resulting plasmid pHD936 carries the intact *faeCDEF-HI* genes and a deleted *faeG* gene. The Eco47III site in the *faeF* gene was mutated without changing the *faeF* reading frame, whereas the Eco47III site at the end of the *faeG* gene was unaffected. This pHD936 plasmid was used to introduce the wild type or mutant $faeG_{ad}$ sequences between the PmeI and EcoR47III sites with the In-Fusion technology. We therefore amplified the wild type or mutant $faeG_{ad}$ sequences using primers K88-168 and K88-169. The resulting plasmids are carrying a complete $F4_{ab}$ encoding gene cluster with mutant $faeG_{ad}$ or wild type $faeG_{ad}$ genes. The capability of the different clones to produce $F4_{ad}$ fimbriae was subsequently analyzed.

Preparation of Heat-sheared F4 Fimbriae—Bacterial strains were grown in 100 ml of LB medium supplemented with ampicillin (100 μg/ml) at 37 °C overnight with gentle shaking. Cells were washed twice by centrifugation for 25 min at 2000 rpm in a table-top centrifuge (Heraeus Megafuge 1.0R) and resuspended in 600 μl of PBS. The bacterial suspension was heated to

TABLE 1
Crystal parameters and data processing statistics

	apo $FaeG_{ntd/dsc}$ variant ab[a]	$FaeG_{ntd/dsc}$ variant ad with lactose[a]
Wavelength	0.98	0.98
Beamline	Soleil - PROXIMA 1	Soleil - PROXIMA 1
Space group	P 1 2_1 1	P 4_1 2_1 2
a, *b*, *c* (Å)	53.6, 49.0, 54.5	93.2, 93.2, 111.1 0
α, β, γ (°)	90, 109.2, 90	90, 90, 90
Resolution (Å)	49.0-1.5 (1.58-1.5)	47.67-2.3 (2.42-2.3)
No. of unique reflections	41977 (5568)	22359 (3192)
CC(1/2)	99.8 (89.9)	99.9 (75.4)
R_{meas}[b] (%)	5.1 (37.2)	8.0 (97.8)
Average $I/\sigma I$	13.8 (2.8)	19.9 (2.1)
Completeness (%)	97.9 (89.1)	99.9 (100)
Multiplicity	3.1 (2.4)	8.7 (8.8)
Wilson B-factor	14.8	35.3
R_{work}/R_{free}[c,d] (%)	17.5 (21.4)	19.2/23.0
Average B-factor (Å^2)	15.2	32.7
r.m.s.[e] deviations		
Bond lengths (Å)	0.021	0.018
Bond angles (°)	2.058	1.942
No. of atoms (except hydrogen)	2238	2052
Protein	1949	1921
Glycan	NA[f]	23
Water	289	108
Residues in allowed regions (% of Ramachandran plot)	99.6	99.6
PDB entry	4WE2	4WEI

[a] Statistics for outer resolution shell are given in parentheses.
[b] $R_{meas} = \Sigma h\ (nh/nh - 1)\ \Sigma l\ |Ihl - \langle Ih\rangle|/\Sigma h\ \Sigma l\ \langle Ih\rangle$, where nh = the number of observations for reflection h, Ihl = the intensity for observation l of reflection h, and $\langle Ih\rangle$ = the average intensity for reflection h.
[c] $R_{work} = \Sigma hkl\ \|F_{obs}| - |F_{calc}\|/\Sigma hkl\ |F_{obs}|$.
[d] R_{free} is defined as above but calculated for 5% of randomly chosen reflections that were excluded from the refinement.
[e] r.m.s., root mean square.
[f] NA, not applicable.

TABLE 2
Primers used to construct the FaeG mutants

Primer name	Primers (5′ → 3′)	Amino acid substitution
K88-64	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAGAAGGAGATATACCATGAAAAAGACTCTGATTGCACTGGCAAT	Wild type $faeG_{ad}$ gene
K88-65	GGGGACCACTTTGTACAAGAAAGCTGGGTATTAGTAATAAGTAATTGCTACGTTCAG	
K88-132	GAGCTGATGTCGCTTGCTGCCGAAGGGTCGCAC	$faeG_{ad}$ F150A
K88-133	GTGCGACCCTTCGGCAGCAAGCGACATCAGCTC	
K88-136	TTGCCGACGAATGTTAAGGCTTCTGAACTCAAGGGTGGGAG	$faeG_{ad}$ N168A
K88-137	CTCCCACCCTTGAGTTCAGAAGCCTTAACATTCGTCGGCAA	
K88-140	GACGAATGTTAAGAATTCTGCACTCAAGGGTGGGAGTG	$faeG_{ad}$ E170A
K88-141	CACTCCCACCCTTGAGTGCAGAATTCTTAACATTCGTC	
K88-152	GGTGTAGATGGAATTCCTCATATTGCATTTACTGACTAT	
K88-153	TACGTTCAGCGGAGCGCTCCACTGAGTGCTGGTAGTTACA	
K88-156	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTTATTATGCAGATCCGGGCAATGAT	
K88-157	GGGGACCACTTTGTACAAGAAAGCTGGGTAAGTCTTTATGTGGTGGTGGCCACAGTC	
K88-164	AAACCTGCAGATAGCACTGCAGGTGGAGGACTGAGGATTA	
K88-165	TACGTTCAGCGGAGCGCTGTTTAAACCCCTTGATATCCACAAATA	
K88-166	GGCACAGCAGCAGATTGAGTTAAGA	
K88-167	CCAAGATGCTGGTCTGTTATCAGA	
K88-168	G ATATCAAGGG GTTTAATTTT ATGAAAAAGACTCTGATTGCAC	
K88-169	TACGTTCAGCGGAGCGCTCCACTGAGTGCTGGTAGTTAC	
K88-168	G ATATCAAGGG GTTTAATTTTATGAAAAAGACTCTGATTGCAC	
K88-169	TACGTTCAGCGGAGCGCTCCACTGAGTGCTGGTAGTTAC	
SeqLA	CTCTCGCGTTAACGCTAGCATGGAT	
SeqLB	GTAACATCAGAGATTTTGAGACAC	

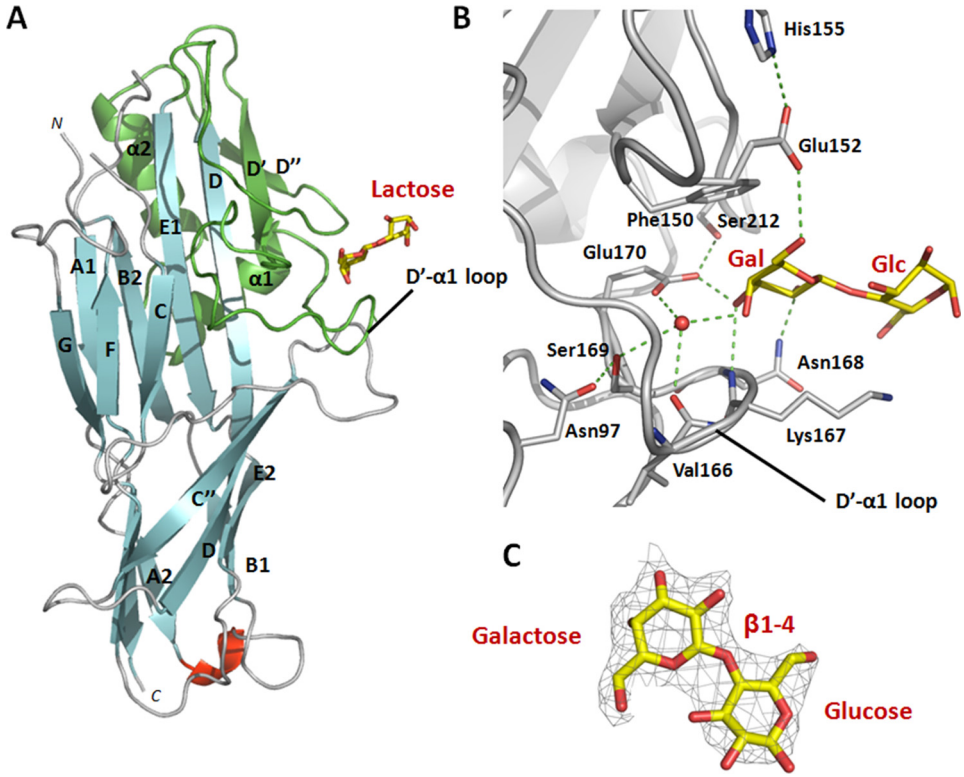

FIGURE 1. **Co-complex structure of the adhesive major subunit FaeG$_{ntd/dsc}$ variant ad with lactose.** *A*, graphic representation of FaeG$_{ntd/dsc}$ variant ad bound to lactose (depicted in stick model). The lactose ligand is bound on the side of the FaeG major subunit by the D′-D″-α1-α2 subdomain (colored in *green*) grafted on the Ig-like core between strands D and E1. In particular, residues located on the D′-α1 loop are involved in complex formation. Strands are colored *cyan*, loops are in *gray*, and a short unassigned α-helix is in *red*. *B*, close-up of the FaeG$_{ntd/dsc}$ variant ad binding site in complex with lactose. Two short amino acid stretches of the D′-α1 loop are involved in receptor binding. Only the galactose monosaccharide of the lactose ligand is interacting with binding site residues. Both side chains and main chain groups of residues on both stretches are coordinating a stabilizing water molecule (depicted as a *red sphere*). The protein backbone is in ribbon representation (*gray*), hydrogen bonds are depicted as *dashed green lines*, and ligand residues are displayed in stick representation with carbon and oxygen atoms in *yellow* and *red*, respectively. *C*, electron density map contoured at 1.3σ and displayed around the lactose ligand (depicted in stick model). Carbon and oxygen atoms are colored in *yellow* and *red*, respectively.

60 °C for 20 min to detach the F4 fimbriae from the cell surface. After centrifugation of the bacterial cells, the released fimbriae were present in the supernatant. Detection of the presence of FaeG$_{ad}$ in the F4$_{ad}$ fimbriae occurred by Western blot analysis with rabbit anti-FaeG polyclonal antiserum.

Hemagglutination Assay—Strains expressing mutant and wild type F4$_{ad}$ fimbriae were grown overnight in LB medium at low shaking speed to ensure that the F4$_{ad}$ fimbriae were not detached. Bacterial cells were washed two times in PBS. Finally, the cell pellet was dissolved in ice-cold PBS (50× less than initial volume), and the A_{600} was determined for each strain, after which they were diluted with PBS until an identical optical density was reached (A_{600} of 50). Guinea pig RBC were employed in the hemagglutination assay as F4$_{ad}$ fimbriae are able to specifically agglutinate these. For the washing step, 3 ml of RBC was added to 30 ml of PBS and subsequently centrifuged at 380 rcf in a fixed angle rotor during 10 min to spin down the RBC. The washing step was repeated at least three times and as often as needed until the supernatant was clear and free of lysed RBC. After the final washing step, 60 ml of PBS was added to obtain 5% RBC. A 2-fold dilution series of 25 μl bacteria (highest concentration A_{600} = 50) was added to a 96-well plate (BD Biosciences) on ice. To each well, 25 μl of PBS and 50 μl of washed RBC were added. The first well of each row was assigned as a negative control with only 50 μl of PBS and 50 μl of RBC added. The 96-well plate was incubated during at least 1 h, after which the hemagglutination pattern became visible and the titer was determined.

RESULTS

Structural Insight into Receptor Binding by the F4 Fimbrial Adhesin FaeG Variant ad—We co-crystallized a stable, self-complemented form of the FaeG subunit of variant ad (FaeG$_{ntd/dsc, ad}$) (38) with the disaccharide lactose (Galβ1-4Glc). The crystals diffracted to a resolution of 2.3 Å, and crystal parameters and data processing statistics are summarized in Table 1. Lactose interacts at the side of the FaeG fold at a shallow groove present on a small subdomain (Fig. 1 *A*). This binding domain is inserted between strands D and E of the immunoglobulin-like fold and comprises two short β-strands (D′ and D″) and two short helices (α1 and α2) (hence termed the D′-D″-α1-α2 subdomain). The D′-D″-α1-α2 subdomain was earlier speculated in being involved in receptor binding (44). FaeG$_{ntd/dsc, ad}$-lactose interactions are restricted to the terminal galactose residue and involve two short amino acid stretches, Phe^{150}–Glu^{152} and Val^{166}–Glu^{170}, located on the D′-α1 loop (Fig. 1, *A* and *B*). The sugar ring of galactose is sandwiched between the side chains of Phe^{150} and Lys^{167}. The presence of an aromatic residue facing the non-polar surface of

the galactopyranose is a recurrent feature of galactose-binding proteins (45).

Specificity for the galactose moiety is further provided by hydrogen bond formation with residues on both continuous stretches. The carboxyl group of Glu^{152} interacts with the C6 hydroxyl group. Glu^{152} is part of a dyad in which the side chain of His^{155} orients and stabilizes Glu^{152} in its interaction with the lactose ligand. A buried water molecule is stabilized and coordinated by the Ser^{169} hydroxyl, the Glu^{170} carboxyl, and the main chain carbonyl group of Val^{166}. Both side chains of Ser^{169} and Glu^{170} are in turn stabilized and oriented by Asn^{97} and Ser^{212}, respectively. These combined interactions position the buried water molecule to interact with the C4 hydroxyl group of galactose. In addition, the side chains of Asn^{168} and Glu^{170} form hydrogen bonds with the C2 and C3 hydroxyl groups, respectively, and the main chain amide group of Asn^{168} is involved in a hydrogen bond with the C3 hydroxyl group. In contrast to the extensive interaction network with the terminal galactose, the lactose Glc moiety does not appear to contribute direct binding interactions. Of note, although positioned in the vicinity of the glucose C6 hydroxyl, the side chain amine of Lys^{167} does not show clear electron density, indicating that it is not involved in a stable hydrogen bond interaction with the sugar.

An overlay of the apo- and lactose-bound $FaeG_{ntd/dsc}$ structures (0.75 Å root mean square deviation for 239 equivalent Cα atoms) shows the absence of conformational changes in the receptor site upon binding the lactose ligand. Similarly, the structures of FimH, PapG, FedF, and F17-G do not differ upon binding of their respective glycan ligand (46–49). For FimH, however, shear-enhanced conformational changes occur in the receptor binding site (50).

$FaeG_{ad}$ Is Responsible for $F4_{ad}$ Fimbriae-mediated Attachment—To determine the physiological relevance of the lactose binding site identified in the $FaeG_{ntd/dsc, ad}$ structure, a mutational analysis was performed. $FaeG_{ad}$ mutants with residues directly interacting with the carbohydrate receptor (Phe^{150}, Asn^{168}, and Glu^{170}) were alternatively substituted by alanine and analyzed for binding capacity. WT and mutant *$faeG_{ad}$* alleles were inserted in a cloned *fae* gene cluster comprising *faeC*, *faeD*, *faeE*, *faeF*, *faeH*, and *faeI* (*e.g.* lacking the regulatory genes *faeA* and *faeB*) under the control of the *lac* promoter. To verify production of WT and mutant F4 fimbriae, fimbrial material was heat-sheared from *E. coli* DH5α cells transformed with the respective plasmids ($pHD949^{WT}$, $pHD945^{F150A}$, $pHD944^{N168A}$, or $pHD948^{E170A}$) and analyzed with anti-F4 immunoblot. For all mutants, the levels of surface-exposed FaeG in mutant F4 fimbriae were found to be similar to wild type (Fig. 2 *B*), demonstrating that the introduced mutations do not affect the stability of FaeG and exposure of the F4 fimbriae on the cell surface. Subsequently, to assess the role of residues Phe^{150}, Asn^{168}, and Glu^{170} in $F4_{ad}$-mediated adherence, cultures expressing WT and mutant FaeG were used to determine the hemagglutination titer of guinea pig red blood (Fig. 2*A*). All of the mutations introduced into $FaeG_{ad}$ reduced binding when compared with wild type $FaeG_{ad}$ (Fig. 2*A*). Mutating Asn^{168} to alanine resulted in 2-fold difference in titer when compared with the wild type, indicating only a partial loss of lactose binding. In contrast, F150A and E170A almost fully abolished the interaction with carbohydrate receptors on the guinea pig red blood cells (8–16-fold difference when compared with wild type $F4_{ad}$). The hemagglutination assay thus validates the galactose binding site captured in the $FaeG_{ad}$-lactose co-complex crystal structure.

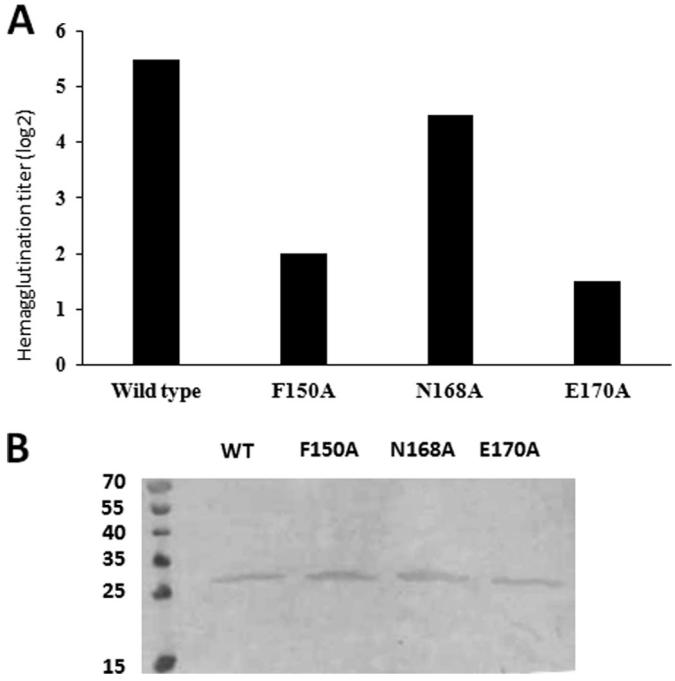

FIGURE 2. **Hemagglutination assay with *E. coli* expressing F4 fimbriae that incorporate either wild type or mutant $FaeG_{ad}$ confirms co-complex structure between $FaeG_{ad}$ and lactose.** *A*, summary of hemagglutination assay experiments of guinea pig red blood cells with *E. coli* expressing F4 fimbriae that incorporate either wild type or mutant $FaeG_{ad}$. The bar graph represents the titer to which extent the corresponding strain is still able to agglutinate the red blood cells. The higher the titer, the better the $F4_{ad}$ fimbriae recognize receptors present on the red blood cells. The Δ*faeG* strain did not demonstrate any hemagglutination capacity. *B*, Western blot of heat shock-detached $F4_{ad}$ wild type and mutant fimbriae to evaluate levels of $FaeG_{ad}$ expression. The intensity of the bands reflects the amount of $FaeG_{ad}$ incorporated in the $F4_{ad}$ fimbriae. The molecular masses of the marker proteins are indicated in kDa.

$F4_{ad}$ Fimbriae Exhibit a Finely Tuned Specificity—Based on our crystal structure, $FaeG_{ad}$ is found to bind the terminal galactose in the lactosyl epitope, but does not indicate an active involvement of the reducing end glucose in the binding interaction. Nevertheless, using TLC overlay assays, Coddens *et al.* (30) observed that $F4_{ad}$ fimbriated bacteria bound lactosylceramide but not galactosylceramide. To further address the role of glucose in $F4_{ad}$ receptor binding, we overlaid available experimentally determined conformers of galactosylceramide (4F7E, 3SDA, 3TNO, 3SDX, 1ZT4, 4EN3, 3HE6, and 3ARG) and lactosylceramide (3SDD, 1SX6, 2EUM, and 2EUD) found in the Protein Data Bank with the $FaeG_{ad}$ lactose structure by superimposition of the terminal galactose residue. In all analyzed galactosylceramide conformers, a kink is found that positions the ceramide group lateral to the FaeG surface. In the context of a lipid bilayer, the galactose epitope in these galactosylceramide conformers would be sterically hindered for $FaeG_{ad}$ binding (Fig. 3). Lactosylceramide on the other hand features an extended conformation that would not lead to a sterically hindered display of the terminal lactose when embedded in a bilayer (Fig. 3).

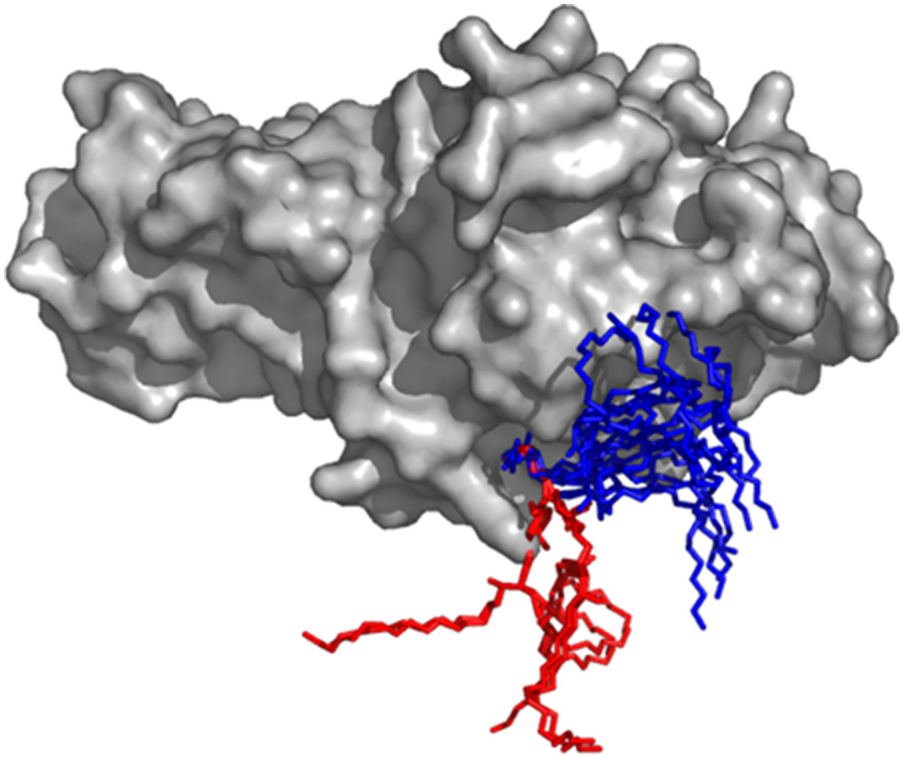

FIGURE 3. **Modeling of galactosylceramide and lactosylceramide into the $FaeG_{ad}$ binding pocket.** Galactosylceramide (4F7E, 3SDA, 3TNO, 3SDX, 1ZT4, 4EN3, 3HE6, and 3ARG) and lactosylceramide (3SDD, 1SX6, 2EUM, and 2EUD) structures as found in the Protein Data Bank were modeled into the $FaeG_{ad}$ binding pocket of the $FaeG_{ad}$-lactose structure by superimposing the galactose monomer of all ligands. Galactosylceramide and lactosylceramide are depicted as stick models and colored *blue* and *red*, respectively. $FaeG_{ad}$ is shown in surface representation and colored *gray*.

The $FaeG_{ad}$-lactose co-complex structure further reveals that substitutions on either the C3 or the C4 position of galactose would result in steric clashes with FaeG. Indeed, sulfatide or sulf-lactosylceramide both have a sulfate group on the C3 position and do not interact with $F4_{ad}$ fimbriae (30). The requirements of unsubstituted C3 or C4 positions also dictate that the galactose receptor residue involved in $FaeG_{ad}$ binding should be located at the non-reducing end and not internally in the glycan chain. Binding studies with chicken erythrocyte and reference glycosphingolipids identified glycans with an *N*-acetyl galactosamine at the non-reducing end to be accepted as $F4_{ad}$ fimbrial receptors (30). Based on the $FaeG_{ad}$-lactose structure, an *N*-acetyl group in the C2 position is predicted to come into close proximity with the side chain of Asn^{168} (1.8 Å distance with the carbon atom of the carbonyl group), although a slight reorientation of the latter would enable the preservation of the hydrogen bond between the C2 substituent and the Asn^{168} side chain. Coddens *et al.* (30) reported that lactosyl moieties with an added Gal/GalNAc in α1-3 linkage, but not α1-4 linkage, maintained $F4_{ad}$ binding. When modeled into the $FaeG_{ad}$ binding pocket, the α1-4-linked Gal in globotriaose (Galα1-4Galβ1-4Glc) results in a kink in the glycan chain such that when present in a bilayer, globotriaosylceramide is expected to be sterically hindered for binding of $FaeG_{ad}$ (Fig. 4 *C*). In the case of isoglobotriaose (Galα1-3Galβ1-4Glc), the α1-3 linkage projects the reducing end glucose further away from the FaeG surface and positions its C1 hydroxyl group perpendicular to the fimbrial shaft (Fig. 4, *A* and *B*) so that polymerized FaeG subunits, as found in the $F4_{ad}$ fimbriae, would not clash with the membrane (Fig. 4*C*). Coddens *et al.* (30) observed enhanced binding interaction when the α1-3 substituting glycan chain was further elongated, for example, as found for the newly discovered glycosphingolipids GalNAcα1-3GalNAcβ1-3 Galβ1-4Glcβ1Cer and GalNAcα1-3GalNAcβ1-3 Galβ1-4 GlcNAcβ1-3Galβ1-4Glcβ1Cer. Based on our structural data, we propose that the more efficient interaction in these extended glycosphingolipids results from the additional glycan units that project the interacting non-reducing end residue further away from the host cell surface.

FaeG Employs a Variable Domain to Diversify Receptor Recognition—Apart from the ad isotype crystallized in complex with lactose (see above), two additional FaeG serotypes are known (ab and ac) that differ in receptor binding profile (17, 18). Co-crystallization of donor strand complemented $FaeG_{ab}$ and $FaeG_{ac}$ in the presence of a selection of previously identified carbohydrate ligands (galactose, lactose, globotriaose (Galα1-4Galβ1-4Glc), galabiose (Galα1-4Gal), *N*-acetyl galactosamine) (30) did not result in the x-ray structures of the binding complexes. Also, soaking experiments with apo crystals of $FaeG_{ab}$ or $FaeG_{ac}$ and the above mentioned sugars did not reveal additional electron density corresponding to the reported ligands. However, these crystals did allow the structure elucidation of apo FaeG serotype ab ($FaeG_{ab}$), which together with the previously published apo $FaeG_{ac}$ structure (PDB identifier 2J6R) (44) (Table 1) enabled us to compare the galactose binding site observed in the $FaeG_{ad}$-lactose structure or its equivalent region among the three FaeG serotypes. In $FaeG_{ad}$, the lactose ligand is bound by residues located on the D′-α1 loop. Strikingly, the different FaeG variants

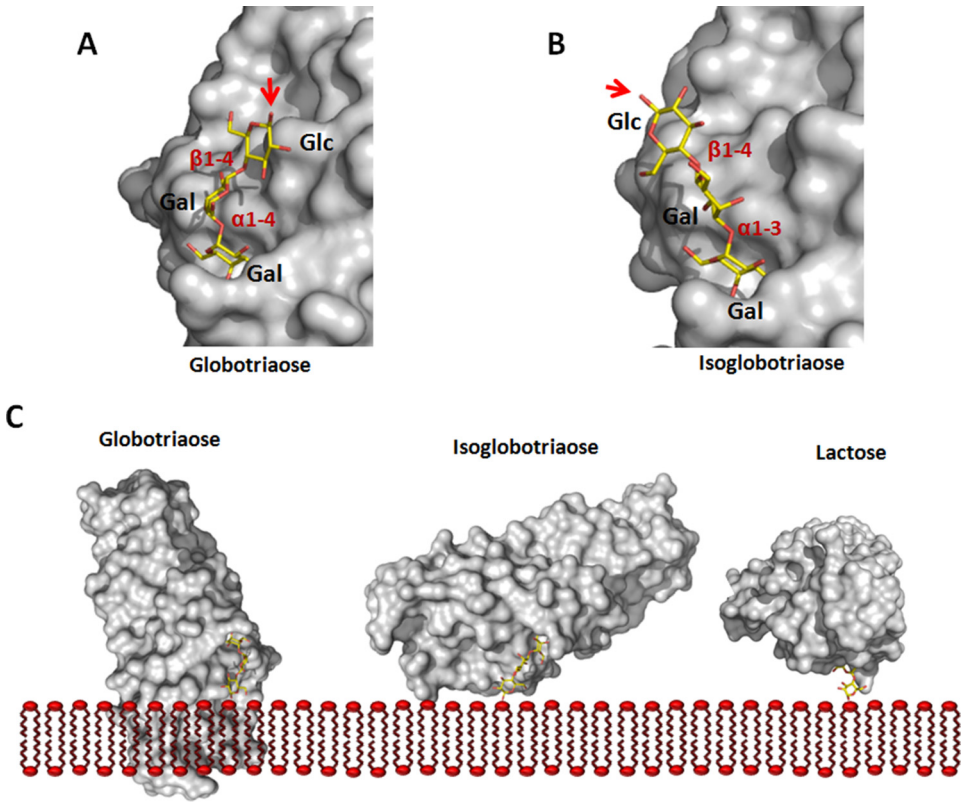

FIGURE 4. **Binding specificity of FaeG$_{ad}$.** Globotriaose (*A*) and isoglobotriaose (*B*) ligands were docked into the FaeG$_{ad}$ carbohydrate binding pocket. Terminal galactose residues substituted with a glycan chain in α1-4 linkage are not functional as FaeG$_{ad}$ receptors, whereas those in α1-3 linkage are able to interact with the FaeG$_{ad}$ major adhesive subunit (30). The α1-4 linkage of globotriaose positions the glycan chain lateral to the long axis of the FaeG$_{ad}$ subunit. A *red arrow* depicts the C1 atom, which connects the glycan with the sphingolipid part. When the glycosphingolipid is embedded in a membrane, the positioning of the reducing end glycan will result in steric hindrance and abolish the interaction between globotriaose and FaeG$_{ad}$. The α1-3 linkage of isoglobotriaose on the other hand positions the C1 group more at an angle to the subunit's long axis, allowing interaction with membrane-embedded glycosphingolipids without steric hindrance (*B* and *C*). A longer glycan chain extends the lipid part further away from the adhesin surface and enhances binding. In the FaeG$_{ad}$-lactose complex, the β1-4 linkage projects the glucose residue perpendicular to the fimbrial shaft and allows the interaction of membrane-embedded lactosylceramide with F4$_{ad}$ fimbriae without steric hindrance (*C*). In *panel C*, the orientation of the reducing end glucose monomers of the three glycans is identical. Glycan structures were generated using the GLYCAM web tool. FaeG$_{ad}$ is depicted as molecular surface and colored *gray*. Glycan structures are depicted as stick models with carbon, oxygen and nitrogen atoms colored *yellow*, *red*, and *blue*, respectively.

show large structural conformational changes in the equivalent loop regions. FaeG$_{ac}$ has a deletion of three residues in and near the amino acid stretch Val^{166}–Glu^{170} (Fig. 5 *A*), which significantly shortens the lower D′-α1 loop (Fig. 5*B*). In the FaeG$_{ad}$-lactose structure, the galactose unit forms direct interactions with the main chain amine group and side chain of residue Asn^{168}. Due to the shortening of the lower D′-α1 loop in FaeG$_{ac}$, galactose would not be able to bind the D′-α1 loop in an equivalent orientation. The shortened D′-α1 loop in FaeG$_{ac}$ also positions Arg^{165} in overlap with the galactose binding pocket as observed in FaeG$_{ad}$. Furthermore, residue Glu^{170}, which in the ad serotype structure interacts with the C3 hydroxyl group of galactose, is replaced by an alanine residue in FaeG$_{ac}$ (Figs. 1 and 5*A*). Together, these amino acid changes in the D′-D″-α1-α2 subdomain of the ac variant abolish the galactose binding site observed in the FaeG$_{ad}$-lactose complex. In the FaeG$_{ab}$ structure, the lower D′-α1 loop is shifted upwards when compared with the FaeG$_{ad}$ structure and overlaps with the lactose binding site seen for FaeG$_{ad}$ (Fig. 5*C*). Clashes between the upper D′-α1 loop and the ligand therefore prevent the accommodation and binding of a ligand in a cleft created by the upper and lower D′-α1 loop of FaeG$_{ab}$. On the edge of the lower D′-α1 loop, a substitution of Thr^{163} in the FaeG$_{ab}$ for proline in the FaeG$_{ad}$ structure seems to be responsible for repositioning the D′-α1 loop from an upward to downward position (Fig. 5*C*). The downward positioned orientation in FaeG$_{ad}$ is further stabilized by the formation of two hydrogen bonds between the side chain of Asn^{165} and the main chain amine and carboxyl groups of Asn^{97}. When comparing the sequences of the different variants, the interacting residues located in the upper D′-α1 loop are conserved (Phe^{150} is conserved in all variants; Glu^{152} in FaeG$_{ad}$ is conserved when compared with similar Asp residue in FaeG$_{ab/ac}$). However, multiple sequence variations accumulated in the lower D′-α1 loop, including the residues Asn^{168}, Ser^{169}, and Glu^{170} of the FaeG$_{ad}$ sequence (Figs. 1 and 5*A*). In both the FaeG$_{ab}$ and the FaeG$_{ac}$ variants, the FaeG$_{ad}$ binding site is significantly distorted, and it seems unlikely that these sites can accommodate galactose or different host glycan receptors unless large structural conformational changes occur upon binding.

DISCUSSION

Pathogenic *E. coli* employ a vast array of cell surface adhesins to establish adherence to host tissue. These adhesins are often presented to host cell receptors by hair-like fimbriae or pili. The majority of fimbriae extend the two-domain tip adhesin from the bacterial cell surface, allowing pathogens to interact with host receptor present on otherwise inaccessible locations. In

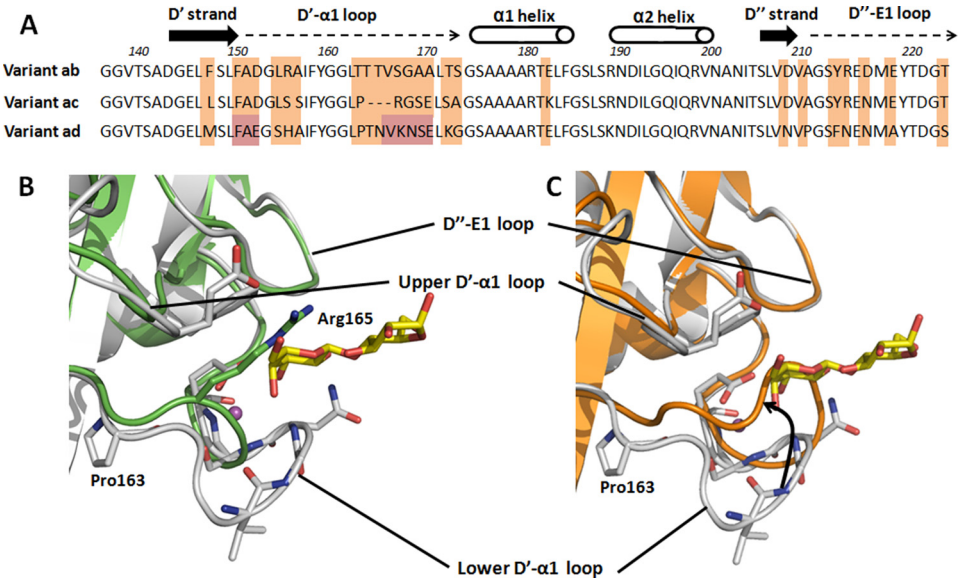

FIGURE 5. **Comparison of the FaeG$_{ad}$ lactose binding site among the different FaeG variants.** *A*, sequence comparison of the D′-D″-α1-α2 subdomain grafted on the Ig-like core between strands D and E among all three natural occurring FaeG variants. Differences between variants are indicated by *orange boxes* and are mainly located on the D′-α1 loop and D″-E1 loop. Amino acid stretches Phe^{150}–Glu^{152} and Val^{166}–Glu^{170} (FaeG variant ad numbering), present on the D′-α1 loop and involved in lactose binding by FaeG$_{ad}$, are designated by *red boxes*. Strands, helices, and loops are indicated and named. *B*, superimposition of the FaeG variant ad (colored *gray*) and variant ac (colored *green*, PDB code 2J6R) crystal structures, both of which are depicted in graphic representation. Comparison of the lactose binding site on the FaeG variant ad surface reveals remarkable differences. Variant ad binds lactose by two amino stretches on the D′-α1 loop, respectively named the upper and lower D′-α1 loop. However, variant ac has a deletion of two amino acid residues in the lower D′-α1 loop, which significantly shortens the lower D′-α1 loop and thereby abolishes the stabilizing interactions between the ligand and the lower D′-α1 loop. Also, the accommodation of the stabilizing water molecule (colored *magenta*, depicted as a *sphere*) in the binding pocket will be hampered. In addition, the presence of arginine 165 on the lower D′-α1 loop of variant ac results in steric hindrance during putative ligand binding. *C*, superimposition of the FaeG variant ad (colored *gray*) and variant ab (colored *orange*) crystal structures, both of which are depicted in graphic representation. The lower D′-α1 loop shifts upwards in the ab variant when compared with the ad variant and obstructs the galactose binding pocket. Residues of variant ad interacting with the lactose ligand are presented in stick model.

some fimbriae, the major subunit that builds up the fimbrial shaft has evolved into the adhesive subunit with multiple receptor binding sites presented along its length. Prototypical examples of this class are the F4 fimbriae, presented on the surface of enterotoxigenic *E. coli*. FaeG, the single domain adhesin of F4 fimbriae, has a dual role combining structural and adhesive properties. Interestingly, FaeG naturally occurs in three antigenic variants, each featuring a distinct hemagglutination profile and interaction pattern with porcine intestinal receptors. Throughout the past two decades, a multitude of studies tried to identify the putative F4 receptor(s), and evidence suggested the existence of both glycoprotein and glycolipid host receptors with a critical role for the carbohydrate part of these glycoconjugates.

In this study, the co-complex structure of FaeG$_{ad}$ with its lactose ligand is presented. This structure represents the first structural insight into carbohydrate binding by poly-adhesive F4 fimbriae. FaeG$_{ad}$ interacts via its D′-D″-α1-α2 binding domain with the minimal galactose binding epitope. The binding domain is grafted onto the immunoglobulin-like core and comprises most of the naturally occurring variations between FaeG variants on two elongated conformational stable loops, D′-α1 and D″-E1. The adhesin-ligand interaction is attained primarily via hydrogen bond formation with amino acid residues located on the D′-α1 loop, but also hydrophobic contacts between Phe^{150} and the nonpolar side of the glycan ring. The involvement of a hydrophobic interaction with an aromatic side chain is a common theme in ligand recognition by fimbrial adhesins. FimH, PsaA, and FedF also make use of a tyrosine residue in their binding pocket to interact with mannose, galactose, and fucose residues, respectively (48, 51, 52). PapG and F17-G employ the hydrophobic interaction of a tryptophan side chain against the carbon atoms of the sugar (47, 49). So far, most co-complex structures that have been determined were between two-domain tip adhesins and their glycan receptors (46–49, 51). The only structural insight of binding of single-domain adhesins thus far comes from the crystal structure of the homopolymeric PsaA adhesin of pH 6 antigen fimbriae of *Yersinia pestis* (52). PsaA features dual receptor recognition of both a β-1-linked galactose and a phosphocholine head group, thereby redirecting binding toward the surface of host tissue.

F4 fimbriae phylogenetically belong to the κ-fimbrial clade of chaperone-usher systems, whose members typically exhibit a flexible fibrillar structure (54). The location of the binding site on the side of the FaeG$_{ad}$ surface requires the fimbrial long axis to be arranged parallel to the supporting surface (either membrane or glycoprotein) on which the carbohydrate receptor is located. A flexible fimbrial structure allows the simultaneous interaction of multiple FaeG subunits of a single F4 fimbrium with surface-located receptors. Poly-adhesive F4$_{ad}$ fimbriae have a stringent constraint regarding the orientation of their carbohydrate host receptors. The FaeG subunits should be able to approach the binding epitope perpendicular to the supporting surface to allow interaction and to avoid steric hindrance. Host carbohydrate receptors positioned parallel to the host membrane (by insertion of an α1-4 linkage) would be, independent of the length of the connecting glycan chain, inaccessible for interaction by the elongated F4$_{ad}$ fimbriae due to steric

clashes between the host membrane and surrounding FaeG subunits. Two-domain tip adhesins are less prone to host receptor orientation as most of them express their carbohydrate-interacting adhesins on the outmost end of a flexible tip or fimbrial structure, which allows them to access a broader array of oriented glycan structures without the concern of steric hindrance with the fimbrial shaft (46, 47).

We were unable to quantify the affinity of the FaeG-lactose interaction using SPR with a surface-bound glycoconjugate and by isothermal titration calorimetry with the soluble carbohydrate receptor (results not shown), most likely due to the limitation in detectable affinity to upper μM/low mM range as the concentration of $FaeG_{ntd/dsc}$ can only be augmented to a certain value. The interaction between two-domain tip adhesins (TDA) and their receptors has been quantified on multiple occasions (16). F17-G, the TDA of F17 fimbriae, has an observed K_D of ~1.2 mM for its monosaccharide *N*-acetyl glucosamine ligand, reflecting the shallow F17-G binding groove (49). FimH, the TDA of type 1 pili, has a much higher affinity for its cellular receptors ranging from 2.3 μM for mannose to up to 20 nM for the oligomannose-3 ligand (55, 56), reflecting the deeply buried mannose binding pocket. PapGII features a K_D value of 540 μM for the disaccharide galabiose and 84 μM for globotetraose ligands (57). However, until now, little information was available on affinities by which single-domain adhesins recognize their carbohydrate receptors. The interaction of galactose toward PsaA has an IC_{50} value of ~40 mM (58), indicating that the single domain adhesins of poly-adhesive fimbria are likely to bind with less affinity to their target receptors. Upon colonization of the intestinal tract, F4 fimbriae avidity effects take place where multiple FaeG subunits simultaneously interact, resulting in a firm interaction with host tissue. A major advantage of the above described binding mode is to avoid clearance by secreted glycoproteins. The latter are secreted in the intestinal tract, and a high affinity of FaeG for glycoproteins with galactosyl residues in the terminal position could saturate the binding sites of F4 fimbriae. However, the observed low affinity enables the F4 fimbriae to shield their binding epitopes in the intestinal lumen from unwanted interactions and drives attachment to cell-attached glycan receptors by avidity effects. The immediate evolutionary advantage for the pathogen to develop such a mechanism is obvious. Also, the requirement of multiple concurrent interactions can only be fulfilled if ligands are correctly presented on the host tissue. The spatial distribution and nature of the immobilization surface therefore likely play a crucial role in determining the susceptibility to ETEC infections expressing F4 fimbriae, which could explain the existence of specific glycoprotein and glycolipid F4 receptors. A recurrent theme in the interaction of chaperone/usher-pili with host tissue manifests there is more than meets the eye. The F18 fimbrial adhesin FedF features a coincidence binding mechanism that targets FedF to glycosphingolipid receptors embedded in the cellular membrane (48). Type 1 pili and Cfa/I employ shear-enhanced adhesion to withstand the hydrodynamic forces acting on biological surfaces and direct binding to surface-bound receptors rather than toward soluble but otherwise identical receptors or soluble inhibitors (53, 59–61).

Zhang *et al.* (19) exchanged variant-specific residues between the different FaeG variants. Mutating Glu^{152} to Asn abolished binding of $F4_{ad}$ fimbriae to the target tissue, whereas mutating residues Ala^{133}, Met^{147}, Ser^{154}, and His^{155} exhibited wild type binding (19). These results are in agreement with our co-complex structure between $FaeG_{ad}$ and lactose. Glu^{152} is involved in a hydrogen bond interaction with the C6 hydroxyl group, whereas none of the other mutated residues are in close proximity to the carbohydrate ligand. In contrast, mutating residues Asn^{152}, Leu^{154}, and Ser^{155} results in loss of binding of $F4_{ac}$ fimbriae, and only the substitution of residue Val^{133} did not alter the binding capabilities. Based on these results, the primary (or secondary) binding sites of $F4_{ac}$ and $F4_{ad}$ fimbriae are expected to overlap, but are not identical.

Comparison of the D′-α1 loop between the three naturally occurring FaeG variants revealed important conformational differences between them and suggested either that the D′-α1 loop of $FaeG_{ab}$ and $FaeG_{ac}$ undergoes a significant conformational change upon interaction with the ligand or that a different binding site is present on the surface of both variants. The first proposed possibility seems less likely as the hereto determined apo and bound structures of fimbrial adhesins did not reveal large structural changes in and around the binding site. Structural changes would require additional energy input and even further reduce the affinity between protein and receptor. Moreover, key residues in the galactose-$FaeG_{ad}$ interaction are lacking in the ab and ac variants. We assume the second possibility to be more likely as a variable surface is present on the FaeG subunit, which allows the accommodation of a distinct binding site. Variants ab and ac share a higher sequence identity and ligand specificity when compared with variant ad. For instance, the D″-E1 loops are almost conserved between the ab and ac variants and would make an ideal candidate to look for the variant-specific binding pocket. The generation of a hypervariable surface with multiple binding sites with a similar, although slightly different binding specificity would allow ETEC strains to attach and colonize to a much wider portion of the piglet population. For example, $F4_{ad}$ fimbriae are unable to interact with galactose residues with a substituent sulfate group in the C3 position, whereas $F4_{ab}$ efficiently interact with sulfated glycans. Also, $F4_{ad}$ fimbriae cannot bind to glycan chains, which contain a α1-4 linkage; however, $F4_{ab}$ fimbriae demonstrate high binding specificity on TLC overlay assays against galabiaosylceramide (Galα1-4Galβ1Cer) and globotriaosylceramide (Galα1-4Galβ1-4Glcβ1Cer) (30).

In summary, we provided the first structural insight into carbohydrate binding by poly-adhesive F4 fimbriae that allowed us to explain their carbohydrate specificity. The crucial D′-α1 loop adopts remarkable different conformations between FaeG variants and led to the hypothesis that either conformational changes occur upon ligand binding or different variant-specific binding pockets exist on the FaeG surface.

Acknowledgments—We are grateful to the beamline staff of the Proxima1 beamline (Soleil, France) for support with data collection and processing. This work was supported by Equipment Grant UABR/09/005 from the Hercules Foundation.

REFERENCES

1. Kaper, J. B., Nataro, J. P., and Mobley, H. L. (2004) Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* **2,** 123–140
2. Fairbrother, J. M., Nadeau, E., and Gyles, C. L. (2005) *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. *Anim. Health Res. Rev.* **6,** 17–39
3. Wennerås, C., and Erling, V. (2004) Prevalence of enterotoxigenic *Escherichia coli*-associated diarrhoea and carrier state in the developing world. *J. Health Popul. Nutr.* **22,** 370–382
4. Nagy, B., and Fekete, P. Z. (2005) Enterotoxigenic *Escherichia coli* in veterinary medicine. *Int. J. Med. Microbiol.* **295,** 443–454
5. Gaastra, W., and Svennerholm, A. M. (1996) Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends Microbiol.* **4,** 444–452
6. Nagy, B., and Fekete, P. Z. (1999) Enterotoxigenic *Escherichia coli* (ETEC) in farm animals. *Vet. Res.* **30,** 259–284
7. Busch, A., and Waksman, G. (2012) Chaperone-usher pathways: diversity and pilus assembly mechanism. *Philos. Trans. R Soc. Lond. B. Biol. Sci.* **367,** 1112–1122
8. 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., Pinkner, J. S., Ford, B., Kay, C. W., Li, H., Hultgren, S. J., Thanassi, D. G., and Waksman, G. (2011) Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature* **474,** 49–53
9. Remaut, H., Tang, C., Henderson, N. S., Pinkner, J. S., Wang, T., Hultgren, S. J., Thanassi, D. G., Waksman, G., and Li, H. (2008) Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell* **133,** 640–652
10. Mol, O., and Oudega, B. (1996) Molecular and structural aspects of fimbriae biosynthesis and assembly in *Escherichia coli*. *FEMS Microbiol. Rev.* **19,** 25–52
11. Oudega, B., de Graaf, M., de Boer, L., Bakker, D., Vader, C. E., Mooi, F. R., and de Graaf, F. K. (1989) Detection and identification of FaeC as a minor component of K88 fibrillae of *Escherichia coli*. *Mol. Microbiol.* **3,** 645–652
12. Valent, Q. A., Zaal, J., de Graaf, F. K., and Oudega, B. (1995) Subcellular localization and topology of the K88 usher FaeD in *Escherichia coli*. *Mol. Microbiol.* **16,** 1243–1257
13. Bakker, D., Vader, C. E., Roosendaal, B., Mooi, F. R., Oudega, B., and de Graaf, F. K. (1991) Structure and function of periplasmic chaperone-like proteins involved in the biosynthesis of K88 and K99 fimbriae in enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* **5,** 875–886
14. Bakker, D., Willemsen, P. T., Willems, R. H., Huisman, T. T., Mooi, F. R., Oudega, B., Stegehuis, F., and de Graaf, F. K. (1992) Identification of minor fimbrial subunits involved in biosynthesis of K88 fimbriae. *J. Bacteriol.* **174,** 6350–6358
15. Bakker, D., Willemsen, P. T., Simons, L. H., van Zijderveld, F. G., and de Graaf, F. K. (1992) Characterization of the antigenic and adhesive properties of FaeG, the major subunit of K88 fimbriae. *Mol. Microbiol.* **6,** 247–255
16. De Greve, H., Wyns, L., and Bouckaert, J. (2007) Combining sites of bacterial fimbriae. *Curr. Opin. Struct. Biol.* **17,** 506–512
17. Guinée, P. A., and Jansen, W. H. (1979) Behavior of *Escherichia coli* K antigens K88ab, K88ac, and K88ad in immunoelectrophoresis, double diffusion, and hemagglutination. *Infect. Immun.* **23,** 700–705
18. Orskov, I., Orskov, F., Sojka, W. J., Wittig W. (1964) K antigens K88ab(L) and K88ac(L) in *E. coli*. A new O antigen: 0147 and a new K antigen: K89(B). *Acta Pathol. Microbiol. Scand.* **62,** 439–447
19. Zhang, W., Fang, Y., and Francis, D. H. (2009) Characterization of the binding specificity of K88ac and K88ad fimbriae of enterotoxigenic *Escherichia coli* by constructing K88ac/K88ad chimeric FaeG major subunits. *Infect. Immun.* **77,** 699–706
20. Bijlsma, I. G., de Nijs, A., van der Meer, C., and Frik, J. F. (1982) Different pig phenotypes affect adherence of *Escherichia coli* to jejunal brush borders by K88ab, K88ac, or K88ad antigen. *Infect Immun.* **37,** 891–894
21. Billey, L. O., Erickson, A. K., and Francis, D. H. (1998) Multiple receptors on porcine intestinal epithelial cells for the three variants of *Escherichia coli* K88 fimbrial adhesin. *Vet. Microbiol* **59,** 203–212
22. Erickson, A. K., Billey, L. O., Srinivas, G., Baker, D. R., and Francis, D. H. (1997) A three-receptor model for the interaction of the K88 fimbrial adhesin variants of *Escherichia coli* with porcine intestinal epithelial cells. *Adv. Exp. Med. Biol.* **412,** 167–173
23. Francis, D. H., Erickson, A. K., and Grange, P. A. (1999) K88 adhesins of enterotoxigenic *Escherichia coli* and their porcine enterocyte receptors. *Adv. Exp. Med. Biol.* **473,** 147–154
24. Grange, P. A., Erickson, A. K., Levery, S. B., and Francis, D. H. (1999) Identification of an intestinal neutral glycosphingolipid as a phenotype-specific receptor for the K88ad fimbrial adhesin of *Escherichia coli*. *Infect Immun.* **67,** 165–172
25. Willemsen, P. T., and de Graaf, F. K. (1992) Age and serotype dependent binding of K88 fimbriae to porcine intestinal receptors. *Microb. Pathog.* **12,** 367–375
26. Erickson, A. K., Willgohs, J. A., McFarland, S. Y., Benfield, D. A., and Francis, D. H. (1992) Identification of two porcine brush border glycoproteins that bind the K88ac adhesin of *Escherichia coli* and correlation of these glycoproteins with the adhesive phenotype. *Infect. Immun.* **60,** 983–988
27. Erickson, A. K., Baker, D. R., Bosworth, B. T., Casey, T. A., Benfield, D. A., and Francis, D. H. (1994) Characterization of porcine intestinal receptors for the K88ac fimbrial adhesin of *Escherichia coli* as mucin-type sialoglycoproteins. *Infect. Immun.* **62,** 5404–5410
28. Grange, P. A., Erickson, A. K., Anderson, T. J., and Francis, D. H. (1998) Characterization of the carbohydrate moiety of intestinal mucin-type sialoglycoprotein receptors for the K88ac fimbrial adhesin of *Escherichia coli*. *Infect. Immun.* **66,** 1613–1621
29. Grange, P. A., and Mouricout, M. A. (1996) Transferrin associated with the porcine intestinal mucosa is a receptor specific for K88ab fimbriae of *Escherichia coli*. *Infect. Immun.* **64,** 606–610
30. Coddens, A., Valis, E., Benktander, J., Ångström, J., Breimer, M. E., Cox, E., and Teneberg, S. (2011) Erythrocyte and porcine intestinal glycosphingolipids recognized by F4 fimbriae of enterotoxigenic *Escherichia coli*. *PLoS One* **6,** e23309
31. Gibbons, R. A., Jones, G. W., and Sellwood, R. (1975) An attempt to identify the intestinal receptor for the K88 adhesin by means of a haemagglutination inhibition test using glycoproteins and fractions from sow colostrum. *J. Gen. Microbiol.* **86,** 228–240
32. Grange, P. A., Mouricout, M. A., Levery, S. B., Francis, D. H., and Erickson, A. K. (2002) Evaluation of receptor binding specificity of *Escherichia coli* K88 (F4) fimbrial adhesin variants using porcine serum transferrin and glycosphingolipids as model receptors. *Infect. Immun.* **70,** 2336–2343
33. Jones, G. W. (1977) The attachment of bacteria to the surfaces of animal cells. in *Microbial Interactions* (Reissig, J. L., ed), pp 141–176, Chapman & Hall, London
34. Laux, D. C., McSweegan, E. F., Williams, T. J., Wadolkowski, E. A., and Cohen, P. S. (1986) Identification and characterization of mouse small intestine mucosal receptors for *Escherichia coli* K-12(K88ab). *Infect. Immun.* **52,** 18–25
35. Meng, Q., Kerley, M. S., Russel, T. J., and Allee, G. L. (1998) Lectin-like activity of *Escherichia coli* K88, *Salmonella choleraesuis*, and *Bifidobacteria pseudolongum* of porcine gastrointestinal origin. *J. Anim. Sci.* **76,** 551–556
36. Sellwood, R. (1980) The interaction of the K88 antigen with porcine intestinal epithelial cell brush borders. *Biochim. Biophys. Acta* **632,** 326–335
37. Payne, D., O'Reilly, M., and Williamson, D. (1993) The K88 fimbrial adhesin of enterotoxigenic *Escherichia coli* binds to β1-linked galactosyl residues in glycosphingolipids. *Infect. Immun.* **61,** 3673–3677
38. Van Molle, I., Moonens, K., Garcia-Pino, A., Buts, L., De Kerpel, M., Wyns, L., Bouckaert, J., and De Greve, H. (2009) Structural and thermodynamic characterization of pre- and postpolymerization states in the F4 fimbrial subunit FaeG. *J. Mol. Biol.* **394,** 957–967
39. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
40. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
41. Kleywegt, G. J., and Jones, T. A. (1998) Databases in protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **54,** 1119–1131

42. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
43. Amann, E., and Brosius, J. (1985) "ATG vectors" for regulated high-level expression of cloned genes in *Escherichia coli*. *Gene* **40,** 183–190
44. Van Molle, I., Joensuu, J. J., Buts, L., Panjikar, S., Kotiaho, M., Bouckaert, J., Wyns, L., Niklander-Teeri, V., and De Greve, H. (2007) Chloroplasts assemble the major subunit FaeG of *Escherichia coli* F4 (K88) fimbriae to strand-swapped dimers. *J. Mol. Biol.* **368,** 791–799
45. Sujatha, M. S., Sasidhar, Y. U., and Balaji, P. V. (2007) MP2/6–311++G(d,p) study on galactose-aromatic residue analog complexes in different position-orientations of the saccharide relative to aromatic residues. *J. Mol. Struct. (Theochem)* **814,** 11–24
46. Hung, C. S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C. G., Strouse, R., Langermann, S., Waksman, G., and Hultgren, S. J. (2002) Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.* **44,** 903–915
47. Dodson, K. W., Pinkner, J. S., Rose, T., Magnusson, G., Hultgren, S. J., and Waksman, G. (2001) Structural basis of the interaction of the pyelonephritic *E. coli* adhesin to its human kidney receptor. *Cell* **105,** 733–743
48. Moonens, K., Bouckaert, J., Coddens, A., Tran, T., Panjikar, S., De Kerpel, M., Cox, E., Remaut, H., and De Greve, H. (2012) Structural insight in histo-blood group binding by the F18 fimbrial adhesin FedF. *Mol. Microbiol.* **86,** 82–95
49. Buts, L., Bouckaert, J., De Genst, E., Loris, R., Oscarson, S., Lahmann, M., Messens, J., Brosens, E., Wyns, L., and De Greve, H. (2003) The fimbrial adhesin F17-G of enterotoxigenic *Escherichia coli* has an immunoglobulin-like lectin domain that binds *N*-acetylglucosamine. *Mol. Microbiol.* **49,** 705–715
50. 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., and Thomas, W. E. (2010) Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like β sheet twisting. *Cell* **141,** 645–655
51. Wellens, A., Garofalo, C., Nguyen, H., Van Gerven, N., Slättegård, R., Hernalsteens, J. P., Wyns, L., Oscarson, S., De Greve, H., Hultgren, S., and Bouckaert, J. (2008) Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the FimH-oligomannose-3 complex. *PLoS One* **3,** e2040
52. Bao, R., Nair, M. K., Tang, W. K., Esser, L., Sadhukhan, A., Holland, R. L., Xia, D., and Schifferli, D. M. (2013) Structural basis for the specific recognition of dual receptors by the homopolymeric pH 6 antigen (Psa) fimbriae of *Yersinia pestis*. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 1065–1070
53. Thomas, W. E., Trintchina, E., Forero, M., Vogel, V., and Sokurenko, E. V. (2002) Bacterial adhesion to target cells enhanced by shear force. *Cell* **109,** 913–923
54. Nuccio, S. P., and Bäumler, A. J. (2007) Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol. Mol. Biol. Rev.* **71,** 551–575
55. Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhrer, M., Hung, C. S., Pinkner, J., Slättegård, R., Zavialov, A., Choudhury, D., Langermann, S., Hultgren, S. J., Wyns, L., Klemm, P., Oscarson, S., Knight, S. D., and De Greve, H. (2005) Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* **55,** 441–455
56. 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., and Knight, S. D. (2006) The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of *Escherichia coli* pathotypes. *Mol. Microbiol.* **61,** 1556–1568
57. Larsson, A., Ohlsson, J., Dodson, K. W., Hultgren, S. J., Nilsson, U., and Kihlberg, J. (2003) Quantitative studies of the binding of the class II PapG adhesin from uropathogenic *Escherichia coli* to oligosaccharides. *Bioorg. Med. Chem.* **11,** 2255–2261
58. Galván, E. M., Chen, H., and Schifferli, D. M. (2007) The Psa fimbriae of *Yersinia pestis* interact with phosphatidylcholine on alveolar epithelial cells and pulmonary surfactant. *Infect. Immun.* **75,** 1272–1279
59. Nilsson, L. M., Thomas, W. E., Trintchina, E., Vogel, V., and Sokurenko, E. V. (2006) Catch bond-mediated adhesion without a shear threshold: trimannose versus monomannose interactions with the FimH adhesin of *Escherichia coli*. *J. Biol. Chem.* **281,** 16656–16663
60. Nilsson, L. M., Thomas, W. E., Sokurenko, E. V., and Vogel, V. (2006) Elevated shear stress protects *Escherichia coli* cells adhering to surfaces via catch bonds from detachment by soluble inhibitors. *Appl. Environ. Microbiol.* **72,** 3005–3010
61. Tchesnokova, V., McVeigh, A. L., Kidd, B., Yakovenko, O., Thomas, W. E., Sokurenko, E. V., and Savarino, S. J. (2010) Shear-enhanced binding of intestinal colonization factor antigen I of enterotoxigenic *Escherichia coli*. *Mol. Microbiol.* **76,** 489–502