

## Samenvatting Kristof Moonens

De *Escherichia coli* bacterie is een belangrijk ziekteverwekkend micro-organisme. Bij zowel mens als dier is *E. coli* verantwoordelijk voor infectie van het gastro-intestinaal kanaal, de urinewegen en de blaas, de nieren en zelfs de hersenen. Vers in het geheugen ligt de uitbraak van de *E. coli* O104:H4-stam in mei 2011 in het noordelijk deel van Duitsland met 36 sterfgevallen. Ook in zoogdieren zijn deze infecties door pathogene *E. coli* wijdverspreid en worden in de varkensindustrie geassocieerd met respectievelijk spreediarree en slingerziekte, veroorzaakt door Enterotoxigene *E. coli* (of ETEC) en Shiga toxine producerende *E. coli* (of STEC) stammen. Deze ETEC en STEC stammen zijn verantwoordelijk voor grote economische verliezen in de varkensindustrie door hoge mortaliteit, morbiditeit, groeivertraging en kost van medicatie. ETEC en STEC stammen bezitten specifieke virulentiefactoren die adhesie aan specifieke gastheerreceptoren bewerkstelligen, gevolgd door de secretie van toxines (oa hitte labiele en stabiele enterotoxines en het Shiga toxine) die de ziektesymptomen veroorzaken. De cruciale aanhechting aan gastheerweefsel, nodig voor bacteriële kolonisatie, gebeurt via lange, haarachtige organellen aanwezig op het bacterieel celoppervlak en worden fimbriae of pili genoemd. De ETEC en STEC stammen die verantwoordelijk zijn voor spreediarree en slingerziekte produceren meestal de F4 en F18 fimbriae uit die deze stammen toelaten de dunne darm te koloniseren. F18 fimbriae beschikken over het FedF tipadhesine, dat bestaat uit twee domeinen en waarbij het N-terminaal domein de adhesie bewerkstelligt. Bij de F4 fimbriae wordt de rol van structurele subeenheid en adhesine gecombineerd in FaeG.

In deze doctoraatsthesis werd de functionele suikerreceptor bepaald van F18 fimbriae met behulp van een natuurlijke glycaan array screening. Bij deze techniek worden glycanen afgesplitst van natuurlijk weefsel, in dit geval van bigenterocyten, en vervolgens geïmmobileerd op een vaste matrix. Een daaropvolgend bindingsexperiment toonde aan de FedF interageerde met bloedgroep A type-1 antigen. We bepaalden vervolgens de co-complex structuur van de FedF en FaeG adhesines met hun respectievelijke suikerreceptoren, bloedgroep A type-1 antigen voor FedF en lactose voor FaeG, door middel van X-straaldiffractie. Grote verschillen bestaan in de manier van binden tussen beide adhesines. Waar FedF een langgerekte bindingssite heeft die vele interacties maakt met de suikerketen heeft FaeG een meer compacte bindingssite die interageert met slechts één monosaccharide. Dit resulteert in een lagere bindingsaffiniteit voor FaeG in vergelijking met FedF, en toont het belang aan van vele simultane interacties bij de aanhechting van poly-adhesieve F4 fimbriae aan gastheerreceptoren. Op basis van de co-complex structuur van beide adhesines waren we in staat de receptor specificiteit van de F4 en F18 fimbriae te verklaren. Daarnaast toonden mutatiestudies het belang aan van een paar geconserveerde lysines aanwezig op een lus die uitsteekt van de tip van het FedF lectine domein, maar verder niet betrokken zijn bij receptor binding. Uit

experimentele analyses bleek dat deze positief geladen lus in de nabijheid van het lipidemembraan komt na interactie van FedF met het bloedgroep A type-1 glycosfingolipide waarna elektrostatische interacties met vetzuren aanwezig in het celmembraan extra bindingsaffiniteit geeft en de FedF-receptor interactie verder stabiliseert. Dit "coincidence binding" mechanisme voorkomt de verwijdering van F18-positieve bacteriën door gesecreteerde glycoproteïnen door de binding te richten naar bloedgroep determinanten aanwezig in het celmembraan. Gesecreteerde glycoproteïnen zijn belangrijk als verdedigingsmechanisme voor de gastheer doordat deze zich gedragen als afleidingsreceptoren welke de interactiesites bestemd voor pathogene adhesie verzadigen. Literatuur suggereert dat dit "coincidence binding" mechanisme belangrijk kan zijn in de interactie van andere glycolipide bindende fimbriële adhesines.

De co-complex structuur van FaeG met lactose bood het eerste structurele inzicht in suikereceptor binding door poly-adhesieve fimbriae. Drie natuurlijk voorkomende serologische varianten van F4 fimbriae bestaan (F4ab, F4ac en F4ad) en de verschillen tussen deze varianten hebben een grote invloed op de conformatie die de bindingssite aanneemt. De co-complex structuur van de F4ad variant met het lactose ligand werd namelijk door ons bepaald en vergeleken met de andere varianten. Dit toonde aan dat de andere varianten niet in staat zijn lactose te binden door een verlies aan stabiliserende interacties. Ofwel zal de bindingssite van deze andere varianten een belangrijke conformationele wijziging ondergaan bij de binding van lactose of er is een verschillende bindingssite aanwezig op het oppervlak van beide varianten. Beide mogelijkheden zijn nieuwe bindingsmechanismen door fimbriële adhesines en bepaling van de co-complex structuren van de FaeG adhesines van F4ab en F4ac met hun respectievelijke suikerliganden zal licht werpen op deze intrigerende vraag.

Als samenvatting heeft deze PhD thesis bijgedragen tot het structureel inzicht in de receptorbinding door fimbriële adhesines van ETEC en STEC stammen die verantwoordelijk zijn voor speendiarree en slingerziekte. We bepaalden de eerste co-complex structuur van een één-domein fimbriële adhesine met zijn suikereceptor. De verkregen co-complex structuren wijzen op het bestaan van verschillende bindingsmechanismen waarmee deze fimbriële adhesines interageren met eukaryote suikereceptoren.

## Summary Kristof Moonens

The *Escherichia coli* bacterium is a normal inhabitant of the gastrointestinal tract, but also an important disease causing microorganism. In humans and animals, *E. coli* is responsible for infections of the gastrointestinal tract, the urinary tract, the kidneys and even the brains. Not long ago an outbreak of the *E. coli* O104:H4 strain in May 2011 in the northern part of Germany resulted in 36 deaths. In mammals as well these infections by pathogenic *E. coli* are widespread and in pig industry are associated with post weaning diarrhea and edema disease, caused by enterotoxigenic *E. coli* (or ETEC) and Shiga-toxin producing *E. coli* (or STEC) strains. These ETEC and STEC strains account for significant economic losses due to an increase in mortality and morbidity, decreased growth rate, and cost of medication. ETEC and STEC strains possess specific virulence factors to attach to host specific receptors, followed by the secretion of toxins (heat labile and stable enterotoxins and Shiga toxins) that cause the disease symptoms. The crucial adherence to host tissue, required for colonization, is most often performed by long, hair-like cell surface proteinaceous appendages designated fimbriae or pili. The ETEC and STEC causing post-weaning diarrhea and edema disease mostly express F4 and F18 fimbriae on their cellular surface to attach and colonize the small bowel. To attain attachment F18 fimbriae employ the FedF tip-adhesin, which consists of two domains with the N-terminal domain being responsible for receptor mediated adhesion. In F4 fimbriae the role of structural subunit and adhesin are combined in FaeG.

In this PhD thesis we first determined the functional carbohydrate receptor of F18 fimbriae by use of a natural glycan array. In this method glycans are separated from natural tissue, in this case piglet enterocytes, and immobilized on a solid matrix. A binding experiment with FedF demonstrated it is interacting with blood group A type-1 antigen. Subsequently, we determined the co-complex structures of the FedF and FaeG adhesins with their respective carbohydrate receptors, blood group A type-1 antigen for FedF and lactose for FaeG, by X-ray diffraction. Important differences exist in the way both adhesins bind their respective target receptors. FedF possesses an extended binding site making many interactions with the glycan chain, whereas FaeG has a more compact glycan binding site which interacts with just one monosaccharide. This results in a lower affinity of binding for FaeG towards its glycan ligand, in comparison to FedF, and demonstrates the importance of many simultaneous interactions in the adhesion of poly-adhesive F4 fimbriae towards host tissue. Based on the co-complex structures we were able to explain the receptor specificity of F4 and F18 fimbriae. Mutational studies also showed the importance of a pair of conserved lysine residues present on a loop protruding from the tip of the FedF lectin domain. This positively charged loop comes into close proximity of the lipid bilayer upon interaction of FedF with blood group ABH type-1 glycosphingolipids. An electrostatic interaction of the loop with acidic lipids in the cell membrane

provides additional binding affinity and would further stabilize the FedF-receptor interaction. This so-called coincidence binding mechanism would avoid clearance by secreted glycoproteins by directing binding preferentially to blood group determinants embedded in the cellular membrane. Secreted glycoproteins have been unveiled to be important as host innate defense mechanism by acting as decoy receptors for fimbrial adhesins. Literature hinted the coincidence binding mechanism might be important in the binding of other glycolipid binding fimbrial adhesins.

The co-complex structure of FaeG with lactose provided the first structural insight in carbohydrate receptor binding by poly-adhesive fimbriae. It shows that FaeG binds its ligand via an additional subdomain, which is no part of the classic fold of fimbrial subunits. Three naturally occurring serological variants of F4 fimbriae exist (F4ab, F4ac and F4ad), each featuring a related but yet different binding and hemagglutination profile. Differences between variants are predominantly located on and around the additional subdomain and have a profound effect on the conformation the binding site adopts. The FaeG-lactose co-complex structure was determined for variant ad and a structural comparison with the other two variants showed them unable to accommodate a lactose ligand in a similar position and orientation due to a loss of stabilizing interactions. Either the binding site of both variants undergoes a significant conformational change upon binding the lactose or a distinct binding site is present on the surface of both variants. Both possibilities represent novel binding mechanisms by fimbrial adhesins and determination of the co-complex structures for the two other variants with their respective carbohydrate ligands would shed light on this intriguing question.

In summary this PhD thesis contributed to the structural insight of receptor binding by fimbrial adhesins expressed by ETEC and STEC causing post weaning diarrhea and edema disease. We provided the first co-complex structure of a single-domain fimbrial adhesin with its carbohydrate ligand. The obtained co-complex structures suggest the existence of different binding mechanisms by which fimbrial adhesins interact with eukaryotic cell-based glycan receptors.