BIOLOGICAL CRYSTALLOGRAPHY
ISSN 1399-0047

Structural and adhesive properties of the long polar fimbriae protein LpfD from adherent-invasive *Escherichia coli*

Fanny Coppens,[a,b] Jegan Iyyathurai,[a,b] Ségolène Ruer,[a,b] Antonella Fioravanti,[a,b] Joemar Taganna,[a,b] Lars Vereecke,[c,d] Henri De Greve[a,b] and Han Remaut[a,b]*

[a]Structural and Molecular Microbiology, Structural Biology Research Center, VIB, Pleinlaan 2, 1050 Brussels, Belgium, [b]Structural Biology Brussels, Vrije Universiteit Brussels, Pleinlaan 2, 1050 Brussels, Belgium, [c]Inflammation Research Center, Unit of Molecular Signal Transduction in Inflammation, VIB, 9052 Ghent, Belgium, and [d]Department of Biomedical Molecular Biology, Ghent University, 9052 Ghent, Belgium. *Correspondence e-mail: han.remaut@vub.ac.be

Received 29 January 2015
Accepted 20 May 2015

Edited by K. Miki, Kyoto University, Japan

Keywords: bacterial adhesion; fimbrial adhesin; Crohn's disease; LpfD; cell adhesion.

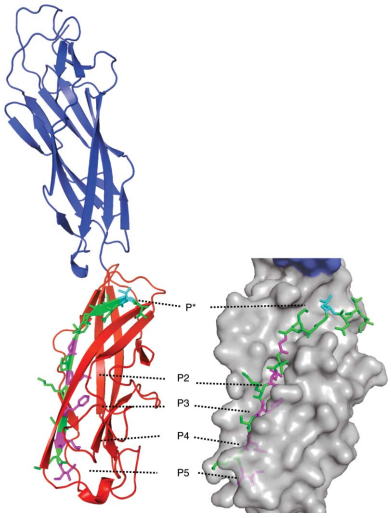
PDB reference: LpfD, 5afo

Supporting information: this article has supporting information at journals.iucr.org/d

Crohn's disease (CD) is an inflammatory bowel disease characterized by an exaggerated immune response to commensal microbiota in the intestines of patients. Metagenomic studies have identified specific bacterial species and strains with increased prevalence in CD patients, amongst which is the adherent-invasive *Escherichia coli* (AIEC) strain LF82. AIEC strains express long polar fimbriae (LPF), which are known to target Peyer's patches in a mouse CD model. Here, the recombinant production of a soluble, self-complemented construct of the LpfD protein of *E. coli* LF82 is reported and it is demonstrated that it forms the adhesive tip subunit of LPF. The LpfD crystal reveals an N-terminal adhesin domain and a C-terminal pilin domain that connects the adhesin to the minor pilus subunit LpfE. Surface topology and sequence conservation in the adhesin domain hint at a putative receptor-binding pocket as found in the *Klebsiella pneumoniae* MrkD and *E. coli* F17-G (GafD) adhesins. Immunohistostaining of murine intestinal tissue sections revealed that LpfD specifically binds to the intestinal mucosa and submucosa. LpfD binding was found to be resistant to treatment with *O*- or *N*-glycosidases, but was lost in collagenase-treated tissue sections, indicating the possible involvement of an intestinal matrix-associated protein as the LpfD receptor. LpfD strongly adhered to isolated fibronectin in an *in vitro* assay, and showed lower levels of binding to collagen V and laminin and no binding to collagens I, III and IV.

1. Introduction

Crohn's disease (CD) is an inflammatory disease of the gastrointestinal tract, the pathogenesis of which is a complex interplay of genetic susceptibility and an excessive immune response to enteric bacteria, leading to tissue injury (Binder & Orholm, 1996; Shanahan, 2002). In healthy individuals, the intestinal mucosa and its immune system act as a sensor of the microenvironment and are in a state of constant, controlled low-grade inflammation (Fiocchi, 2005). However, genetically predisposed individuals can exhibit an altered regulation of the mucosal immune response, leading to uncontrolled inflammation as a reaction to otherwise harmless, commensal microorganisms (Macpherson *et al.*, 1996; Cario & Podolsky, 2000; Eckburg & Relman, 2007). In recent years, increased attention has been paid to the role of dysbiosis in the intestinal microbiota and the possible involvement of specific, presumptively infectious, components in initiating and/or exacerbating inflammatory bowel diseases such as CD (Manichanh *et al.*, 2006; Sartor, 2010; Willing *et al.*, 2010).



Although the aetiology of intestinal dysbiosis is as yet unclear, a series of gut metagenome studies have consistently pointed to decreased microbial diversity in patients with CD, often with a concomitant increase in Enterobacteriaceae, including *Escherichia coli* and *Ruminococcus gnavus* (Manichanh *et al.*, 2006; Frank *et al.*, 2007). Darfeuille-Michaud and coworkers previously observed an abnormal increase in *E. coli* associated with resected ileal CD lesions (Darfeuille-Michaud *et al.*, 1998). The strains isolated from such CD lesions were originally believed to lack the typical virulence factors known from *E. coli* pathovars that cause acute enteric disease [*e.g.* enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), diffuse adherent *E. coli* (DAEC) or enteroaggregative *E. coli* (EAEC)], but in most cases demonstrated *in vitro* mannose-resistant (*i.e.* independent of type 1 pili) adherence to intestinal epithelial cells. In addition, such strains invade and replicate within intestinal epithelial cells and macrophages (Boudeau *et al.*, 1999). Genome analysis of LF82, a reference CD-associated isolate, showed that this adherent-invasive *E. coli* (AIEC) most closely resembles group B2 extraintestinal pathogenic *E. coli* (ExPEC) such as the avian and uropathogenic strains APEC01 and UTI89 (Miquel *et al.*, 2009). The enrichment of AIEC in CD patients has been suggested to be associated, at least in part, with abnormal overexpression of CEACAM6 (carcinoembryonic antigen-related cell adhesion molecule 6) as a result of bowel inflammation, which then acts as a mannosylated receptor for type 1 pili (Barnich *et al.*, 2007). More recently, long polar fimbriae (LPF) have been shown to be responsible for mannose-insensitive adherence of AIEC to murine Peyer's patches (PP; Chassaing *et al.*, 2011). LPF-dependent adherence to M cells on the surface of human and murine PP leads to efficient invasion and translocation of AIEC across the epithelial barrier (Chassaing *et al.*, 2011).

The name long polar fimbriae was coined following the detection by electron microscopy of long (2–10 µm) filaments at one or both poles of nonfimbriated *E. coli* (*fim* mutant) transformed with a cosmid carrying the *lpf* operon from *S. enterica* serovar Typhimurium (Bäumler & Heffron, 1995). The typical *lpf* gene cluster consists of five genes, which based on homology to other fimbrial operons can be annotated as a periplasmic chaperone (LpfB) and an outer membrane usher (LpfC) that together make up the pilus-assembly machinery, as well as the predicted fimbrial subunits LpfA, LpfD and LpfE (Bäumler & Heffron, 1995). The term 'long polar fimbriae' is used generically, as the LPF in different species or pathotypes exhibit somewhat different properties and show at least two paralogous lineages (Fig. 1). The *lpf* operon in *Salmonella* is involved in adhesion to the murine intestinal tract, more specifically to the PP. Lpf-mutant *Salmonella* exhibits an increased LD_{50} in mice, as well as a delayed action (Bäumler *et al.*, 1996). Two *lpf*-like operons were subsequently also identified in enterohaemorrhagic *E. coli* (EHEC) and AIEC (Torres *et al.*, 2002; Fig. 1). EHEC strain O157:H7 contains two non-identical operons, *lpf1* and *lpf2*, with the former being the closer homologue to the LPF found in *S. enterica* serovar Typhimurium (Fig. 1). The *lpf1* operon contains *lpfABCC′DE* and is located on O-island 141, between *yhjX* and *yhjW* (of the *E. coli* K12 chromosome; Torres *et al.*, 2002). The *lpf2* operon in EHEC is located on O-island 154 of the O157:H7 chromosome, in the same locus as the *lpf* operon of *E. coli* O113:H21, and contains the genes *lpfABCDD′* (Torres *et al.*, 2004). Both gene clusters have been shown to influence the intestinal colonization and persistence of EHEC strain O157:H7 (Jordan *et al.*, 2004; Torres *et al.*, 2007). EHEC O157:H7 adheres to PP follicle-associated epithelium of *in vitro*-cultured human intestines and causes localized attaching/effacing (A/E) lesions. In the case of *lpf1* or *lpf2* single mutants, as well as the *lpf1*/*lpf2* double-knockout

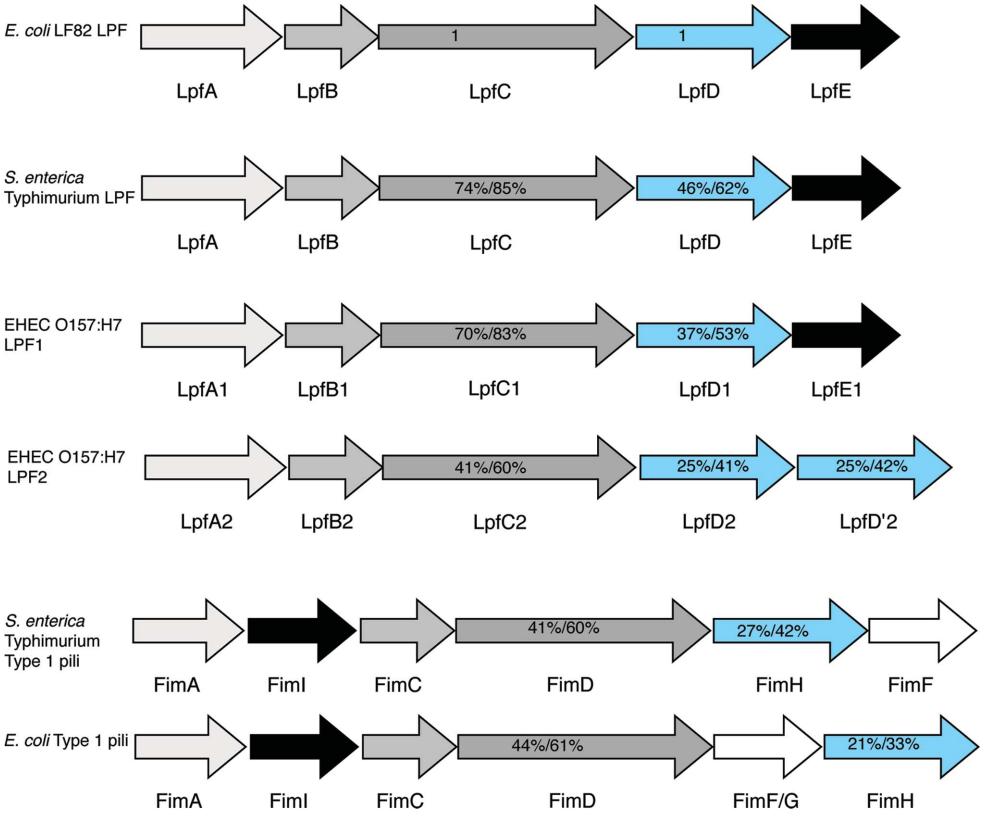


Figure 1
Schematic representations of the *lpf* and *fim* operons in *S. enterica* Typhimurium and *E. coli*, with colour fillings according to (predicted) homology and numerical values representing the percentage of amino-acid sequence identity and similarity to the corresponding protein in *E. coli* LF82.

mutant, the bacteria still caused A/E lesions, but lost tropism to the follicle-associated epithelium (Fitzhenry *et al.*, 2006). LPF in EHEC have also been reported to play a role in the inflammatory response of infected cells by upregulating the induction of pro-inflammatory cytokines (Farfan *et al.*, 2013). However, the exact role as well as the receptor-binding profile of either *lpf* operon in *E. coli* O157:H7 remains unclear. On top of the split into *lpf1* and *lpf2* paralogues, analysis of *lpfA1* sequences (major fimbrial subunit) has shown a phylogenetic subgrouping of five *lpfA1* variants which correlated with the intimin types expressed by the various pathogenic *E. coli* strains (Torres *et al.*, 2009). AIEC possesses *Lpf* variant 1–2, with the closest similarity to EHEC O111:H8, and is distinct from EHEC O157:H7 (*Lpf*1–3) and EPEC O127:H6 (*Lpf*1–1) (Torres *et al.*, 2009). The functional significance of the *Lpf1* subgrouping is as yet unclear.

To obtain a better understanding of the molecular action of LPF-mediated adherence to enteric epithelial cells, we hypothesized that LpfD is the LPF adhesin, and produced a soluble, self-complemented construct, determined its three-dimensional structure and analysed its adhesive properties. We report the X-ray structure of LfpD from AIEC strain LF82 fused to the N-terminal extension peptide of the LpfE subunit and, using immunohistostaining of murine ileal tissue sections and *in vitro* binding studies, we demonstrate that LpfD encompasses the LPF adhesin and specifically binds to a matrix component in the enteric epithelium.

2. Materials and methods

2.1. Cloning, expression and purification

The LpfD coding sequence was PCR-amplified from *E. coli* LF82 (NCBI RefSeq YP_002558392.1), with an addition through primer extension at the 3′ end of a sequence encoding a ten-glycine linker as well as the 20 N-terminal residues of the LpfE minor subunit (ATTDLGAKGTLKFSLKISQG) or the 17 N-terminal residues of LpfA (ADAGDGSVKFTGEIVDA), followed by a 6×His tag, and cloned in the pDEST14 plasmid for periplasmic expression using the Gateway technique (Invitrogen). The resulting plasmids (pJI1 and pJI2, respectively) were transformed into *E. coli* C43(DE3) cells, which were grown in 12 l LB broth supplemented with 100 µg ml^{-1} ampicillin under constant aeration at 37°C. Expression was induced by the addition of a 0.5 m*M* final concentration of isopropyl *β*-D-1-thiogalactopyranoside (IPTG) when the cells reached an OD of 0.7 and was allowed to continue for 18 h at 28°C. The cells were harvested by centrifugation at 4000*g* for 15 min and were resuspended in 3 ml sucrose buffer per gram of wet cells (20% sucrose, 20 m*M* Tris pH 8.5, 5 m*M* EDTA, 0.1 mg ml^{-1} lysozyme) and stirred for 20 min at 4°C. Centrifugation at 13 000*g* for 45 min yielded the periplasmic extract, which was dialysed overnight against dialysis buffer (20 m*M* Tris pH 8.5, 150 m*M* NaCl) at 4°C. 20 m*M* imidazole and 10 m*M* $MgCl_2$ (to chelate the EDTA) were added to the periplasmic extract before ultracentrifugation (Beckman Coulter Optima L-90K, 45 Ti rotor, 30 000 rev min^{-1}, 30 min, 4°C) to remove the insoluble particles prior to chromatography.

LpfD fused to the first 20 residues of LpfE ($LpfD_{ENte}$) was purified by Ni-affinity chromatography followed by hydrophobic interaction chromatography: the cleared periplasmic extract was loaded onto a 5 ml HisTrap FF column (GE) equilibrated in buffer *A* (20 m*M* Tris pH 8.5, 150 m*M* NaCl, 20 m*M* imidazole), after which the column was washed with buffer *A*. The recombinant protein was eluted with a gradient to elution buffer (20 m*M* Tris pH 8.5, 150 m*M* NaCl, 500 m*M* imidazole) and the fractions containing the protein of interest were pooled and supplemented with 1.5 *M* ammonium sulfate. After filtration through a 0.45 µm pore-size filter, the sample was loaded onto a 1 ml Phenyl FF low-sub (GE) column equilibrated in buffer *C* (20 m*M* Tris pH 8.5, 1.5 *M* ammonium sulfate). Elution was with 20 m*M* Tris pH 8.5 buffer using a gradient over 30 column volumes. The protein was concentrated to 3.6 mg ml^{-1} using a 10K MWCO Amicon concentrator.

2.2. Crystallization and data collection

The concentrated protein was screened against sparse-matrix crystallization screens using the sitting-drop vapour-diffusion method in 96-well MRC plates (Molecular Dimensions) incubated at 20° containing 70 µl well solution and drops consisting of 0.4 µl protein solution and 0.2 µl well solution. 29 hits were observed using JBScreen PACT++ (Jena Bioscience) and were optimized by modifying the precipitant concentration and using 48-well pre-greased hanging-drop plates with crystallization drops consisting of 2.0 µl protein solution and 1.0 µl well solution. Crystals obtained in the condition 0.1 *M* Tris pH 8.0, 0.18 *M* sodium citrate, 20% PEG 3350 were used to solve the structure of $LpfD_{ENte}$.

Native crystals were mounted on a cryoloop and passed through a Paratone layer for cryoprotection before flash-cooling.

Experimental phases were obtained through the iodination of Tyr. Since iodination of proteins is a well known labelling technique and iodination of tyrosines has already been achieved after crystal formation and successfully used for phasing (Ghosh *et al.*, 1999), a slightly modified protocol, using only iodine and no iodide, was attempted and was found to be successful. A tiny chunk of I_2 was fixed with silicone grease onto the side of the well (hanging drop) containing the selected crystal and allowed to sublime and diffuse overnight in the closed well. Following overnight incubation the crystal was mounted on a cryoloop, soaked in well solution supplemented with 15% glycerol for cryoprotection and flash-cooled in liquid N_2.

X-ray diffraction data were collected on the PROXIMA 1 beamline at SOLEIL, Saint-Aubin, France tuned to 6.8 keV and equipped with a Pilatus 6M detector.

One data set was collected from an $LpfD_{ENte}$–I_2 crystal using an inverse-beam protocol and indexed and scaled using *XDS* and *XSCALE* (Kabsch, 2010). The space group was determined to be $C222_1$, and *SHELXC* and *SHELXD* were

used to locate anomalous scattering sites (Sheldrick, 2008). The *AutoSol* and *AutoBuild* workflows from *PHENIX* (Adams *et al.*, 2010) were run successively, followed by manual rebuilding and maximum-likelihood refinement using *Coot* and *REFMAC*, respectively (Emsley *et al.*, 2010; Murshudov *et al.*, 2011).

Two data sets were also collected from the $LpfD_{ENte}$ native crystal at 12.65 keV and were processed and scaled as described above. The model obtained from the $LpfD_{ENte}$–I_2 data sets was further refined using the native data, ultimately resulting in a model with 1.8 Å resolution and an R_{work} and R_{free} of 0.17 and 0.19, respectively, which was deposited in the PDB with accession code 5afo.

2.3. Fluorescence microscopy

Purified $LpfD_{ENte}$ in 20 m*M* HEPES pH 7.5, 150 m*M* NaCl, 5% glycerol was fluorescently labelled with DyLight 650 amine-reactive dye (Thermo Scientific). Unbound dye was removed by three successive dialyses against 20 m*M* HEPES pH 7.5, 150 m*M* NaCl, and a sample of the last dialysis buffer after equilibration was used as the blank in the microscopy experiments.

The protocol used for the *in situ* binding assay is based on that of Johansson & Hansson (2012). C57BL/6WT mice were sacrificed according to Belgian regulations and the bladder and small intestine were dissected and fixed with methanol-Carnoy mixture for a minimum of 3 h at room temperature. The fixed tissues were embedded in paraffin and 4 µm thin sections were made and put on glass slides. Prior to labelling, deparaffinization and rehydration of the sections was performed by incubation of the slides for 2 × 3 min in xylene, 2-propanol, 100% ethanol and 70% ethanol, respectively, followed by rinsing in tap water and PBS. Some of the slides were pre-treated overnight at 37°C under shaking conditions with either *O*-glycosidase (NEB Enzyme; 40 000 units in 5 ml 50 m*M* sodium phosphate buffer pH 7.5), PNGase (NEB Enzyme; 5 U in 5 ml 250 m*M* sodium phosphate buffer pH 7) or 250 µl collagenase mix in 5 ml PBS (10 U ml^{-1} collagenase type I, 400 U ml^{-1} collagenase type IV, 30 U ml^{-1} DNase I). The PNGase and *O*-glycosidase activity was assessed by the observation of markedly reduced binding of FimH and peanut lectin, respectively. After several washes with PBS, the slides were incubated in blocking buffer (PBS + 5% FCS) at room temperature for 30 min before counter-staining with Hoechst stain (1:1000 in PBS) for 10 min at room temperature. After three washes with PBS, the sections were incubated with anti-Muc2 mucin primary antibody [1:200; mucin 2 (H-300) rabbit polyclonal IgG, Santa Cruz Biotechnology] for 2 h. After three washes with PBS, the slides were incubated with goat anti-rabbit DyLight 488 secondary antibody (1:1000 dilution) in blocking buffer for 1 h at room temperature. Incubation with 10 µg of the labelled $LpfD_{ENte}$ in blocking buffer alone, in blocking buffer with mouse naive serum or in blocking buffer with mouse serum raised against $LpfD_{ENte}$ took place at 4°C overnight. Slides were finally washed with PBS five times before mounting with fluorescence mounting medium (*n*-propyl gallate in glycerol) and were observed with a Leica

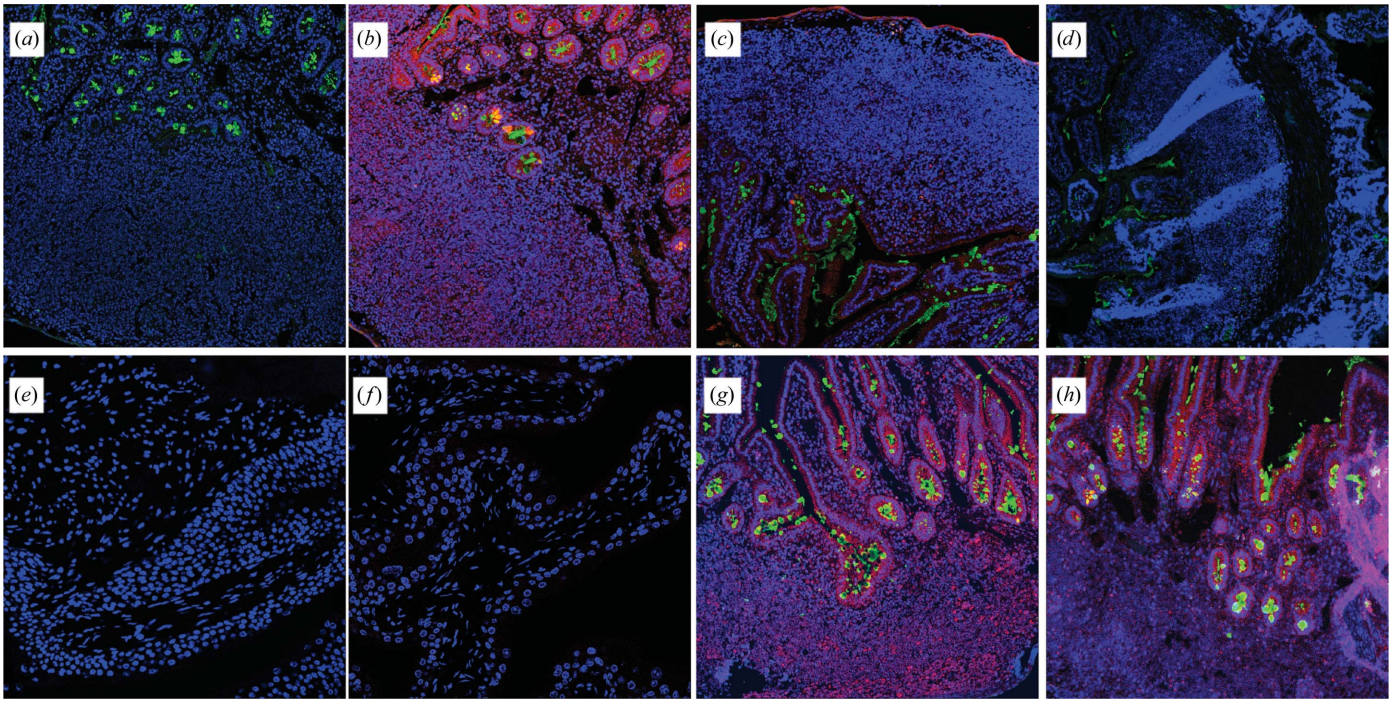


Figure 2
Fluorescence microscopy experiments on murine small intestine (*a*, *b*, *c*, *d*, *g*, *h*) and bladder (*e*, *f*) sections. Sections were incubated with (*a*, *e*) the last dialysis buffer after DyLight 650 labelling as a blank, (*b*, *f*, *g*, *h*) DyLight 650-labelled $LpfD_{ENte}$, (*c*) DyLight 650-labelled $LpfD_{ENte}$ with murine naive serum and (*d*) DyLight 650-labelled $LpfD_{ENte}$ with mouse polyclonal anti-LpfD serum. Slides (*g*) and (*h*) were pre-treated prior to incubation with $LpfD_{ENte}$ with (*g*) *O*-glycosidase or (*h*) PNGase. Blue is Hoechst stain of DNA, green is Muc2 antibody and red is $LpfD_{ENte}$ labelled with DyLight 650.

confocal microscope (Leica Microsystems LAS-AF-TCS SP5) with a 20 × 125 objective. All of the pictures were taken using the same microscope settings and the *Fiji* software was used to process the data. The selected images are representative of multiple experiments on nonconsecutive tissue sections. Where nonconsecutive intestinal regions are shown in different panels of Fig. 2, these were individually compared with positive and negative controls (*i.e.* + or − $LpfD_{ENte}$) on near sections (not shown) to minimize nonspecific staining differences.

2.4. ELISA assay

LpfD binding to ECM proteins was evaluated as described previously (Farfan *et al.*, 2011). Briefly, 96-well microtitre plates (Nunc) were coated with a solution of 10 µg ml^{-1} collagen type I (collagen from rat tail, Bornstein and Traub type I; Sigma), collagen type III (collagen from human placenta, Bornstein and Traub type III; Sigma), collagen type IV (collagen from human placenta, Bornstein and Traub type IV; Sigma), collagen type V (collagen from human placenta, Bornstein and Traub type V; Sigma), laminin (from human placenta; Biopur) and fibronectin (from human plasma; Sigma) per well overnight at 4°C. Unbound protein was removed and blocked with 1% bovine serum albumin (BSA; Sigma) for 2 h at 37°C. $LpfD_{ENte}$ at 10 µg ml^{-1} was added to the ECM protein-coated wells and incubated for 90 min at room temperature. As negative and positive controls, wells were coated with PBS or $LpfD_{ENte}$ (10 µg ml^{-1}), respectively. After three washes with PBS and Tween, wells were incubated with the primary antibody, a 1:500 dilution of anti-His mAB (Serotec). Wells were subsequently washed (three times) and bound antibodies were detected by incubation with a goat anti-mouse IgG alkaline phosphatase-conjugated secondary antibody (1:500 dilution; Sigma–Aldrich). Binding was revealed with *p*-dinitrophenylphosphatase (p-DNPP; Sigma) as a substrate. Absorbance values were measured at 405 nm after 15 and 60 min. The values after 60 min are plotted in Fig. 7(*b*).

2.5. Bioinformatics tools

To map the conservation of residues onto the LpfD structure, 205 *E. coli* LpfD protein sequences were obtained from the PiliomeDB database (http://www.piliomedb.org). By filtering out truncates and selecting only one representative for each group of identical sequences, 13 sequences remained (*E. coli* strains/GenBank accession Nos. O55:H7/AEZ42654.1, KO11FL/AFH15989.1, O157:H7/AAG58691.1, 55989/CAV00473.1, IAI1/CAR00505.1, O127:H6/CAS11343.1, O103:H2/BAI33386.1, O111:H-/BAI38125.1, O55:H7/ADD58754.1, O26:H11/BAI28212.1, O83:H1/ADR28936.1, W/ADT77155.1 and LF82/CAP78003.1)

A multiple alignment of these sequences was used to compute residue frequencies for each column, which were then mapped onto the LpfD surface-filling model using a custom Python script developed in-house. Multiple sequence alignments were performed with *MUSCLE* (Edgar, 2004) and visually represented with *WebLogo* (Crooks *et al.*, 2004), and structural alignments were performed with the *DALI* server (Holm & Rosenström, 2010). Identity and similarity values were calculated by *EMBOSS Needle* (McWilliam *et al.*, 2013), interface buried surface area was calculated with *PDBePISA* (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html; Krissinel & Henrick, 2007) and secondary-structure prediction was performed with *PSIPRED* (Buchan *et al.*, 2013). Figures were prepared with *MacPyMOL* (*PyMOL* v.1.3; Schrödinger).

3. Results and discussion

3.1. LpfD is responsible for the adhesion of LPF

The adhesive properties of chaperone–usher pili are either found integrated into polymerizing single-domain subunits that are distributed throughout the length of the fibre (single-domain adhesins or SDAs) or as two-domain adhesins (TDA) found as a single copy located at the distal tip of the pilus (De Greve *et al.*, 2007). Based on secondary-structure predictions, its localization in the LPF gene cluster (Fig. 1) and its homology to FimH [LpfD from *E. coli* LF82 (NCBI RefSeq YP_002558392.1) and FimH from *S. enterica* Typhimurium share 27 and 42% sequence identity and sequence similarity, respectively], we hypothesized that mature LpfD would be a putative TDA. In known TDAs, the adhesive properties are located in an N-terminal β-sandwich domain, whilst the C-terminal domain forms an incomplete Ig domain that connects the subunit to the pilus through a noncovalent donor-strand complementation reaction with an N-terminal extension (Nte) of the next subunit in the row (Choudhury *et al.*, 1999). As such, isolation of a stable TDA can only be performed (i) as a binary complex with its chaperone (Choudhury *et al.*, 1999), (ii) as a fragment corresponding to the N-terminal adhesin domain only (Dodson *et al.*, 2001) or (iii) in a donor-strand exchanged (DSE) complex with its cognate N-terminal extension peptide, either noncovalently or as a self-complemented construct in which a complementing Nte is fused to the C-terminus of the adhesin (Barnhart *et al.*, 2000; Sauer *et al.*, 2002). For LPF, attempts to produce a stable complex of LpfD and the LpfB chaperone were unsuccessful. Similarly, the expression of constructs corresponding to the predicted N-terminal domain did not result in a stably folded protein. Therefore, we followed a route to generate Nte-complemented LfpD. As it is presently unknown whether LpfD is attached to the LpfE or LpfA subunits inside LPF, we made self-complemented constructs by fusing the protein to either the first 20 residues of LpfE (hereafter referred to as $LpfD_{ENte}$) or the first 17 residues of LpfA (hereafter referred to $LpfD_{ANte}$), in both cases preceded by a ten-glycine linker to allow the Nte to attain its binding conformation and followed by a 6×His tag for metal-ion affinity purification. Although $LpfD_{ANte}$ was found to be expressed in the periplasm (Supplementary Fig. S1), it did not represent a stable protein that was amenable to purification. By comparison, $LpfD_{ENte}$ formed a stable, soluble fusion product that could be purified from the periplasm in high yields. Together, these data

suggested that the LpfE Nte forms the natural complementing strand of LpfD.

To determine whether LpfD represents the LPF adhesin and has glycan-binding properties, as observed for most other fimbrial tip adhesins, the purified $LpfD_{ENte}$ was sent to the Consortium for Functional Glycomics (CFG). The recombinant protein was subjected to a protein–glycan interaction test by exposure to version 5.1 of the glycan array, which consists of 610 immobilized glycans in six replicates. This failed to show statistically significant binding to any of the glycans present in the CFG 5.1 array (the data are publicly available at http://www.functionalglycomics.org, accession code cbp_2960). We next examined whether the LpfD subunit indeed confers the adhesive properties of LPF. Therefore, purified $LpfD_{ENte}$ was fluorescently labelled with DyLight 650 and the binding of the protein to mouse intestinal tissue sections was monitored by fluorescence microscopy (Fig. 2). The purified, DyLight 650-labelled $LpfD_{ENte}$ was found to abundantly bind the intestinal epithelium and PP of mouse small intestine tissue sections (Figs. 2*a* and 2*b*). A co-incubation with mouse polyclonal anti-LpfD serum, but not naive mouse serum, abolished $LpfD_{ENte}$ binding to the ileal tissue section (Figs. 2*c* and 2*d*). To assess whether $LpfD_{ENte}$ binding was specific to intestinal tissue, murine bladder epithelium sections were incubated with DyLight 650-labelled $LpfD_{ENte}$ but failed to show binding of the protein (Figs. 2*e* and 2*f*). Previously reported binding studies obtained with whole *E. coli* LF82 expressing LPF showed that bacterial binding to mouse ileal loops or isolated human PPs was localized to the PP-associated M cells only (Chassaing *et al.*, 2011). By comparison, using the purified $LpfD_{ENte}$ protein, we find a broader binding profile that, in addition to the M-cell binding in the PP, includes goblet and columnar epithelium cells in the intestinal epithelium (Figs. 2*b* and 2*c*). Co-staining with a Muc2 antibody shows that the $LpfD_{ENte}$ binding is cell-associated and is absent in the luminal mucin secretions. We propose that this apparent discrepancy in LPF binding tropism stems from the inability of the whole bacteria to penetrate the overlaying mucus layer in the context of an intact mucosa. Our results suggest the LpfD receptor is present throughout the ileal epithelium, but becomes accessible to bacteria only in places with reduced mucus shielding, as is the case at the PP.

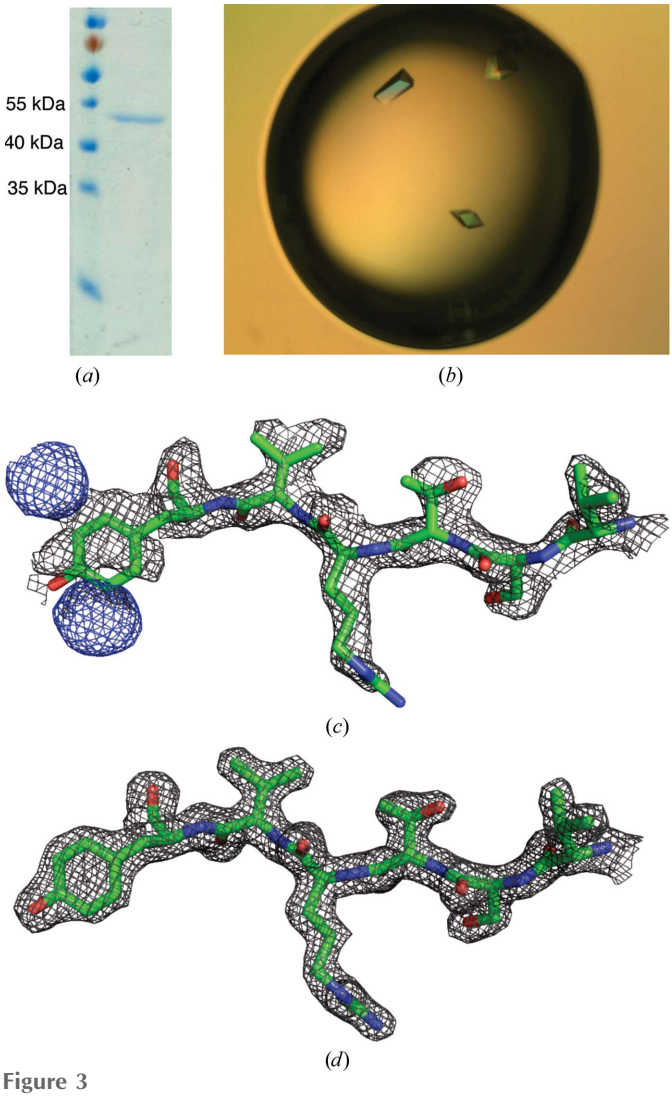


Figure 3
(*a*) SDS–PAGE showing the purity of the crystallized recombinant $LpfD_{ENte}$ obtained after nickel-affinity chromatography followed by hydrophobic interaction chromatography. (*b*) A hanging-drop vapour-diffusion crystallization experiment with condition 0.1 *M* Tris pH 8.0, 0.18 *M* sodium citrate, 20% PEG 3350 containing crystals of $LpfD_{ENte}$. (*c*, *d*) Detail of the refined structure model (residues 90–95) overlaid with (*c*) the density-modified experimental electron-density map contoured at 1.5σ (grey) and the anomalous density (blue) contoured at 3.0σ or (*d*) the refined $2F_o - F_c$ electron-density map contoured at 1.5σ.

Most fimbrial adhesins bind to specific glycoreceptors presented in the form of glycoproteins, glycolipids or matrix glycans (De Greve *et al.*, 2007). To evaluate whether the negative glycan-array analysis of $LpfD_{ENte}$ could stem from the absence of the natural LpfD glycan receptor in the version 5.1 array, we subjected mouse ileal tissue sections to *O*-glycosidase or PNGase treatment to remove O- or N-linked glycans, respectively, prior to incubation with $LpfD_{ENte}$. Neither treatment abolished $LpfD_{ENte}$ binding (Figs. 2*g* and 2*h*), although PNGase-treated ileal sections seemed to show a reduced binding in the underlying PP but not at the intestinal epithelium.

3.2. Structural analysis of LpfD

To obtain structural insight into the adhesive properties of LpfD, purified $LpfD_{ENte}$ protein (Fig. 3*a*) was concentrated to 3.6 mg ml^{-1} and subjected to crystallization and structure determination through X-ray crystallography. The polyhedral crystals (Fig. 3*b*) obtained in the crystallization condition 0.1 *M* Tris pH 8.0, 0.18 *M* sodium citrate, 20% PEG 3350 diffracted to 2.2 Å resolution and were experimentally phased using a single anomalous dispersion (SAD) experiment on a tyrosine-iodinated crystal. The structure was built and further refined to 1.8 Å resolution using data collected from a native crystal, resulting in a final model with an R_{work} and R_{free} of 0.17 and 0.19, respectively (Figs. 3*c* and 3*d*, Table 1), a *MolProbity* score of 2.00 and with 98.4% of the residues lying in the most favoured regions of the Ramachandran plot.

The structure of LpfD comprises two independent domains, similar to the architecture observed for fimbrial tip adhesins such as FimH (Choudhury *et al.*, 1999). Both the N-terminal adhesin and the C-terminal pilin domain of LpfD show a β-sandwich pattern connected by a short linker (residues 180–183 of the mature protein; Figs. 4*b* and 4*d*). The pilin domain has an incomplete Ig-like fold, lacking the last β-strand, which is complemented by 17 of the 20 added residues corresponding to the first β-strand of LpfE (Figs. 4*b* and 4*c*, green and magenta). Residue Ile17 of the N-terminal extension has its side chain pointing towards the hydrophobic core of the pilin domain of LpfD, thereby filling the cavity known as the P5 pocket, in an analogy to the numbering convention used for the Pap system (Sauer *et al.*, 2002) and other chaperone–usher systems undergoing DSE. Residues Leu15, Phe13 and Leu11 fill the P4, P3 and P2 pockets, respectively, stabilizing the pilin domain through hydrophobic interactions. This pattern is somewhat different from that observed in subunits of the Pap and Fim systems, in which the P4 forms a shallow pocket that accepts a Gly residue in the N-terminal extension (Sauer *et al.*, 2002; Puorger *et al.*, 2008) and is more reminiscent of the pilin–Nte interactions observed in the *Yersinia pestis* Caf and *S. enterica* Saf pilus systems (Zavialov *et al.*, 2003; Remaut *et al.*, 2006). Similar to the latter, LpfE reveals an extended Nte that inserts a bulky residue, Leu5, into a hydrophobic pocket outside the Nte acceptor groove, referred to as the P* pocket in the case of the Saf system (Remaut *et al.*, 2006). As expected, no electron density was observed for the ten-glycine linker used to covalently link the C-terminus of LpfD to the first N-terminal residue of the Nte of LpfE in $LpfD_{ENte}$. The binding interactions of the LpfE Nte in the LpfD acceptor groove provide another pointer to the subunit order found in the LPF. Alignment of the LpfA and LpfE N-terminal sequences (Fig. 4*a*) shows that the former would form a suboptimal donor strand for interaction with the LpfD acceptor groove, as amino-acid substitutions at the P4 (Leu→Gly), P2 (Leu→Val) and P* (Leu→Asp) positions

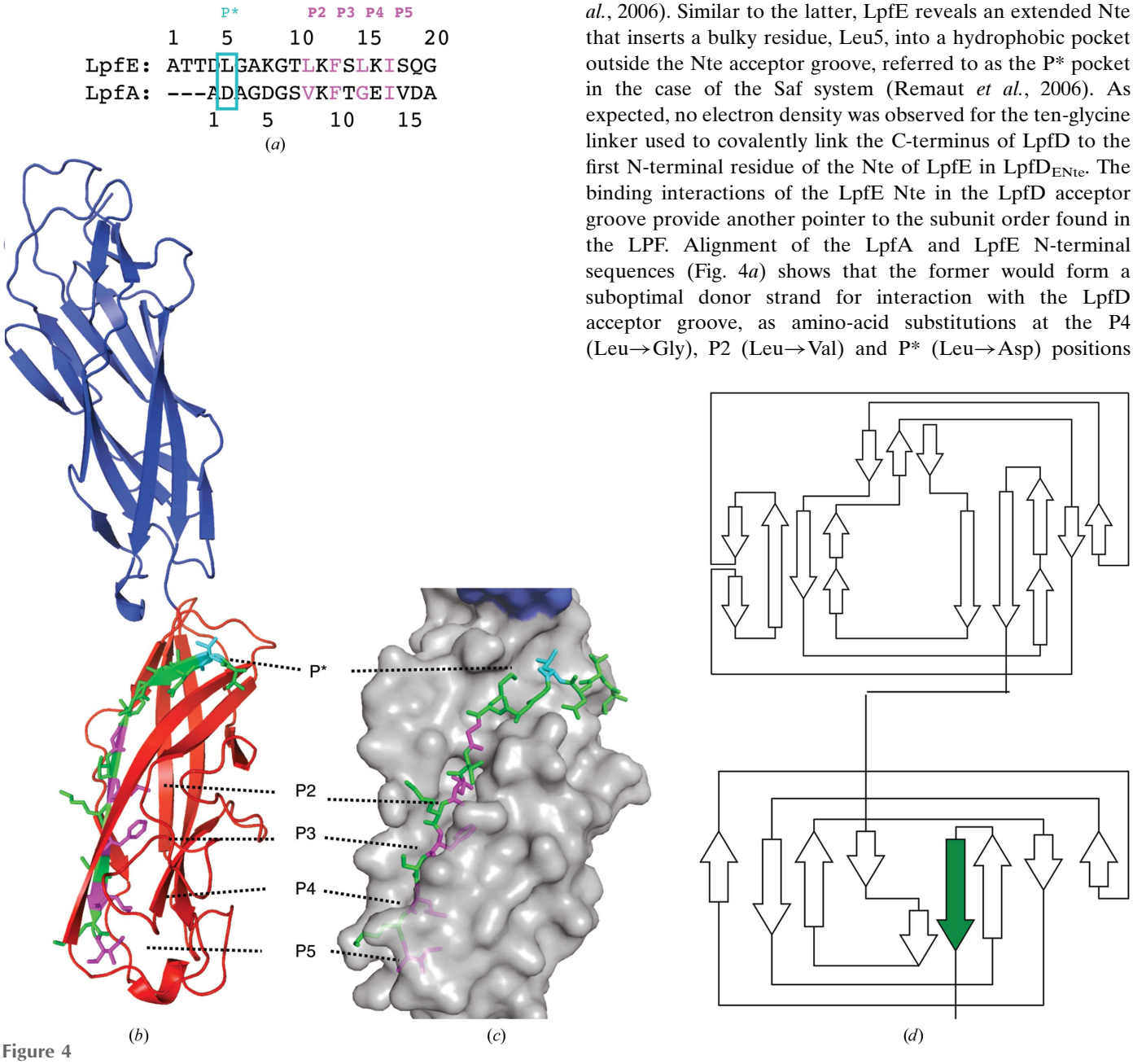


Figure 4
(*a*) Alignment of the N-terminal extensions of LpfE and LpfA. The P2, P3, P4 and P5 residues are indicated in magenta and the P* residue in orange. (*b*) Cartoon model of the structure of LpfD; the adhesin domain is shown in dark blue and the pilin domain is in red, with the complementing first β-strand of LpfE in green and magenta (P5–P2 residues). (*c*) Hydrophobic interactions between the N-terminal extension of LpfE and the core of the pilin domain of LpfD. Residues filling the P2–P5 pockets are shown in magenta and Leu5 filling the P* pocket is shown in cyan. (*d*) Topology diagrams of the adhesin domain (top) and the pilin domain (bottom) of LpfD. The complementing N-terminal extension of LpfE is shown in green.

Table 1
Data-collection and refinement statistics.

Beamline	PROXIMA 1
Wavelength (Å)	0.98011
Space group	$C222_1$
Unit-cell parameters (Å)	a = 49.01, b = 372.13, c = 99.03
Molecules per asymmetric unit	2
Resolution range (Å)	48.59–1.82 (1.883–1.818)
Unique reflections	82152 (8034)
Completeness (%)	99.88 (98.86)
Multiplicity	11.6
R_{meas} (%)	6.7 (83.3)
Mean $I/\sigma(I)$	24.04 (3.26)
Wilson B factor (Å^2)	23.15
R_{work}/R_{free}	0.172/0.193
R.m.s.d., bond lengths (Å)	0.021
R.m.s.d., bond angles (°)	1.95
Ramachandran plot statistics (%)	
Favoured	98.0
Allowed	2.0
Outliers	0.0

would result in a destabilizing loss of hydrophobic contacts in the DSE interaction. Indeed, when complemented with the LpfA Nte, LpfD did not form a stable soluble fusion protein (see above). Together, these results suggest an LPF architecture in which LpfD forms the capping tip adhesin that is connected to the pilus rod *via* the LpfE subunit.

To date, experimental structures of six fimbrial TDAs are available, F17-G (GafD; PDB entry 1o9w; Buts *et al.*, 2003), FimH (PDB entry 3jwn; Le Trong *et al.*, 2010), PapG (PDB entry 1j8s; Dodson *et al.*, 2001), FedF (PDB entry 4b4p; Moonens *et al.*, 2012), MrkD (PDB entry 3u4k; Rêgo *et al.*, 2012) and CfaE (PDB entry 3vac; Liu *et al.*, 2013), of which just two comprise the full-length protein including the pilin domain (Le Trong *et al.*, 2010; Liu *et al.*, 2013). In different tip adhesins, the interdomain region forming the contact between the lectin and pilin domains of the adhesin has been found to play an important role in conformation and ligand-binding affinity (Thomas *et al.*, 2002; Yakovenko *et al.*, 2008; Liu *et al.*, 2013; Stahlhut *et al.*, 2013). The interaction mode and role of the interdomain region has been characterized structurally for the *E. coli* adhesins FimH and CfaE (Le Trong *et al.*, 2010; Liu *et al.*, 2013). The FimH lectin domain contacts the pilin domain through two loops situated at the base of the lectin domain. Although 35 Å away from the mannose-binding pocket, the absence of the pilin domain or disruption of the interdomain contact by the application of mechanical force causes the lectin domain to adopt an elongated higher-affinity conformation (Le Trong *et al.*, 2010). This elongated conformation is propagated throughout the lectin domain *via* the distortion of a β-sheet, ultimately influencing the loops of the binding pocket and resulting in the shear-enhanced binding properties of the adhesin (Le Trong *et al.*, 2010). The interdomain interface of CfaE also appears to play a role in ligand-binding affinity, as was shown by introducing bulkiness in this interdomain plane (Liu *et al.*, 2013). In LpfD, the interdomain region corresponds to a 383 Å^2 solvent-buried surface area (Krissinel & Henrick, 2007) encompassing residues 24–32, 139–143 and 179 in the lectin domain, residues 184, 227–230 and 288–290 in the pilin domain and residues 180–183 of the linker. The LpfD interdomain region shows a high degree of sequence conservation compared with that observed in the lectin domain (Fig. 5). This includes contacts by residue Val142 (conservatively mutated to Ile142 in 20% of LpfD polymorphs) and a defined hydrophobic pocket on the nearby surface of the pilin domain. Whether LpfD shows a shear-enhanced binding profile is presently unknown; sequence conservation in the interdomain contact zone argues at least for a role in maintaining a rather fixed angle between both domains in the absence of a pulling force (Fig. 5).

3.3. In search of the receptor and binding site of LpfD

In the absence of an experimentally identified LpfD receptor, we sought structure-guided clues pointing to the LpfD receptor-binding site and candidate receptors by comparison with fimbrial adhesins of known structure. When looking at sequence conservation of surface-exposed residues in the LpfD adhesin (Fig. 6*a*), a striking difference is seen between the N- and C-terminal domains. Whilst the pilin domain exhibits a high degree of sequence conservation, the putative adhesin domain reveals a high degree of sequence variation for the most part of its surface. We postulate that this increased sequence divergence is indicative of a selective pressure to escape neutralizing immune responses directed to the receptor-binding domain. Contrary to the residues in the immediate receptor-binding site, which are expected to be under a strong negative selection to maintain epitope recognition, residues surrounding the binding site can be expected to undergo episodic sweeps of

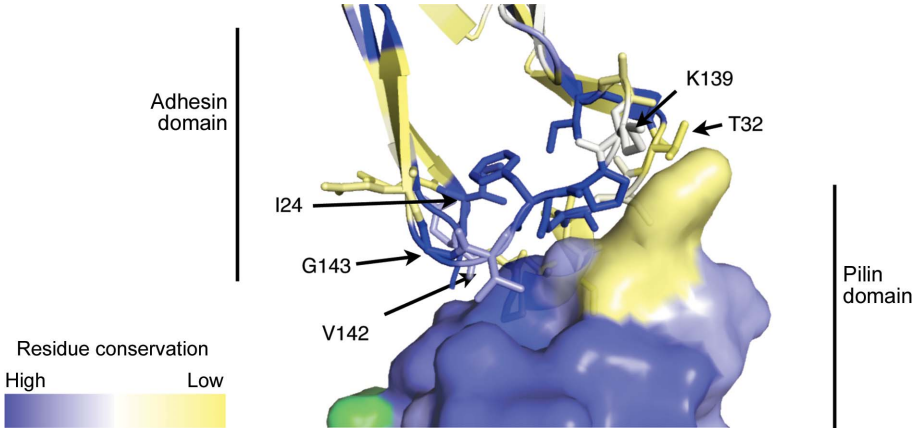


Figure 5
Close-up view of the interaction between the adhesin and pilin domains of LpfD. The pilin domain is shown as a surface-filling model and the adhesin domain as a cartoon model, with the residues corresponding to the 'swing' loop shown as stick models. Residues are coloured blue/white/yellow according to sequence conservation, with blue and yellow the highest and the lowest level of conservation, respectively. The N-terminal extension of LpfE is shown in green.

positive selection of non-synonymous mutations that evade the binding of neutralizing antibodies which create a steric blockage to receptor binding. A more detailed analysis of the surface residues of the LpfD adhesin domain reveals two discrete zones with increased sequence conservation, pointing to the putative receptor-binding site(s). A first high-similarity region, 'HSR1', is located near the distal tip of the N-terminal domain and is composed of the highly conserved surface-localized residues Ala1, Asp2, Glu53, Pro55 and Pro57 (Fig. 6*b*; Supplementary Fig. S3). This region contains both disulfide bridges of the adhesin domain: between Cys6 and Cys52 and between Cys54 and Cys118. A second region, 'HSR2', coincides with a crevice in the molecular surface on the side of the N-terminal domain and is delineated by residues Tyr67, Ala97, Gln113, Glu117, Trp125 and Thr126 (Fig. 6*c*). Other than these two regions, conserved surface-localized residues map to the interdomain contact region, as discussed above.

To further evaluate the role of HSR1 or HSR2 as candidate receptor-binding sites, we compared their location and structural topology with those of confirmed ligand-binding sites in fimbrial adhesins of known structure. A structural homology search using the *DALI* server and the LpfD adhesin domain (residues 1–183) found the *Klebsiella pneumoniae* adhesin MrkD, the *E. coli* adhesin F17-G (GafD) and *E. coli* FimH as the best-scoring homologues (*Z*-scores of 11.1, 10.8 and 6.8, respectively). Strikingly, despite their low sequence similarity (31 and 33%, respectively), MrkD and F17-G (GafD) show a high structural homology to the LpfD adhesin. Superimposition of the MrkD (PDB entry 3u4k), F17f-G (GafD; PDB entry 1zk5) and FimH (in both low-affinity and high-affinity conformations; PDB entries 3jwn and 2vco, respectively) lectin domains with $LpfD^{1–183}$ corroborates the overall match of LpfD with MrkD, F17f-G and FimH (Supplementary Fig. S2). The FimH adhesin binds terminal mannose residues in epithelial high-mannose glycoproteins. Mannose binding is determined by a defined receptor-binding pocket located at the distal tip of the lectin domain formed by the N-terminus and loops consisting of residues 45–54 and 135–142 (Fig. 6*b*; Supplementary Fig. S3). This region coincides with HSR1 in the LpfD adhesin (Fig. 6*b*). However, the local conformation in this region is markedly different in both adhesins (Supplementary Fig. S3). The equivalent secondary-structure elements in LpfD (L1, L2, L3 and the N-terminal β-strand) are disulfide-bonded and close off the binding pocket, as seen in FimH. Instead, they give rise to a shallow molecular surface that would be more compatible with the formation of protein–protein interactions or with the binding of more extended oligosaccharides, as is seen for the binding interactions of PapG or FedF in their globoside (GbO4) or type 1 blood group A interactions, respectively (Dodson *et al.*, 2001; Moonens *et al.*, 2012). To evaluate the role of HSR1 in LPF-mediated adhesion,

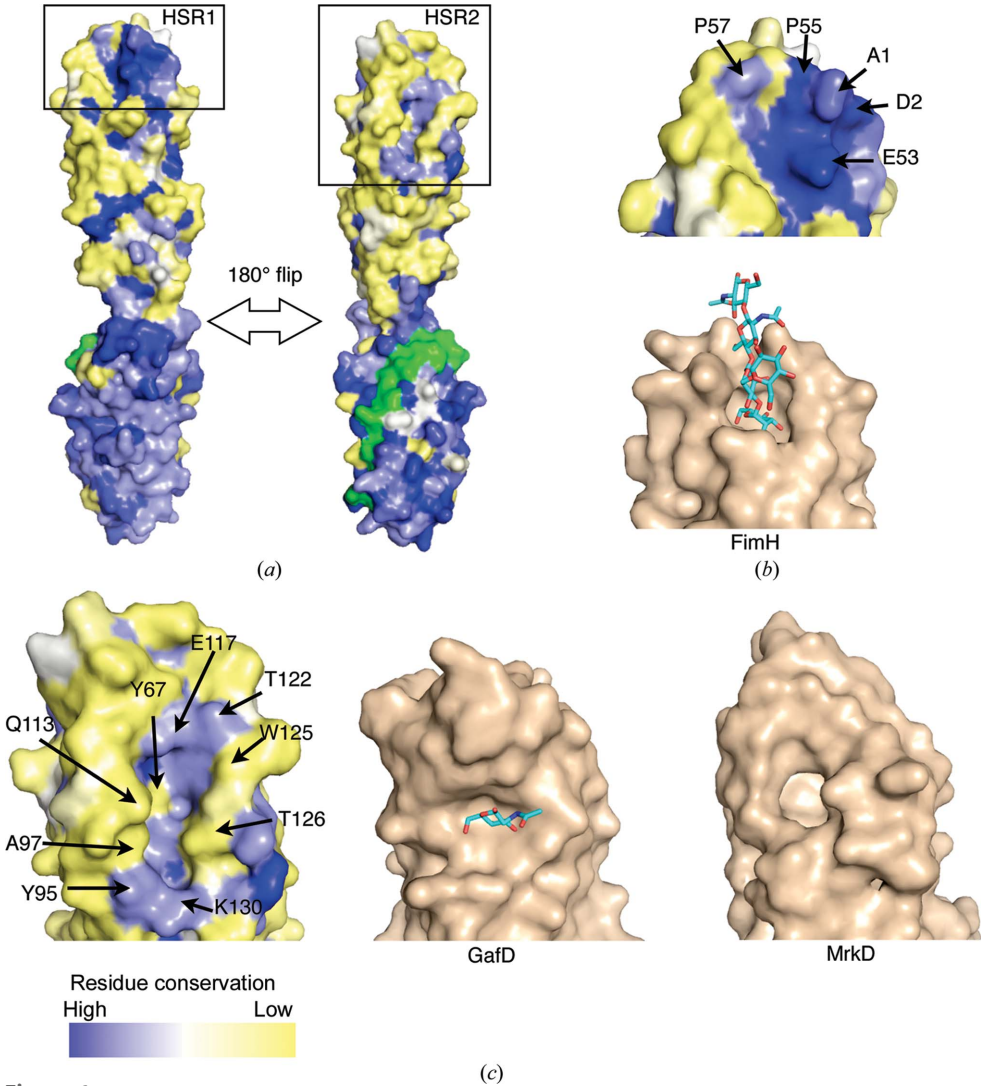


Figure 6
Sequence conservation in LpfD and comparison with the fimbrial adhesins FimH, F17-G (GafD) and MrkD. (*a*) Surface representation of full-length $LpfD_{ENte}$. The boxed regions are enlarged in (*b*) and (*c*). (*b*) Detailed view of HSR1 in the adhesin domain of LpfD and the corresponding binding pocket of FimH with oligomannose-3 (PDB entry 2vco). (*c*) Detailed view of HSR2 in the adhesin domain of LpfD and the corresponding binding pockets of F17-G (GafD) with GlcNAc (PDB entry 1zk5) and of MrkD (PDB entry 3u4k). All models are shown in an equivalent orientation after structural alignment. $LpfD_{ENte}$ is coloured blue/white/yellow according to residue conservation, with blue and yellow the highest and the lowest level of conservation, respectively. The LpfE N-terminal extension is coloured green.

we followed the binding of two LpfD mutants, E53A and E53K (Glu53 was chosen for its central location in the HSR1 region and associated surface pocket), to mouse ileal tissue sections (Supplementary Fig. S4). However, neither mutant was found to alter LpfD binding in our experimental setup.

In addition to HSR1, we evaluated the possible role of HSR2 in LPF adhesion. Remarkably, the overall structural homology between LpfD and the MrkD and GafD adhesin domains extends to the topology of the solvent-accessible surface of the molecules near HSR2. In particular, F17-G (GafD) and MrkD contain a glycan-binding pocket along the side of the lectin domain (Buts *et al.*, 2005; Rêgo *et al.*, 2012) that is in an equivalent position to the pocket delimited by Tyr67, Ala97, Gln113, Glu117, Trp125 and Thr126 revealed in the $LpfD_{ENte}$ structure. In F17-G (GafD) a co-crystal structure showed this pocket to bind a GlcNAc residue (Buts *et al.*, 2003, 2005) and in MrkD mutational disruption showed the pocket to be involved in type V collagen binding by the adhesin, in conjunction with a second collagen-binding region on the surface of the adhesin (Sebghati *et al.*, 1998; Rêgo *et al.*, 2012). However, a mutant $LpfD_{ENte}$ in which Gln113 was substituted by Ala failed to show a significant reduction in a fluorescence-based binding assay using mouse ileal tissue sections (Supplementary Fig. S4). Our observations of native-like ileal binding of the tested HSR1-localized (Glu53) and HSR2-localized (Gln113) mutants may argue against a critical role of either region in LPF-mediated adherence, although it cannot be excluded that mutational disruption of a single residue in the putative LpfD binding site is insufficient to give a phenotype in the immunofluorescence microscopy assay, where receptor-binding sites are present in high concentration. Of note, at least for MrkD, receptor binding was also found to be localized in two binding zones.

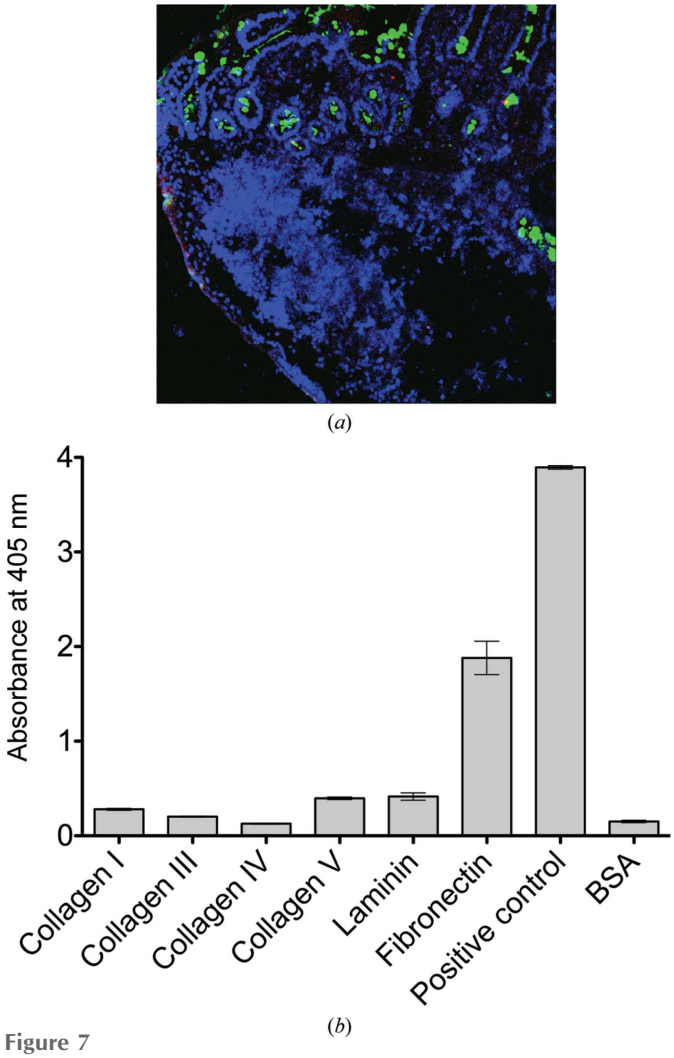


Figure 7
Assessment of the putative collagen-binding properties of $LpfD_{ENte}$. (*a*) Fluorescence microscopy of a murine small intestine section following collagenase pre-treatment prior to incubation with DyLight 650-labelled $LpfD_{ENte}$ (an untreated control is shown in Fig. 2*b*). (*b*) Binding of $LpfD_{ENte}$ to collagen types I, III, IV and V, laminin and fibronectin. The negative control was a microtitre plate well coated with PBS and blocked with BSA and the positive control was a microtitre plate well coated with $LpfD_{ENte}$. The columns and error bars represent the mean and the standard error of the mean, respectively, from three separate experiments (each performed in triplicate).

Given the striking structural similarity of LpfD to MrkD from *K. pneumoniae*, we sought to evaluate the possibility that LPF could bind intestinal collagen. Therefore, slides of mouse small intestine sections were pre-incubated with collagenase prior to the addition of DyLight 650-labelled $LpfD_{ENte}$ and fluorescence microscopy analysis. This pre-treatment with collagenase obliterated binding of $LpfD_{ENte}$ to the ileal tissue sections (Fig. 7*a*). Of note, collagen types I, III and V have previously been shown to be implicated in the symptomatic presentation of CD (Alexakis *et al.*, 2004). To determine which extracellular matrix (ECM) protein or proteins represent the LPF receptor in the ileal epithelium, *in vitro* binding of $LpfD_{ENte}$ to isolated human collagen types I, III, IV and V (type II collagen is restricted to cartilage), laminin and fibronectin was assessed by ELISA. LpfD strongly adhered to fibronectin, and showed weaker levels of binding to collagen V and laminin (Fig. 7*b*). No binding above the background was observed for collagens I, III or IV (Fig. 7*b*).

4. Conclusion

Despite the high level of interest in the study of CD pathogenesis, little is known about the molecular interplay with the bacteria implicated in CD onset and/or sustenance. AIEC have been identified as a commonly enriched *E. coli* strain in lesions. Long polar fimbriae in the model AIEC LF82 were found to play an important role in its attachment and invasion at the PP sites. Therefore, we sought a molecular and structural understanding of LPF-mediated adherence. We recombinantly produced the putative LPF adhesin LpfD as a soluble, self-complemented construct by C-terminal fusion with the N-terminal extension peptide of the LpfE pilus subunit. Using fluorescence microscopy, we show that LpfD is responsible for LPF binding to mouse intestinal tissue. No LpfD binding was observed in mouse bladder sections, arguing for an LPF-mediated tropism to the intestinal mucosa. Ileal LpfD binding was found to be sensitive to mouse anti-LpfD serum, but proved resistant to pre-treatment of the tissue sections with *O*- or *N*-glycosidases. In addition, interrogation of the CFG

glycan array version 5.1 with $LpfD_{ENte}$ did not reveal the presence of a glycan receptor in the array. Whereas LPF-mediated binding of AIEC LF82 to intact mouse or human ileal resections was previously found to show tropism to PP-localized M cells (Chassaing *et al.*, 2011), our binding studies with the purified LPF adhesin LpfD reveal a broader binding profile that includes the epithelial and subepithelial tissue throughout the ileum. Co-staining with anti-Muc2 antibodies showed that this LpfD binding is excluded from the mucus layer. In the healthy gut, a thick mucus layer forms a shielding barrier that prevents direct contact between the intestinal epithelium and the gut microbiota. At the PP, the mucus barrier is reduced and specialized epithelial cells, called M cells, now come into direct contact with the luminal microbiota.

Based on our data, we propose that the LpfD receptor is present throughout the ileal mucosa and that the apparent M-cell tropism for LPF-mediated bacterial adherence is a result of the reduced mucus shielding at the PP. Therefore, a compromised mucus layer could result in the exposure of the LPF receptor and lead to AIEC binding at sites of mucosal injury or depletion of goblet cell secretory function owing to inflammation or transformation (McCormick *et al.*, 1990; Kim & Ho, 2010). Recent studies showed that a compromised mucin layer in Muc2 mucin-deficient mice led to increased bacterial adherence and the spontaneous development of colitis (Van der Sluis *et al.*, 2006; Johansson *et al.*, 2008). LPF could thus play a major role in the increased AIEC adherence during inflammatory bowel disease. Our *in vitro* binding studies and LpfD staining patterns of mouse intestinal histo tissue sections suggest that one or more ECM proteins might be receptors for LPF-mediated bacterial adherence. The ELISA assays performed in this work using several major ECM proteins present in the intestinal epithelium showed a strong and direct *in vitro* binding of LpfD to fibronectin, as well as weaker binding to laminin and collagen type V. These proteins, as well as other ECM proteins, are generally localized in the epithelial basement membrane, where they are not available for interaction with luminal bacteria. However, during infection ECM components can be found partially degraded and exposed to the intestinal lumen, now becoming targets for the adherence of pathogenic bacteria (Vanlaere & Libert, 2009). Also, an apical secretion of fibronectin has been observed in human polarized intestinal epithelial cells exposed to physiological concentrations of adenosine, a ubiquitous pro-inflammatory signalling molecule. Under these conditions, fibronectin was accessible from the intestinal lumen and mediated the adherence and invasion of epithelial cells by *S. enterica* serovar Typhimurium and enteroaggregative *E. coli* (Walia *et al.*, 2004; Farfan *et al.*, 2008).

An earlier study by Fitzherny and coworkers found that LPF in EHEC O157:H7 (*i.e. Lpf*1–3) targeted the bacteria to the PP in human *in vitro* gut tissue cultures, but noted that deletion of either or both of the *lpfA* genes maintained PP binding and additionally expanded the tropism of the bacteria to include the small intestine (Fitzhenry *et al.*, 2006). The potential ability of *lpf* expression to upregulate or downregulate the expression of other fimbrial operons is suggested as a possible explanation of this feature. Here, using purified LpfD adhesin ($LpfD_{ENte}$) from AIEC strain LF82, we were able to take a first step in the identification of a receptor for AIEC LPF (*i.e. Lpf*1–2). We found that collagenase pre-treatment of ileal tissue sections resulted in a complete loss of $LpfD_{ENte}$ binding. Direct binding assays showed that LpfD binds to fibronectin and to a lesser extent to laminin and collagen type V. Strikingly, the LpfD crystal structure reveals a close structural homology to the *K. pneumoniae* adhesin MrkD, which is known to bind type V collagen (Tarkkanen *et al.*, 1990). The structural homology between MrkD and LpfD includes the presence of a pocket on the side of the adhesin domain. In MrkD, collagen binding was found to depend on the combined presence of the conserved binding pocket and a second region on the side of the adhesin domain (Rêgo *et al.*, 2012). The results presented here suggest further investigations into the identity of the LpfD receptor should include the direction of the ECM proteins, especially since collagen accumulation has been observed in the intestine during the course of CD (Alexakis *et al.*, 2004).

Acknowledgements

We thank the late Arlette Darfeuille-Michaud for providing us with the LF82 *E. coli* strain; and Amanda Gonçalves at the VIB imaging core for her assistance with the microscopy experiments. We also wish to acknowledge the support staff of beamlines PROXIMA 1 at SOLEIL and I24 at the Diamond synchrotron, as well as the participation of the Protein–Glycan Interaction Resource of the CFG and the supporting grant R24 GM098791. This work was supported by the FWO under grant G030411N and the Hercules Foundation through equipment grant UABR/09/005.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Alexakis, C., Caruelle, J. P., Sezeur, A., Cosnes, J., Gendre, J. P., Mosnier, H., Beaugerie, L., Gallot, D., Malafosse, M., Barritault, D. & Kern, P. (2004). *Gut*, **53**, 85–90.
Barnhart, M. M., Pinkner, J. S., Soto, G. E., Sauer, F. G., Langermann, S., Waksman, G., Frieden, C. & Hultgren, S. J. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 7709–7714.
Barnich, N., Carvalho, F. A., Glasser, A.-L., Darcha, C., Jantscheff, P., Allez, M., Peeters, H., Bommelaer, G., Desreumaux, P., Colombel, J.-F. & Darfeuille-Michaud, A. (2007). *J. Clin. Invest.* **117**, 1566–1574.
Bäumler, A. J. & Heffron, F. (1995). *J. Bacteriol.* **177**, 2087–2097.
Bäumler, A. J., Tsolis, R. M. & Heffron, F. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 279–283.
Binder, V. & Orholm, M. (1996). *Neth. J. Med.* **48**, 53–56.
Boudeau, J., Glasser, A. L., Masseret, E., Joly, B. & Darfeuille-Michaud, A. (1999). *Infect. Immun.* **67**, 4499–4509.
Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K. & Jones, D. T. (2013). *Nucleic Acids Res.* **41**, W349–W357.
Buts, L., Bouckaert, J., De Genst, E., Loris, R., Oscarson, S., Lahmann, M., Messens, J., Brosens, E., Wyns, L. & De Greve, H. (2003). *Mol. Microbiol.* **49**, 705–715.
Buts, L., Wellens, A., Van Molle, I., Wyns, L., Loris, R., Lahmann, M., Oscarson, S., De Greve, H. & Bouckaert, J. (2005). *Acta Cryst.* D**61**, 1149–1159.
Cario, E. & Podolsky, D. K. (2000). *Infect. Immun.* **68**, 7010–7017.

Chassaing, B., Rolhion, N., de Vallée, A., Salim, S. Y., Prorok-Hamon, M., Neut, C., Campbell, B. J., Söderholm, J. D., Hugot, J.-P., Colombel, J.-F. & Darfeuille-Michaud, A. (2011). *J. Clin. Invest.* **121**, 966–975.
Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J. & Knight, S. D. (1999). *Science*, **285**, 1061–1066.
Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. (2004). *Genome Res.* **14**, 1188–1190.
Darfeuille-Michaud, A., Neut, C., Barnich, N., Lederman, E., Di Martino, P., Desreumaux, P., Gambiez, L., Joly, B., Cortot, A. & Colombel, J.-F. (1998). *Gastroenterology*, **115**, 1405–1413.
De Greve, H., Wyns, L. & Bouckaert, J. (2007). *Curr. Opin. Struct. Biol.* **17**, 506–512.
Dodson, K. W., Pinkner, J. S., Rose, T., Magnusson, G., Hultgren, S. J. & Waksman, G. (2001). *Cell*, **105**, 733–743.
Eckburg, P. B. & Relman, D. A. (2007). *Clin. Infect. Dis.* **44**, 256–262.
Edgar, R. C. (2004). *Nucleic Acids Res.* **32**, 1792–1797.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Farfan, M. J., Cantero, L., Vergara, A., Vidal, R. & Torres, A. G. (2013). *Vet. Immunol. Immunopathol.* **152**, 126–131.
Farfan, M. J., Cantero, L., Vidal, R., Botkin, D. J. & Torres, A. G. (2011). *Infect. Immun.* **79**, 3744–3750.
Farfan, M. J., Inman, K. G. & Nataro, J. P. (2008). *Infect. Immun.* **76**, 4378–4384.
Fiocchi, C. (2005). *Inflammatory Bowel Disease: From Bench to Bedside*, pp. 101–120. Boston: Springer.
Fitzhenry, R., Dahan, S., Torres, A. G., Chong, Y., Heuschkel, R., Murch, S. H., Thomson, M., Kaper, J. B., Frankel, G. & Phillips, A. D. (2006). *Microbes Infect.* **8**, 1741–1749.
Frank, D. N., St Amand, A. L., Feldman, R. A., Boedeker, E. C., Harpaz, N. & Pace, N. R. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 13780–13785.
Ghosh, D., Erman, M., Sawicki, M., Lala, P., Weeks, D. R., Li, N., Pangborn, W., Thiel, D. J., Jörnvall, H., Gutierrez, R. & Eyzaguirre, J. (1999). *Acta Cryst.* D**55**, 779–784.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
Johansson, M. E. V. & Hansson, G. C. (2012). *Methods Mol. Biol.* **842**, 229–235.
Johansson, M. E. V., Phillipson, M., Petersson, J., Velcich, A., Holm, L. & Hansson, G. C. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 15064–15069.
Jordan, D. M., Cornick, N., Torres, A. G., Dean-Nystrom, E. A., Kaper, J. B. & Moon, H. W. (2004). *Infect. Immun.* **72**, 6168–6171.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kim, Y. S. & Ho, S. B. (2010). *Curr. Gastroenterol. Rep.* **12**, 319–330.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
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. (2010). *Cell*, **141**, 645–655.
Liu, Y., Esser, L., Interlandi, G., Kisiela, D. I., Tchesnokova, V., Thomas, W. E., Sokurenko, E., Xia, D. & Savarino, S. J. (2013). *J. Biol. Chem.* **288**, 9993–10001.
Macpherson, A., Khoo, U. Y., Forgacs, I., Philpott-Howard, J. & Bjarnason, I. (1996). *Gut*, **38**, 365–375.
Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P., Roca, J. & Dore, J. (2006). *Gut*, **55**, 205–211.
McCormick, D. A., Horton, L. W. & Mee, A. S. (1990). *J. Clin. Pathol.* **43**, 143–146.
McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013). *Nucleic Acids Res.* **41**, W597–W600.
Miquel, S. *et al.* (2009). *PLoS One*, **5**, e12714.
Moonens, K., Bouckaert, J., Coddens, A., Tran, T., Panjikar, S., De Kerpel, M., Cox, E., Remaut, H. & De Greve, H. (2012). *Mol. Microbiol.* **86**, 82–95.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Puorger, C., Eidam, O., Capitani, G., Erilov, D., Grütter, M. G. & Glockshuber, R. (2008). *Structure*, **16**, 631–642.
Rêgo, A. T., Johnson, J. G., Geibel, S., Enguita, F. J., Clegg, S. & Waksman, G. (2012). *Mol. Microbiol.* **86**, 882–893.
Remaut, H., Rose, R. J., Hannan, T. J., Hultgren, S. J., Radford, S. E., Ashcroft, A. E. & Waksman, G. (2006). *Mol. Cell*, **22**, 831–842.
Sartor, R. B. (2010). *Gastroenterology*, **139**, 1816–1819.
Sauer, F. G., Pinkner, J. S., Waksman, G. & Hultgren, S. J. (2002). *Cell*, **111**, 543–551.
Sebghati, T. A., Korhonen, T. K., Hornick, D. B. & Clegg, S. (1998). *Infect. Immun.* **66**, 2887–2894.
Shanahan, F. (2002). *Lancet*, **359**, 62–69.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Stahlhut, S. G., Chattopadhyay, S., Kisiela, D. I., Hvidtfeldt, K., Clegg, S., Struve, C., Sokurenko, E. V. & Krogfelt, K. A. (2013). *J. Bacteriol.* **195**, 5602–5613.
Tarkkanen, A. M., Allen, B. L., Westerlund, B., Holthöfer, H., Kuusela, P., Risteli, L., Clegg, S. & Korhonen, T. K. (1990). *Mol. Microbiol.* **4**, 1353–1361.
Thomas, W. E., Trintchina, E., Forero, M., Vogel, V. & Sokurenko, E. V. (2002). *Cell*, **109**, 913–923.
Torres, A. G., Blanco, M., Valenzuela, P., Slater, T. M., Patel, S. D., Dahbi, G., López, C., Barriga, X. F., Blanco, J. E., Gomes, T. A. T., Vidal, R. & Blanco, J. (2009). *J. Clin. Microbiol.* **47**, 2442–2451.
Torres, A. G., Giron, J. A., Perna, N. T., Burland, V., Blattner, F. R., Avelino-Flores, F. & Kaper, J. B. (2002). *Infect. Immun.* **70**, 5416–5427.
Torres, A. G., Kanack, K. J., Tutt, C. B., Popov, V. & Kaper, J. B. (2004). *FEMS Microbiol. Lett.* **238**, 333–344.
Torres, A. G., Milflores-Flores, L., Garcia-Gallegos, J. G., Patel, S. D., Best, A., La Ragione, R. M., Martínez-Laguna, Y. & Woodward, M. J. (2007). *Int. J. Med. Microbiol.* **297**, 177–185.
Van der Sluis, M., De Koning, B. A. E., De Bruijn, A. C. J. M., Velcich, A., Meijerink, J. P. P., Van Goudoever, J. B., Büller, H. A., Dekker, J., Van Seuningen, I., Renes, I. B. & Einerhand, A. W. C. (2006). *Gastroenterology*, **131**, 117–129.
Vanlaere, I. & Libert, C. (2009). *Clin. Microbiol. Rev.* **22**, 224–239.
Walia, B., Castaneda, F. E., Wang, L., Kolachala, V. L., Bajaj, R., Roman, J., Merlin, D., Gewirtz, A. T. & Sitaraman, S. V. (2004). *Biochem. J.* **382**, 589–596.
Willing, B. P., Dicksved, J., Halfvarson, J., Andersson, A. F., Lucio, M., Zheng, Z., Järnerot, G., Tysk, C., Jansson, J. K. & Engstrand, L. (2010). *Gastroenterology*, **139**, 1844–1854.e1.
Yakovenko, O., Sharma, S., Forero, M., Tchesnokova, V., Aprikian, P., Kidd, B., Mach, A., Vogel, V., Sokurenko, E. & Thomas, W. E. (2008). *J. Biol. Chem.* **283**, 11596–11605.
Zavialov, A. V., Berglund, J., Pudney, A. F., Fooks, L. J., Ibrahim, T. M., MacIntyre, S. & Knight, S. D. (2003). *Cell*, **113**, 587–596.