The crystal structure of the FaeG adhesin of F4 fimbriae from enterotoxigenic *Escherchia coli* reveals its putative receptor binding site

I. Van Molle[1], J.J. Joensuu[1,2], L. Buts[1], S. Panjikar[3], J. Bouckaert[1], L. Wyns[1] and H. De Greve[1]

[1]Vrije Universiteit Brussel and Vlaams Interuniversitair Instituut voor Biotechnologie, Pleinlaan 2, B-1050 Brussel, Belgium

[2]Department of Applied Biology, Latokartanonkaari 7, PO -box 27, FIN-00014 University of Helsinki, Finland

[3] EMBL Hamburg c/o DESY, Notkestraße 85, 22603 Hamburg, Germany

Binding of bacterial pathogens to host cells is typically mediated by adhesins. These are located on the bacterial surface in polymeric, proteinaceous appendages called fimbriae or pili, or in non-pilus structures [1]. The most prevalent assembly pathway for these adhesive structures is the chaperone/usher pathway [2]. Regardless of their ultrastructure, all known adhesive structures dependent on the chaperone/usher pathway are assembled via a donor strand complementation/exchange mechanism [3-5].

F4 fimbriae encoded by the *fae*-operon are the major colonization factor associated with porcine neonatal and post-weaning diarrhea caused by enterotoxigenic *E. coli* (ETEC). FaeG is the major structural subunit of the F4 fimbriae and the adhesin. There are three naturally occurring serological variants of F4 fimbriae: F4ab, F4ac and F4ad [6,7]. Each variant exhibits a distinctive binding profile to porcine enterocytes. The amino acid sequence of FaeG in the F4 variants shares a conserved epitope designated “a” and each variant has a specific “b”, “c” or “d” epitope [8].

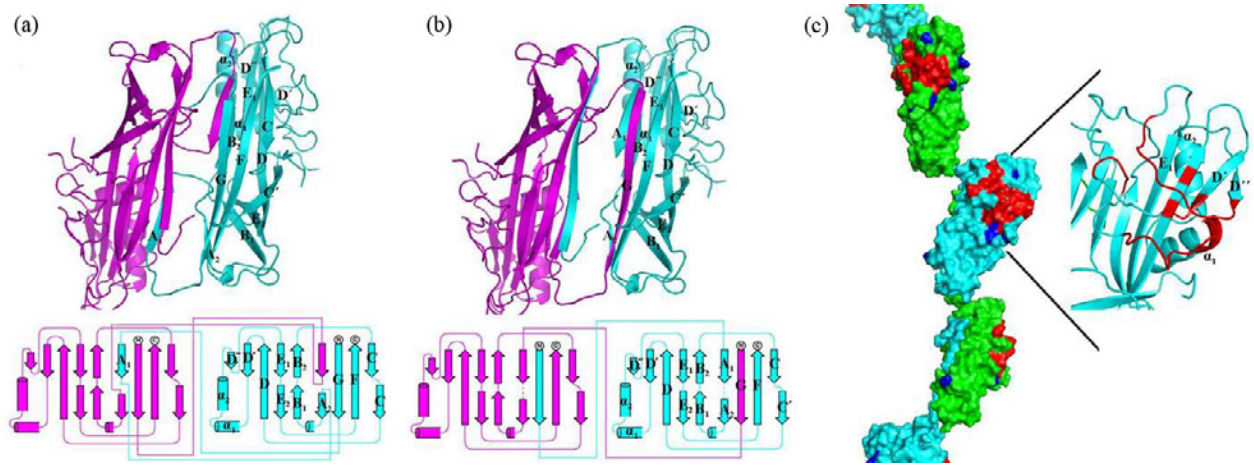

Figure 1. Three dimensional structure of the chloroplast-targeted FaeG. (a) Ribbon presentation and topology diagram of the crystal structure of the chloroplast-targeted FaeG. The molecules are differentiated by color, topology is annotated to the cyan molecule. (b) Ribbon presentation and topology diagram of the unswapped dimer model of FaeG. Residues 31 and 32 are missing in this model (dashed line in the topology diagram). Color and annotations as in (a). (c) Model of F4 fimbriae, based on the structure of FaeG. Molecules are coloured cyan and green, alternating. Residues proposed to be involved in receptor binding are indicated in red. Amino acids differing in the FaeG sequence of ETEC strains 5/95 and GIS26 are indicated in blue. The inset shows a close up view of the mapping of the residues suggested to be involved in receptor binding in a ribbon presentation.

In order to develop a cost-efficient edible vaccine against porcine F4 ETEC infections the major subunit/adhesin of the fimbriae, FaeG, was expressed in plant chloroplasts. Here, we pesent the cystal structure of this chloroplast-derived fimbrial subunit.

In the chloroplast, FaeG was assembled into a strand-swapped dimer (Figure 1(a)). The chloroplast-targeted FaeG protein has an Ig-core made up of strands A_1, A_2, B_1, B_2, C, D, E_1, E_2, F and G, named according to their place in the Ig-fold and in analogy to the nomenclature used for other fimbrial subunit structures. These show an incomplete Ig-like structure, lacking the last β-strand G. In bacteria, in the growing pilus, this G-strand is provided by the Nte of the adjacent subunit. In the chloroplast-targeted FaeG dimer, strand-swapping occurs between strands A_1 and A_2. As a result, the A_1 strand of monomer 1 is inserted in the Ig-core of monomer 2 and the Nte of the chloroplast-targeted FaeG monomer becomes self-complementing.

The Ig protein family shows a large degree of variation. The Ig-fold contains a common structural core of only four strands (B, C, E and F) [9]. The addition of three to five strands determines topological subclasses of this large protein family. In the Ig-fold of the chloroplast-targeted FaeG a short helical turn occurs between strands A_2 and B_1 and an extra strand C' is inserted between strands C and D. In addition to the Ig-folded core, the FaeG structure contains an extra domain introduced between strands D and E1. This domain is made up of strands D' and D", linked by two α-helices α_1 and α_2.

The domain swapped dimeric structure of FaeG is incompatible with the donor strand complementation/exchange in fimbrial assembly. We rearranged the FaeG structure in a way that strand A_1 is retained within the Ig-core of one monomer. In the resulting unswapped dimer the monomers mutually donor strand complement each others fold (Figure 1(b)). Although this unswapped dimer is still not functional in fiber formation, the FaeG structure in the model is closer to the one presumed to be present in F4 fimbriae, where the hydrophobic groove of one subunit is shielded by the Nte of the adjacent subunit.

Conserved as well as variable epitope regions have been proposed to be involved in F4 receptor binding [10]. The structure of FaeG allows us to localize these residues on the surface of FaeG (Figure 1(c)). This suggests that the receptor binding site of FaeG is not located within the Ig-core of the protein but rather in the extra domain made up of strands D' and D" and α-helices α_1 and α_2. Especially the long loop between D' and α_1 and the loop connecting the extra domain to the Ig-core (between D" and E_1) are indicated as part of the binding site.

References

[1] G.E. Soto and S.J. Hultgren, J. Bacteriol. 181, 1059-1071(1999)

[2] D.G.Thanassi, E.T. Saulino and S.J. Hultgren, Curr. Opin. Microbiol. 1, 223-231 (1998)

[3] A.V. Zavialov, J. Kersley, T. Korpela, V.P. Zavyalov, S. MacIntyre, and S.D. Knight, Mol. Microbiol. 45, 983-995 (2002)

[4] F.G. Sauer, H. Remaut, S.J. Hultgren and G. Waksman, Biochim. Biophys. Acta 1694, 259-267 (2004)

[5] R. Piatek, B. Zalewska, K. Bury and J. Kur, Acta Biochim. Pol. 52, 639-646 (2005)

[6] I. Orskov, F. Orskov, W.J. Sojka and W. Wittig, Acta Pathol. Microbiol. Scand. 62, 439-447 (1964)

[7] P.A. Guineeb and W.H. Jansen, Infect. Immun. 23, 700-705 (1979)

[8] F. G. van Zijderveld, J. Anakotta, R.A. Brouwers, A. M. van Zijderveld, D. Bakker and F.K. de Graaf, Infect. Immun. 58, 1870-1878 (1990)

[9] P. Bork, L. Holm and C. Sander, J. Mol. Biol. 242, 309-320 (1994)

[10] D. Bakker, P.T. Willemsen, L.H. Simons, F.G. van Zijderveld and F.K. de Graaf, Mol. Microbiol. 6, 247-255 (1992)