The structure of *E. coli* IgG-binding protein D suggests a general model for bending and binding in trimeric autotransporter adhesins

Jack C. Leo[1‡], Andrzej Lyskowski[2‡#], Katarina Hattula[2], Marcus D. Hartmann[3], Heinz Schwarz[3], Sarah J. Butcher[2], Dirk Linke[3], Andrei N. Lupas[3] and Adrian Goldman[1*]

[1]Macromolecular crystallography group, Institute of Biotechnology, [2]Structural virology group, Institute of Biotechnology
University of Helsinki, 00014 Helsinki, Finland

[3]Max-Planck Institute for Developmental Biology, Department of Protein Evolution, D-72076 Tübingen, Germany

[‡]These authors made an equal contribution to the manuscript

[#]Present address: ACIB c/o Institute of Molecular Biosciences, University of Graz, A-8010 Graz, Austria

manuscript

[*]Corresponding author, tel: +358 9 191 58923, fax: +358 9 191 59940, e-mail: adrian.goldman@helsinki.fi

Running title: Crystal structure of EibD

SUMMARY

The *Escherichia coli* Ig-binding (Eib) proteins are trimeric autotransporter adhesins (TAAs) and receptors for IgG Fc. We present the structure of a large fragment of the passenger domain of EibD, the first TAA structure to have both a YadA-like head domain and the entire coiled-coil stalk. The stalk begins as a right-handed superhelix, but switches handedness halfway down. An unexpected, novel β-minidomain joins the two and inserts a ~120° rotation such that there is no net twist between the beginning and end of the stalk. This may be important in folding and autotransport. The surprisingly large cavities we found in EibD and other TAAs may explain how TAAs bend to bind their ligands. We identified how IgA and IgG bind and modelled the EibD-IgG Fc complex. We further show that EibD promotes autoagglutination and biofilm formation and forms a fibrillar layer covering the cell surface making zipper-like contacts between cells.

Highlights:

- The first TAA structure containing both a head domain and a complete stalk
- A β-minidomain mediates the change from right-handed to left-handed supercoiling
- The binding sites for IgG and IgA are in the stalk
- Large cavities in the stalk may explain how bending of TAAs occurs

INTRODUCTION

Trimeric autotransporter adhesins (TAAs) are a family of virulence-related, secreted proteins in Gram-negative bacteria that form the type V_c secretion pathway (Dautin and Bernstein, 2007; Linke et al., 2006). They are all obligate homotrimers, and the conserved C-terminal sequence forms the translocation unit consisting of a 12-stranded β-barrel outer membrane anchor and an α-helical linker (Meng et al., 2006). The more varied N-terminal portions, the passenger domains, are secreted into the extracellular space, and typically consist of a globular head domain connected to the anchor by a coiled-coil stalk (Hoiczyk et al., 2000).

The immunoglobulin (Ig)-binding proteins from *E. coli* (Eibs) are a subgroup of TAAs. Though first detected in commensal strains (Sandt et al., 1997), they are relatively common in intimin-negative, shiga toxin-producing enterohaemorrhagic *E. coli* (EHEC) strains (Merkel et al., 2010). The group currently comprises six characterised members: EibA, EibC, EibD, EibE, EibF, EibG and EibH (Sandt and Hill, 2000; Sandt and Hill, 2001; Lu et al., 2006; Merkel et al., 2010) (we use the notation EibACDEFG to designate multiple members of the group). The Eibs are receptors for IgAs and IgGs and bind to the Fc portion of Igs non-immunologically (Leo and Goldman, 2009). They also increase the serum resistance of the bacteria and are involved in adherence to epithelial cells (Sandt and Hill, 2000; Sandt and Hill, 2001; Lu et al., 2006). We have shown that the Eibs bind IgG Fc with affinities in the 100 nM range (Leo and Goldman, 2009), while only EibCDFG (*i.e.* not EibAE) bind IgA with affinities between 70 nM (EibF) and 500 nM (EibC) (Sandt and Hill, 2001; Lu et al., 2006; Leo and Goldman, 2009).

This study presents, in two overlapping parts, the crystal structure of the head and stalk (residues 161-418 and 391-440) of the EibD passenger domain. This region, containing the binding activity towards both IgG and IgA, consists of a YadA-like β-roll (Nummelin et al., 2004) followed by the extended coiled-coil stalk connecting the head to the membrane anchor. It is the first example of a head domain combined with a complete stalk and includes the striking transition from right-handed to canonical left-handed supercoiling characteristic of many TAA stalks (Hernandez Alvarez et al., 2010). By detailed mutagenesis analysis, we have pinpointed the IgG and IgA-binding activities in the stalk. Staphylococcal protein A (SpA) could block EibD binding to IgG Fc, suggesting that they bind to the same region of IgG Fc. This has allowed us to model the complex. The model suggests that EibD can simultaneously bind up to three IgG molecules, which is in good agreement with experimental results. We further examined the expression of EibD on the cell surface by electron microscopy (EM) and demonstrated two new activities for EibD suggesting a role for the Eibs in virulence: autoagglutination and biofilm formation.

RESULTS

Structure of EibD$_{161-440}$. We crystallised two overlapping fragments of EibD, EibD$_{161-418}$ and EibD$_{391-440}$, as described in more detail in Supplemental Experimental Procedures. The smaller fragment was produced with GCN4 adaptor sequences at either end to promote stable trimerisation (Hernandez Alvarez et al., 2008). R_{work}/R_{free} for the two structures were 15.9/19.6% and 26/30.8% (Table 1).

EibD$_{161-440}$ is an elongated lollipop-like structure with a globular head domain, a neck and an extended stalk (Figure 1A; see Figure S1A for a stereo image). The N-terminal part of EibD$_{161-440}$ is a left-handed parallel β-roll (LPBR) domain very similar to the head of YadA (Figure S1B). In YadA, the LPBR consists of repeats with the consensus motif NSVAIGXXS (Nummelin et al., 2004), and EibA has eight such repeats (Figure 1). These motifs have a dual structural role, forming both the hydrophobic core of the β-roll itself as well as the shared hydrophobic core of the trimer (Nummelin et al., 2004). However, in contrast to YadA, the EibA motifs have the consensus sequence XSTALGXXA (Figure S1C). A further difference is the large loop protruding from EibD between the last two LPBR repeats (Figure S1B). This loop is a part of the neck, a connecting element between the LBPR domain and the stalk (Nummelin et al., 2004) that rotates the rest of the polypeptide by 120° clockwise (Figure S1D) so that the following helix in the stalk is located directly beneath the adjacent monomer.

The N-terminal part of the stalk is a right-handed coiled-coil (RHcc) with hendecad periodicity, *i.e.* 11 residues per three turns (Lupas and Gruber, 2005). It terminates in

a β-sheet minidomain (residues 350-371) that we call the saddle (Figure 1). The saddle is followed by a canonical left-handed coiled-coil (LHcc) with heptad periodicity (3.5 residues per turn). The saddle is responsible for a second 120° rotation that is in the opposite direction to the one introduced by the neck (Figure S1E). The saddle is only 22 residues long and consists of three anti-parallel β-sheets, without any significant similarity to other known structures. Its presence was not obvious from sequence analysis, as the residue pattern is compatible with a continuous structure that retains coiled-coil periodicity. The saddle is stabilized by an ion network at the 'exit' and 're-entry' points (Figure 1B), similar to the YadA neck region (Nummelin et al., 2004).

The core of the LHcc is surprisingly hydrophilic. It contains five chloride ions, a water molecule and three glutamines that point inwards. In four cases, the chloride is coordinated by asparagine residues in position *d* of the coiled-coil heptad (the seven positions of the heptad are designated *a*-*g*) (Figure S1F). Though hydrophobic residues are normally expected, Asn in the *d* position ("N@*d*") occurs in a number of trimeric autotransporters (Hartmann et al., 2009), and binds anions in both SadA (Hartmann et al., 2009) and UspA1 (Conners et al., 2008). Novel hydrophilic motifs in the EibD stalk are the fifth chloride, which is coordinated by threonine residues in an RTD motif where the T is in position *a* (Figure S1G), the water molecule coordinated by H421 in position *d* (Figure S1H), and the Q428s at position *d* that form a three-way hydrogen bond (Figure S1I). RxD motifs are relatively common in LHccs of TAAs and form polar networks at the interfaces between monomers (Hernandez Alvarez et al., 2010). Usually, the middle position (*a* in the coiled-coil) is occupied by a hydrophobic aliphatic residue, so threonine in this position is somewhat

surprising. In EibD, in addition to the chloride-coordinating ^{382}RTDRID^{387} motif, there are two more RxD motifs between residues 393-397 and 421-426, the second including H421. In RHccs, YxD motifs play an analogous role to RxD, with threonine commonly in the middle position (Hernandez Alvarez et al., 2010). There are two YxD motifs in the RHcc: ^{317}YDDVK^{321} and ^{334}YTDQK^{338}. YxD motifs normally bracket the *l* position of pentadecad coiled-coils, but in EibD the middle residue of the second YxD is in the *a* position, the geometric equivalent in hendecad coiled-coils. Surprisingly, the core residue of the first motif (V) is between the second D and K, in position *h*. In contrast to the usual Yx(*l*)D arrangement, abundant in TAA stalks (Szczesny and Lupas, 2008), this second configuration is rare.

In addition to the water- and chloride-binding cavities in the LHcc, there are others in both the right- and left-handed part of the stalk. Most of these are small (solvent-accessible volume $\leq$ 5 Å^3), but two are significantly larger, with solvent-accessible volumes of 11 Å^3 and 46 Å^3 (calculated using the CASTp server (Dundas et al., 2006)). Both large cavities, located in the RHcc (Figure 1C), are lined by hydrophobic groups and devoid of any extra electron density at one σ. We discuss the potential implications of this for function below.

Immunoglobulin binding studies. EibD$_{161\text{-}418}$ bound both IgA and IgG (Figure S2A). To map the Ig-binding regions more precisely, we produced four C-terminal truncations of EibD$_{161\text{-}418}$ and a number of double and triple point mutants, where surface residues were changed to alanine (Table 2). Ig binding to all of the point mutants was measured by dot blot and by surface plasmon resonance (SPR), and

conformational changes were monitored by far-UV circular dichroism (CD) spectroscopy (Figure S2C).

The truncations indicate that the binding site for IgG is at the C-terminus of the stalk between residues 384 and 418, and that the binding site for IgA lies between residues 329 and 344 (Figure S2B). Among the point mutants (Table 2) we included some residues outside the putative binding sites, as it was also possible that the truncations deleted only part of the binding regions. Qualitatively (Figure 2A), it is clear that mutant K332A-K333A-Y334A had a severe effect on IgA binding, practically abolishing the signal. The only mutant to have a large effect on IgG binding was D387A-Y388A.

The apparent affinities for Ig (K_d(app)) determined for the mutants by SPR were mostly similar to wild type (wt), *i.e.* 50-100 nM for IgG and 100-200 nM for IgA (Table 2). However, we were unable to estimate a reliable K_d(app) for the mutant K332A-K333A-Y334A binding to IgA (Figure 2B). For D387-Y388, we obtained an K_d(app) of 288 nM for IgG. This, however, cannot be considered a reliable estimate, because of the peculiar shape of the curves, which are far from equilibrium (Figure 2B). The true affinity is therefore likely lower. These results are consistent with the results from the dot blots, suggesting that these mutations have a deleterious effect on the binding sites for IgA and IgG, respectively. These mutants also abolished binding to the respective Igs (IgA for K332A-K333A-Y334A and IgG for D387A-Y388A) when expressed on the cell surface (Figure S3B). The K_d(app) for the T415A-R416A-T17A mutant had a large SEM, but it also bound weakly by dot blot and had a somewhat altered CD spectrum (Figure S2C), suggesting that it causes some subtle

structural change that may be propagated along the coiled-coil. These data show that IgA and IgG binding have separate binding sites on the coiled-coil, one in the RHcc and the other in the LHcc.

To determine the role of the saddle in Ig binding, we produced two deletion mutants: Δ350-372 and Δ352-372. The first retains the periodicity of the coiled-coil, while the second introduces a two-residue insertion, which could cause more prominent structural alterations (Lupas and Gruber, 2005). Neither deletion caused any major conformational changes, though Δ352-372 had a larger effect than Δ350-372 (Figure S2C). Δ350-372 seemed to improve binding to Igs and Δ352-372 was comparable to the wt by dot blot (Figure 2A) – but both weakened binding in the SPR experiments (Table 2, Figure 2B). The most likely explanation for this paradox is the differences in experimental set up: in the dot blot experiments, the EibD derivatives were immobilised, whereas in the SPR assays the antibodies were. The lack of the saddle may destabilise the structure of the EibD stalk, consistent with our earlier results showing that the stalk alone binds poorly to Igs (Figure S2). If the stalk is to some degree labile, as suggested by sensitive size exclusion chromatography coupled to static light scattering and refractive index measurements (Figure S2E), the absence of the saddle may cause the coiled-coil to have a heightened tendency to dissociate in solution. Immobilising the stalk deletion mutants would stabilise the coiled-coil.

Model of EibD binding to Fc. SpA and streptococcal protein G (SpG) are IgG-binding proteins which both target the same epitope: the exposed connector between the constant domains C_H2 and C_H3 in IgG Fc (Tashiro and Montelione, 1995). We showed (Figure 3A) that SpA had a concentration-dependent inhibitory effect on the

binding of HRP-labelled IgG to immobilised EibD and SpG, suggesting that EibD also binds to the same site in IgG Fc. SpA is a three-helix bundle, and the stalk of EibD is similarly composed of three helices. Therefore, we modelled EibD bound to IgG Fc based on the SpA-Fc crystal structure (Figure 3B). The model suggests that the EibD IgG binding site could accommodate the connector region of Fc. Furthermore, three molecules of Fc can bind to EibD simultaneously (Figure 3C), consistent with previous experimental data, which suggested Eibs can bind at least two IgGs at the same time (Leo and Goldman, 2009).

EibD expressed on the cell surface. In transmission electron micrographs (TEM), EibD-expressing cells clearly show an additional layer on the cell surface (Figure 4, upper panel). This is reminiscent of YadA, which also coats the cell at high density and forms a "quasi-periplasmic space" on the outside of the cell between the headgroups and the outer membrane (Hoiczyk et al., 2000). In scanning electron micrographs (SEM), the surface of EibD-expressing bacteria appeared ruffled, in contrast to the smooth surfaces of control cells (Figure 4, lower panel insets).

When culturing EibD-expressing bacteria, we noticed that they tended to flocculate (Figure 5A). Autoagglutination is typical for YadA and other TAAs (Linke et al., 2006). To show quantitatively that EibD mediates autoagglutination, we measured the reduction in optical density at the surface of static cultures (Figure 5B). EibD-expressing bacteria settled rapidly at the bottom of the tubes, whereas the vector control culture remained turbid for much longer. The self-associating activity of EibD was also evident in electron micrographs (Figure 4). The cell surface protrusions form tight zipper-like contacts (Figure 4, upper panel inset, and Figure S3A), similar to the

arrangement seen for autoagglutinating YadA (Hoiczyk et al., 2000). These contacts were strong enough to tear off outer membrane fragments from neighbouring cells leading to lysis (upper panels in Figure 4 and Figure S3A). In SE micrographs, the EibD-expressing bacteria clearly formed aggregates (Figure 4, lower panels). We believe that EibD also promotes aggregation under physiological conditions as well, as EibG has been shown to mediate chain-like aggregation (Lu et al., 2006).

Aggregation and adhesion are among the first steps in biofilm formation (Dunne, 2002). We therefore tested whether EibD could mediate the formation of biofilms when expressed at the cell surface. In a crystal violet staining assay, EibD expression led to intense staining of adherent bacteria, whereas the vector control was barely above background (Figure 5C). EibD - and presumably other Eibs like EibG - does promote the formation of biofilm, which may be biologically important.

Comparison of EibD with other members of the Eib family. Alignment of the Eib sequences (Figure 6) shows that the saddle is conserved in all Eibs, as is the LHcc containing the IgG binding site (Figure 6A). Interestingly, Ala frequently replaces the Tyr (Y388 in EibD) at the IgG binding site, suggesting that the crucial residue for binding is the completely conserved D387. In EibCD, the IgA and IgG binding sites are situated relatively close to each other (~50 Å), explaining why we observed that IgG binding interferes with IgA binding (Leo and Goldman, 2009).

The IgA binding regions of EibCD are identical (Figure 6B). Intriguingly, a region with 80% identity is present in EibEG, though EibE does not bind IgA (Sandt and Hill, 2001). However, the IgA binding motif ^{332}KKY^{334} of EibCD is replaced by REY

in EibEG. The change K→E is presumably why EibEG bind IgA very poorly (Sandt and Hill, 2001; Lu et al., 2006).

Conversely, the IgA binding motif KKY is similar to a previously identified repeat in EibF, which contains the sequence NKY (Figure 6B) (Sandt and Hill, 2001). This is also consistent with the central K being essential for IgA binding and explains why EibF binds IgA tightest of all the Eibs (Leo and Goldman, 2009).

Discussion

TAA architecture and the mechanism of bending. All of the current TAA structures are of pieces; this is the first one to contain both a head domain and a complete stalk. It validates our earlier work (Hernandez Alvarez et al., 2008) that structures of individual fragments of TAA head domains and stalks can be combined to generate a complete model – if the granularity of the fragments is correct. However, it also shows the limits of such an approach; you find only what you expect to find. The switch from a RHcc into a left-handed one in EibD was expected to be continuous as in YadA (Hernandez Alvarez et al., 2010), where the transition proceeds *via* local unwinding of the supercoil. There was no sequence evidence for the saddle; we would not have been aware of its existence had we not solved the entire structure.

What, then, is the saddle's role? It does not seem to be required for binding to either IgA or IgG, though deleting it does have some effect (Table 2; Figure 2). An intriguing feature of the region (Figure 1) is that the ~120° rotation is in the opposite

sense to that imposed by the neck. As a result of this and of the equivalent superhelical twists of the right-handed and left-handed regions, there is essentially no twist between the top and bottom of the stalk. This may be an advantage in folding, if the passenger domain, as is generally believed, has to pass through the constricted pore of the β-barrel translocator domain (Dautin and Bernstein, 2007). Overall, there are two 120° rotations of opposite hand in the EibD structure (Figure 1); the YadA head-stalk (Nummelin et al., 2004) has exactly the same arrangement, but with the rotations above the LPBR and in the neck region. Having an even number of opposite-handed 120° rotations could reduce torsional strain during autotransport and folding because there would be no superhelical twisting that would have to be unwound by counter-rotation of the barrel in the membrane.

TAAs are extended, rod-like structures, and bending of TAAs appears to be important for function. The stalk of UspA1 is deformable (Conners et al., 2008); kinks and bends are seen in electron micrographs of EmaA and BadA (Szczesny et al., 2008; Yu et al., 2008); and bending has been postulated to be important in the binding of YadA to collagen (Nummelin et al., 2004) and of Hia to its cellular receptor (Meng et al., 2008). The question is, how? To our surprise, we noticed a large number of cavities in the EibD stalk, two of them of significant size (>10 $Å^3$ accessible-surface volume). We therefore examined the only other TAA stalk structures, UspA1 (2QIH; Conners et al., 2008) and YadA (3H7X; Hernandez Alvarez et al., 2010; Koretke et al., 2006). Both contained large cavities: the proposed bending site of UspA1 is at the stutter in the coiled-coil stalk (Conners et al., 2008), and there is a 78 $Å^3$ cavity at this position. YadA has a 14 $Å^3$ cavity at the RHcc to LHcc transition (three times the volume of the largest cavity in the EibD saddle.) We propose that deformability induced by

cavities may be a universal mechanism used by TAAs to allow bending. The result also implies that different TAAs bend at different places.

Ig binding and biological implications. The binding of IgG by SpA and SpG prevents the activation of the classical pathway of complement by blocking the Fc binding site for C1q (Lambris et al., 2008; Nitsche-Schmitz et al., 2007). Fc binding by Eibs presumably performs the same function and contributes to serum resistance. The serum resistance activity of YadA maps to the coiled-coil stalk (Biedzka-Sarek et al., 2008), and the stalks of other trimeric autotransporters also may harbour protective activities against host defences (Linke et al., 2006), suggesting this is optimal position for protecting the bacterium. Our model of the EibD:Fc complex shows that three Fc monomers can bind to EibD simultaneously, consistent with previous data (Leo and Goldman, 2009). This is also only the second example where there has been enough structural and interaction data to propose a model for how a TAA might interact with its cognate ligand, the previous one being YadA bound to a collagen-like peptide (Nummelin et al., 2004; Leo et al., 2010). In contrast to YadA, where the collagen-binding activity is located in the LPBR domain, our EibD:Fc complex model is the first to show binding in the stalk. The model has various plausible interactions, chief of which is that sidechain rotations alone of Fc H310 and H435 bring them within 4-5 Å of the conserved EibD D387 (Figure 3D).

Biological function of Eibs. Binding of Igs, autoaggregation and biofilm formation all suggest that Eibs are virulence factors. Five of the characterised members of the group are from apparently harmless commensal ECOR strains (Sandt and Hill, 2000; Sandt and Hill, 2001), and as yet uncharacterised *eib* genes have been detected in

other commensal *E. coli* strains (*e.g.* ED1a; Touchon et al., 2009). In these strains, the *eib* genes are located in prophage sequences (Sandt and Hill, 2000). The commensal strains may thus act as a reservoir for these potential virulence factors from which the genes, possibly after recombination to produce new TAA variants, are passed on to more pathogenic strains. The properties conferred by Eibs would be advantageous for commensal strains as well, suggesting the Eibs could be "commensalism factors" that promote stable colonisation of the host intestinal tract.

Consistent with this idea, *eib* genes occur in pathogenic and multidrug-resistant *E. coli* strains (Bardiau et al., 2010; Touchon et al., 2009; Fricke et al., 2008). EibG is present in 15% of *eae*-negative EHEC strains (Merkel et al., 2010). It mediates an unusual chain-like adhesion phenotype (Lu et al., 2006) that has been suggested to contribute to persistence in the host (Merkel et al., 2010). In addition, the conservation of IgG binding in particular also implies that the selection pressure on the Eib family has been exposure to serum – and consequently that they are virulence factors. EibD shares 58-62% sequence identity with the various EibG variants (Merkel et al., 2010), so our results provide a first structural model for a non-fimbrial adhesin of EHEC O:91 strains. Experimental validation of our EibD:Fc model, as well investigation of the mechanism of EibG in EHEC, are our chief goals.

EXPERIMENTAL PROCEDURES

Cloning and mutagenesis of EibD constructs, protein expression and purification. The cloning, production and expression of the full-length EibD passenger domain ($EibD_{27-441}$) have been described before (Leo and Goldman, 2009). Cloning, production and purification of other EibD fragments was carried out using standard procedures as detailed in Supplemental Experimental Procedures. Mutants were produced using the QuikChange® protocol from Stratagene. Primer sequences are listed in Table S1.

Crystallisation and data collection. Crystallisation of the fragment $EibD_{161-418}$ (which bound both IgA and IgG (Figure S1) and included the YadA-like region) was performed by vapour diffusion at 20 °C. For initial screening for crystallisation conditions for $EibD_{161-418}$, crystallisation drops of 200 nl (100 nl of protein solution and 100 nl of well solution) were set up using a Cartesian Microsys nanopipettor. The screens used were modifications of commonly used sparse matrix screens (Helsinki Random Screens I & II; http://www.biocenter.helsinki.fi/bi/xray/automation/services.htm). For optimisation, we constructed grids around the most promising initial hit conditions and set up drops of 500 nl (250 nl protein solution plus 250 nl well solution). $EibD_{161-418}$ (concentrated to 10 mg/ml) crystallised readily in many conditions, and the best crystals came from 1 M ammonimum phosphate containing 0.1 M sodium citrate at pH 5.6. For optimisation, we constructed grids around the most promising initial hit conditions and set up drops of 500 nl (250 nl protein solution plus 250 nl well solution). For data collection, glycerol was added to the well solution to 10% (v/v) and this solution was

used to cryoprotect the crystals when vitrifying in liquid nitrogen. Although the crystals were small (~100 μm in the largest dimension), they diffracted to 2.0 Å at the European Synchrotron Radiation Facility (ESRF). The structure was solved by molecular replacement using YadA (PDB ID: 1P9H) as the search model.

For the GCN4-EibD$_{391-440}$ fusion protein, crystallization trials of EibD$_{391-440}$ were performed at 22 °C with 384 conditions by mixing 400 nl of reservoir solution with 400 nl of protein solution at a concentration of 2 mg/ml on Corning 3550 96-well sitting drop plates, using a Honeybee 961 crystallization robot (Genomic Solutions). This fragment crystallised in 20% isopropanol containing 200 mM sodium citrate and 100 M HEPES at pH 7.5. For cryoprotection, crystals were transferred into drops containing 25 % (v/v) glycerol in addition to the reservoir solution, loop-mounted and vitrified in liquid nitrogen. Datasets of different cystals were collected at beamlines X10SA and X06DA of the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland) at 100 K. All crystals were strongly anisotropic with diffraction limits of up to 2.1 Å along the *c* axis but at best 2.7 Å along *a* and *b*. The dataset used for structure solution and refinement was collected on a mar225 CCD detector (Marresearch) at the beamline X06DA.

Data processing, structure solution and refinement. For EibD$_{161-418}$, the images were processed in Mosflm (Leslie, 1992) and analysed with the programme Pointless in CCP4 version 6.1.1 (1994). Afterwards the data were scaled in Scala (Evans, 2006) of the CCP4 program suite.

Initial molecular replacement was done with the CaspR web server (Claude et al., 2004) using YadA (1P9H) (Nummelin et al., 2004) as the search model, and manual model building was done in Coot (Emsley and Cowtan, 2004) with refinement rounds performed in REFMAC (Vagin et al., 2004). The final refinement round and automated water picking were performed in Phenix (Adams et al., 2002). The structure quality was assessed with Molprobity (Davis et al., 2004).

For GCN4-EibD$_{391\text{-}440}$, data were processed and scaled to 2.8 Å using the *XDS* program suite (Kabsch, 1993). The crystals belong to space group P321, with one monomer in the asymmetric unit. The structure was solved by molecular replacement using the program MOLREP (Vagin and Teplyakov, 2010) and a trimmed monomer of the SadAK3 structure 2WPQ (Hartmann et al., 2009) as a search model. The structure was completed in cyclic manual modeling with Coot (Emsley and Cowtan, 2004) and refinement with REFMAC5 (Murshudov et al., 1999). All residues of the final model fall into the right handed alpha helix core region in the Ramachandran plot.

Binding and blocking experiments. Dot blot experiments were carried out essentially as described before (Leo and Goldman, 2009). SPR experiments were conducted according to standard procedures. To determine apparent dissociation constants, we analyzed the binding responses using a steady-state affinity model. For blocking experiments, we immobilised EibD$_{27\text{-}441}$ or SpG onto microtiter plate wells. We then added horseradish peroxidase-conjugated IgG that had previously been incubated with varying concentrations of SpA. To detect bound IgG, we added substrate solution and measured absorbance at 450 nm and compared the responses to

wells with no added SpA. Details of SPR and blocking experiments are given in Supplemental Experimental Procedures.

EM, autoagglutination and biofilm formation assays. For TEM, bacteria were high-pressure frozen, freeze-substituted and embedded in Epon. The sectioned samples were imaged using a FEI Tecnai F20 electron microscope. For SEM, fixed bacteria were coated with gold/palladium and imaged using a Hitachi S-800 scanning electron microscope. Biofilm formation was assayed using crystal violet staining. We followed autoagglutination by measuring the OD_{600} at the surfaces of static cultures. Full details are given in Supplemental Experimental Procedures.

Molecular modelling. The model of the $EibD_{161-511}$ was built by structure-based alignment of overlapping amino acid fragments of the available crystal structures and the theoretical model of the EibD membrane anchor. The membrane anchor model was generated automatically by the SWISS-MODEL server (http://swissmodel.expasy.org) (Arnold et al., 2006) based on the crystal structure of the Hia transmembrane domain (3EMO). The overlapping segments were identified by external alignment of the amino acid sequences. Finally, the model was constructed by structure-based alignment and stepwise extension in the program PyMOL. In order to generate single polypeptide chains, overlapping ends were trimmed with the preference being given to the experimentally determined structures.

The EibD-IgG Fc model was constructed based on the crystal structure of the human Fc fragment in complex with fragment B of SpA (1FC2). The Cα atoms of the a-helical part of SpA were aligned with an equal number of Cα atoms of the putative

IgG binding site from EibD. The EibD-IgG Fc monomer model was then generated by removal of the SpA structure. The functional trimers of the EibD model and EibD-Fc complex were then generated by applying symmetry operators derived from the $EibD_{161-418}$ crystal structure.

ACCESSION NUMBERS

Atomic coordinates and structure factors have been deposited in the PDB with the accession numbers 2XQH ($EibD_{161-418}$) and 2XZR (GCN4-$EibD_{391-440}$).

ACKNOWLEDGEMENTS

We thank S. Mäki at the Biocenter Finland Protein Crystallisation Facility, Institute of Biotechnology, University of Helsinki for expert help with protein crystallisation, D. Bansfield for help with protein purification, R. Kolodziejczyk (Institute of Biotechnology, University of Helsinki) for help with the SEC experiments, and M. Aatonen (Division of Biochemistry, University of Helsinki) for assistance with the SPR measurements. We thank B. Hernandez Alvarez, O. Ridderbusch, S. Deiss and R. Albrecht (MPI for Developmental Biology, Tübingen) for cloning, purification and assistance with crystallization of GCN4-$EibD_{391-440}$. We also acknowledge the beamline staff at the ESRF and SLS, and the Electron Microscopy Unit, Institute of Biotechnology, University of Helsinki. The plasmid pCS6364 was a kind gift from Prof. C. Hill (Pennsylvania State College of Medicine, Hershey, PA), and the strain BL21(DE3)omp8 was provided by R. Koebnik (IRD, Montpellier, France). This study was funded by the Sigrid Jusélius Foundation, the Academy of Finland (grant 1114752 to AG and 1129684 to SJB), and the EU (Marie Curie fellowship FP7 219889 to AL).

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. *58*, 1948-1954.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. *22*, 195-201.

Bardiau, M., Grégoire, F., Muylaert, A., Nahayo, A., Duprez, J. N., Mainil, J., and Linden, A. (2010). Enteropathogenic (EPEC), enterohaemorragic (EHEC) and verotoxigenic (VTEC) *Escherichia coli* in wild cervids. J. Appl. Microbiol. *109*, 2214-2222.

Biedzka-Sarek, M., Salmenlinna, S., Gruber, M., Lupas, A. N., Meri, S., and Skurnik, M. (2008). Functional mapping of YadA- and Ail-mediated binding of human factor H to *Yersinia enterocolitica* serotype O:3. Infect. Immun. *76*, 5016-5027.

Bond, C. S., and Schüttelkopf, A. W. (2009). ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr. D Biol. Crystallogr. *65*, 510-512.

Claude, J. B., Suhre, K., Notredame, C., Claverie, J. M., and Abergel, C. (2004). CaspR: a web server for automated molecular replacement using homology modelling. Nucleic Acids Res. *32*, W606-9.

Conners, R. et al. (2008). The *Moraxella* adhesin UspA1 binds to its human CEACAM1 receptor by a deformable trimeric coiled-coil. EMBO J. *27*, 1779-1789.

Dautin, N., and Bernstein, H. D. (2007). Protein secretion in Gram-negative bacteria *via* the autotransporter pathway. Annu. Rev. Microbiol. *61*, 89-112.

Davis, I. W., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic Acids Res. *32*, W615-9.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. *34*, W116-8.

Dunne, W. M. J. (2002). Bacterial adhesion: seen any good biofilms lately? Clin. Microbiol. Rev. *15*, 155-166.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126-2132.

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. *62*, 72-82.

Fricke, W. F., Wright, M. S., Lindell, A. H., Harkins, D. M., Baker-Austin, C., Ravel, J., and Stepanauskas, R. (2008). Insights into the environmental resistance gene pool from the genome sequence of the multidrug-resistant environmental isolate Escherichia coli SMS-3-5. J. Bacteriol. *190*, 6779-6794.

Hartmann, M. D., Ridderbusch, O., Zeth, K., Albrecht, R., Testa, O., Woolfson, D. N., Sauer, G., Dunin-Horkawicz, S., Lupas, A. N., and Alvarez, B. H. (2009). A coiled-coil motif that sequesters ions to the hydrophobic core. Proc. Natl. Acad. Sci. USA. *106*, 16950-16955.

Hernandez Alvarez, B., Gruber, M., Ursinus, A., Dunin-Horkawicz, S., Lupas, A. N., and Zeth, K. (2010). A transition from strong right-handed to canonical left-handed

supercoiling in a conserved coiled-coil segment of trimeric autotransporter adhesins. J. Struct. Biol. *170*, 236-245.

Hernandez Alvarez, B., Hartmann, M. D., Albrecht, R., Lupas, A. N., Zeth, K., and Linke, D. (2008). A new expression system for protein crystallization using trimeric coiled-coil adaptors. Protein Eng. Des. Sel. *21*, 11-18.

Hoiczyk, E., Roggenkamp, A., Reichenbecher, M., Lupas, A., and Heesemann, J. (2000). Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. EMBO J. *19*, 5989-5999.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. *26*, 795–800.

Koretke, K. K., Szczesny, P., Gruber, M., and Lupas, A. N. (2006). Model structure of the prototypical non-fimbrial adhesin YadA of *Yersinia enterocolitica*. J. Struct. Biol. *155*, 154-161.

Lambris, J. D., Ricklin, D., and Geisbrecht, B. V. (2008). Complement evasion by human pathogens. Nat. Rev. Microbiol. *6*, 132-142.

Larkin, M. A. et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics. *23*, 2947-2948.

Leo, J. C., Elovaara, H., Bihan, D., Pugh, N., Kilpinen, S. K., Raynal, N., Skurnik, M., Farndale, R. W., and Goldman, A. (2010). First analysis of a bacterial collagen-binding protein with collagen Toolkits: promiscuous binding of YadA to collagens may explain how YadA interferes with host processes. Infect. Immun. *78*, 3226-3236.

Leo, J. C., and Goldman, A. (2009). The immunoglobulin-binding Eib proteins from *Escherichia coli* are receptors for IgG Fc. Mol. Immunol. *46*, 1860-1866.

Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography.

Linke, D., Riess, T., Autenrieth, I. B., Lupas, A., and Kempf, V. A. (2006). Trimeric autotransporter adhesins: variable structure, common function. Trends. Microbiol. *14*, 264-270.

Lu, Y., Iyoda, S., Satou, H., Satou, H., Itoh, K., Saitoh, T., and Watanabe, H. (2006). A new immunoglobulin-binding protein, EibG, is responsible for the chain-like adhesion phenotype of locus of enterocyte effacement-negative, shiga toxin-producing *Escherichia coli*. Infect. Immun. *74*, 5747-5755.

Lupas, A. N., and Gruber, M. (2005). The structure of alpha-helical coiled coils. Adv. Protein Chem. *70*, 37-78.

Meng, G., St Geme, J. W. r., and Waksman, G. (2008). Repetitive architecture of the *Haemophilus influenzae* Hia trimeric autotransporter. J. Mol. Biol. *384*, 824-836.

Meng, G., Surana, N. K., St Geme, J. W. r., and Waksman, G. (2006). Structure of the outer membrane translocator domain of the *Haemophilus influenzae* Hia trimeric autotransporter. EMBO J. *25*, 2297-2304.

Merkel, V. et al. (2010). Distribution and phylogeny of immunoglobulin-binding protein G in Shiga toxin-producing *Escherichia coli* and its association with adherence phenotypes. Infect. Immun. *78*, 3625-3636.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr. D Biol. Crystallogr. *55*, 247-255.

Nitsche-Schmitz, D. P., Johansson, H. M., Sastalla, I., Reissmann, S., Frick, I. M., and Chhatwal, G. S. (2007). Group G streptococcal IgG binding molecules FOG and

protein G have different impacts on opsonization by C1q. J. Biol. Chem. *282*, 17530-17536.

Nummelin, H., Merckel, M. C., Leo, J. C., Lankinen, H., Skurnik, M., and Goldman, A. (2004). The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. EMBO J. *23*, 701-711.

Prilipov, A., Phale, P. S., Van Gelder, P., Rosenbusch, J. P., and Koebnik, R. (1998). Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from *E. coli*. FEMS Microbiol. Lett. *163*, 65-72.

Sandt, C. H., and Hill, C. W. (2000). Four different genes responsible for nonimmune immunoglobulin-binding activities within a single strain of *Escherichia coli*. Infect. Immun. *68*, 2205-2214.

Sandt, C. H., and Hill, C. W. (2001). Nonimmune binding of human immunoglobulin A (IgA) and IgG Fc by distinct sequence segments of the EibF cell surface protein of *Escherichia coli*. Infect. Immun. *69*, 7293-7303.

Sandt, C. H., Wang, Y. D., Wilson, R. A., and Hill, C. W. (1997). Escherichia coli strains with nonimmune immunoglobulin-binding activity. Infect. Immun. *65*, 4572-4579.

Szczesny, P., Linke, D., Ursinus, A., Bär, K., Schwarz, H., Riess, T. M., Kempf, V. A., Lupas, A. N., Martin, J., and Zeth, K. (2008). Structure of the head of the *Bartonella* adhesin BadA. PLoS Pathog. *4*, e1000119.

Szczesny, P., and Lupas, A. (2008). Domain annotation of trimeric autotransporter adhesins--daTAA. Bioinformatics. *24*, 1251-1256.

Tashiro, M., and Montelione, G. T. (1995). Structures of bacterial immunoglobulin-binding domains and their complexes with immunoglobulins. Curr. Opin. Struct. Biol. *5*, 471-481.

(1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760-763.

Touchon, M. et al. (2009). Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths. PLoS Genet. *5*, e1000344.

Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. Acta Crystallogr. D Biol. Crystallogr. *66*, 22-25.

Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004). REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr. D Biol. Crystallogr. *60*, 2184-2195.

Yu, C., Ruiz, T., Lenox, C., and Mintz, K. P. (2008). Functional mapping of an oligomeric autotransporter adhesin of *Aggregatibacter actinomycetemcomitans*. J. Bacteriol. *190*, 3098-3109.

Table 1

Crystallographic statistics.

Name	**EibD $_{161\text{-}418}$**	**GCN4-EibD $_{391\text{-}440}$**
Space group (number)	H 3 (146)	P 3 2 1 (150)
Unit cell [Å]	a=b=48.95, c=409.60	a=b=36.67, c=228.58
Data collection		
Wavelength [Å]	1.039	1.000
Resolution limits [Å]	41.51-1.99 (2.04-1.99)	38.10-2.80 (2.97-2.80)
Completeness [%]	97.9 (96.7)	98.8 (99.1)
Observations		
Total	80792 (5554)	13879 (2188)
Unique	24632 (1838)	4931 (750)
Redundancy	3.3 (3.0)	2.8 (2.9)
I/σ(I)	29.16 (4.98)	9.77 (4.09)
R_{merge} [%]	3.4 (12.2)	7.4 (25.1)
R_{meas} [%]	6.0 (22.3)	9.1 (30.7)
Refinement		
R_{cryst}	0.1585	0.2600
R_{free}	0.1855	0.3080
R_{free} test set size [%]	5.0	11.0
R_{free} set count	1232	541
Final structure		
RMSD from ideal:		
Bond length [Å]	0.008	0.013
Bond angle	1.027	1.061
Chirality	0.070	0.051
Planarity	0.003	0.004
Dihedral	13.177	21.496
Ramachandran statistics		
Favoured [%]	97.7	100.0
Outliers [%]	0.8	0.0
Bad rotamers [%]	4.2	4.3

$$R_{meas} = \frac{\sum_{hkl}\sqrt{\frac{N}{N-1}}\sum_{i}\left|I_i(hkl) - \overline{I(hkl)}\right|}{\sum_{hkl}\sum_{i} I_i(hkl)}$$

$$R_{merge} = \frac{\sum_{hkl}\sum_{i}\left|I_i(hkl) - \overline{I(hkl)}\right|}{\sum_{hkl}\sum_{i} I_i(hkl)}$$

$$R = \frac{\sum |F_{obs} - F_{calc}|}{\sum |F_{obs}|}$$

Table 2

Apparent affinities (in nM ± standard error of the mean) of $EibD_{161-418}$ mutants to IgA and IgG as measured by SPR

Mutant	IgA	IgG
$EibD_{161-418}$	89 ± 23	54 ± 7
E327A-E329A-K330A	70 ± 14	146 ± 48
K332A-K333A-Y334A	> 600	85 ± 15
D336A-K338A	187 ± 33	87 ± 11
E341A-E344A	86 ± 15	119 ± 21
D384A-R385A	169 ± 42	109 ± 23
D387A-Y388A	184 ± 65	288 ± 53
D394A-R396A	99 ± 27	62 ± 6
R399A-T401A-Q402A	156 ± 18	58 ± 1
E405A-K406A	181 ± 20	95 ± 6
S408A-K409A	250 ± 41	171 ± 25
T415A-R416A-T417A	456 ± 260	242 ± 96
Δ350-372	504 ± 65	485 ± 12
Δ352-372	258 ± 62	130 ± 23

Figure legends

Figure 1. Structure of EibD. The experimental structures of $EibD_{161-418}$ and the $EibD_{391-440}$ are shown in ribbon representation with the three monomers coloured differently. The smaller fragment ($EibD_{391-440}$) is demonstrated by one monomer in brown (GCN4 adapters (Hernandez Alvarez et al., 2008) omitted for clarity). The experimental structures were aligned in the overlapping fragment (residues 391-418, in red) with an RMSD of 0.243 Å (145 atoms). The chloride ions in the stalk are shown as green spheres, and the water oxygen as a red sphere. B) The EibD saddle (residues 350-371) showing the ionic and hydrogen bonding network stabilising the structure. C) Cavities (the two largest highlighted in yellow circles) in the RHcc. Two monomers are shown in space-filling representation. All figures were prepared with PyMOL (Schrödinger Corp.). See also Figure S1.

Figure 2. Effect of double and triple point mutations and saddle deletions on the Ig-binding activities of $EibD_{161-418}$. A) Dot blot of point and deletion mutants using IgG Fc-HRP and IgA-HRP. B) Representative sensorgrams for the binding of $EibD_{161-418}$ and mutants to IgG and IgA. Curves for the wt, two point mutants (K332A-K333A-Y334A and D387A-Y388A) and the saddle deletion mutant Δ352-372 are shown. In addition, the fitting of the data to the one-site affinity model is shown below the sensorgrams. See also Figure S2.

Figure 3. Binding of EibD to IgG Fc. A) Blocking of antibody binding to EibD by SpA. Wells of a microtitre plate were coated with either EibD (solid line) or SpG (dashed line). HRP-conjugated IgG diluted into solutions containing SpA at the

indicated concentrations was added to the wells and binding detected using TMB substrate. The binding level is shown as a percentage of the signal given by antibody incubated without SpA. B) Model of EibD (blue/gray) bound to IgG Fc (orange). The EibD model contains the experimental structure (residues 161-440) and a homology model of the membrane anchor (residues 441-511). IgG Fc monomers were docked onto the EibD structure based on the mutagenesis and blocking experiments. See Experimental Procedures for details. C) The EibD:IgG Fc complex model viewed from the N-terminus of EibD. This view demonstrates that three IgG Fc monomers (I, II, III), each containing two domains, can be accommodated around the EibD structure without steric hindrance. The colouring is the same as for panel B. D) Close-up of the binding interface showing histidines 310 and 435 of Fc close to D387 of EibD (4.7 Å and 6.0 Å in this configuaration, respectively). EibD is in blue, IgG Fc in yellow.

Figure 4. Electron micrographs. The upper panel shows TEM images of cell sections and the lower panel SEM images of whole cells. The images show control BL21(DE3)omp8 cells containing the vector pET22 or BL21(DE3)omp8 cells (Prilipov et al., 1998) expressing full-length wild type EibD (pEibD9; this constuct includes an N-terminal His tag) or control cells. Insets show selected regions at higher magnification. Expression of EibD leads to clumping of cells, altered surface morphology and the formation of zipper-like contacts between cells (inset, top panel). See also Figure S3.

Figure 5. EibD-conferred surface properties. A) Photograph of a 3-hour static culture of EibD-expressing bacteria (left tube), showing strong autoagglutination. The control

tube is on the right. B) Quantitative autoagglutination assay. BL21(DE3)omp8 (Prilipov et al., 1998) expressing EibD (pEibD10, continuous curve) and the vector control (pET22, dashed curve) were grown statically after IPTG induction and the OD_{600} value at the surface of the culture measured at 15 minute intervals. C) Biofilm formation assay. Adherent bacteria were stained with crystal violet, the dye was released from the cells with ethanol, and the absorbance at 595 nm was recorded. The autoagglutination and biofilm formation assays were performed on three independent cultures.

Figure 6. Comparison of Eib Ig-binding regions. A) Alignment of the C-terminal half of the Eib stalks, including the saddle and the left-handed coiled-coil. Conserved residues are coloured yellow (saddle) or green (coiled-coil). The *a* and *d* positions of the heptad periodicity are indicated, as are the secondary structure elements based on the EibD structure. The positions of chlorides in the core of the coiled coil are shown with red circles; the critical IgG binding residues are marked with green stars. B) Alignment of the right-handed region of the EibCDEFG stalks. The *a*, *d* and *e* positisions of the hendecad periodicity of the coiled-coil based on the EibD structure are shown. The EibF repeats (Sandt and Hill, 2001) are shown in purple. Conserved residues in the repeats are also coloured purple. The position of the KKY motif of EibD is marked with purple triangles. Similar motifs in repeats 1 and 2 of EibF are marked with yellow triangles. Conserved residues outside the repeats are shown in cyan. The alignments were generated in ClustalW (Larkin et al., 2007) and manually edited in the program ALINE (Bond and Schüttelkopf, 2009).

Figure 1

Click here to download high resolution image

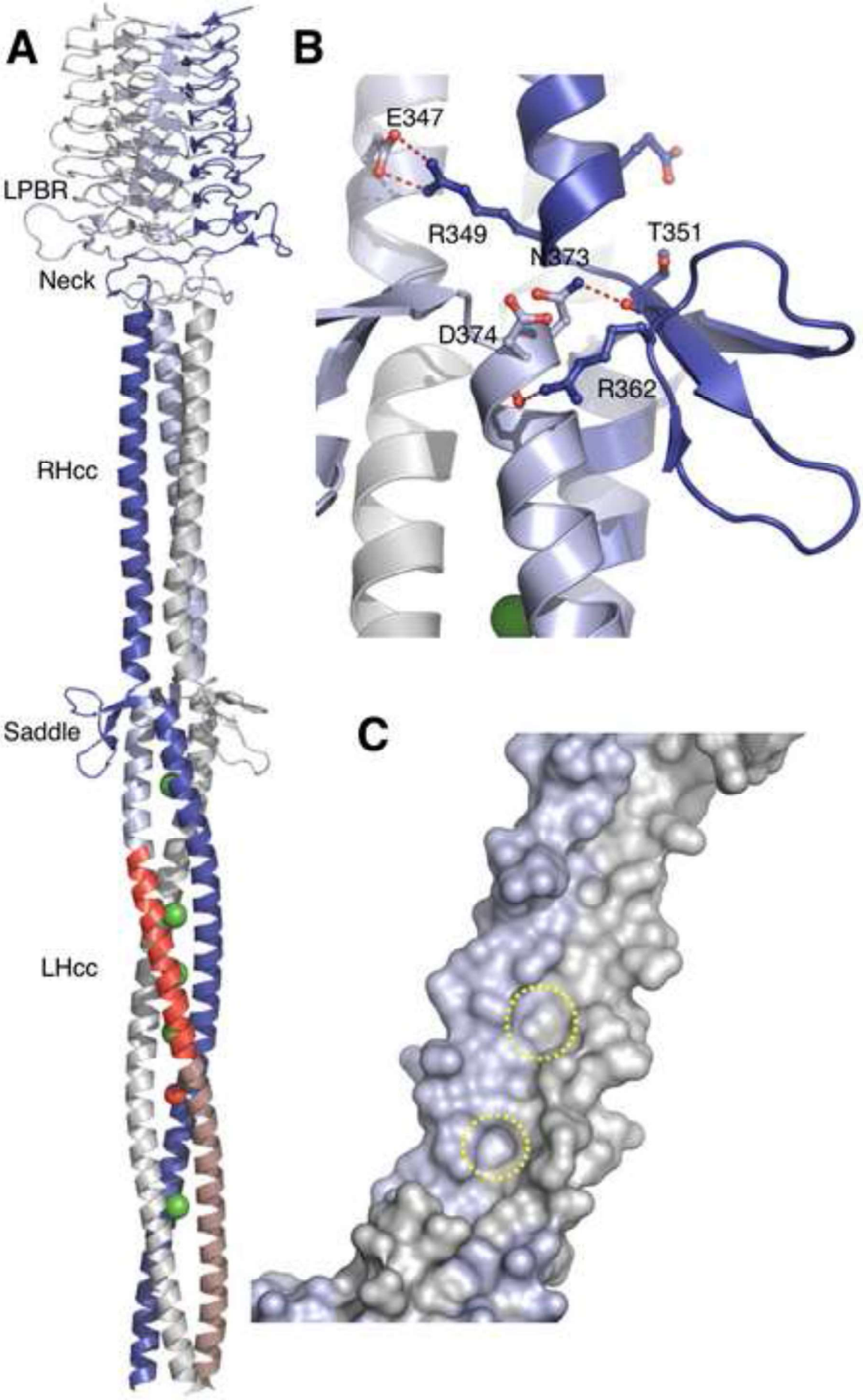


Figure 2

Click here to download high resolution image

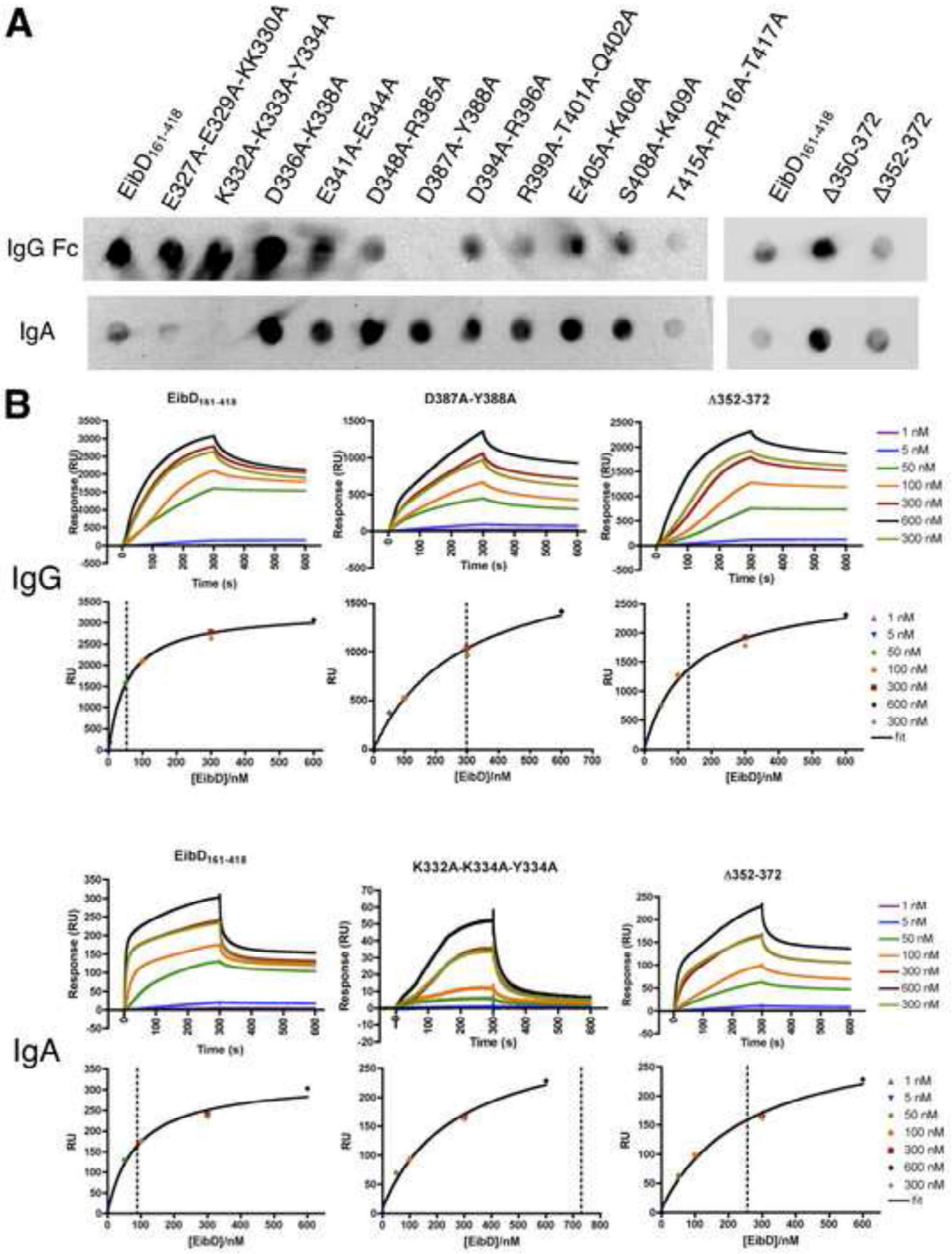


Figure 3
Click here to download high resolution image

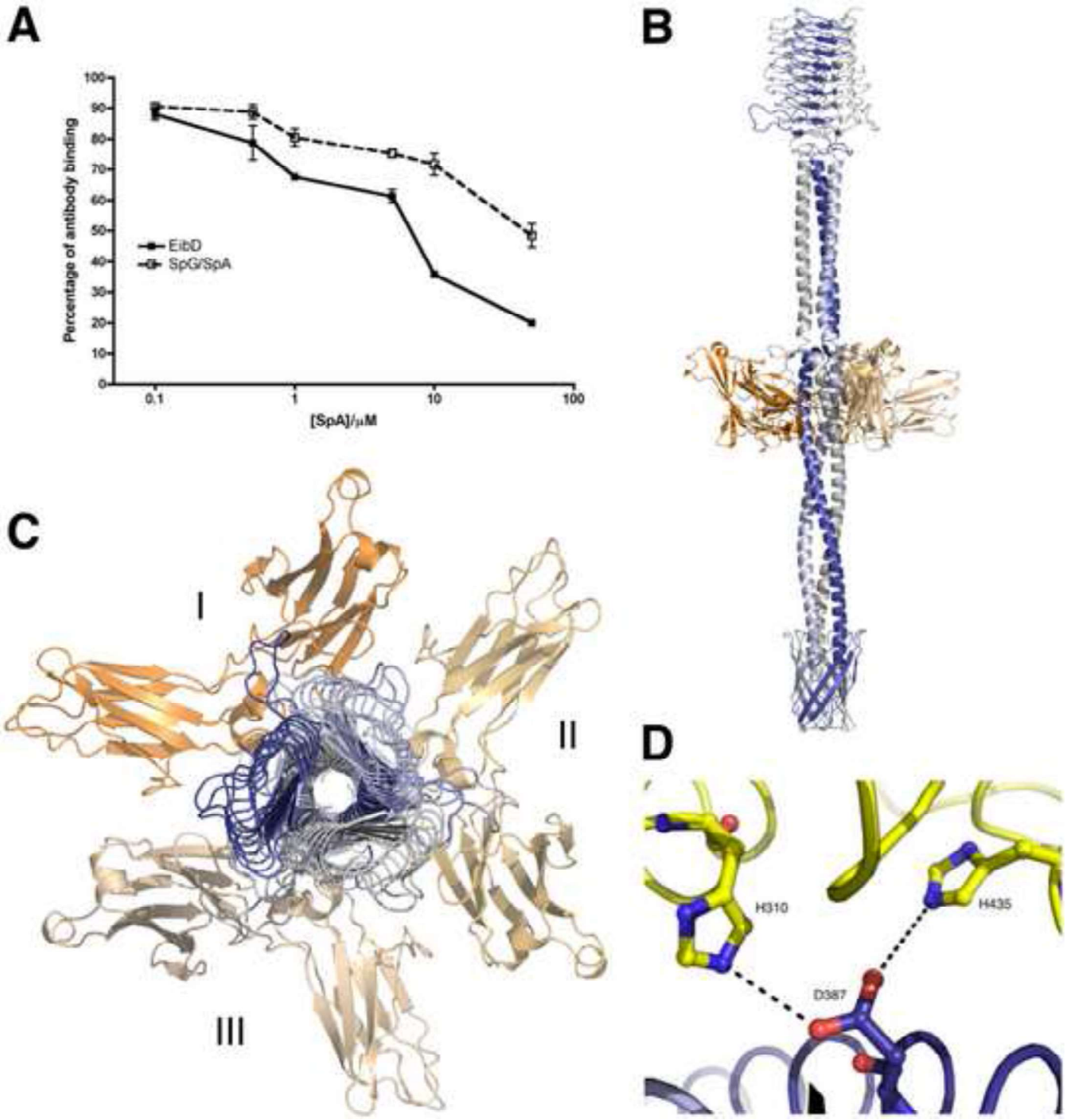


Figure 4

Click here to download high resolution image

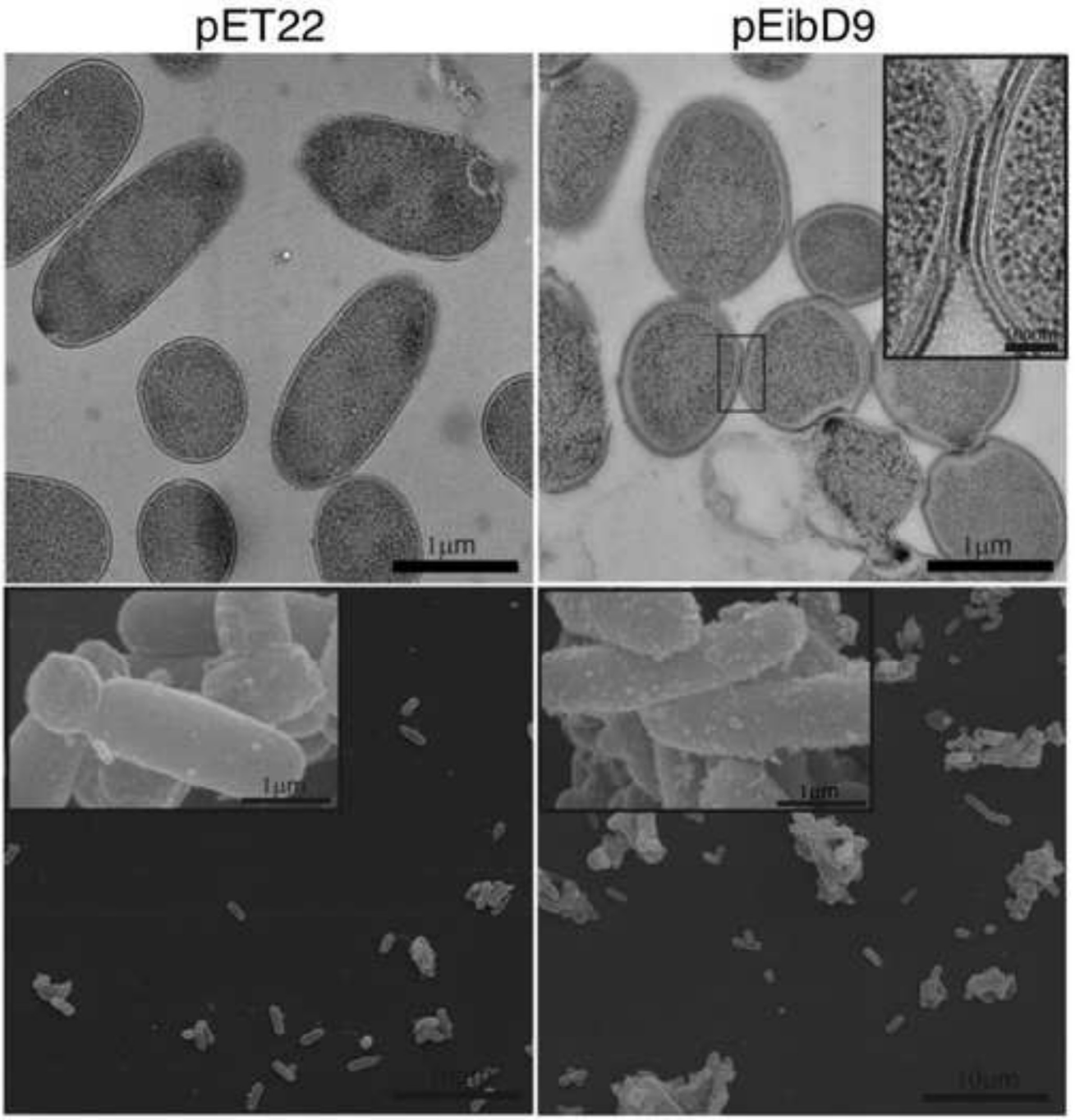


Figure 5

Click here to download high resolution image

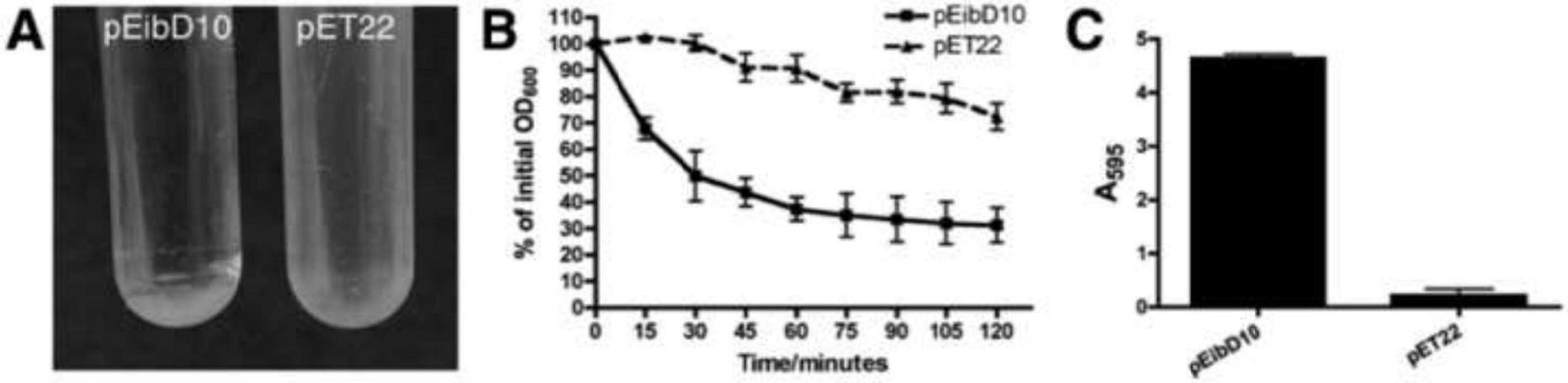
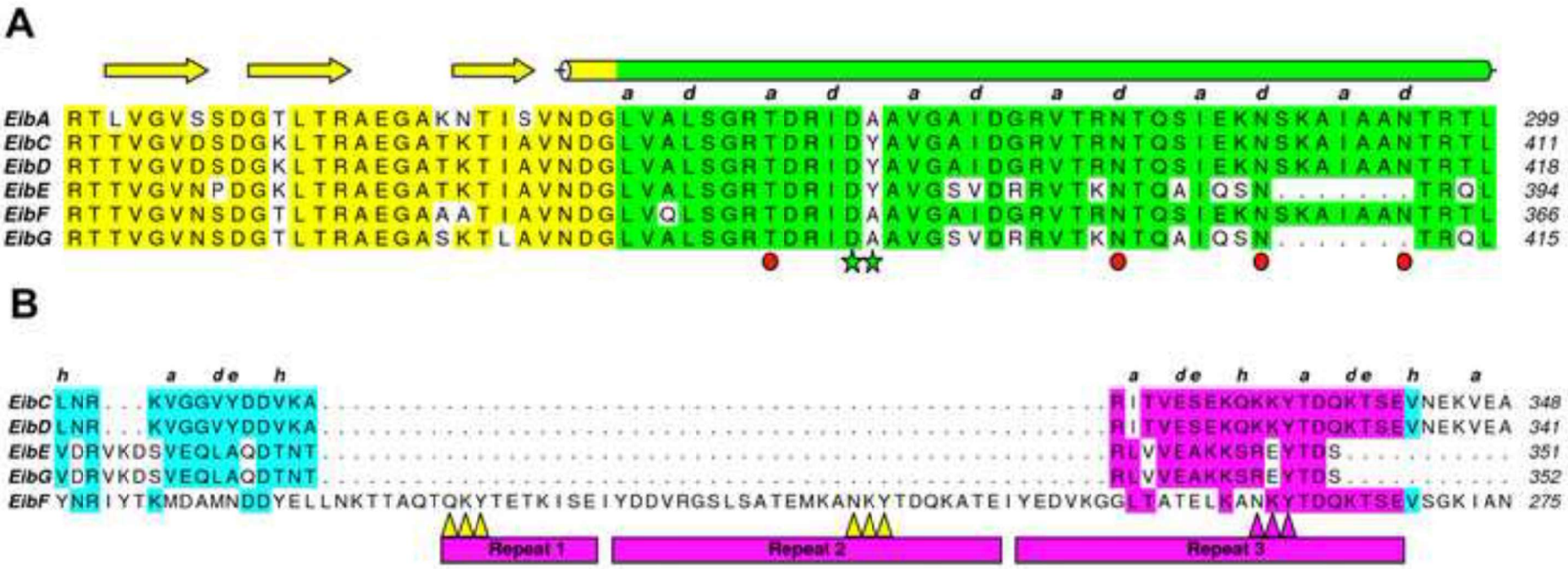


Figure 6

Click here to download high resolution image



Supplemental information to Leo *et al*.: Bending, binding and biofilm formation explained by the first complete TAA headgroup-stalk structure

Inventory

SUPPLEMENTAL DATA

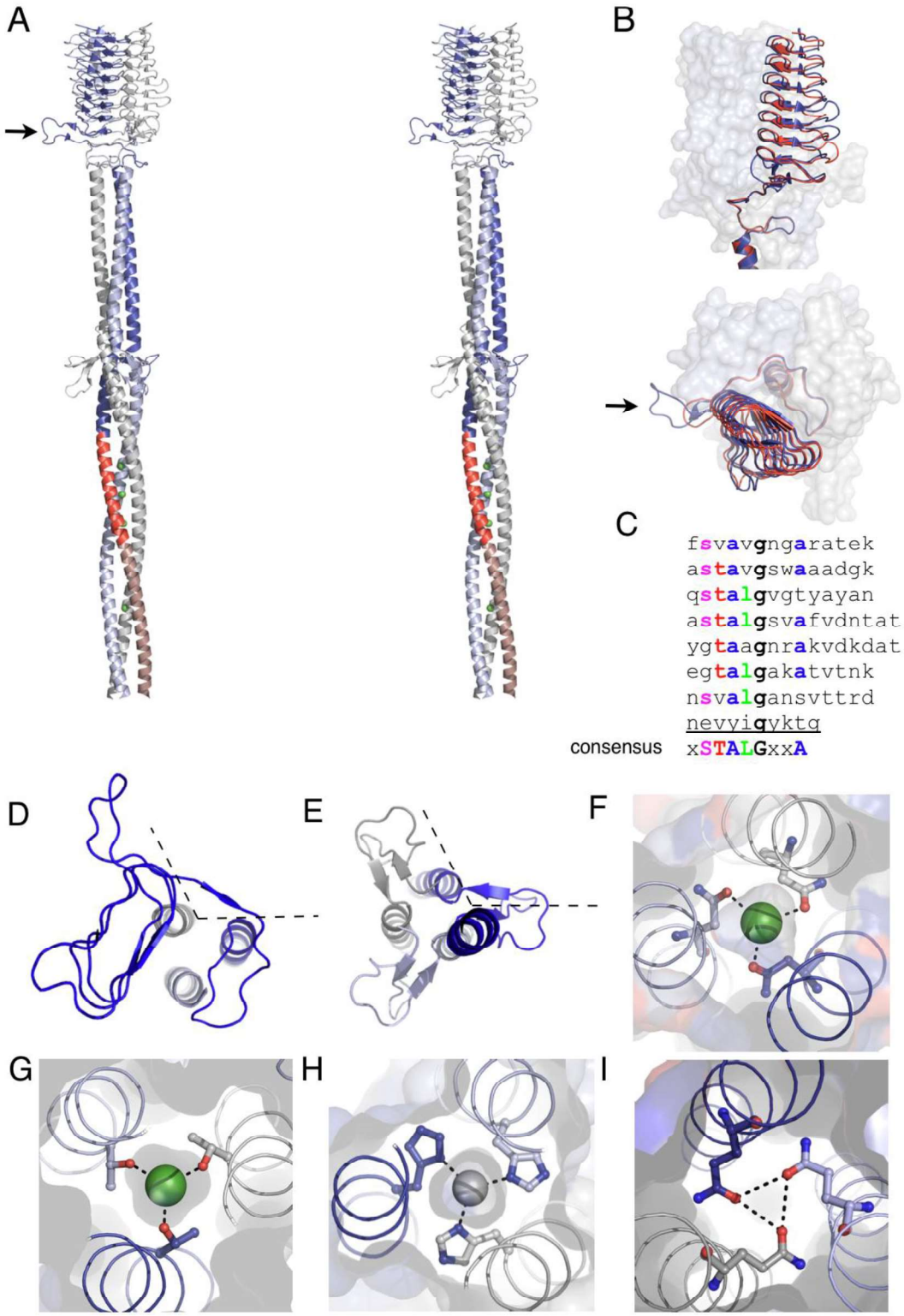


Figure S1, related to Figure 1. Structural features of EibD. A) Stereo image of $EibD_{161-440}$. The colouring of the individual elements is the same as in Figure 1. The arrow indicates the large loop in the neck of EibD. B) The YadA-like LPBR domain of EibD (blue) in comparison to the YadA (PDB ID: 1PH9, red). Chains B and C of the EibD are shown in surface representation. The surperimposition is shown from two different angles, from the side and viewed from the N-terminus. The large loop of EibD is indicated by an arrow. C) The β-roll repeats of EibD LPBR domain. Conserved residues are indicated in colour. The LPBR repeats of EibD have the consensus sequence XSTALGXXA. D) The 120° right-handed rotation introduced by the neck. The rotation is shown for one monomer (in blue) and viewed from the N-terminus. The other two monomers are shown only from the beginning of the stalk, *i.e.* after the neck. E) The 120° left-handed rotation at the saddle. The rotation is shown viewed from the N-terminus for the monomer in blue; the other two monomers are shown from the saddle onwards. The dotted lines in panels D and E represent a 120° angle. F) Coordination of a chloride ion in the hydrophobic core of the stalk by asparagine residues (N407), N@*d*. G) Coordination of a chloride ion in the hydrophobic core of the stalk by threonine (T382) residues in the RTD motif, T@*a*. H) Coordination of a water oxygen by His residues (H421, H@*d*). I) Hydrogen bonding between glutamines (Q428) in the hydrophobic core (Q@*d*). H)

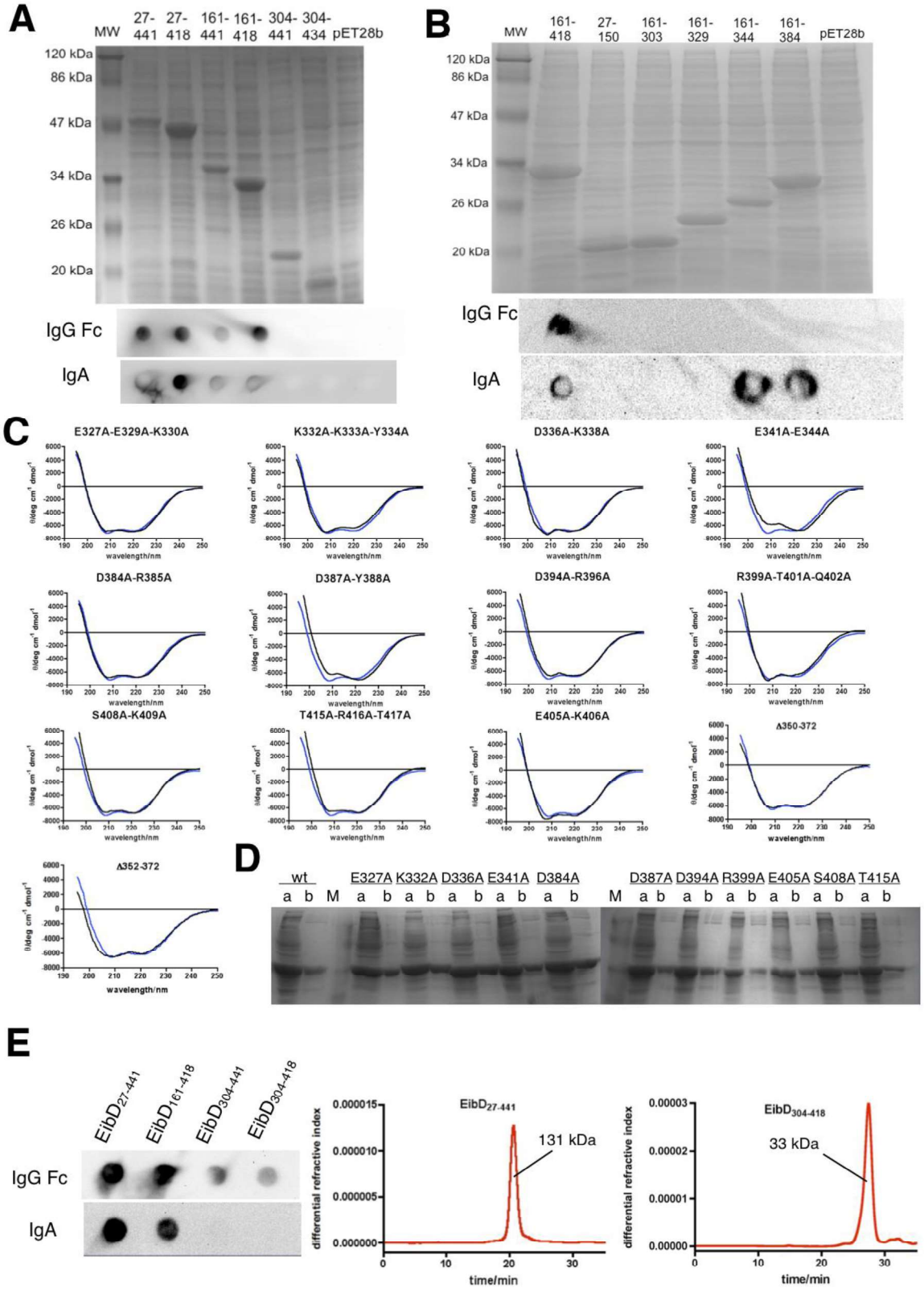


Figure S2, related to Figure 2. EibD cloning, truncation and point mutagenesis.

A) Binding of EibD fragments to IgG Fc and IgA. The upper panel shows an SDS-PAGE of cell lysates prepared from strains expressing different lengths of the EibD passenger domain. The lower panels show dot blots of the lysates probed with IgG Fc-HRP or IgA-HRP (Jackson Immunoresearch). As controls, we included $EibD_{27-441}$, which is known to bind both Ig classes (Leo and Goldman, 2009), and a lysate prepared from a strain containing only the vector (pET28b). The order of the samples in dot blots is the same as in the SDS-PAGE figures.
Residue 161 represents the predicted beginning of the YadA-like region, and 303 the beginning of the coiled-coil stalk. Residue 418 lies well outside the predicted translocation unit, whereas 434 is located just outside the translocation unit and 441 is predicted to be part of the α-helical linker within the transmembrane β-barrel pore. We then cloned these fragments into an expression vector and produced the corresponding polypeptides in *E. coli*. To determine the minimal fragment still retaining Ig-binding activity, we prepared cell lysates of overnight cultures grown in autoinducing medium. Running the lysates in SDS-PAGE showed that all the fragments are produced as soluble protein in *E. coli*.

B) Effect of C-terminal truncations on the Ig-binding activities of EibD. The upper panel shows an SDS-PAGE of lysates prepared from $EibD_{161-418}$ and C-terminal truncations. The lower panels show dot blot of the lysates probed wiht IgG Fc-HRP or IgA-HRP (Jackson Immunoresearch). As controls, we included $EibD_{27-150}$ and a lysate prepared from a strain containing only the vector (pET28b). The order of the samples in dot blot panels is the same as for the SDS-PAGE. The truncations $EibD_{161-344}$ and $EibD_{161-384}$ bind IgA, whereas the other two truncations do not, suggesting that binding site for IgG is at the very end of the stalk between residues 384 and 418. In our earlier experiments, $EibD_{303-418}$ and $EibD_{303-441}$, which lack the LPBR domain, did not give a detectable signal (Supplementary Figure 1) although both constructs should contain the IgA and IgG binding sites. This is possibly due to the stalks not forming stable trimers without the LPBR domain.

C) Far-UV CD spectra of double and triple point mutants and stalk deletion mutants of $EibD_{161-418}$. We diluted $EibD_{161-418}$ and mutant proteins to a concentration of approximately 0.1 mg/ml in CD buffer (5 mM sodium phosphate pH 7.4, 5 mM NaCl). Far-UV CD spectra between 195 and 250 nm were recorded using a Jasco J-715 instrument and a 1 mm cuvette at +20 °C, with the final spectrum for each protein representing the average of five measurements. We converted the CD signal to molar ellipticity (θ) values using the formula $\theta = \mathrm{CD}/(n \times c \times l \times 10)$, where the CD signal is in millidegrees, n the number of residues, c the molar concentration of protein and l the length of the light path in centimetres. The spectrum of wt $EibD_{161-418}$ is shown in blue.
In addition to T415A-R416A-T417A, the only other mutants to show changes in their CD spectra were E341A-E344A and D387A-Y388A. The former showed wt binding, while the latter bound IgA like the wt, suggesting that the changes are local. For D387A-Y388A, the change in CD signal is probably due to the removal of the Tyr residue.

D) SDS-PAGE of purification of $EibD_{161-418}$ point mutants. For clarity, only the first mutated residue of each mutant is marked. a = crude cell extract, b = eluate from Ni column, M = molecular weight marker (same as in panels A & B).

E) The stalk of EibD alone is unstable. We purified four EibD proteins ($EibD_{27-441}$, $EibD_{161-418}$, $EibD_{304-441}$ and $EibD_{303-418}$) and measured their Ig-binding activities by dot blot (left-hand panels). The larger proteins, containing the LPBR domain, bound both IgG Fc and IgA strongly, whereas the shorter proteins consisting only of stalk, bound Igs poorly – the binding to IgA is barely detectable. The poor binding is probably due to the stalk dissociating. To verify this, we performed size exclusion chromatography coupled to multi-angle static light scattering and refractive index measurements on the largest and shortest EibD constructs. $EibD_{27-441}$ eluted with an apparent molecular weight of 131.4 ± 0.26 kDa; the expected size of the trimer is 139.7 kDa. In contrast, $EibD_{304-418}$ had an apparent molecular weight of only 33.2 ± 0.17 kDa, suggesting the protein is mostly in a dimeric form (expected dimer size 29.8 kDa). Furthermore, the peak is asymmetric with tailing at the leading edge, indicating that there is an equilibrium between the dimer and trimer and that $EibD_{304-418}$ is predominantly dimeric.

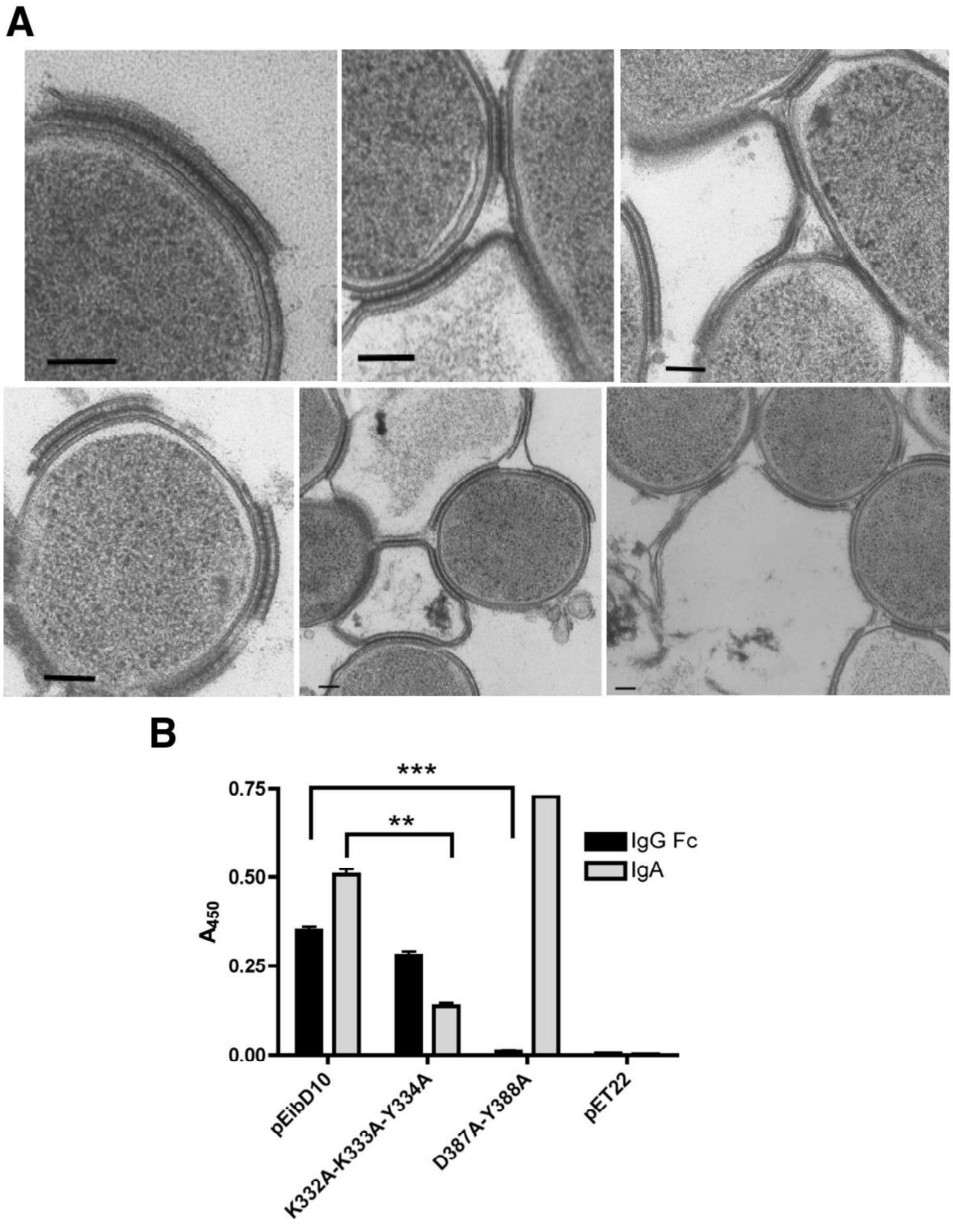


Figure S3, related to Figure 4. Binding of EibD expressing cells to each other and immunoglobulins.

A) TEM images showing zipper-like connection between EibD-expressing cells and intact cells binding membranes and membrane fragments from lysed cells. The scale bar corresponds to 100 nm in each image.
B) Whole-cell binding assay of EibD10 and point mutants to HRP-conjugated IgA and IgG Fc. pET22 is the vector control. The point mutations K332A-K333A-Y334A and D387A-Y388A significanlty reduce binding to IgA and IgG Fc, respectively, compared to the wt. The Igs binding to the non-mutated site in the mutants gave strong signals, showing that the proteins are surface exposed and correctly folded. IgG Fc gave a weaker signal than IgA in these experiments, probably due to the binding site for IgG, located closer to the membrane, being less accessible. Error bars show the SEM from three replicate experiments. Statistical significance was estimated with Student's *t*-test (** $p < 0.001$, *** $p < 0.001$).

Table S1
Primers for generation of EibD passenger domain fragments and used in truncating and site-directed mutagenesis

EibD fragment/mutant	**Primer sequences**
$EibD_{27-418}$	Forward: 5'-gcgGCTAGCcagaatggtacatattcagtgttgca-3' Reverse: 5'-gcgAAGCTTttacagcgtacgggtattggctgc-3'
$EibD_{161-418}$	Forward: 5'-gcggGCTAGCgatgcgaaagcctctggtgagt-3' Reverse: 5'-5'-gcgAAGCTTttacagcgtacgggtattggctgc-3'
$EibD_{161-441}$	Forward: 5'-gcggGCTAGCgatgcgaaagcctctggtgagt-3' Reverse: 5'- gcgAAGCTTttagttacgcgtaccatcgcttaag-3'
$EibD_{303-418}$	Forward: 5'-gcgaGCTAGCgcgg caactgttgg gcagttaa-3' Reverse: 5'-gcgAAGCTTttacagcgtacgggtattggctgc-3'
$EibD_{303-441}$	Forward: 5'-gcgaGCTAGCgcgg caactgttgg gcagttaa-3' Reverse: 5'- gcgAAGCTTttagttacgcgtaccatcgcttaag-3'
Eib-p1	5`-gacccatggtctccgattggcgcgattgatggccgcgtgacccgcaacacccagagc-3'
Eib-p2	5`-gtgcgggtgttcgccgcaatcgctttgctgtttttttcaatgctctgggtgttgcgggtc-3'
Eib-p3	5`-gattgcggcgaacacccgcaccctgcagcagcatagcgcgcgcctggatagccagcagcgccag-3'
Eib-p4	5`-gaccatggtctcctcatttctttatggttttcgttaatctggcgctgctggctatcc-3'
EibDStop150	Forward: 5'-gtttcagcgtcgaagatTAAggggagaatg ggaagg-3' Reverse: 5'-ccttcccattctccccTTAatcttcgacgctgaaac-3'
EibDStop308	Forward: 5'-ggcaactgttgggcagTAAttaaaccggaaagttgg-3' Reverse: 5'-ccaactttccggtttaaTTActgcccaacagttgcc-3'
EibDStop329	Forward: 5'-taactgttgagtcagagTAAaagcagaaaaaatacac-3' Reverse: 5'-gtgtatttttttctgcttTTActctgactcaacagtta-3'
EibDStop344	Forward: 5'-gcgaggttaacgaaTAAaaagttgaagcccgcac-3' Reverse: 5'-gtgcgggcttcaactttTTAttcgttaacctcgc-3'
EibDStop384	Forward: 5'-ctggcagaacagacTAAcgtattgattatgcagttgg-3' Reverse: 5'-ccaactgcataatcaatacgTTAgtctgttctgccag-3'
E327A-E329A-K330A	Forward: 5'-gcccgtataactgttGCGtcaGCGGCGcagaaaaaatacacagacc-3' Reverse: 5'-ggtctgtgtatttttttctgCGCCGCtgaCGCaacagttatacgggc-3'
K332A-K333A-Y334A	Forward: 5'-ctgttgagtcagagaagcagGCAGCAGCCacagaccagaagaccagcg-3' Reverse: 5'-cgctggtcttctggtctgtGGCTGCTGCctgcttctctgactcaacag-3'
D336A-K338A	Forward: 5'-gcagaaaaaatacacaGCCcagGCGaccagcgaggttaacg-3' Reverse: 5'-cgttaacctcgctggtCGCctgGGCtgtgtatttttttctgc-3'
E341A-E344A	Forward: 5'-ccagaagaccagcGCGgttaacGCAaaagttgaagcc-3' Reverse: 5'-ggcttcaactttTGCgttaacCGCgctggtcttctgg-3'
D384A-R385A	Forward: 5'-tctggcagaacaGCCGCTattgattatgcagttggtgccattg-3' Reverse: 5'-caatggcaccaactgcataAGCaatAGCgtctgttctgccaga-3'
D387A-Y388A	Forward: 5'-ggcagaacagaccgtattGCTGCTgcagttggtgccattgatgg-3' Reverse: 5'-ccatcaatggcaccaactgcAGCAGCaatacggtctgttctgcc-3'
D394A-R396A	Forward: 5'-gcagttggtgccattGCTggcGCTgtcacccgtaacacgc-3' Reverse: 5'-gcgtgttacgggtgacag*c*gccagcaatggcaccaactgc-3'
R399A-T401A-Q402A	Forward: 5'-ggccgtgtcaccGCTaacGCGGCAtccattgagaaaaacagcaaggcg-3' Reverse: 5'- cgccttgctgtttttctcaatggaTGCCGCgttAGCggtgacacggcc-3'

E405A-K406A	Forward: 5'-cgtaacacgcaatccattGCGGCAaacagcaaggcgattgcag-3' Reverse: 5'-ctgcaatcgccttgctgttTGCCGCaatggattgcgtgttacg-3'
S408A-K409A	Forward: 5'-gcaatccattgagaaaaacGCCGCGgcgattgcagccaataccc-3' Reverse: 5'-gggtattggctgcaatcgcCGCGGCgtttttctcaatggattgc-3'
T415A-R416A-T417A	Forward: ggcgattgcagccaatGCCGCTGCGctgtaaaagcttgcgggc-3' Reverse: 5'-gcccgcaagcttttacagCGCAGCGGCattggctgcaatcgcc-3'

Restriction sites (GCTAGC for *Nhe*I and AAGCTT for *Hind*III) are capitalised. The introduced termination codon in the reverse primer is underlined. Introduced terminator and mutated codons are capitalised.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cloning of EibD fragments. We amplified the DNA fragments corresponding to the fragments $EibD_{27-418}$, $EibD_{161-418}$, $EibD_{161-441}$, $EibD_{303-418}$ and $EibD_{303-441}$ by polymerase chain reaction (PCR) using the plasmid pCS6364 (Sandt and Hill, 2000) as a template, and appropriate primers (Table S1). We then digested the PCR fragments with *Nhe*I and *Hind*III and ligated with the hexahistidine tag-encoding expression vector pET28b (Novagen), which had been digested with the same enzymes and dephosphorylated (enzymes from New England Biolabs). The ligation mixture was chemically transformed into *E. coli* DH5α and plated onto Luria-Bertani (LB) medium containing kanamycin at 25 μg/ml. The correctness of all clones was determined by sequencing.

To produce the GCN4-$EibD_{391-440}$ fusion, the sequence corresponding to $EibD_{391-438}$ was amplified by a two-stage PCR: the central part of the fragment was first amplified using primers Eib-p2 and Eib-p3 and subsequently used as template in the next primer extension reaction using Eib-p1 and Eib-p4 as primers (primer sequences in Table S1). The produced fragment was purified, digested with *Bsa*I and cloned into the vector pIBA-GCN4tri-His (Hernandez Alvarez et al., 2008). The resulting construct comprises residues 391-438 of EibD fused in-register to GCN4pII-adaptors at its termini and a C-terminal $(His)_6$ linker. As the first two residues of the C-terminal GCN4pII adaptor are identical to residues 439 and 440 of EibD, the fused construct covers residues 391-440 of EibD in total.

We were unable to see any expression of EibD on the cell surface using the plasmid pCS6364 (Sandt and Hill, 2000). Strong expression of the Eibs from their native promoter requires the regulatory factor *ibrAB* (Sandt et al., 2002); the strains we had lacked this factor. Therefore, we overexpressed EibD from the vector pET22. The coding sequence of EibD, excluding the signal peptide (residues 1-26) was amplified by polymerase chain reaction (PCR) using the plasmid pCS6364 as the template. We prepared two constructs: one including an N-terminal hexahistidine tag and one producing the native sequence. To produce the His-tagged version, we included 6 histidine codons in the 5‘ primer (Table S1). The primers used to produce the untagged construct (Table S2) were designed so that the 5’ regions were complementary to sequences in the vector pET22b (Novagen). The product of this PCR reaction was inserted into pET22b by a PCR-based, restriction-free cloning method (van den Ent and Löwe, 2006). Insert-containing clones were identified by colony PCR, and the insert was sequenced to verify its correctness. The resulting constructs contained the EibD passenger and translocation domains fused to the PelB signal sequence already present in pET22b. pEibD9 had the N-terminal His tag inserted directly after the signal peptide cleavage site, so as to include the tag in the surface-expressed protein. pEibD10 had the EibD passenger fused directly to the signal peptide.

Site-directed mutagenesis. Mutants were produced using the Quick Change™ protocol from Stratagene. Briefly, we designed overlapping primers (Table S1) containing the desired nucleotide changes to produce either truncating or surface residue mutations of $EibD_{161-418}$ or pEibD10. The entire plasmid, *i.e.* pET28b containing the sequence

corresponding to $EibD_{161-418}$ (or $EibD_{27-441}$ to produce the truncation ending at residue 150), was amplified by PCR using mutation primer pairs, *Pfu* polymerase (Fermentas) and an extension temperature of 68 °C. After PCR, the template plasmid was removed by digestion with *Dpn*I (New England Biolabs). 5 µl of the digested mix was directly transformed chemically into the cloning strain XL1-Blue and plated on LB + kanamycin (25 µg/ml). The mutagenesis was verified by sequencing.

Protein production and purification. For the expression of EibD fragments and surface residue mutants, verified plasmid DNA was transformed into *E. coli* BL21(DE3) and plated on LB + kanamycin (25 µg/ml) supplemented with 0.5% (w/v) D-glucose. For small-scale expression cultures, we inoculated plasmid-containing bacteria into 5 ml of the autoinducing ZYP-5052 medium (Studier, 2005), containing 100 µg/ml of kanamycin, and grew the cultures overnight at +37 °C, during which time the protein was expressed.

For medium-scale expression, we inoculated bacteria in 50 ml of autoinducing ZYP-5052 medium (Studier, 2005) containing kanamycin as above in a 250 ml shake flask. The cultures were grown overnight at +37 °C and harvested the following morning.

For large-scale expression to produce protein for crystallisation, we first inoculated bacteria in 5 ml of the defined, protein expression-inhibiting medium PA-0.5G (Studier, 2005), supplemented with 100 µg/ml of kanamycin, and allowed the cultures to grow for approximately 7 hours. We then diluted the cultures 1:100 into the autoinducing ZYP-5052 medium containing antibiotics as above into a total volume of 1 litre divided into two 2-litre shake flasks. The cultures were grown overnight as above.

For surface expression of EibD for EM and autoagglutination assays, we transformed pEibD9 or pEibD10 into the the strain BL21(DE3)omp8 (Prilipov et al., 1998), which is optimised for high-level expression of outer membrane proteins. Cultures were grown in 5 ml LB + ampicillin (100 µg/ml) overnight at +37 °C and diluted the following morning 1:50 into fresh LB medium. The cultures were grown to OD_{600} ~0.5, at which time protein expression was induced by the addition of isopropylthiogalactoside to a final concentration of 1 mM (for EM, to ensure a high occupancy of EibD) or 0.2 mM (for autoagglutination assays, to minimise plasmid loss during incubation). Expression was continued for two hours before samples were taken for experiments.

To produce crystallisable protein, we followed a purification protocol that has been described in detail before (Leo and Goldman, 2009). To purify mutant proteins, we used prepacked Protino® Ni-IDA 2000 columns (Macherey-Nagel) and performed cell lysis and purification according the manufacturer's instructions.

The GCN4-$EibD_{391-440}$ fusion was expressed and purified under denaturing conditions using a Ni-NTA column (GE Healthcare) as described previously for His-tagged SadAK3 (Hernandez Alvarez et al., 2008). Refolding was performed at 4°C by dialysis in 20 mM Tris/HCl, 400 mM NaCl, 5% glycerol, pH 7.5. Subsequently, the protein was concentrated and further purified using a size exclusion chromatography step on a

Superdex 75 column (GE Healthcare) equilibrated with 20 mM Tris/HCl, 400 mM NaCl, pH 7.5.

Blocking experiments. We immobilised 1 μg of either $EibD_{27\text{-}441}$ or SpG (recombinant protein from Sigma, catalogue no. 19459) in 100 μl of PBS for 1 hour onto microtitre plate wells. After blocking for 30 minutes with 5% bovine serum albumin (BSA; from Sigma) in PBS, we added a solution containing goat anti-mouse HPR-conjugated antibody (Santa Cruz Biotechnology) diluted to 1:8000 in 5% BSA in PBS that had previously been incubated for 30 minutes with a varying concentration of SpA (recombinant protein from Sigma, catalogue no. P7837). After washing four times with 150 μl of PBS-T, binding was detected by adding 100 μl of substrate solution (TMB; from Thermo Scientific). Once colour had developed the reaction was stopped by the addition of 100 μl of 2.5 M H_2SO_4, and absorbance was read at 450 nm. All steps were carried out at RT. For plotting, we took the signal from wells with no added SpA to be 100% of binding and the background given by the BSA control to be 0% and calculated the percentage of antibody binding accordingly.

SPR experiments. We examined the binding of Eibs to antibodies and antibody fragments with a BIAcore T100™ instrument (GE Healthcare). We coupled IgA and IgG to a CM5 sensor chip using an amine coupling kit (both from GE Healthcare) according to the manufacturer's instructions. To measure the binding of EibD mutants to immobilised antibodies, we prepared dilutions of EibD in running buffer (PBS + 0.05% Tween20) at six concentrations between 1 nM and 600 nM of trimeric EibD. One of the concentrations (300 nM) was run in duplicate, and an additional zero-concentration sample was included. The dilutions were run over the chip surfaces at a flow rate of 25 μl/min at +25 °C. The injection and dissociation steps were both 5 minutes, after which the surface was regenerated with 0.2% SDS. A wash step of 5 minutes with running buffer preceded the next injection. The results were analysed in the Biacore T100 evaluation software (GE Healthcare). Kinetic analysis was hampered by the multivalent interactions between EibD and Igs, resulting in poor curve fits. Therefore, the data were fitted to the steady-state affinity model and K_d(app) values were calculated from the resulting Scatchard plots generated by the Biacore T100 software. Though the resulting fits to the model are not optimal, the data for each mutant were treated in the same way and are thus comparable.

Electron microscopy. For TEM, EibD-expressing bacteria bacteria were sucked into cellulose capillary tubes, transferred to 150 μm planchettes filled with hexadecane, and high-pressure frozen in a Bal-Tec *HPM 010* freezing device. The samples were freeze substituted (starting at -90 °C down to -40 °C) in 2% osmium tetroxide, 0.5% uranyl acetate, and 0.5% glutaraldehyde in acetone and finally embedded in Epon. Thin sections (90 nm) were cut and post-stained with uranyl acetate and lead citrate. The samples were imaged in a FEI Tecnai F20 at 200 kV and images captured using a Gatan Ultrascan 4000 camera. For SEM, samples were fixed with 2.5% glutaraldehyde in PBS, postfixed with 1% osmium tetroxide in phosphate-buffered saline, dehydrated in a graded series of ethanol and critical-point-dried from CO_2. Finally the samples were sputter-coated with a

layer of 7 nm gold/palladium (Bal-Tec MED 010) and examined at 20 kV accelerating voltage in a Hitachi S-800 field emission scanning electron microscope.

Whole-cell binding assay. To measure binding of whole cells to IgG and IgA, we used a modified solid-phase binding assay. We were unable to get a good signal from cells that were adsorbed onto a microtitre plate, probably because of the occlusion caused by EibD overexpression. Therefore, we used the following protocol: an overnight culture of EibD expressing cells or the vector control were diluted 1:10 in fresh medium and grown at +37 °C until the OD_{600} value reached ~0.5. We then added BSA to 1% to block unspecific binding and continued growing for 15 minutes. We added HRP-conjugated IgG Fc (1:5000) or HRP-IgA (1:2500, both from Jackson Immunoresearch) and IPTG to 0.2 mM and continued growing for 30 minutes. We then collected the cells (10 min 4000 x *g*), washed them twice with PBS to remove free antibody and resuspended in PBS at OD_{600} = 0.2. 100 µl of this suspension was added to microtitre plate wells (Immulon 2 HB, Thermo Scientific). The bacteria were allowed to adsorp for 1 hour at RT. After this, the wells were washed four times with 150 µl PBS and binding was detected using TMB substrate as above.

Autoagglutination and biofilm formation assays. For autoagglutination assays, BL21(DE3)omp8 cultures containing either pEibD10 or pET22 were grown as described above. After two hours of IPTG induction (see above), the cultures were taken out of the shaker and grown statically at +37 °C. Every 15 minutes, a 60 µl sample was taken from the surface of the culture and the OD_{600} value was measured. We repeated this procedure for a total of two hours.

To assay for biofilm formation, EibD-expressing and control strains were inoculated in 1 ml of ZYP-5052 (+ ampicillin 150 µg/ml) in wells of a 24-well culture plate and grown overnight with gentle agitation (60 rpm) at +37 °C. The following day, we removed the medium and washed the wells three times with 2 ml PBS. Following this, 2 ml of crystal violet solution (1% w/v) was added to the wells and incubated for 5 minutes at RT. After removing the crystal violet, we again washed three times as above. To elute the crystal violet from cells in biofilms, we added 2 ml of ethanol, incubated for two minutes with agitation, and then measured absorbances at 595 nm.

Size exlusion chromatography with multi-angle static light scattering. To examine the dissociation of EibD, we performed size exclusion chromatography using a Shimadzu High Performance Liquid Chromatography unit with a Superdex 200 10/300 column (GE Healthcare). The column was equilibrated with TBS buffer and runs were performed at RT. Static light scattering and refractive index changes were measured using a Wyatt miniDAWN Treos and a Wyatt Optilab rEX detector, respectively. ~100 ug of protein was used for each run.