STRUCTURE AND FUNCTION OF THE LECTIN FIMH FROM ESCHERICHIA COLI

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

JONATHAN ERAS

M.Sc. in Biochemistry, TU Munich

born on 17.03.1986

citizen of Germany

accepted on the recommendation of

Prof. Dr. Rudi Glockshuber, examiner
Prof. Dr. Eilika Weber-Ban, co-examiner
Prof. Dr. Julia Vorholt, co-examiner
Prof. Dr. med. Olivier Devuyst, co-examiner

2017
# TABLE OF CONTENTS

SUMMARY ........................................................................................................................................ 1

ZUSAMMENFASSUNG .......................................................................................................................... 3

1. INTRODUCTION ........................................................................................................................... 7
   1.1 Surface organelles of Gram-negative bacteria ............................................................................ 7
   1.2 Pili: non-flagellar bacterial surface appendages .......................................................................... 10
   1.3 Type 1 pili from Escherichia coli ............................................................................................... 14
   1.4 The type 1 pilus adhesin FimH .................................................................................................. 16
   1.5 Catch-bond behavior in biomolecules ....................................................................................... 18
   1.6 Urinary tract infections in humans ............................................................................................ 21
   1.7 The glycoprotein uromodulin ..................................................................................................... 25
   1.8 Goals of the thesis ..................................................................................................................... 27

2. RESULTS ......................................................................................................................................... 29
   2.1 Catch-bond mechanism of the bacterial adhesin FimH ............................................................... 29
      2.1.1 Introduction .......................................................................................................................... 30
      2.1.2 Results .................................................................................................................................. 32
      2.1.3 Additional results .................................................................................................................. 46
      2.1.4 Discussion ............................................................................................................................ 51
      2.1.5 Materials and Methods ....................................................................................................... 55
   2.2 Ligand binding is coupled to domain interaction in FimH ........................................................... 63
      2.2.1 Introduction .......................................................................................................................... 64
      2.2.2 Results .................................................................................................................................. 67
      2.2.3 Discussion ............................................................................................................................ 75
      2.2.4 Materials and Methods ....................................................................................................... 77
   2.3 A natural ligand of FimH – The glycoprotein uromodulin .......................................................... 81
      2.3.1 Introduction .......................................................................................................................... 82
      2.3.2 Results .................................................................................................................................. 86
      2.3.3 Discussion ............................................................................................................................ 95
      2.3.4 Supplementary information ............................................................................................... 97
      2.3.5 Materials and Methods ....................................................................................................... 102

3. GENERAL DISCUSSION AND OUTLOOK .................................................................................. 107

4. REFERENCES .................................................................................................................................. 113

Acknowledgements .......................................................................................................................... 129
This page was intentionally left blank.
SUMMARY

Adhesive type 1 pili of uropathogenic *Escherichia coli* strains are filamentous, extracellular protein complexes that enable bacterial attachment to epithelial cells of the urinary tract. Adhesion to the epithelium is mediated by the lectin FimH at the tip of type 1 pili. FimH is a two-domain protein. It consists of a lectin domain (FimHL) that binds terminal mannoses in N-linked glycans on epithelial target receptors, and a pilin domain (FimHP) that interacts with the lectin domain and anchors FimH to the rest of the pilus. The subunits in the pilus are linked non-covalently by a mechanism in which the incomplete immunoglobulin-like fold of the pilin domain is complemented by an N-terminal peptide extension, called donor strand (Ds), of the successive subunit. FimH shows the remarkable ability to increase its apparent affinity to target glycans when its domains become separated under tensile mechanical force. The main goal of the thesis was to elucidate this mechanism by a comprehensive structural analysis and by comparing the ligand binding dynamics of the isolated lectin domain (FimHL) with those of the full-length protein using *n*-heptyl α-D mannopyranoside (HM) as a model ligand.

We first established an experimental system to produce a stable FimH monomer with the properties of FimH in the context of the assembled pilus. For this purpose, the complex between FimH and the pilus assembly chaperone FimC was purified and mixed with an excess of a synthetic peptide (DsG) that corresponds to the N-terminal extension of the subsequent pilus subunit FimG. After displacement of FimC by DsG, the resulting FimH-DsG complex could be purified in milligram quantities. Comparison of the kinetics of ligand binding and dissociation of FimHL (representing the domain-separated state of FimH under tensile mechanical force) and FimH-DsG revealed that the isolated lectin domain binds HM with a 3300-fold higher affinity compared to full-length FimH. This dramatic difference in affinity primarily resulted from a more than 100’000-fold slower dissociation of HM from FimHL compared to full-length FimH, with dissociation half-lives of 58 minutes for FimHL compared to 30 milliseconds for full-length FimH. The solved X-ray structures of FimH-DsG in the absence and presence of the ligand HM demonstrated that the interactions between the pilin and the lectin domain are not affected by ligand binding and intramolecularly catalyze spontaneous ligand release in full-length FimH. In contrast, the X-ray
structure of the FimHₖ-HM complex showed that the FimHₖ surface loops interacting with the pilin domain underwent large conformational changes that no longer allow the interaction between FimHₖ and the pilin domain. This was confirmed by the X-ray structure of the domain separated state of a variant of full-length FimH in complex with HM. The results showed that domain separation in FimH only occurs under mechanical stress, allowing bacterial movement along the urinary epithelium in the absence of mechanical force.

The results of the first part predicted that FimH variants with weakened interdomain-interactions should show higher ligand binding affinities. In the second part of the thesis, mutations were introduced into FimH to generate variants of full-length FimH in which the domain-separated state is favored. This was achieved with two strategies. First, by introducing single amino acid replacements into the FimH pilin domain that were predicted to disturb the domain interactions, and second, by amino acid insertions in the linker region between the domains. As predicted, two FimH variants with the amino acid substitutions Ala188Asp or Ala188Trp in the pilin domain showed a 10- to 20-fold increase in affinity to HM, while the affinity increased up to 2100-fold with increasing number of inserted residues in the linker between the FimH domains. The results could be confirmed qualitatively by recording the free energy of domain interaction in the FimH variants, revealing that affinity for HM increased with decreasing interaction energy.

In the third part of the thesis it could be shown that the glycoprotein uromodulin (Tamm-Horsfall protein), the most abundant protein in human urine, is a natural ligand of FimH. Uromodulin thus most likely contributes to protection against urinary tract infections by competing with urinary epithelium cells for binding to FimH. Kidney cells secrete uromodulin as a filamentous oligomer that is highly glycosylated. The uromodulin fibers were purified and showed a binding stoichiometry of two to three FimHₖ molecules per uromodulin monomer with an apparent dissociation constant of \((4.0 \pm 1.0) \cdot 10^{-7}\) M. The three-dimensional architecture of uromodulin filaments was investigated with cryo-electron tomography. The results indicate that the fibers adopt a regular zigzag architecture in which successive monomers are rotated by 180° relative to each other and the C-terminal ZP domain of uromodulin forms the core of the fibril.
ZUSAMMENFASSUNG


Im dritten Teil der Arbeit konnte gezeigt werden, dass das Glykoprotein Uromodulin (Tamm-Horsfall-Protein), das in großen Mengen im menschlichen Urin vorkommt, ein natürlicher Ligand von FimH ist. So beugt Uromodulin möglicherweise Harnwegsinfektionen vor, indem es mit Epithelzellen im Harntrakt um die Bindung von FimH konkurriert. Nierenzellen sekretieren Uromodulin als hoch-glykosiliertes Filament. Diese Uromodulin Fibrillen wurden gereinigt und eine Bindungsstöchiometrie von drei bis vier FimH Molekülen pro Uromodulin-Monomer mit einer apparenten Dissoziationskonstante von \((4.0 \pm 1.0) \cdot 10^{-7} \text{M}\) bestimmt. Der dreidimensionale Aufbau von Uromodulin Filamenten wurde mittels Elektronenkryotomographie untersucht. Die Ergebnisse deuten darauf hin, dass die Fibrillen eine regelmäßige Zickzack Anordnung zeigen, bei der Monomere um 180° entlang der Achse gegeneinander rotieren und dass die C-terminalen ZP Domänen von Uromodulin den Kern der Fibrillen bilden.
This page was intentionally left blank.
1. INTRODUCTION

1.1 Surface organelles of Gram-negative bacteria

A remarkable array of diverse macromolecular assemblies to display or secret virulence factors across the bacterial cell envelope has been evolved by bacteria. These protein complexes are used in many different processes. For example, to deliver toxins, find nutrient sources, transport proteins and nucleic acids or assemble surface organelles such as pili for enhancing binding to eukaryotic cells. They can be divided into different groups, based on their architecture, function, and specificity (Green et al. 2016). Here, they will be classified into three groups (Figure 1) according to their anchoring in the bacterial membrane (Costa et al. 2015): (i) assemblies spanning both the inner membrane and the outer membrane, (ii) those that span the outer membrane only, and (iii) single inner membrane spanning transporters. The first group is built by the currently known type 1 secretion system (T1SS), T2SS, T3SS, T4SS, and T6SS. The bacterial multidrug efflux pumps from the RND (Resistance-Nodulation-Division) superfamily can be added to this group. To the second group belong the T5SS auto-transporter, the chaperone-usher pathway assemblies, and the curli secretion machinery. The majority of secretory proteins pass across the inner membrane via the Sec or Tat pathway, which belongs to the third group.

Figure 1. Surface organelles of Gram-negative bacteria. Schematic representation of secretion systems spanning both the inner and outer membrane (blue), the outer membrane only (red) and the inner membrane only (yellow). C/U: chaperone/usher pathway, T1-6SS: Type 1-6 secretion system. (Adapted from Green et al. 2016).
Due to the two-membrane system in Gram-negative bacteria, the secretion of substrates involves either a one-step or two-step secretion mechanism (Rego et al. 2010). A one-step mechanism is used by almost all (apart from T2SS) double-membrane spanning secretion systems. Here, substrates are transported directly from the cytoplasm into the extracellular space or injected into a target cell. Two-step mechanisms, involving above mentioned group ii and group iii systems, must be used by all outer-membrane anchored secretion systems and the T2SS, such that substrates are first translocated into the periplasm by inner-membrane transporters (SecYEG translocon or Tat system) and afterwards transferred and secreted by the second component.

Currently it is known that most of group i and group ii systems secret unfolded or partially folded substrates. In contrast, folded and partially folded substrates are secreted by the T2SS, T6SS and the chaperone-usher pathway (Costa et al. 2015). Analogous in the inner membrane, unfolded proteins are primarily translocated by the Sec pathway, whereas folded proteins are secreted by the Tat pathway (Robinson et al. 2004).

Substrates, ATP-dependently secreted by the T1SS are diverse in size and function but often associated with nutrient acquisition (such as the iron binding protein HasA) and virulence (for example the pore-forming toxin HlyA) (Kanonenberg et al. 2013). The above mentioned and closely related RND system uses a proton gradient to pump small molecules out of the cell, including synthetic antibacterial compounds, thus contributing to antibiotic resistance mechanisms (Piddock 2006). The T2SS secretes folded proteins for bacterial survival in the host or environment (hydrolyzing enzymes) or toxins (for example cholera toxin) from the periplasm into the extracellular space in an ATP-dependent manner (Nivaskumar et al. 2014). The T3SS are specialized protein delivery machines spanning both bacterial membranes and deliver effector proteins from the cytoplasm to modulate host cell functions for bacterial invasion and colonization (Buttner 2012, Cornelis 2006). T3SS substrates are recognized by specific chaperones and the secretion mechanism is catalyzed by ATP hydrolysis. On top of the ability to mediate the secretion of many different proteins involved in bacterial pathogenesis, the T4SS has the unique function to translocate single-stranded DNA, thus enabling the conjugation of plasmid DNA with other bacteria (Cascales et al. 2003). The energy for this process is delivered by three ATPases. The T6SS is a huge machine composed of modules spanning from
1. INTRODUCTION

the cytoplasm to the outer membrane (Mougous et al. 2006). From its structure it resembles an intracellular membrane-associated contractile phage tail (Pukatzki et al. 2007), punctures prokaryotic and eukaryotic cells and delivers toxic effector proteins important for pathogenesis and competition between bacteria (Ho et al. 2014). It is speculated that substrate molecules might be fused as N-terminal or C-terminal extension or bind non-covalently to the spike of the T6SS (Shneider et al. 2013). Additionally, the tube of the T6SS could be filled with effector molecules which are released after contraction (Silverman et al. 2013). Recycling of the contracted structure happens in an ATP-dependent process (Basler et al. 2012).

The T5SS is also referred to as the autotransporter system, as the substrate is fused to the secretion pore. It is composed of a secreted domain and a transmembrane domain embedded in the outer membrane. Mainly virulence factors from the periplasm, which are involved in cell-to-cell adhesion and biofilm formation, are secreted. Transferred domains or substrates might either remain anchored with the outer-transmembrane domain, for example adhesins, or be cleaved by its peptidase domain and released into the extracellular environment (Leyton et al. 2012). The chaperone/usher pathway was extensively studied and bacteria use this mechanism to assemble and secrete heteroologomeric surface appendages, which enable the bacteria to recognize and attach to host cells, thereby contributing to bacterial pathogenicity (Wright et al. 2007). Best characterized are the type I and P pili (Lillington et al. 2014), which are used by uropathogenic bacteria to colonize and infect the urinary tract. These and other pili types will be discussed in the following (Chapter 1.2). Like almost all outer membrane transport systems, no ATP or proton gradient is required for the assembly or transport. As there is no ATP in the periplasmic space, the energy might be derived from the folding of super-stable, low-energy structures after being translocated from the secretion channel. Curli are displayed on the bacterial surface as protein fibers or amyloids and protect bacteria from being detected by host immune defense mechanisms. It was speculated that the energy for their assembly comes from a diffusion-based, entropy driven assembly mechanism (Goyal et al. 2014).
1. INTRODUCTION

1.2 Pili: non-flagellar bacterial surface appendages

Gram-negative bacteria have the ability to present a repertoire of different oligomeric protein complexes on their cell surface for a variety of functions. Due to their morphology under the electron microscope they were traditionally grouped into two categories: flagella, which are long, lash-like structures, required for bacterial movement; or pili which are thinner, hair-like structures, usually required for adhesion (Ottow 1975). The focus in this chapter will be on pili. Various schemes have been proposed over the years to classify the different types (Duguid et al. 1966, Orskov et al. 1990, Ottow 1975, Thanassi et al. 2012). Here, they will be classified according to their assembly mechanism into five groups: chaperone- usher pili, curli, type IV pili, type III secretion pili, and type IV secretion pili (Fronzes et al. 2008, Rego et al. 2010, 2010, Steadman et al. 2014). These adhesive pili play a crucial role during infection by pathogenic bacterial strains. A schematic representation of the structures and assembly mechanism from Fronzes et al. 2008 is shown in Figure 2.

![Figure 2. Pili and their assembly machineries in Gram-negative bacteria. Schematic representation of the different types of pili displayed at the surface of Gram-negative bacteria. Chaperone/usher pili and curli are fibers attached to the outer membrane. Exemplarily for the class of chaperone/usher pili, the type 1 and P pilus subunits and assembly proteins are shown. Differences between this examples](image-url)
are marked with *. It indicates (i) that there are 5–10 PapE subunits in the tip fibrillum of the P pilus, whereas the equivalent subunit in the type 1 pilus system, FimF, is present in only one copy; and (ii) that the PapK adaptor between the PapE and PapA polymers does not have an equivalent in type 1 pili. Chaperone/usher pili and curli assembly is catalyzed by simple systems located at the outer membrane. Type IV pili, type III secretion and type IV secretion pili are assembled by large multisubunit machines crossing the whole bacterial cell envelope. Copied, with the approval of the journal, from (Fronzes et al. 2008).

The first class consists of pili assembled by the chaperone/usher pathway. They are the most abundant bacterial surface appendages and around 50 members have been reported belonging to this class (Vetsch et al. 2005, Zav'yalov et al. 2010). Chaperone/usher pili enable attachment of bacterial cells to host cell glycoproteins and glycolipids mainly in the urinary tract via carbohydrate-specific lectins, termed adhesins. For their biogenesis unfolded pilus subunits are transported via the Sec machinery into the periplasm where they are oxidized and bound by a chaperone, which enables folding and prevents premature aggregation. The chaperone/subunit complexes are transported to and recognized by an outer membrane assembly protein, the so-called usher. Here, subunits are secreted into the extracellular space and incorporated in the growing pilus. Most extensively studied examples of this class are type 1 and P pili with specific functions (Busch et al. 2015, Lillington et al. 2014). The function of type 1 pili is to promote attachment to epithelial cells in the bladder. In the next step bacterial cells can invade into deeper layers underneath the epithelial cells. After some days, biofilm formation can occur. These processes happen in the lower urinary tract (cystitis) (Mulvey et al. 1998). In contrast, the target of P pili is the upper urinary tract. These pili are associated with the development of acute kidney infections (pyelonephritis) (Roberts et al. 1994). The structural architecture of type 1 and P pili is very similar, consisting of a homo-multimeric rod, forming a right handed helix and a fibrillar tip containing an adhesin (lectin). These two parts are connected by a linker consisting of a variable number of adaptor subunits. The subunit composition and assembly mechanism of type 1 pili will be discussed in more detail in Chapter 1.3.

Curli were first discovered in the late 1980s having a thin and coiled appearance in electron micrographs (Olsen et al. 1989). They built the major proteinaceous component of a complex extracellular matrix produced by certain bacteria. For their assembly, partially folded subunits are secreted in a monomeric form to the cell surface. Here, they polymerize into amyloid-like fibers (Chapman et al. 2002)
facilitated by accessory proteins. Curli bind to many host proteins including extracellular matrix proteins (such as Fibronectin and Laminin) and components of the immune system (such as MHC class I and TLR2) (Barnhart et al. 2006). But the molecular mechanism of the interaction has not been fully understood.

Type IV pili are long, flexible, and very stable helical polymers produced by many Gram-negative bacteria, including several human pathogens. They are essential for bacterial virulence and involved in many processes including attachment to host cells, bacterial cell motility, biofilm formation or invasion in epithelial cells (Craig et al. 2008, Piepenbrink et al. 2016). In contrast to chaperone/usher pili and curli, type IV pili are assembled by a huge ATP driven machinery, spanning both inner and outer membranes. Interestingly, type IV pili have the unique feature to retract. The disassembly is catalyzed by specific ATPases (Misic et al. 2010, Satyshur et al. 2007).

Type III secretion pili are needle-like surface structures. Their assembly involves over 20 proteins and is related to the type III secretion system. It is composed of a basal body that spans across both outer and inner membranes and the periplasm, an external needle on the surface of bacteria and the translocon, a pore-forming protein complex that is inserted into the membrane of the host cell (Worrall et al. 2011). On the cytoplasmic side other accessory proteins as well as ATPases are associated. At the distal end of the external needle specific proteins are located involved in pore formation (Sato et al. 2011). Therefore, these pili have also been called injectosomes enabling the entry of bacteria into host epithelial cells (Galan et al. 2006).

Type IV secretion pili are evolutionarily derived from proteins involved in bacterial conjugation. They are important for establishing a stable and specific contact between cells. Thereby they can transfer many different substrates. From single proteins to protein-protein and protein-DNA complexes (Wallden et al. 2010). This system is both used by plant and human pathogens to secrete virulence factors into the host (Llosa et al. 2009). Little is known about the assembly mechanism of type IV secretion pili. Best studied are the type IV secretion pili of *Agrobacterium tumefaciens*. They are composed of 12 proteins, including several subunits forming the inner and outer membrane channel, two subunits forming the extracellular pilus and three ATPases delivering the energy for the assembly and disassembly of the pili (Fronzes et al. 2009).
In recent years, the structural and molecular mechanism of pili has been studied in great detail and considerable progress has been made. Due to their important function for the pathogenicity of bacteria, these advances will hopefully be used and fuel the development of new antimicrobial strategies (Steadman et al. 2014). Examples and recent advances in this direction will be described in Chapter 1.6.
1. INTRODUCTION

1.3 Type 1 pili from *Escherichia coli*

Type 1 pili are 0.5-2 µm long, very stable protein-filamentous cell surface organelles anchored to the outer membrane of Gram-negative bacteria and are required for bacterial attachment to host cells (Eden et al. 1977). The pilus is composed of five structural subunits. The structure mainly visible in electron micrographs forms a right-handed, helical pilus rod, which is made by up to several thousand copies of the main structural subunit FimA (Hahn et al. 2002) (Figure 3A). A single copy of the functional subunit FimH sits at the tip of the pilus and is connected to the pilus rod with the subunits FimG and FimF forming the short tip fibrillum (Figure 3B). FimH is the only two domain structural subunit composed of an N-terminal lectin domain (FimHL) that mediates binding to terminal mannose residues of glycoprotein receptors and a C-terminal pilin domain (FimHP) which connects FimH to the remaining pilus.

Structurally homologous proteins are the subunits FimA, FimF, FimG, and FimHP. They all share a non-canonical immunoglobulin-like fold without the C-terminal β-strand (Choudhury et al. 1999). But with the exception of FimH these subunits possess an N-terminal extension of 15-20 residues, termed donor strand. This small peptide inserts into the preceding subunit in an anti-parallel orientation relative to its C-terminal β-strand and thereby completes its fold in the assembled pilus. All donor strand peptides contain a pattern of alternating hydrophobic residues which point inside an open hydrophobic groove of each subunit. This was experimentally shown by crystal structures of subunit/subunit complexes and the tip fibrillum. It was termed “donor strand complementation” mechanism and explained the assembly from monomeric subunits to a supramolecular protein complex. The assembly follows a general mechanism, the so-called chaperone usher pathway (Nishiyama et al. 2008, Waksman et al. 2009) which is known for more than 50 other different adhesive filaments (Vetsch et al. 2005). During type 1 pilus biogenesis the structural subunits are transported into the periplasm and the formation of a single disulfide bond is catalyzed by DsbA. The periplasmic chaperone FimC catalyzed subunit folding by delivering the missing β-strand. It forms stoichiometric complexes with unfolded, oxidized subunits. These chaperone/subunit complexes are recognized by the usher FimD, located in the outer membrane. Independent of cellular energy sources, FimD catalyzes pilus growth. Here the donor strand exchange reaction occurs, where the
β-strand donated by FimC is replaced by the respective N-terminal extension of the next subunit. A schematic representation of a current model for assembly of type 1 pili is shown in Figure 3B. The assembled subunits are very stable against denaturation. Complete dissociation can only be achieved under very harsh conditions (boiling or incubation in saturated GdmCl solutions for more than two hours at 37 °C). Fortunately it was demonstrated that stable monomeric subunits, for example for crystallographic analysis, can be produced by complementing the lacking donor strand with a synthetic peptide (Sauer F. G. et al. 2002, Vetsch et al. 2004).

Figure 3. Structure of type 1 pili from E. coli. (A) Visualization with negative-stain electron microscopy of a single E. coli cell with type 1 pilus rods extending from the cell surface (E. Hahn, University of Zurich) (B) Subunit composition and biogenesis of type 1 pili showing the quaternary structure of the tip fibrillum, containing the functional subunit FimH, and the pilus rod. The periplasmic chaperone FimC catalyzes subunit folding and mediates transport to the usher FimD, which catalyzes subunit assembly. Adapted from (Puorger et al. 2008).
1. INTRODUCTION

1.4 The type 1 pilus adhesin FimH

The focus of the thesis lies on the study of the terminal subunit FimH, the only two-domain Fim protein in the assembled pilus. The name of the N-terminal lectin domain (FimHL) indicates its function. It specifically binds to terminal mannoses of N-linked glycans. More precisely, FimHL binds to the cell surface receptor uroplakin Ia (Min et al. 2002, Zhou et al. 2001) of epithelium cells of the urinary tract. The pilin domain (FimHP) is connected via a short linker to FimHL and anchors FimH to the pilus. The FimHP fold is homologous to the other pilus subunits, interacting with the neighboring subunit FimG via donor strand complementation (Waksman et al. 2009). The two-domain architecture of FimH may explain its ability to increase its affinity to target carbohydrates under shear stress e.g. occurring during urine excretion: A model has been proposed in which FimH undergoes a conformational transition to a state of higher affinity when tensile mechanical force leads to domain separation (Le Trong et al. 2010). This was derived from two different x-ray structures of FimH and is known as the catch-bond phenomenon (Thomas W. E. et al. 2008). The first structure of FimHL was derived from the entire tip fibrillum structure, which is ligand-free and was later defined as the low-affinity state (open binding pocket). In this state, FimHL, and FimHP specifically interact via loop segments at the domain interface. In the second structure of the isolated lectin domain, butyl mannose was found in the closed binding pocket (Le Trong et al. 2010), with a more extended domain conformation. A comparison of the two mentioned states is shown with the corresponding crystal structures in Figure 4. It was additionally demonstrated that the isolated lectin domain shows an about two orders of magnitude higher ligand binding affinity compared to tip-incorporated FimH (Le Trong et al. 2010). This conformation was defined as the high-affinity state. The interpretation of these data was that the interaction between FimHL and FimHP stabilizes the low affinity conformation and an open ligand binding pocket. The disruption of the domain interface, occurring under mechanical force, would favor the high affinity state of FimHL. This interpretation was supported by mutagenesis studies in which the specific inter-domain contacts in FimH were partially disrupted (Aprikian et al. 2007, Schembri et al. 2000, Tchesnokova et al. 2008, Yakovenko et al. 2008).
1. INTRODUCTION

Figure 4. Crystal structures of low-affinity and high-affinity states of the lectin domain of FimH. (A) 2.7 Å crystal structure of the type 1 pilus tip fibrillum with FimC in dark green, two molecules of FimF in light green, FimG in blue, the pilin domain of FimH in yellow (FimHp) and the lectin domain of FimH in red (FimHl). (Le Trong et al. 2010) (pdb-code: 3JWN). (B) The lectin domain (red) in context of the pilin domain (yellow) from the fimbrial tip structure. (C) The isolated lectin domain (red) in complex with n-butyl α-D-mannopyranoside (BM) (yellow) shown in stick-representation (Bouckaert et al. 2005) (pdb-code: 1UWF). Adapted from (Le Trong et al. 2010).

All cartoons of crystal structures in this thesis were made with Pymol (Schrödinger).
1.5 Catch-bond behavior in biomolecules

In several biological processes, where cell receptors bind to immobilized ligands, the resulting interaction can be exposed to tensile mechanical force. With increasing force this intuitively might be expected to disrupt the interaction. Nevertheless, there are several exceptions. For example, receptors of blood and muscle cells and even bacterial receptors show increasing bond lifetimes when subjected to tensile force. These bonds are referred to as catch-bonds.

In 1988 the term catch-bond was first introduced by scientist trying to find appropriate quantitative models for fitting data of biological adhesion and detachment events (Dembo et al. 1988). Here, the authors mentioned also two practical examples, where the lifetimes of biological adhesive bonds were enhanced by tensile mechanical force. In short, catch-bonds are receptor-ligand bonds showing increasing average bond lifetime upon mechanical stretching of the bond, in contrast to the more conventional slip bonds (Bell 1978), whose lifetime is shortened by tensile mechanical force. Just above a critical force point, catch-bonds behave like slip bonds. This is graphically illustrated in Figure 5.

![Figure 5. Comparison between catch-bond and slip bond models. The effect of tensile mechanical force on average bond lifetime. Receptors are shown in red and black, ligands in grey. Adapted from (Sokurenko et al. 2008, Thomas W. E. et al. 2008).](image)

Today, numerous biomolecules have been reported to exhibit catch-bond behavior. But directly measured they were just recently for the first time with atomic force...
microscopy (AFM) for the dissociation of P-selectin and its ligand (PSGL-1) (Marshall et al. 2003). Later it was demonstrated with biomembrane force probes and flow chamber assays, that also the other two classes of selectins, L-selectin (Sarangapani et al. 2004) and E-selectin (Snook et al. 2010) show catch-bond behavior. Selectins are lectins expressed on different biological surfaces, e.g. on the vascular endothelium (E-, P-selectin), platelets (P-selectin), or leukocytes (L-selectin) and bind to sugars on cell-surface glycoproteins (Preston et al. 2016). The main function of selectins is the guidance of leukocytes to specific tissues by binding to free moving leukocytes from the blood stream to endothelial cells during inflammation. This process can finally lead to wound healing (Ley et al. 2007). Another binding event in the vascular system which is only possible under a defined shear force is the attachment of platelets to the plasma protein von Willebrand factor (Fredrickson et al. 1998). Indeed, catch-bond binding was later demonstrated with AFM experiments between the von Willebrand factor and the platelet glycoprotein Ib alpha (Yago et al. 2008).

Catch-bonds are also important during bacterial adhesion events. Particularly, during urinary tract infections by uropathogenic E. coli (UPEC) strains. This was first experimentally shown with flow chamber assays (Thomas W. E. et al. 2002). Later it became evident that the lectin FimH, located at the tip of type 1 pili bearing UPEC, is responsible for catch-bond binding to terminal mannoses of glycoproteins (Yakovenko et al. 2008). This mechanism enables bacteria to bind tighter to host cells under high shear stress conditions, for example during urine excretion, and in the absence of mechanical force favors bacterial movement on tissue surfaces and host tissue colonization.

Integrins are another class of molecules exhibiting catch-bond formation, demonstrated with AFM measurements (Kong et al. 2009). This was additionally supported by Förster resonance energy transfer experiments, where different conformational states of integrin could be detected (Chigaev et al. 2003). Cell adhesion to the extracellular matrix or to another cell is bridged by integrins. Thereby extracellular signals are transmitted to intracellular structures for many cellular processes, including fibroblast and T cell migration (Akiyama et al. 1989, Shimizu et al. 1990). In this context, biomembrane force probe experiments showed that cancer cell adhesion to the vascular endothelium is mediated by mechanical force activated integrins, an important step in tumor metastasis (Fiore et al. 2014).
1. INTRODUCTION

Just recently it was shown with AFM experiments that the cadherin family of cell-cell adhesion proteins, which play key roles in embryogenesis and in maintaining tissue structure, form catch-bonds. Interestingly, it was demonstrated that the formation of different adhesive conformational states and therefore catch-bond formation is Ca^{2+}-dependent (Manibog et al. 2014). Another study using optical tweezers demonstrated how mechanical force influences interactions of the cadherin/catenin complex with single actin filaments. The experimental data show typical catch-bond behavior, in which force shifts the equilibrium from a weakly bound state to a strongly bound state. A model was derived which could explain how cells influence and stabilize tissue structure while at the same time being able to move and change their form (Buckley et al. 2014).
1.6 Urinary tract infections in humans

Urinary tract infections (UTIs) are one of the major bacterial infections in humans, affecting 150 million people each year worldwide (Stamm et al. 2001). It was estimated that the social costs of these infections, including health care costs and work losses, are approximately 3.5 to 5 billion USD per year in the United States alone (Flores-Mireles et al. 2015, Nielubowicz et al. 2010, Spaulding et al. 2016). The frequency of UTIs is about 4-fold higher in women compared to men. This is mainly due to an anatomical difference. Women have a shorter urethra of 4-5 cm, compared to 13-20 cm in men (Hickling et al. 2015). Therefore, the entry of uropathogenic *Escherichia coli* (UPEC) into the urinary tract is facilitated and bacterial infections happen more often in women. More than 85% of UTIs are caused by urophathogenic *Escherichia coli* strains (UPEC) (Ronald 2002). It was estimated that more than 50% of women in western countries are believed to develop at least one urinary tract infection during their life span; 20-40% will suffer from reoccurring infections (Foxman 2002, 2014, Foxman et al. 2000). Today, UTIs are treated empirically with antibiotics. This strongly contributes to the emergence of antibiotic-resistant strains (Foxman et al. 2013).

The discovery of penicillin in 1928 can be taken as the starting point of the antibiotic era. In the second half of the 19th century many pharmaceutical companies invested in the screening for additional antibiotics from natural sources. Later this shifted to rational design of compounds by synthetic chemist. Several, mainly practical problems made this research direction unattractive and the interest dropped in the development of antimicrobial compounds. One reason for this was that in contrast to the relatively short several day treatment course for most antibiotic infections, chronic diseases require continuous medication, more beneficial for pharmaceutical companies. Therefore, the focus of research and development shifted to these more lucrative therapeutic areas and research projects for new antimicrobials became less. In parallel, resistance to existing antibiotics was increasing. This resulted in a lack of innovative solutions to target pathogenic bacteria (Cusumano Z. T. et al. 2016). In the last decade, the search for alternative therapeutic strategies to target and disrupt pathways related to virulence but not to general bacterial viability has increased (Aminov 2010).
UTIs are classified in the clinic as complicated or uncomplicated. The majority of UTIs are uncomplicated, typically occurring in patients without abnormalities in the urinary tract and who are otherwise healthy. Complicated UTIs describe individuals having additional defects, for example urinary retention caused by kidney dysfunction or catheterization (Nielubowicz et al. 2010).

Currently, uncomplicated UTIs are usually treated with oral intake of antibiotics such as Nitrofurantoin or a combination of Trimethoprim/Sulfamethoxazole (Salvatore et al. 2011). In complicated cases, antibiotics are administered intravenously for a longer time period. But as graphically illustrated in Figure 6 antibiotic resistance to many of the drugs used to treat UTIs emerged relatively fast after their introduction. Additionally, antibiotic treatment can influence the gut microbiota and provoke other side effects (diarrhea, vomiting) (Foxman et al. 2013).

Figure 6. Timeline of antibiotic introduction and emergence of resistance. Adapted from (Foxman et al. 2013).
The first critical step during an UTI infection cycle is the bacterial adherence to the host tissue, necessary for invasion and colonization. After this first step, toxins or other bacterial effectors are delivered to the host cell. Adhesion is mediated by bacterial pili, making them promising drug targets (Steadman et al. 2014). Mainly advances in understanding the molecular mechanism of pilus assembly and function have permitted the development of novel therapeutics, also referred to as anti-adhesive drugs. They can be grouped into compounds inhibiting pilus biogenesis and receptor analogs, blocking the first step of infection through competitive inhibition of binding (Krachler et al. 2013). The first described small molecules inhibiting the chaperone usher pilus assembly were pilicides (Pinkner et al. 2006). They belong to a class of molecules known as pyrisides. By mimicking a β-strand, they were originally designed to disrupt the formation of chaperone/subunit complexes. But experimental data confirmed later that pilicides inhibited P and type 1 pilus growth by preventing binding of chaperone/subunit complexes to the usher (Pinkner et al. 2006). Interestingly, another group of molecules, referred to as curlicides, inhibited curli and type 1 pilus formation in vitro but their mode of action on a molecular level is not clearly understood (Cegelski et al. 2009). The first potential competitive inhibitors, binding to the lectin domain of FimH and therefore inhibiting bacterial adhesion, were oligomannosides and aryl α-D-mannosides, reported already more than 30 years ago (Aronson et al. 1979, Firon et al. 1982, 1983). Their effectiveness was demonstrated in mouse models. This initiated the research on several monovalent mannose-based FimH mannosides. Here, the non-mannose component that remains after hydrolysis of a glycoside (aglycon) was varied. This led to reports on several FimH anti-adhesive molecules containing aglycons, such as n-alkyl, phenyl, dioxocyclobutenylaminophenyl, umbelliferyl, biphenyl, indol(in)ylyphenyl, triazolyl, or thiazolylamino (Kleeb et al. 2016). Furthermore, several multivalent presentations of mannose derivates have been investigated (Hartmann et al. 2011). To sum up, currently many monovalent and oligovalent antagonists with nanomolar affinity have been reported. But for FimH-antagonists which are administered orally and active in the urinary tract, high solubility in aqueous solutions and fast renal excretion rates are needed (Ernst et al. 2009). Therefore the prodrug principle (Albert 1958), in which a drug is delivered and activated at the target site has been transferred to the development of anti-adhesive drugs. The first example is a FimH-antagonist which was demonstrated to undergo
intestinal absorption and fast renal clearance (Klein et al. 2010). In a second example the aqueous solubility could be improved up to 140-fold using a phosphate prodrug (Kleeb et al. 2016). The potency of these drugs has been verified in mouse models, but today they are not commercially available.

The fact that FimH is highly specific for terminal mannoses, the simplest idea of consuming D-Mannose as a prophylaxis for recurrent UTIs was just recently tested in humans (Kranjcec et al. 2014). In this study around 300 women, with history of acute cystitis and initial antibiotic treatment were split into three groups: One group received 2 g of D-mannose powder daily, the second received 50 mg Nitrofurantoin daily, and the third group received just water as a control. 15 % in the D-mannose group, 20 % in Nitrofurantoin group, and 61 % in the control group suffered under recurrent UTIs. The authors summarize that there was no difference between the intake of D-mannose and the antibiotic. But patients taking D-mannose had fewer side effects, compared to the antibiotics group. D-mannose powders are currently sold by companies for prevention of UTIs in humans (Falcento or Nutrin, Germany).
1. INTRODUCTION

1.7 The glycoprotein uromodulin

In the 1950s, Tamm and Horsfall isolated a high molecular weight glycoprotein from the urine of several species inhibiting viral hemagglutination (Porter et al. 1955, Tamm et al. 1950, 1952). Due to its immunosuppressive function the so-called Tamm-Horsfall protein was later also termed uromodulin (UMOD) (Muchmore et al. 1985). Its gene is evolutionary conserved and was found in all vertebrates (Badgett et al. 1998). In human urine it is the most abundant protein, secreted about 20-200 mg per day (Hunt et al. 1985). It is only produced in the thick ascending limb of the loop of Henle and secreted via proteolytic cleavage by the protease hepsin (Brunati et al. 2015). In the urine it is present as a high molecular weight polymer ($M_r 1 - 10 \cdot 10^6$ Da) that appears at the electron microscope as a linear filament with a width of about 10 nm and an average length of 2500 nm (Fletcher et al. 1970, Porter et al. 1955). UMOD forms gel-like structures under high ionic strength conditions. This property was used to purify the protein (Serafini-Cessi et al. 1989, Tamm et al. 1952).

Numerous reports exist about the function of UMOD. It was speculated to have a protective function against UTI by capturing uropathogenic bacteria and facilitating their clearance (Orskov et al. 1980). In this context it was shown that UMOD can encapsulate bacteria (Kuriyama et al. 1986), inhibit their adhesion to surfaces (Parkkinnen et al. 1988) or bind to type 1 piliated $E.coli$ (Cavallone et al. 2004, Pak et al. 2001). Additional evidence for this function comes from the observation that UMOD knockout mice suffer under UTI more significantly compared to WT mice (Bates et al. 2004). But all studies were done using bacterial cells and no direct evidence of a complex between UMOD and a pilus protein was shown so far. Furthermore, it is described that UMOD is linked to the water/electrolyte balance and to kidney innate immunity, as well as protect against renal stones. Therefore, it plays a key role in chronic kidney diseases and is a promising therapeutic target for hypertension (Kottgen et al. 2010, Rampoldi et al. 2011, Trudu et al. 2013).

In kidney cells UMOD is synthesized as a 640 amino-acid precursor and consists of 563 amino acids in its mature form. It has a high number of cysteine residues (48, 7 % of amino acid content), all forming intramolecular disulfide bonds and a high content of acidic residues (theoretical $p_I = 4.8$). During the maturation process of the protein it enters the secretory pathway where it is glycosylphosphatidylinositol
anchored (position 614), glycosylated and later transported to the apical plasma membrane of epithelial cells.

The domain composition of human UMOD includes a signal peptide of 24 amino acids for its entry in the secretory pathway, four epidermal growth factor (EGF)-like domains, followed by a domain of unknown function (named D8C as it contains eight conserved cysteines) and at the C-terminus a zona pellucida (ZP) domain. The ZP-domain by itself is composed of two domains. A common feature of proteins containing ZP-domains (for example in egg coats) is the formation of filaments or matrices, nicely reviewed in (Jovine et al. 2005). Several models for the polymerization mechanism have been proposed, mainly derived from crystal structures of ZP-domains (Han L. et al. 2010, Lin et al. 2011, Monne et al. 2008).

Recently, the crystal structure of the ZP domain of UMOD was solved (Bokhove et al. 2016) (Figure 7). Fused to maltose binding protein it crystalized as a homodimer, with the ZP-N domains interacting via a hydrophobic interface. Still, this conformation could not explain the assembly mechanism from the monomer to a filamentous structure.

**Figure 7. Crystal structure of homodimers of the ZP domain of human UMOD.** (Bokhove et al. 2016) (pdb-code: 4WRN). Homodimer with one ZP domain in cyan and the second ZP domain divided in N-terminal domain (ZP-N, red) and C-terminal domain (ZP-C, green). In the homodimer the ZP-N domains interact via a hydrophobic interface. ZP = zona pellucida.
1. INTRODUCTION

1.8 Goals of the thesis

Despite numerous structural and biochemical data, mainly on the isolated lectin domain, a full understanding of the function and mechanism of FimH remains elusive. For example, there is no information on the rate constants of carbohydrate binding and dissociation of full-length FimH. Furthermore, the question of whether ligand binding automatically leads to domain separation has not been investigated.

The first aim of this thesis was the construction of a soluble FimH molecule as a model of FimH in the tip fibrillum, representing the physiologically relevant target in UTI infections. Since no electron density was detected in the ligand binding pocket of FimH in the tip fibrillum structure, it is not known if this so called low-affinity conformation is able to bind carbohydrate ligands at all. The present model assumes that ligand binding is coupled to domain separation, meaning that full-length FimH adopts the same conformation with bound ligand as observed for the isolated lectin domain. The experimental support for this model was the observation that crystals of the tip fibrillum cracked upon addition of carbohydrate ligand. Therefore, one major aim of this thesis was to study conformational differences upon ligand binding in full-length FimH. Another aim was the comprehensive analysis of the ligand binding properties of FimH, mainly a comparison of the rate constants of ligand binding and dissociation of full-length FimH relative to isolated FimH\textsubscript{L}. A better understanding of the mechanism underlying force dependent bacterial adhesion is expected to provide a solid basis for better strategies for the development of anti-adhesive drugs in the future.

In the second part of the thesis, a mutational study on FimH was done with the idea to get deeper insights into the catch-bond mechanism of FimH. The rational was to introduce mutations in the domain interface which would disrupt the interaction between FimH\textsubscript{L} and FimH\textsubscript{P}. And finally, analyze if there is a general correlation between inter-domain interaction energy and the thermodynamics of ligand binding.

In the third part, the glycoprotein UMOD was investigated. The first goal was to show if UMOD is a natural ligand of FimH. And if so, determine the binding stoichiometry and affinity of the UMOD-FimH complex. Finally, another goal of the thesis was to get detailed insights in the architecture of UMOD filaments, thus they were studied with electron cryotomography.
This page was intentionally left blank.
2. RESULTS

2.1 Catch-bond mechanism of the bacterial adhesin FimH

Figures and text passages were used in their original form, with the approval of the journal, or adapted for this chapter and are published in:


**Contribution of the author:** the entire experimental work on FimH from *E.coli* strain K12 with the exception of x-ray data collection and structure determination and the analytical gel filtration runs presented in Figure 1B.
2. RESULTS

2.1.1 Introduction

"Cell-cell adhesion often occurs under dynamically varying conditions and mechanical stress. In many cell-cell adhesion systems, the lifetime of adhesin-receptor complexes is increased under tensile mechanical force via “catch-bonds”, which permit capture or retention of cells under flow conditions while still allowing for release under reduced mechanical force. Catch-bond interactions are prominent in vascular systems and are formed e.g. by selectins for leukocyte recruitment (Evans et al. 2004, Marshall et al. 2003), by cadherins controlling tissue integrity (Akiyoshi et al. 2010, Buckley et al. 2014) in the epithelial adhesion of cancer cells (Fiore et al. 2014), and by the interactions between T cell receptors (TCR) and peptide-bound major histocompatibility complexes (MHC) on antigen presenting cells (Das et al. 2015, Liu et al. 2014). Catch-bonds also play a major role in bacterial adhesion and infection by uropathogenic *Escherichia coli* (*E. coli*) strains, which are responsible for the vast majority of urinary tract infections (UTI) in humans (Ronald 2002). A first critical step in the establishment of infection is bacterial adhesion to urothelial cells under flow conditions, which is mediated by 0.1–2 µm long, proteinaceous filaments on the bacterial surface termed type 1 pili (Hahn et al. 2002, Jones et al. 1995). Type 1 pili are composed of up to 3'000 copies of the subunit FimA building the pilus rod, as well as the subunits FimF, FimG, and FimH forming the distal tip fibrillum (Waksman et al. 2009). The adhesin FimH at the fimbrial tip specifically binds in a catch-bond mode (Thomas W. E. et al. 2002) to terminal α-D-linked mannoses of N-linked glycans of the receptor uroplakin 1a (UPIa) on urinary epithelial cells (Zhou et al. 2001). Due to its important role in establishing infection, FimH is an attractive target for the development of anti-adhesive drugs for UTI treatment (Ofek et al. 2003, Sharon 2006).

FimH is a two-domain protein, composed of an N-terminal, mannoside-binding lectin domain (FimHL) and a C-terminal pilin domain (FimHP). FimHP possesses an incomplete immunoglobulin-like fold that is completed by insertion of an N-terminal donor strand of FimG, the subsequent subunit in pilus assembly (Waksman et al. 2009). The two-domain architecture of FimH is a prerequisite for catch-bond formation, because the interactions between FimHL and FimHP determine the conformational state and ligand binding properties of FimHL (Buckley et al. 2014, Evans et al. 2004, Guo et al. 2006, Le Trong et al. 2010, Liu et al. 2014, Marshall et
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

al. 2003). A “compressed” FimH₇ conformation was observed in the crystal structure of FimH in the context of the type 1 pilus fibrillum in the absence of ligands, with an open binding site and interactions to FimH₈ mediated via three loop segments, the swing (amino acids 27–33), linker (aa. 154–160), and insertion loops (aa. 112–118) (Le Trong et al. 2010). In contrast, an “extended” FimH₇ conformation was observed in crystal structures of the isolated, ligand-bound FimH₇ domain (Bouckaert et al. 2005, Brument et al. 2013, Han Z. et al. 2010, Vanwetswinkel et al. 2014, Wellens et al. 2008, Wellens et al. 2012) and in the complex between FimH and the pilus assembly chaperone FimC, where FimC prevents the interactions between FimH₇ and FimH₈ (Choudhury et al. 1999). This extended form of FimH₇ is characterized by a closed ligand binding pocket and rearranged swing, linker, and insertion loops.

Notably, isolated FimH₇ was reported to show a ligand binding affinity about two orders of magnitude higher than that of full-length FimH in the tip fibrillum (Le Trong et al. 2010, Schwartz et al. 2013). Together with mutagenesis experiments disrupting the inter-domain interface (Yakovenko et al. 2008), these data indicated that ligand binding is linked to domain separation in FimH, and that mechanical force shifts the ligand binding affinity towards that of the isolated FimH₇. However, fundamental aspects of the mechanism underlying the force-dependent binding of FimH remained unknown: (i) How is domain-associated, full-length FimH interacting with ligands? (ii) Does ligand binding directly induce domain separation? (iii) How are interdomain interactions linked to the ligand binding affinity of FimH and the kinetics of ligand binding and dissociation?

To address these questions, a stable, soluble variant of full-length FimH, that is equivalent in its structural and functional properties to those of FimH in the assembled fimbrial tip, is designed. This variant allowed to obtain high-resolution structural snapshots of all functional states of FimH and to obtain a complete characterization of ligand-binding kinetics in solution. These data reveal a three-state mechanism of FimH catch-bond formation. FimH₈ accelerates ligand release from FimH₇ via dynamic allostery by 100’000-fold. This shows that the modulation of ligand affinity by FimH₈ is not only required for adhesion under mechanical stress, but also for efficient bacterial surface motility in the absence of shear force. These results provide a first complete structural and kinetic description of a catch-bond system.” (Sauer M. M. et al. 2016).
2. RESULTS

2.1.2 Results

Construction of a soluble, peptide-complemented FimH

Isolated FimH with its non-complemented pilin domain is only marginally stable and shows aggregation tendency under physiological conditions (Vetsch et al. 2002). To establish a stable, isolated FimH molecule with all properties of FimH in the tip fibrillum, FimHP was complemented with the donor strand peptide of FimG (FimG residues 1–14; termed DsG). The FimH-DsG complex was obtained in good yields and purified after an in vitro reaction mimicking the first donor strand exchange (DSE) reaction during pilus assembly in vivo. In this reaction, the FimG donor strand displaces the pilus assembly chaperone FimC from FimH (Figure 1).

Figure 1. Preparation of the FimHK12–DsG complex by donor strand exchange (DSE). (A) Reaction scheme of the DSE reaction, in which DsG displaces the FimC chaperone from the FimH pilin domain. (B) Kinetics of FimHK12–DsG complex formation at 37 °C and pH 7.0, monitored by analytical gel filtration. DSE was initiated by mixing the FimC–FimH K12 complex (15 µM) with excess DsG peptide (50 µM). The reaction can be followed by the decrease in FimC–FimH K12 complex concentration and the simultaneous increase in the concentrations of FimH K12–DsG and free FimC (FimC and FimH K12–DsG co-elute as a single peak at ~ 12 mL). The chromatogram at the bottom shows that the FimC–FimH K12 complex is stable against dissociation/aggregation under the chosen conditions. The rate constant of donor strand exchange estimated from these data is about 0.5 M⁻¹s⁻¹.
2. RESULTS

The experiments described in the following were performed with FimH from the wild-type *E. coli* strain K12 (FimH\textsuperscript{T12}) by myself and from the fecal *E. coli* strain F18 (FimH\textsuperscript{F18}) (Schwartz et al. 2013) by Maximilian Sauer, which differ in three amino acids in FimH\textsubscript{L} (K12→F18: Val27Ala, Asn70Ser, Ser78Asn). The isolated lectin domain (residues 1-159) was produced by direct expression in the *E. coli* periplasm and purified as described (Vetsch et al. 2002).

**Ligand-free FimH·DsG resembles FimH in the fimbrial tip**

The crystal structure of the binary complex FimH·DsG was determined at atomic resolution by molecular replacement (Table 1). FimH·DsG comprises the jellyroll fold FimH\textsubscript{L} and the immunoglobulin-like FimH\textsubscript{P} domain complemented with the FimG donor strand (Figure 2A). It closely resembles unliganded FimH in the fimbrial tip complex (Figure 2B) (Le Trong et al. 2010), with a root-mean-square deviation of C\textalpha{} positions (C\textalpha{} rmsd) of 1.1 Å. The individual FimH\textsubscript{P} and FimH\textsubscript{L} domains are even more closely resembling unliganded, fimbrial FimH (rmsd 0.45 Å and 0.55 Å, respectively) and undergo only a minimal hinge-bending rotation of 4° (Figure 2B).

---

**Figure 2.** FimH·DsG resembles fimbrial tip FimH. (A) Structure of FimH\textsuperscript{F18}·DsG (lectin domain FimH\textsubscript{L}, red; pilin domain FimH\textsubscript{P}, yellow; DsG, blue; circle and square indicate N- and C-terminus, respectively). (B) FimH from the fimbrial tip structure (left, PDB ID: 3NJW (Le Trong et al. 2010); FimG, blue; FimF, green) is superposed onto FimH\textsuperscript{F18}·DsG based on their pilin domains (aa. 160-279), in the superposition (right) fimbrial FimH is shown in grey. (C) Close-up on the DsG peptide (stick representation) bound to FimH\textsuperscript{F18}·DsG with 2F\textsubscript{o} - F\textsubscript{c} electron density map. Backbone hydrogen bonds of the DsG peptide and β-strands 2 (β2) and 9 (β9) of FimH\textsubscript{P} are indicated.
2. RESULTS

The DsG peptide in FimH-DsG is in identical position as compared to the N-terminal FimG extension in the fimbrial tip structure; it interacts with β-strand 2 and 9 of FimH\(\text{\textsubscript{P}}\) (Figure 2C). All contacts in the FimH\(\text{\textsubscript{L}}\) - FimH\(\text{\textsubscript{P}}\) interdomain region (Le Trong et al. 2010) as well as the conformation of the empty ligand-binding pocket observed in FimH in the fimbrial tip are preserved in FimH-DsG. Thus, FimH-DsG represents the ligand-free state of fimbrial FimH, with Associated FimH\(\text{\textsubscript{L}}\) and FimH\(\text{\textsubscript{P}}\) (A\textsubscript{free} state) and is an elegant minimal system to analyze the crosstalk between ligand-binding and inter-domain interactions underlying formation of catch-bonds by FimH.

**Persistence of domain association in ligand-bound FimH-DsG**

To test whether ligand binding causes domain separation in FimH, the co-crystal structure of the ternary complex of FimH-DsG with \(n\)-heptyl α-D-mannoside (HM), an established model ligand of FimH (Bouckaert et al. 2005), as well as a crystal structure of the isolated lectin domain in complex with HM (Table 1) were determined. FimH-DsG-HM adopts the same, closed conformation of the ligand binding site as previously observed in other FimH\(\text{\textsubscript{L}}\)-ligand complexes (Figure 3A,B) (Gouin et al. 2014): The mannopyranose moiety of HM is coordinated by the side chains of Asp54, Gln133, Asn135, and Asp140, and the main chain of Phe1 and Asp47 and the \(n\)-heptyl aglycone of HM is sandwiched between Tyr48 and Tyr137. Compared to the A\textsubscript{free} form, all loops surrounding the binding pocket close down onto the HM ligand. The most substantial conformational difference to A\textsubscript{free} is observed for the clamp loop (aa. 8-16), whose tip moves almost 6 Å towards HM.

Besides the closing of the ligand-binding pocket, the overall conformation of ligand-free FimH\(\text{\textsubscript{L}}\) in A\textsubscript{free} and HM-bound FimH-DsG is closely similar (C\textsubscript{α} rmsd 1.1 Å) (Figure 3A,B). Unexpectedly, the structural change in the ligand-binding site in FimH-DsG-HM was not transmitted to the domain interface, where the inter-domain contacts and the conformations of the swing, linker, and insertion loops remained intact. The lectin domain in the FimH-DsG-HM complex thus differs drastically from HM-bound isolated FimH\(\text{\textsubscript{L}}\) domains with respect to the swing, linker, and insertion loop conformations (Figure 3C). Based on the persistence of the domain Association in the ligand-bound form, this state of FimH was termed A\textsubscript{bound}. 
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

Figure 3. Crystallographic analysis of FimH conformational states. (A) FimH<sup>F18</sup>·DsG in the A<sub>free</sub> (left) and in the A<sub>bound</sub> state (FimH<sup>F18</sup>·DsG·HM), in comparison to the S<sub>bound</sub> state of FimH<sup>K12</sup>·DsF·HM and the isolated FimH<sub>L</sub>F<sup>18</sup>·HM (right). The FimH<sub>L</sub>, FimH<sub>P</sub>, DsF, and DsG are colored in red, yellow, green, and blue. The experimentally in crystallo trapped orientation of FimH<sub>P</sub> in FimH<sup>K12</sup>·DsF·HM and a modeled position based on a hinge motion stretching around Gly157 is indicated. A schematic representation for each crystal structure, similar to Fig. 1a, is given. The tip of the clamp loop and the C-terminus of FimH<sub>L</sub> are indicated as a circle and diamond, respectively. (B) Comparison of the conformation of the ligand binding site in the A<sub>bound</sub> (red) and S<sub>bound</sub> (orange) states with the isolated lectin domain FimH<sub>L</sub> (grey) and (C) comparison of the interdomain interface of the lectin domain (all FimH<sup>F18</sup>).
Trapping of a domain-separated state of full-length FimH

The increase in apparent affinity of FimH to its target glycans under tensile mechanical forces (Thomas W. E. et al. 2004, Thomas W. E. et al. 2002) has previously been linked to a separation of the FimH_L and FimH_P (Le Trong et al. 2010). To trap a potential domain-separated state of FimH for structural characterization in the absence of tensile force variants with weakened inter-domain interactions were investigated. It was previously shown that FimH_P also accepts the donor strand of the non-cognate subunit FimF (DsF). However, FimH_P is slightly less stabilized by complementation with DsF than with the natural donor strand DsG (Puorger et al. 2008). Therefore, it was hypothesized that such complementation with DsF instead of DsG could also result in a mild destabilization of the interdomain interface in full-length FimH.

The co-crystal structure of FimH-DsF with HM (FimH-DsF-HM) at 3.0 Å resolution with four molecules in the asymmetric unit was determined. Three FimH-DsF-HM molecules closely resembled the A_bound state (rmsd of 0.6 Å to FimH-DsG-HM) with a preserved interdomain interface. In the fourth molecule, however, the FimH_L and FimH_P domains were separated and adopted a drastically different relative orientation with an angle between the domains of about 45° instead of ~150° in the other three molecules (Figure 3A). FimH_P is virtually identical in all four FimH molecules in the crystal (rmsd 0.4 Å). In contrast, the FimH_L domain differs significantly between the fourth, domain-separated and the three full-length FimH molecules in the crystal. It shows closest similarity to the isolated FimH_L-HM (rmsd 0.45 Å), in particular all interdomain loops adopt identical conformations, which are incompatible with domain association (Figure 3C and 4). Remarkably, in the bent fourth molecule, no interactions between FimH_L and FimH_P other than the direct covalent linkage are detected, equivalent to a breakdown of the total 500 Å² interdomain interface of the A_bound state (Figure 4). This molecule thus represents a third state, the domain-Separated, ligand-bound state of FimH, S_bound. The complete absence of non-covalent inter-domain interactions indicates that the S_bound state does not possess a defined relative domain orientation in solution and that the observed, kinked conformation has been selected only by crystal packing.
Figure 4. The interdomain region in the $S_{\text{bound}}$ state. Close-up of the interdomain region of FimH$^{K12}$-DsF-HM in the $A_{\text{bound}}$ form (left) and FimH$^{K12}$-DsF-HM in the $S_{\text{bound}}$ state (right). A cartoon representation for each crystal structure, similar to Fig. 3A, is given. Key residues in the interface are shown as sticks. FimH$_P$, FimH$_L$, and DsF are colored in red, yellow, and green, respectively.
### 2. RESULTS

**Catch-bond mechanism of the bacterial adhesin FimH**

Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>FimH$^{K12}$-DsG</th>
<th>FimH$^{K12}$-DsG-HM</th>
<th>FimH$^{K12}$ DsF-HM</th>
<th>FimH$^{K12}$-HM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>C 1 2 1</td>
<td>P 1</td>
<td>I 21 21 21</td>
<td>P 21 21 21</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>99.5, 35.6, 72.9</td>
<td>56.5, 77.6, 78.1</td>
<td>94.5, 147.1, 250.8</td>
<td>63.0, 68.4, 95.9</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 105, 90</td>
<td>101.5, 111.1, 96.3</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Resolution (Å) *</td>
<td>48.1-1.14 (1.2-1.14)</td>
<td>52-2.54 (2.63-2.54)</td>
<td>48.1-3.0 (3.19-3.0)</td>
<td>52-1.7 (1.76-1.7)</td>
</tr>
<tr>
<td><strong>R_{merge}</strong></td>
<td>3.8 (69.1)</td>
<td>16.1 (85.9)</td>
<td>42.7 (174.1)</td>
<td>12.9 (127.8)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>100.0 (84.2)</td>
<td>99.2 (82.2)</td>
<td>95.8 (46.7)</td>
<td>99.9 (57.7)</td>
</tr>
<tr>
<td>I/σI</td>
<td>17.8 (2.8)</td>
<td>8.79 (2.0)</td>
<td>14.4 (1.2)</td>
<td>15.1 (1.7)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.7.0 (81.5)</td>
<td>90.7 (89.4)</td>
<td>99.9 (98.7)</td>
<td>96.9 (79.3)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.2 (2.5)</td>
<td>3.4 (3.5)</td>
<td>6.6 (4.1)</td>
<td>11.5 (6.3)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>47.9-1.14</td>
<td>52.6-2.54</td>
<td>48.13-3.0</td>
<td>52.6-1.7</td>
</tr>
<tr>
<td>No. reflections</td>
<td>277806 (85217)</td>
<td>122641 (11994)</td>
<td>172553 (35275)</td>
<td>518524 (23173)</td>
</tr>
<tr>
<td>R_{work}/R_{free} (%)</td>
<td>15.3/17.5</td>
<td>22.7/27.6</td>
<td>22.0/25.1</td>
<td>17.1/19.6</td>
</tr>
<tr>
<td>No. atoms</td>
<td>3086</td>
<td>8855</td>
<td>8751</td>
<td>3232</td>
</tr>
<tr>
<td>Protein</td>
<td>2590</td>
<td>8690</td>
<td>8652</td>
<td>2407</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>0</td>
<td>0</td>
<td>86</td>
<td>38</td>
</tr>
<tr>
<td>Water</td>
<td>496</td>
<td>165</td>
<td>13</td>
<td>471</td>
</tr>
<tr>
<td>B-factors (Å$^2$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>13.3</td>
<td>43.7</td>
<td>53.2</td>
<td>21.3</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>0</td>
<td>32.3</td>
<td>20.7</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>23.6</td>
<td>22.4</td>
<td>25.9</td>
<td>40.6</td>
</tr>
<tr>
<td>R.m.s deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.01</td>
<td>0.019</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.45</td>
<td>1.78</td>
<td>0.82</td>
<td>1.02</td>
</tr>
<tr>
<td><strong>PDB code</strong></td>
<td>4XO9</td>
<td>4XOA</td>
<td>4XOB</td>
<td>4XO8</td>
</tr>
</tbody>
</table>

*Highest resolution shell is shown in parenthesis.*
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

Domain association alters FimH ligand-binding kinetics

To analyze the ligand-binding properties of FimH·DsG, the increase in intrinsic tryptophan fluorescence in the FimH·DsG complexes of about 10 % upon HM binding was exploited (Figure 5A). This difference was used to measure the dissociation constant of HM binding by equilibrium titration (Figure 5B) and the rates of HM binding and dissociation by stopped-flow fluorescence kinetics (Figure 5C,D). The FimH·DsG construct showed uniform binding and dissociation kinetics, consistent with the view that domain-separated states of FimH are not significantly populated in the absence of shear force. The results revealed equilibrium dissociation constants ($K_d$) of 3.6 μM for FimH·DsG (Table 2). HM binding to FimH·DsG is extremely dynamic and was characterized by fast association rates ($k_{on}$) of $5.0 \cdot 10^6$ M$^{-1}$s$^{-1}$, and rapid dissociation reactions. The rates of HM dissociation ($k_{off}$) of 22 s$^{-1}$ for FimH·DsG, translate into dissociation half-lives of only 32 ms.
Figure 5. Kinetics of HM binding and release by full-length FimH. (A) Fluorescence spectra (excitation at 280 nm) of FimH\textsubscript{L} (2 µM) (red lines) and FimH-DsG (2 µM) (black lines) in the absence (solid lines) or presence of 200 µM n-heptyl α-D-mannoside (HM) (dashed lines). (B) Equilibrium titration of FimH\textsuperscript{K12}.DsG (2 µM) with HM, recorded via the fluorescence increase at 320 nm. The total concentration of HM is plotted against the recorded fluorescence signal. Data were fitted (solid line) according to equation (1) (cf. experimental section) and yielded a \( K_d \) value of 3.6 ± 0.3 µM. (C) Stopped-flow fluorescence kinetics of HM binding to FimH-DsG (2.0 µM), recorded via the fluorescence change above 320 nm. The HM concentration was varied between zero and 40 µM. Five representative traces are shown. The fluorescence traces were globally fitted according to a second-order binding and first-order dissociation reaction (solid lines) (Table 1). (D) Amplitudes of the reactions monitored in (C), plotted against the total HM concentration. Data were fitted (solid line) according to equation (1), yielding a \( K_d \) value of 4.2 ± 0.3 µM.
In contrast to full-length FimH, isolated FimH\textsubscript{L} showed no change in tryptophan fluorescence upon HM binding (Figure 5A). Therefore, the HM affinity of isolated FimH\textsubscript{L} was determined indirectly by a competition experiment based on a newly designed fluorescent ligand, the fluorescein-labeled α-D-mannoside GN-FP-4 (Figure 6A,B). Displacement of GN-FP-4 from FimH\textsubscript{L} by increasing HM concentrations under equilibrium conditions showed that FimH\textsubscript{L} binds HM with 3300-fold higher affinity compared to the FimH·DsG complex ($K_d$-value 1.1 nM) (Figure 6E and Table 3).

In an inverse competition experiment (Figure 6F), in which HM in preformed FimH\textsubscript{L}·HM complexes was displaced by GN-FP-4, off-rates of $2.0 \cdot 10^{-4}$ s\textsuperscript{-1} were determined for FimH\textsubscript{L}, corresponding to a dissociation half-live of 58 minutes. Based on these measured off-rates and equilibrium dissociation constant, a $k_{on}$ rate of $1.8 \cdot 10^5$ M\textsuperscript{-1}s\textsuperscript{-1} was calculated for FimH\textsubscript{L} (Table 2). The on-rate for the isolated FimH\textsubscript{L} domain is thus 30-fold lower than the one of the full-length FimH·DsG complex (Table 3).
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

**A**

**B**

\[ A_{\text{max}} = 497 \text{ nm} \]
\[ \varepsilon = 54.897 \text{ M}^{-1} \text{ cm}^{-1} \]

\[ 10 \mu M \text{ GN-FP-4} \]

**C**

**D**

\[ [\text{GN-FP-4}]_0 = 1 \text{ nM} \]
\[ [\text{GN-FP-4}]_0 = 2 \text{ nM} \]

**E**

**F**

<table>
<thead>
<tr>
<th>[GN-FP-4] ((\mu M))</th>
<th>(k_{\text{obs}} \text{ (s}^{-1}\text{)})</th>
<th>(t_{1/2} \text{ (min)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 (•)</td>
<td>2.0 ± 0.1 (10^4)</td>
<td>58 min</td>
</tr>
<tr>
<td>20 (x)</td>
<td>2.0 ± 0.3 (10^4)</td>
<td>58 min</td>
</tr>
<tr>
<td>30 (□)</td>
<td>2.1 ± 0.3 (10^4)</td>
<td>56 min</td>
</tr>
<tr>
<td>mean</td>
<td>2.0 ± 0.4 (10^4)</td>
<td>57 min</td>
</tr>
</tbody>
</table>
Figure 6. Synthesis and spectroscopic properties of the fluorescent ligand GN-FP-4 and its application to determine ligand affinity and ligand binding kinetics. (A) Structure and synthesis reaction of GN-FP-4. (B) Absorbance spectrum of 10 µM GN-FP-4. (C) Fluorescence spectra of GN-FP-4 (1 µM) before (black line) and after (red line) addition of one molar equivalent of FimH<sub>L</sub>. The results, together with the high affinity of the FimH<sub>L</sub>·GN-FP-4 complex (0.08 nM), show that the GN-FP-4 fluorescence decreases about 2-fold upon binding. (D) Equilibrium titration of GN-FP-4 (1.0 nM) with FimH<sub>L</sub> at pH 7.4 and 25 °C, recorded via the decrease in GN-FP-4 fluorescence at 520 nm upon binding to the FimH lectin domain (excitation at 497 nm). Titrations were performed at constant GN-FP-4 concentrations of 1 nM (black diamonds) or 2 nM (black circles) and globally fitted according to equation 1 (solid lines). Samples contained 0.001 % Tween 20 to prevent unspecific adsorption of GN-FP-4 to tubes or cuvettes. (E) HM binding to FimH<sub>L</sub> analyzed by displacement of GN-FP-4 from FimH<sub>L</sub>. An equimolar mixture of FimH<sub>L</sub> was preincubated with GN-FP-4 (1 µM each) and different HM concentrations (10 nM – 3.2 mM) for >18 h. GN-FP-4 displacement is monitored by a decrease in fluorescence polarization at 528 ± 20 nm (excitation at 485 nm). Data were fitted (solid lines) according to a mechanism in which two ligands compete for the same binding site, with fixed $K_d$ values for GN-FP-4 binding. (F) Determination of the rate constant of dissociation of HM from FimH<sub>L</sub> by displacement of HM with GN-FP-4 (pH 7.4, 25 °C). Shown are superimposed, normalized fluorescence kinetics. In all experiments, the 1:1 complex between FimH<sub>L</sub> and HM (3 µM each, guaranteeing >95 % occupancy with HM) was mixed with excess GN-FP-4 (10, 20 and 30 µM), and the decrease in GN-FP-4 fluorescence at 520 nm as a consequence of HM dissociation and GN-FP-4 binding was recorded. The obtained first-order kinetics are independent of the GN-FP-4 concentration and thus directly monitor HM dissociation. The decrease in GN-FP-4 fluorescence thus coincided with the dissociation of the FimH<sub>L</sub>·HM complex. (F, inlay) Rate constants and half-lifes of HM association at different concentrations of excess GN-FP-4.
2. RESULTS

Together, these results demonstrate that the 3’300-fold higher affinity of the isolated FimH<sub>L</sub> compared to full-length FimH results from a more than 100’000-fold lower ligand dissociation rate in isolated FimH<sub>L</sub>, combined with a ligand binding rate reduced by only 30-fold (Table 3). The 3300-fold higher affinity for HM of FimH<sub>L</sub> relative to FimH·DsG translates into a free energy of 20 kJ mol<sup>-1</sup> for the interaction between FimH<sub>L</sub> and FimH<sub>P</sub> in full-length FimH. This corresponds very well with the mechanical work required for domain separation, as a displacement of FimH<sub>L</sub> from FimH<sub>P</sub> by 11 Å for complete domain separation (Le Trong et al. 2010) and a force of 40 pN required to populate the domain-separated state of FimH (Thomas W. et al. 2006) yields a value of 26.5 kJ mol<sup>-1</sup>.

**Table 2.** Kinetics and thermodynamics of n-heptyl α-D-mannoside (HM) binding to FimH<sub>L</sub> or FimH·DsG.

<table>
<thead>
<tr>
<th>Protein</th>
<th>k&lt;sub&gt;on&lt;/sub&gt; (M&lt;sup&gt;-1&lt;/sup&gt;s&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>k&lt;sub&gt;off&lt;/sub&gt; (s&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>k&lt;sub&gt;off&lt;/sub&gt;/k&lt;sub&gt;on&lt;/sub&gt; (M)</th>
<th>K&lt;sub&gt;d&lt;/sub&gt; (amplitude analysis) (M)</th>
<th>K&lt;sub&gt;d&lt;/sub&gt; (equilibrium titration) (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FimH&lt;sub&gt;L&lt;/sub&gt;</td>
<td>1.8 ± 0.6 · 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>2.0 ± 0.4 · 10&lt;sup&gt;-4&lt;/sup&gt;</td>
<td>n.a.</td>
<td>n.a.</td>
<td>1.1 ± 0.1 · 10&lt;sup&gt;-9&lt;/sup&gt;</td>
</tr>
<tr>
<td>FimH·DsG</td>
<td>5.0 ± 0.1 · 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>2.2 ± 0.1 · 10&lt;sup&gt;1&lt;/sup&gt;</td>
<td>4.3 ± 0.1 · 10&lt;sup&gt;-6&lt;/sup&gt;</td>
<td>4.2 ± 0.3 · 10&lt;sup&gt;-6&lt;/sup&gt;</td>
<td>3.6 ± 0.3 · 10&lt;sup&gt;-6&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

The K<sub>d</sub> values for the complex formation between HM and FimH were determined (a) from the ratio of rate constants (k<sub>off</sub>/k<sub>on</sub>), (b) from the analysis of the amplitudes and (c) from equilibrium titration experiments. (d) Values of K<sub>d</sub> were deduced from competition equilibria experiments with a fluorescently labeled mannoside (FimH<sub>L</sub> K<sub>d</sub> = 7.0 ± 0.1·10<sup>-11</sup> M). (e) k<sub>on</sub> was calculated from k<sub>off</sub>/K<sub>d</sub>. Experiments were performed at pH 7.4 and 25°C. n. a.: not applicable.

**Table 3.** Comparison of n-heptyl α-D-mannoside (HM) binding of FimH<sub>L</sub> vs. FimH·DsG.

<table>
<thead>
<tr>
<th>k&lt;sub&gt;on&lt;/sub&gt; (FimH·DsG) / k&lt;sub&gt;on&lt;/sub&gt; (FimH&lt;sub&gt;L&lt;/sub&gt;)</th>
<th>k&lt;sub&gt;off&lt;/sub&gt; (FimH·DsG) / k&lt;sub&gt;off&lt;/sub&gt; (FimH&lt;sub&gt;L&lt;/sub&gt;)</th>
<th>K&lt;sub&gt;d&lt;/sub&gt; (FimH·DsG) / K&lt;sub&gt;d&lt;/sub&gt; (FimH&lt;sub&gt;L&lt;/sub&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>110’000</td>
<td>3300</td>
</tr>
</tbody>
</table>
This page was intentionally left blank.
2. RESULTS  Catch-bond mechanism of the bacterial adhesin FimH

2.1.3 Additional results

2.1.3.1 Competition assay between FimH-DsG and FimH₇ for HM

The affinity of HM to the isolated lectin domain (FimH₇) was determined in a competition assay with the ligand GN-FP-4 (Figure 6E). To prove the higher affinity of FimH₇ to HM without using another ligand the following titration experiment was performed. Here, the property that FimH₇ doesn’t show an intrinsic fluorescence increase upon HM binding was used (Figure 5A). The fluorescence change is due to the 10 % signal increase of FimH-DsG upon HM binding but at the same time HM is still consumed by FimH₇. A set of three experiments were prepared. The concentrations of the full-length protein FimH-DsG and FimH₇ were kept constant (2 µM and 2, 5 and 10 µM, respectively) and HM was titrated in from zero to 200 µM. The samples were equilibrated overnight and measured at 25 °C in 20 mM MOPS-NaOH pH 7.4 on a QM 7/2003 spectrofluorometer (PTI, USA) by the increase in fluorescence at 320 nm upon HM binding to FimH-DsG (excitation at 280 nm). The fluorescence intensities were recorded for 30 s and averaged. These data were plotted against total HM concentration and fitted with Dynafit (Kuzmic 1996) according to an equilibrium competition mechanism, with the total concentrations of FimH-DsG, FimH₇ and HM (variable) and the respective Kd of FimH-DsG to HM (3.6 ± 0.3 x 10⁻⁶ M) as input, and Kd of FimH₇ to HM and the fluorescence intensities at zero and infinite HM concentration as open parameters (Figure 7). The result of the global fit for the affinity between FimH₇ and HM was \( K_d = 2.4 \pm 0.2 \times 10^{-9} \) M, qualitatively confirming the much higher affinity of FimH₇ for HM compared to FimH-DsG. The about two-fold lower affinity of FimH₇ for HM obtained from the fit compared to the more accurate experiment shown in Figure 6 D,E is a result of the large error of equilibrium titration experiments at protein concentrations several orders of magnitude above the \( K_d \).
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

Figure 7. Equilibrium fluorescence titration of mixtures of FimH·DsG and FimHₖ at different ratios with HM. A mixture of FimHₖ and FimH·DsG was incubated overnight with different HM concentrations (0 - 200 µM). Since FimHₖ is spectroscopically silent upon HM binding only HM binding to FimH·DsG is monitored after saturation of FimHₖ by HM. The fluorescence intensities at 320 nm were measured (excitation at 280 nm). Data were fitted (solid lines) according to a mechanism in which two proteins compete for the same ligand, with a fixed $K_d$ value ($3.6 \times 10^{-6}$ M) of the FimH·DsG-HM complex.
2. RESULTS

2.1.3.2 Carbohydrate specificity of FimH

The fact that the intramolecular interactions between the lectin and pilin domain in FimH-DsG lowered the affinity of FimH-DsG for HM by three orders of magnitude relative to FimH_L predicted that the same difference in affinity would be observed for other α-linked mannoside ligands. To test this, and to investigate the specificity of FimH for terminal mannoses, the affinity of the following ligands (Figure 8) to FimH-DsG and FimH_L was measured. For this purpose, we made use of the strong increase in fluorescence polarization of the fluorescein group of GN-FP-4 upon binding to FimH, and performed equilibrium competition experiments in which GN-FP-4 was displaced from FimH-DsG or FimH_L by increasing concentrations of the respective carbohydrate ligand.

Figure 8. Molecular structures of investigated ligands.

Experiments were done at 25 °C in 20 mM MOPS-NaOH pH 7.4. A mixture of FimH_L or FimH-DsG and GN-FP-4 (1 µM each) was incubated with different concentrations of the ligands and incubated overnight. The displacement of GN-FP-4 by the ligands was recorded by the decrease in the fluorescence polarization at 528 ± 20 nm (excitation at 485 ± 20 nm) on a microplate reader (Biotek, USA), using flat black bottom 96 well microtiter plates (Greiner, Austria). The fluorescence polarization data were fitted with Dynafit (Kuzmic 1996) according to an equilibrium competition
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

mechanism, with the total concentrations of FimHₐ or FimH-DsG, GN-FP-4 and the ligands (variable) and the respective $K_d$ of GN-FP-4 as input, and the $K_d$ of the different ligands and the fluorescence polarization at zero and infinite ligand concentrations as open parameters (Figure 9).

Figure 9. Comparison of FimH-DsG and FimHₐ binding to different carbohydrate ligands. Ligand binding to FimH-DsG (A) FimHₐ (B) was analyzed by displacement of GN-FP-4 from the proteins. An equimolar mixture of the proteins was preincubated with GN-FP-4 (1 µM) and different ligand concentrations were added. GN-FP-4 displacement was monitored by a decrease in fluorescence polarization at 528 ± 20 nm (excitation at 485 nm). Data were fitted (solid lines) according to a mechanism in which two ligands compete for the same binding site, with fixed $K_d$ values for GN-FP-4 binding. (C) Comparison of ligand binding of FimHₐ vs. FimH-DsG. (D) Equilibrium titration of GN-FP-4 (10 nM) with FimH-DsG at pH 7.4 and 25 °C, recorded via the decrease in GN-FP-4 fluorescence at 528 ± 20 nm upon binding to FimH-DsG (excitation at 485 ± 20 nm). Samples contained 0.001 % Tween 20 to prevent unspecific adsorption of GN-FP-4 to tubes or cuvettes.
The experiments show that FimH has a high specificity for α-linked D-mannosides as no binding could be detected for glucose, galactose, maltose, sucrose and lactose. The result suggests that FimH indeed evolved towards specific recognition of terminal mannoses in high-mannose-type surface glycoproteins.

Furthermore, the result shows that the three orders of magnitude higher affinity to ligands of FimH relative to FimH-DsG is conserved and independent of the ligand. The results are thus consistent with a thermodynamic cycle in which the A and S-states of FimH differ in ligand affinity by a factor of 1500-3000.

![Thermodynamic cycle](image)

**Figure 10.** Thermodynamic cycle of the coupling between domain separation and ligand (HM) binding in the absence of tensile mechanical force. As no experimental evidence for a significant population of $S_{\text{bound}}$ could be obtained, the lower limit for the energy required for domain separation in the ligand-bound state was set to at least 10 kJ/mol. This translates into at least 30 kJ/mol for domain separation in the ligand free-state of FimH, corresponding to an $A_{\text{free}}/S_{\text{free}}$ ratio of > 200'000/1. $A = \text{associated state}, S = \text{separated state}, K_d = \text{dissociation constant}$, energy-values are given for the solid dashes.

**Table 4.** Summary and comparison of $K_d$ values for different ligands.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>$K_d$ (FimH$_L$)</th>
<th>$K_d$ (FimH-DsG)</th>
<th>$K_{\text{A-S}}$ $K_d$ (FimH-DsG) / $K_d$ (FimH$_L$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D(+)-Mannose (Man)</td>
<td>1.6 ± 0.1 x 10$^{-7}$ M</td>
<td>4.0 ± 0.1 x 10$^{-4}$ M</td>
<td>2.5 ± 0.2 x 10$^3$</td>
</tr>
<tr>
<td>1,5-anhydro D-mannitol (1,5-AM)</td>
<td>1.8 ± 0.3 x 10$^{-7}$ M</td>
<td>2.7 ± 0.1 x 10$^{-4}$ M</td>
<td>1.5 ± 0.3 x 10$^3$</td>
</tr>
<tr>
<td>n-methyl α D-mannopyranoside (MM)</td>
<td>1.9 ± 0.1 x 10$^{-7}$ M</td>
<td>3.7 ± 0.1 x 10$^{-4}$ M</td>
<td>1.9 ± 0.1 x 10$^3$</td>
</tr>
<tr>
<td>n-heptyl α D-mannopyranoside (HM)</td>
<td>1.1 ± 0.1 x 10$^{-9}$ M</td>
<td>3.4 ± 0.1 x 10$^{-6}$ M</td>
<td>3.1 ± 0.3 x 10$^3$</td>
</tr>
<tr>
<td>α1-3, α1-6 Mannotriose (OM-3)</td>
<td>6.3 ± 0.3 x 10$^{-8}$ M</td>
<td>9.2 ± 0.3 x 10$^{-5}$ M</td>
<td>1.5 ± 0.1 x 10$^3$</td>
</tr>
<tr>
<td>GN-FP-4</td>
<td>7.0 ± 0.1 x 10$^{-11}$ M</td>
<td>1.3 ± 0.1 x 10$^{-7}$ M</td>
<td>1.9 ± 0.1 x 10$^3$</td>
</tr>
</tbody>
</table>
2.1.4 Discussion

"The characterization of full-length FimH had so far been restricted to the analysis of the adhesive properties of piliated E. coli cells and binding studies with the purified type 1 pilus tip fibrillum. With the FimH-DsG complex, a model system for quantitative studies of the interaction of FimH with carbohydrate ligands was established: Soluble FimH-DsG efficiently mimics FimH in the context of the assembled tip fibrillum, is readily available in milligram quantities and permits the determination of ligand binding and release kinetics in solution. Using FimH-DsG, high-resolution snapshots of three functionally relevant states of FimH (Figure 11) were obtained: In the absence of ligands, FimH adopts the $A_{\text{free}}$ state with associated FimH$_L$ and FimH$_P$ and an open conformation of the ligand-binding site, which is responsible for the 30-fold faster ligand-binding of full-length FimH as compared to the isolated FimH$_L$ domain. Ligand binding in the absence of shear force induces the $A_{\text{bound}}$ state with a closed binding site. In contrast to earlier hypotheses (Le Trong et al. 2010), the transition from $A_{\text{free}}$ to $A_{\text{bound}}$ is restricted to the ligand-binding site, while all interactions between FimH$_L$ and FimH$_P$ observed in the $A_{\text{free}}$ state remain preserved in $A_{\text{bound}}$. The $A_{\text{free}}$$\rightarrow$$A_{\text{bound}}$ transition most likely follows an induced fit mechanism, in which the formation of an encounter complex between FimH-DsG and HM is rate limiting and followed by a fast, unimolecular rearrangement to the $A_{\text{bound}}$ state, in agreement with the observation that binding of the model ligand HM is rate-limiting for the formation of $A_{\text{bound}}$ even at the highest HM concentrations used. Stopped-flow binding kinetics indicate that the lifetime of the proposed encounter complex prior to $A_{\text{bound}}$ formation is below 1 ms (Figure 5C). Under tensile mechanical force applied to the FimH-ligand complex, mimicked here by the destabilized variant FimH-DsF and crystal packing forces, the domain-separated state of FimH, $S_{\text{bound}}$, is formed. In the $S_{\text{bound}}$ state, FimH$_L$ and FimH$_P$ no longer interact specifically and are only connected via the linker segment comprising FimH residues 154-160. In this $S_{\text{bound}}$ state, FimH$_L$ adopts a conformation closely resembling isolated FimH$_L$ with bound ligand. Notably, ligand dissociation from FimH-DsG is 100’000-fold faster than from the isolated FimH$_L$ domain. This is striking, because the respective crystal structures revealed indistinguishable ligand interactions and binding site conformations in the FimH-DsG-HM and FimH$_L$-HM complexes (Figure 3B). The allosteric communication
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

from the FimH<sub>p</sub>-FimH<sub>L</sub> interface to the ligand-binding site reaches over 40 Å and is mediated via changes in protein dynamics rather than in static structure, in line with a general model of dynamic allostery (Cooper et al. 1984, Tsai et al. 2008). These data demonstrate that the interdomain-interactions in FimH (i) maintain the open conformation of the binding pocket, guarantee rapid ligand binding, and (ii) intramolecularly catalyze ligand dissociation by more than 100'000-fold. Rapid ligand binding and short lifetimes of the FimH ligand complex allow for rapid dissociation of individual pili from their ligands in the absence of shear force.

Different mechanistic models, like the two-pathway (Pereverzev et al. 2005), the deformation (Pereverzev et al. 2006), and the sliding re-binding model (Lou et al. 2007) have been developed to describe catch-bond interactions, often based on powerful single-molecule atomic force measurements. These models included the principle of allosteric control of ligand binding affinity (Thomas W. et al. 2006, Yakovenko et al. 2008), which was clearly fully confirmed in our present study. However, these conceptual models did not reveal the underlying atomic scale mechanisms in different catch-bond systems. For most catch-bond systems, including the cadherin-catenin binding to actin filaments (Buckley et al. 2014, Manibog et al. 2014), integrin epithelial cell adhesion (Choi et al. 2014, Kong et al. 2009) and TCR-MHC interactions (Das et al. 2015, Depoil et al. 2014, Liu et al. 2014), structural information is, if at all, available only for one state or from computer simulations. One exception are the selectins, which employ catch-bond binding for leukocyte recruitment. Selectins are multi-domain cell surface receptors, which consist of a lectin domain for complex carbohydrate binding, linked via an EGF-like domain to a variable number of short-consensus repeat domains and a transmembrane anchoring helix. Selectins exist in two conformations, a bent and an extended one, which differ in the angle between their lectin and EGF-like domain. Ligand-binding and conformational changes in the ligand-binding site are directly linked via a complex allosteric coupling mechanism to the adoption of the extended conformation (Preston et al. 2015, Somers et al. 2000). Tensile mechanical force under flow conditions acts along the axis of the ligand-binding site and the Lec-EGF-interface resulting in a stabilization of the extended conformation and thus increased ligand complex lifetimes (Evans et al. 2004, Preston et al. 2015). Also in FimH, catch-bond behavior is mediated by the interplay of a lectin and an anchoring domain that does not interact with ligand. Ligand binding by FimH in the absence of
shear force results in a closing of the ligand-binding site, but, in contrast to selectins, is not directly linked to altered interdomain interactions. Here, mechanical force promotes domain-separation and completely releases FimH\(_L\) from FimH\(_P\), which acts as an activator of ligand release via dynamic allostery. Remarkably, selectins and the fimbrial adhesin FimH thus employ entirely different mechanisms for establishing catch-bond behavior by crosstalk between a lectin and an anchoring domain that provides tethering to a shaft. In both systems, the selectins and fimbrial adhesion, the shaft structures linking the terminal lectin/coupling domains to the cell surface, may contribute to the overall catch-bond behavior, either via directly influencing coupling domain behavior or via their general elastic properties (Whitfield et al. 2014, Zakrisson et al. 2013).

Rapid ligand release from the high-mannose type glycoprotein receptor uroplakin 1a in the lower urinary tract (Xie et al. 2006) or flagellar motility of piliated bacteria and hence their ability to colonize new tissue areas under certain conditions during infection (Anderson et al. 2007, Thomas W. E. et al. 2004, Thomas W. E. et al. 2002) are important for bacterial survival. This provides a plausible explanation for the fact that low-affinity FimH variants were preserved in numerous uropathogenic E. coli strains. Binding of terminal mannoses with low affinity in the absence of shear force may also play a role in facilitating clearance of uropathogenic E. coli from the urinary tract by competitive binding to the Tamm-Horsfall protein in the urine (Pak et al. 2001). In turn, populating the \(S_{\text{bound}}\) state with extremely low dissociation rate ensures tight bacterial adhesion under the mechanical forces of urine excretion. FimH is a promising target for anti-adhesive therapy of UTI, because FimH antagonists, in contrast to antibiotics, are not exerting selection pressure towards resistance formation (Cusumano C. K. et al. 2011, Han Z. et al. 2010, Klein et al. 2010). Previous ligand binding studies on the isolated FimH\(_L\) domain mimic the domain-separated \(S_{\text{bound}}\) state of FimH. This state is characterized by extremely low off-rates and promoted \textit{in vivo} only after ligand binding and the onset of flow conditions. Our kinetic data on ligand dissociation from full-length FimH demonstrate that rapid, competitive displacement of FimH from its carbohydrate ligands by FimH antagonists is well possible. Thus, full-length FimH (e.g. in the form of the FimH-DsG complex established in this study) instead of the isolated FimH\(_L\) domain is the relevant target for the development of anti-adhesive drugs. In combination with the
novel fluorescent GN-FP-4 ligand, this model system paves the way for efficient screening for anti-adhesive drug candidates."


Figure 11. Catch-bond mechanism of FimH-mediated cell adhesion. (A) In the absence of tensile mechanical force, formation of the FimH-Uroplakin 1a (UPIa) complex comprises the highly dynamic transition of the \( \text{A}_{\text{free}} \) to the \( \text{A}_{\text{bound}} \) state. The reaction likely proceeds via a transient encounter complex (indicated in square brackets). The reaction of the encounter complex to \( \text{A}_{\text{bound}} \) is not rate limiting and must have a half-life of less than 1 ms. Dissociation of the receptor from the FimH lectin domain in the \( \text{A}_{\text{bound}} \) state is promoted via dynamic allostery by the pilin domain, that acts as a negative allosteric regulator. The reaction from \( \text{A}_{\text{bound}} \) to the encounter complex corresponds to \( k_{\text{off}} \). Fast binding and release of UPIa by FimH enables bacterial motility on the bladder epithelium. (B) Shear force increases the population of the \( \text{S}_{\text{bound}} \) state of FimH, in which the pilin and lectin domain are separated. The dissociation of \( \text{S}_{\text{bound}} \) under shear force is slowed down 100'000-fold compared to \( \text{A}_{\text{bound}} \). The indicated rate constants and half-lifes correspond to the interaction between FimH and the model ligand HM. Rate limiting reactions are indicated by solid arrows, and fast, non-limiting reactions by dotted arrows.
2. RESULTS

2.1.5 Materials and Methods

Materials

The synthetic DsG (sequence: ADVTITVNGKVVKAR) and DsF peptide (sequence: ADSTITIRGYVRDNG) (>95 % purity) was purchased from JPT (Germany). Guanidinium chloride ("AA-Grade" for spectroscopy) was obtained from NIGU Chemie (Germany). Chemicals of highest available purity were purchased from Sigma, Merck, or AppliChem. If not mentioned otherwise, chromatography media for protein purification were purchased from GE Healthcare (U.K.). Oligonucleotides were from Microsynth (Switzerland).

Expression plasmids

Expression plasmids for the periplasmic production of the *E. coli* FimH lectin domain (FimH<sub>L</sub>) and for the periplasmic co-expression of full-length FimH with FimC were used as previously described. Protein production is under control of the trc-promotor/lac operator.

Protein production and purification

For purification of the complexes FimC∙FimH, *E. coli* HM125 harboring the corresponding co-expression plasmid were grown at 30 °C in 2YT medium containing ampicillin (100 µg/mL). At an OD<sub>600</sub> of 1.5, isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 1 mM. The cells were further grown for 12–18 hours, harvested by centrifugation, suspended in cold 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 1 mg/mL polymyxin B sulfate (18 mL per liter of culture) and stirred at 4 °C for 1.5 h. After centrifugation, the supernatant (periplasmic extract) was dialyzed against 20 mM Tris-HCl pH 8.0 and applied to a QA52 (Whatman, Maidstone, U.K.) column equilibrated with the same buffer. The flow-through containing the respective FimC∙FimH complex was dialyzed against 20 mM MOPS-NaOH pH 7.0, loaded onto a Resource S column equilibrated with the same buffer, and the complexes were eluted with a linear NaCl gradient (0–400 mM). Fractions containing FimC∙FimH were pooled and loaded onto a Superdex 75 (HiLoad 26/60) column equilibrated with 20 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH pH 7.4, 50 mM NaCl. Fractions containing the pure complex were pooled and stored at
4 °C until further use. Typically, 3-5 mg of the purified complex were obtained per liter of bacterial culture.

For expression of the isolated *E.coli* FimH<sub>L</sub>, *E. coli* HM125 transformed with the respective expression plasmid were grown at 30 °C in M9 medium containing ampicillin (100 µg/mL) to an OD<sub>600</sub> of 1.0, and expression was induced with 1 µM IPTG. After further growth for 12 h, cells were subjected to periplasmic extraction (see above). The extracts were mixed with 0.11 volumes of 1 M acetic acid-NaOH pH 4.5, dialyzed against 10 mM acetic acid-NaOH pH 4.5 and then loaded onto a SP-Sepharose column equilibrated with the same buffer. The flow-through was collected and its pH was adjusted to 8.0 by addition of 1 M Tris-HCl pH 8.2. This solution was then applied to a Q-Sepharose column equilibrated with 20 mM Tris-HCl pH 8.0. The flow-through containing FimH<sub>L</sub> was loaded onto a Resource S column dialyzed against 20 mM formic acid-NaOH pH 4.0. The protein was eluted with a linear NaCl gradient (0-1 M). Fractions containing pure FimH<sub>L</sub> were pooled, dialyzed against water and stored at -20 °C. The identity of the purified proteins was confirmed by ESI mass spectrometry (FimH<sub>L</sub>: calculated: 16963.0 Da; measured: 16962.8 Da). About 11 mg of the pure FimH<sub>L</sub> was obtained per liter of bacterial culture.

**Production of FimH·DsG and FimH·DsF complexes**

The respective FimC·FimH complex (40 µM) was incubated with a 3-fold molar excess of the DsG peptide and incubated in 20 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH, pH 7.0, 50 mM NaCl for 48 h at 37 °C. The reaction mixture containing isolated FimC, the FimH·DsG complex and excess DsG was dialyzed against 20 mM acetic acid-NaOH pH 4.5 and loaded onto a Resource S (6 mL) column equilibrated with the same buffer. The FimH·DsG complex was eluted with a linear NaCl gradient (0-400 mM). Fractions containing the pure complex were pooled, dialyzed against water and stored at 4 °C. The FimH·DsG partially dissociated during ESI-mass spectrometry analysis, so that masses of the intact complexes and free FimH were obtained: FimH·DsG: calculated mass: 30635.3 Da; measured mass: 30636.0 Da. The overall yields of the purified FimH·DsG complexes relative to the initial amount of FimC·FimH were in the range of 50-55 %. The FimH·DsF complex was generated and purified as described for the FimH·DsG complexes (FimH·DsF: calculated mass:
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

30702.2 Da; measured mass: 30702.5 Da). The FimH-DsF complex was prepared from the FimC-FimH complex after mixing with DsF exactly according to the protocol described above for FimH-DsG and obtained in similar yields. Despite the non-natural interaction between FimH and DsF, the FimH-DsF complex was formed four times faster than the FimH-DsG complex at pH 7.0 and 37 °C, with a rate constant of $2.2 \pm 0.5 \text{ M}^{-1}\text{s}^{-1}$. The FimH-DsF complex was stable against dissociation and unspecific aggregation.

**Determination of protein concentrations**

Protein concentrations were measured via the specific absorbance at 280 nm, using the following extinction coefficients: FimC-FimH (59090 M$^{-1}$cm$^{-1}$), FimH$_L$ (24670 M$^{-1}$cm$^{-1}$), FimH-DsG (35090 M$^{-1}$cm$^{-1}$), FimH-DsF (36580 M$^{-1}$cm$^{-1}$). The concentrations of DsG and DsF were determined via their absorbance at 205 nm (42650 M$^{-1}$cm$^{-1}$ and 49700 M$^{-1}$cm$^{-1}$, respectively).

**Synthesis of the fluorescent-labeled FimH ligand GN-FP-4**

(The synthesis of GN-FP-4 was done by Giulio Navarra in the lab of Beat Ernst, Institute of Molecular Pharmacy, University of Basel)

To a stirred solution ofmannoside 1 (25 mg, 0.061 mmol) (Klein et al. 2010) in dry DMF (1 mL), N-hydroxysuccinimide (21 mg, 0.183 mmol) was added, followed by N,N'-dicyclohexylcarbodiimide (9.2 mg, 0.073 mmol). The mixture was stirred at room temperature (RT) for 2 h, then N-Boc-ethylendiamine (10.7 mg, 0.067 mmol) was added and the reaction was stirred for additional 10 h. After cooling to 0 °C, the reaction mixture was diluted with water and concentrated. Chromatography on silica gel (CH$_2$Cl$_2$/MeOH) yielded 23 mg (0.042 mmol, 68 %) of tert-butyl (3'-chboro-4'-(α-D-mannopyranosyloxy)-biphenyl-4-yl-carboxamido)ethyl)carbamate. This product was dissolved in CH$_2$Cl$_2$ (3 mL) and trifluoroacetic acid (TFA, 1 mL) was added. The solid dissolved during addition of TFA. After 10 min at RT the reaction was complete. The mixture was evaporated and excess TFA was removed in high vacuum. The intermediate N-(2-aminoethyl)-3'-chboro-4'-(α-D-mannopyranosyloxy)-biphenyl-4-carboxamide TFA salt (23 mg, 0.042 mmol, quant.) was used without purification in the next step. After dissolution in dry dimethylformamide (DMF) (0.5 mL), triethylamine (12.8 mg, 0.127 mmol) was added. The mixture was cooled to 0 °C, then fluorescein isocyanate (14.8 mg, 0.038 mmol) was added and the mixture was
stirred for 3 h in the dark. After the addition of water, DMF was removed azeotropically, the residue dissolved in MeOH/10 % acetic acid and evaporated. Chromatography on silica gel (CH$_2$Cl$_2$/MeOH) yielded compound 2 (Figure 6A), contaminated with triethylammonium acetate. Therefore, after dissolution in MeOH, 0.5 N HCl in MeOH was added, the mixture evaporated and chromatographed on silica gel, to yield pure compound GN-FP-4 (Kleeb et al. 2015) (15 mg, 47 %). 

\[
[\alpha]_D^{20} +12.1 \text{ (c 0.3, MeOH); } \text{^1H NMR (500 MHz, CD$_3$OD): } \delta = 8.12 \text{ (s, 1H), } 7.92 \text{ (d, } J = 8.3 \text{ Hz, 2H, Ar-H)}, 7.70 \text{ (dd, } J = 5.0, 13.1 \text{ Hz, 2H, Ar-H}), 7.64 \text{ (d, } J = 8.3 \text{ Hz, 2H, Ar-H}), 7.54 \text{ (dd, } J = 2.2, 8.6 \text{ Hz, 1H, Ar-H}), 7.46 \text{ (d, } J = 8.7 \text{ Hz, 1H, Ar-H}), 7.09 \text{ (d, } J = 8.2 \text{ Hz, 1H, Ar-H}), 6.74 \text{ (s, 2H), } 6.69 \text{ (d, } J = 1.4 \text{ Hz, 2H, Ar-H}), 6.55 \text{ (d, } J = 8.4 \text{ Hz, 2H, Ar-H}), 5.63 \text{ (d, } J = 1.3 \text{ Hz, H-1}), 4.15 \text{ (dd, } J = 1.8, 3.1 \text{ Hz, H-2}), 4.03 \text{ (dd, } J = 3.4, 9.5 \text{ Hz, H-3}), 3.94 \text{ (s, 2H, } CH_2), 3.86-3.64 \text{ (m, 6H, H-4, H-5, H-6, CH$_2$); } \text{^13C NMR (126 MHz, CD$_3$OD): } \delta = 153.21, 143.84, 136.41, 129.66, 129.18, 127.66, 127.76, 127.70, 125.37, 118.64, 103.62 \text{ (Ar-C), } 100.75 \text{ (C-1), } 76.00 \text{ (C-5), } 72.41 \text{ (C-3), } 71.86 \text{ (C-2), } 68.24 \text{ (C-4), } 62.69 \text{ (C-6), } 40.76 \text{ (CH$_2$); } \text{ESI-MS: } m/z: \text{ Calcd for C$_{42}$H$_{37}$ClN$_3$O$_{12}$S [M+H]$^+$: 842.2, found: 842.2.}

Fluorescence spectroscopy

Fluorescence emission spectra of FimH variants were recorded between 300 and 450 nm (excitation at 280 nm) at 25 °C in 1.0 x 0.4 cm quartz cuvettes on a QM 7/2003 spectrofluorimeter (PTI, USA) equipped with a magnetic stirrer. Protein concentrations were 1–2 µM in 20 mM MOPS-NaOH pH 7.4. Fluorescence spectra of GN-FP-4 (ε$_{495 \text{ nm}} = 54900 \text{ M}^{-1}\text{ cm}^{-1}$) were recorded between 500 and 650 nm (excitation at 497 nm) in the same buffer.

Kinetic parameters of the FimH-DsG to HM interaction

The rate constants of binding ($k_{on}$) and dissociation ($k_{off}$) for the complex between FimH-DsG and HM were measured at 25 °C in 20 mM MOPS-NaOH pH 7.4 in a SX20 stopped-flow instrument (Applied Photophysics, U.K.). A constant FimH-DsG concentration of 1 or 2 µM was used. FimH-DsG was mixed with different concentrations of HM (2-100 µM) and binding was monitored by the increase in fluorescence above 320 nm (excitation at 280 nm). The fluorescence traces were
globally fitted with *Dynafit* (Kuzmic 1996) according to a second order binding and first order dissociation reaction. As an additional control, the fluorescence amplitudes of the individual reactions were plotted against the total HM concentration and fitted according to equation (1). The deduced dissociation constants reproduced the *K*ₐ values obtained with equilibrium titration within experimental error.

**Equilibrium titration of FimH-DsG with HM**

The binding equilibrium between FimH-DsG and HM was followed at 25 °C in 20 mM MOPS-NaOH pH 7.4 on a QM 7/2003 spectrofluorometer (PTI, USA) by the increase in fluorescence at 320 nm upon HM binding (excitation at 280 nm). Measurements were performed with a stirred 1x0.4 cm quartz cuvette. The concentration of FimH-DsG was kept constant at 2 µM and the concentration of HM was varied between zero and 200 µM. The samples were equilibrated overnight, and their fluorescence intensities were recorded for 30 s and averaged. The fluorescence intensities were plotted against total HM concentration and fitted according to equation (1)

$$ F = \left( F_\infty - F_0 \right) \cdot \frac{[P]_0 + [L]_0 + K_D - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4 \cdot [P]_0 \cdot [L]_0}}{2 \cdot [P]_0} + F_0 \quad (1) $$

where *F* is the monitored fluorescence signal, *F₀* is the fluorescence signal in absence of ligand, *F_∞* is the fluorescence signal at full saturation with ligand, *K_D* is the dissociation constant, *[P]₀* is the total concentration of FimH-DsG and *[L]₀* is the total concentration of HM.

**Equilibrium titration of FimH₇ with GN-FP-4**

The binding equilibrium between FimH₇ and GN-FP-4 at 25 °C in 20 mM MOPS-NaOH pH 7.4, supplemented with 0.001 % Tween 20 to prevent unspecific adsorption effects at nanomolar concentrations, was recorded by the decrease in GN-FP-4 fluorescence at 520 nm (excitation at 497 nm). Measurements were performed with a stirred 1x0.4 cm quartz cuvette. The concentration of GN-FP-4 was kept constant at 1.0 or 2.0 nM and the concentration of FimH₇ was varied between zero and 10 nM. The samples were equilibrated overnight, and their fluorescence intensities at 520 nm were recorded for 30 s and averaged. The experimental data were fitted according to equation (1).
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

Displacement of HM from the FimH L by GN-FP-4

The rate constant of dissociation ($k_{\text{off}}$) for HM from FimH L at 25 °C in 20 mM MOPS-NaOH pH 7.4 was measured indirectly by binding of excess GN-FP-4 to FimH L after dissociation of HM, recorded with the decrease in GN-FP-4 fluorescence at 520 nm (excitation at 497 nm). A mixture of FimH L and HM (3 µM each), pre-incubated for at least 18 h, was mixed with different amounts of excess GN-FP-4 (final concentrations: 10−40 µM) and GN-FP-4 fluorescence was recorded every 10 min for 10 s, averaged and fitted according to first-order kinetics. The obtained rate constants were independent of GN-FP-4 concentration and thus identical to the dissociation rate of HM from FimH L.

Determination of the FimH L·HM dissociation constant

The affinity of FimH L for HM at 25 °C in 20 mM MOPS-NaOH pH 7.4 was determined by the competition between HM and GN-FP-4 for binding to FimH L. A mixture of FimH L and GN-FP-4 (1 µM each) was incubated with different concentrations of HM (10 nM - 3.2 mM) and incubated for at least 18 h. The displacement of GN-FP-4 by HM was recorded by the decrease in the fluorescence polarization at 528 ± 20 nm (excitation at 485 ± 20 nm) on a microplate reader (Biotek, USA), using flat black bottom 96 well microtiter plates (Greiner, Austria). The fluorescence polarization data were fitted with Dynafit (Kuzmic 1996) according to an equilibrium competition mechanism, with the total concentrations of FimH L, GN-FP-4 and HM (variable) and the respective $K_d$ of GN-FP-4 (Table 2) as input, and $K_d$ of HM and the fluorescence polarization at zero and infinite HM concentration as open parameters.

Crystallization of FimH variants

All crystallization experiments were performed at 4 °C with the sitting drop vapor diffusion method. For crystallization, FimH$^{F18}$-DsG and FimH$^{K12}$-DsG (0.1-0.2 µl, 15 mg/mL in H$_2$O) was mixed with 0.1−0.2 µl of precipitant (25 % (w/v) polyethylene glycol (PEG) 3350, 0.2 M magnesium chloride, 0.1 M BisTris-HCl pH 5.5 at 4°C. Crystals of FimH$^{F18}$-DsG and FimH$^{K12}$-DsG grew within 4–6 weeks and are of the space group C2, respectively, with one molecule per asymmetric unit. FimH$^{K12}$-DsG crystals in space group P1 grew at 0.2M NaMalonate, 20 % PEG3350 within 2 months at 4 °C. For crystallization of the FimH$^{F18}$-DsG·HM complex, a 3-fold excess over FimH-DsG was used (protein concentration and protein/precipitant ratios were
as described for FimH-DsG). Crystals of the space group P2₁\textsubscript{1}3 appeared after one month in 30\,\% (v/v) 2-Methyl-2,4-pentanediol (MPD), 0.1 M sodium cacodylate, 0.2 M magnesium acetate pH 6.5 at 4°C. FimH\textsuperscript{K12-}\textsuperscript{DsF-HM} crystals appeared after two months in 30\,\% w/v PEG 5'000, 0.1 M MES monohydrate, 0.2 M Ammonium sulfate pH 6.5 at 20 °C (2.5-fold excess of ligand over protein). FimH\textsubscript{L}\textsuperscript{F18-}\textsuperscript{HM} crystallized in 17\,\% PEG2000MME, 0.1 M Hepes-NaOH pH 7.5 at 4 °C. FimH\textsubscript{L}\textsuperscript{K12-}\textsuperscript{HM} crystallized in 1.5 M (NH₄)₂SO₄, 0.2 M NaAcetate pH 5.5 at 20 °C.

**Crystallographic data collection**

(This work was done by Roman Peter Jakob, Biozentrum, University of Basel)

All crystals, except for FimH\textsuperscript{F18-}\textsuperscript{DsG-HM}, were cryo-preserved by addition of ethane-1,2-diol (EG) to a final concentration of 20\,\% (v/v). The precipitant solution used for the crystallization of FimH\textsuperscript{F18-}\textsuperscript{DsG-HM} already contained 30\,\% (v/v) methyl-2,4-pentanediol (MPD) which acts as cryoprotectant. Crystals were flash cooled in liquid nitrogen. All measurements were done at the SLS beamline X06DA and X06SA (Swiss Light Source, Paul Scherrer Institute, Switzerland) at 100 K. All data were integrated, indexed and scaled using the XDS software package (Kabsch 2010) (5\,\% of the reflections were set aside as test set, respectively.) Data collection statistics are summarized in Table 1.

**Crystallographic structure determination**

(This work was done by Roman Peter Jakob, Biozentrum, University of Basel)

All structures were solved by molecular replacement using structures of isolated FimH\textsubscript{L} (AA1-158, PDB ID: 3MCY, (Han Z. et al. 2010) and the pilin domain of FimC-FimH (AA160-297, PDB ID: 1QUN, (Choudhury et al. 1999) as search models with the program Phaser (McCoy et al. 2007). Model building and structure refinement were performed with Coot (Emsley et al. 2004) and PHENIX (Adams et al. 2002). Twelve out of thirteen residues could be built for the FimG donor strands in the crystal structures and only the C-terminal lysine residue had weak electron density. Refinement statistics are summarized in Table 1.
2. RESULTS

Catch-bond mechanism of the bacterial adhesin FimH

Accession Numbers

The atomic coordinates and structure factors have been deposited in the Protein Data Bank under the accession codes:

4XO8 (FimH\textsubscript{L\textsuperscript{K12}}·HM)

4XO9 (FimH·DsG\textsubscript{K12})

4XOA (FimH·DsG\textsubscript{K12}·HM)

4XOB (FimH·DsF\textsubscript{K12}·HM)

4XOC (FimHl\textsubscript{F18}·HM)

4XOD (FimH·DsG\textsubscript{F18})

4XOE (FimH·DsG\textsubscript{F18}·HM)
2.2 Ligand binding is coupled to domain interaction in FimH

J. Eras, R. Glockshuber.

Contribution of the first author: the idea and the entire experimental work.
2. RESULTS  

2.2.1 Introduction

The catch-bond mechanism of FimH is mediated by the two-domain architecture of the protein. There are numerous studies in the literature where FimH variants were investigated. The mutations were mainly derived from clinical isolates of UTI patients (Bouckaert et al. 2006, Hommais et al. 2003, Schwartz et al. 2013), computationally designed using molecular dynamic simulations (Rodriguez et al. 2013, Thomas W. E. et al. 2002) or selected from gene libraries (Schembri et al. 2000).

Certainly most often used and compared is the FimH mutant A188D, which is characterized by a destabilized interaction between FimH_L and FimH_P. This variant was generated in bacterial strains and used in \textit{in vivo} assays with whole bacterial cells always as a model for the high-affinity state of FimH (Anderson et al. 2007, Aprikian et al. 2007, Tchesnokova et al. 2008, Yakovenko et al. 2008). But a quantitative \textit{in vitro} analysis of the ligand affinity or the domain interaction energy of this variant is missing.

Furthermore, there is no report of a full-length FimH variant, which can bind ligands with the same order of magnitude high affinity as the isolated lectin domain (FimH_L). This potential variant could be used in \textit{in vivo} assays representing the domain separated state, occurring under shear stress conditions.

Therefore, the idea was to generate mutants with a disturbed or even abrogated domain interface. Determine their dissociation constants for ligands and compare their stabilities. The free energy of domain interactions could be determined by making use of the difference between the stability against unfolding of the isolated lectin domain and the lectin domain in the context of full-length FimH. Importantly, unfolding of FimH_L in the context of the pilin domain can be selectively monitored by tryptophan fluorescence, because the only two tryptophans in FimH are located in its lectin domain. Beneficial is also the fact that the fluorescence in the native state is highly quenched and that the pilin domain in FimH does not unfold prior to unfolding of the lectin domain (Erilov et al. 2007, Vetsch et al. 2002).
For single amino acid replacements residue 188 in FimHₚ was chosen, which is opposite of one loop of FimHₗ and was replaced by a charged (A188D) or bulky (A188W) residue (Figure 1).

**Figure 1. The interdomain region in FimH.** (A) Overall structure of the FimH molecule with the lectin domain of FimH (Hₗ) in red, the pilin domain of FimH (Hₚ) in yellow, and the donor strand of subunit FimG (DsG) in blue. (B) Close-up of the interdomain region of FimH-DsG. Residues relevant for the investigated FimH variants are labeled and all side chains in a radius of 5 Å relative to A188 shown as sticks (Sauer M. M. et al. 2016) (pdb-code: 4XO9).

Additionally, a set of variants was constructed with amino acid insertions between residue 159 and 160 in FimH with the purpose to mimic the physical separation which is happening for example under flow conditions. The most drastic separation was the insertion of a short and stable model α-helix (159-AAAAKA-160) (Marqusee et al. 1989). The prediction was that this construct would represent the domain separated state with comparable affinities for ligands as the isolated lectin domain (FimHₗ). Another set of variants was designed with different flexible linker lengths from one to seven amino acids between the domains. Cartoon representations of the variants with the corresponding nomenclature are summarized in Figure 2.
Figure 2. Set of FimH variants. Cartoon representation of the respective FimH molecules. With the lectin domain (FimH\textsubscript{L}) in red and the pilin domain (FimH\textsubscript{P}) in yellow which is complemented by the synthetic peptide DsG (donor strand of FimG) displayed in blue. The site of the single amino acid mutation is indicated with a green symbol (X). S and G symbols stand for amino acids and a model helix inserted between position 159 and 160 is shown in grey.

Dissociation constants for the ligands GN-FP-4 and \textit{n}-heptyl \textalpha-D-mannopyranoside (HM) for all FimH variants were determined and compared. Figure 3 shows the molecular structure of the ligands used in binding assays.

Figure 3. Structure of relevant ligands. Fluorescein-labeled mannoside GN-FP-4 and reference compound \textit{n}-heptyl \textalpha-D-mannopyranoside (HM).
2. RESULTS

2.2.2 Results

Ligand binding affinity increase upon domain separation

The experimental data measuring the affinities to HM and GN-FP-4 for all FimH variants are shown in Figure 4 and Figure 5, respectively. The previously described FimH variant with introduction of a charged amino acid in the domain interface (A188D) showed a 10- to 20-fold increase in affinity as well as the variant with a bulky amino acid (A188W, data not shown) for both ligands. The constructs with a small helix between the domains (159-AAAAKA-160) showed the biggest gain in affinity (ca. 2600-fold for HM). Interestingly, the affinity of constructs with a flexible linker between the domains could be tuned by the linker length with increased affinities for HM of 400-fold (159-S-160) to 800-fold (159-SGS-160) and 2100-fold (159-SGSGSGS-160) compared to FimH-DsG (WT). All results are summarized and compared in Table 1.
2. RESULTS

Ligand binding is coupled to domain interaction in FimH

\[ \text{[FimH-DsG A188D]} = 2 \, \mu\text{M} \]
\[ \text{[FimH-DsG A188D]} = 0.2 \, \mu\text{M} \]

A

Fraction of sites occupied

mol. equiv. of added ligand (HM)

B

Fluorescence polarization

\[ \text{[HM]} (\mu\text{M}) \]

\[ \text{[GN-FP-4]} = 1.0 \, \mu\text{M} \]

FimH-DsG 159-S-160

C

\[ \text{[GN-FP-4]} = 1.0 \, \mu\text{M} \]

FimH-DsG A188W

D

\[ \text{[GN-FP-4]} = 1.0 \, \mu\text{M} \]

FimH-DsG 159-SGS-160

E

\[ \text{[GN-FP-4]} = 1.0 \, \mu\text{M} \]

FimH-DsG 159-AAAAKA-160

F

\[ \text{[GN-FP-4]} = 1.0 \, \mu\text{M} \]

FimH-DsG 159-SGSGSGS-160
Figure 4. Determination of HM binding affinity of FimH-DsG interface variants. (A) Equilibrium titration of FimH-DsG A188D with HM at pH 7.4 and 25 °C, recorded via the increase in FimH-DsG A188D fluorescence at 320 nm upon binding of HM (excitation at 280 nm). Titrations were performed at two constant protein concentrations of 0.2 µM (black diamonds) or 2 µM (red circles) and globally fitted according to equation 1 (solid lines). (B, C, D) Equilibrium titration of HM binding to FimH-DsG variants analyzed by displacement of GN-FP-4 from the proteins. An equimolar mixture of protein and GN-FP-4 (1 µM each) was mixed and different HM concentrations (10 nM - 1 mM) were added. Samples were incubated o.n. at 25 °C. GN-FP-4 displacement is monitored by a decrease in fluorescence polarization at 528 ± 20 nm (excitation at 485 ± 20 nm). Data were fitted (solid lines) according to a mechanism in which two ligands compete for the same binding site, with fixed $K_d$ values for GN-FP-4 binding.
2. RESULTS

Ligand binding is coupled to domain interaction in FimH

A

B

C

D

E

F

Norm. fluorescence at 520 nm

Fluorescence at 520 nm (a.u.)
Figure 5. Determination of GN-FP-4 binding affinity of FimH·DsG interface variants. (A-F) Equilibrium titration of GN-FP-4 with FimH variants at pH 7.4 and 25 °C, recorded via the decrease in GN-FP-4 fluorescence at 520 nm upon binding to the FimH lectin domain (excitation at 497 nm). (A,B) measured on a plate-reader. (C-F) measured on a PTI spectrofluorimeter. Titrations were performed at constant GN-FP-4 concentrations of 1 nM or 25 nM (black diamonds) or 2 nM and 50 nM (red circles) and globally fitted according to the standard binding equation 1 (solid lines). Samples contained 0.001 % Tween 20 to prevent unspecific adsorption of GN-FP-4 to tubes or cuvettes.

Table 1. Summary and comparison of n-heptyl α-D-mannoside (HM) and GN-FP-4 binding of FimH variants.

<table>
<thead>
<tr>
<th>FimH·DsG (WT)</th>
<th>$K_d$ (HM) (M)</th>
<th>$K_d$ (HM) (WT/mutant)</th>
<th>$K_d$ (GN-FP-4) (M)</th>
<th>$K_d$ (GN-4-FP) (WT/mutant)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FimH-L</td>
<td>$1.1 \pm 0.1 \times 10^9$</td>
<td>$3.3 \pm 0.4 \times 10^3$</td>
<td>$7.0 \pm 0.1 \times 10^{-11}$</td>
<td>$2.1 \pm 0.1 \times 10^3$</td>
</tr>
<tr>
<td>FimH·DsG A188D</td>
<td>$1.6 \pm 0.4 \times 10^{-7}$</td>
<td>$2.3 \pm 0.6 \times 10^1$</td>
<td>$1.3 \pm 0.1 \times 10^{-8}$</td>
<td>$1.2 \pm 0.1 \times 10^1$</td>
</tr>
<tr>
<td>FimH·DsG A188W</td>
<td>$2.4 \pm 0.1 \times 10^{-7}$</td>
<td>$1.5 \pm 0.1 \times 10^1$</td>
<td>$1.1 \pm 0.2 \times 10^{-8}$</td>
<td>$1.4 \pm 0.3 \times 10^1$</td>
</tr>
<tr>
<td>FimH·DsG 159-AAAAKA-160</td>
<td>$1.4 \pm 0.1 \times 10^{-9}$</td>
<td>$2.6 \pm 0.3 \times 10^3$</td>
<td>$1.5 \pm 0.2 \times 10^{-10}$</td>
<td>$1.0 \pm 0.1 \times 10^3$</td>
</tr>
<tr>
<td>FimH·DsG 159-S-160</td>
<td>$9.1 \pm 0.6 \times 10^{-9}$</td>
<td>$4.0 \pm 0.4 \times 10^2$</td>
<td>$3.6 \pm 0.3 \times 10^{-10}$</td>
<td>$4.2 \pm 0.4 \times 10^2$</td>
</tr>
<tr>
<td>FimH·DsG 159-SGS-160</td>
<td>$4.6 \pm 0.4 \times 10^{-9}$</td>
<td>$7.8 \pm 0.9 \times 10^2$</td>
<td>$1.5 \pm 0.1 \times 10^{-10}$</td>
<td>$1.0 \pm 0.1 \times 10^3$</td>
</tr>
<tr>
<td>FimH·DsG 159-SGSGGS-160</td>
<td>$1.7 \pm 0.1 \times 10^{-9}$</td>
<td>$2.1 \pm 0.2 \times 10^3$</td>
<td>$8.8 \pm 1.1 \times 10^{-11}$</td>
<td>$1.7 \pm 0.2 \times 10^3$</td>
</tr>
</tbody>
</table>
2. RESULTS

The pilin domain is stabilizing the lectin domain in FimH

Two set of experiments were done to compare the affinities of the variants with the domain interaction energy. First, equilibrium Guanidinium chloride (GdmCl) dependent unfolding and refolding transitions were measured with fluorescence spectroscopy (Figure 6B). And secondly, unfolding kinetics were recorded using fluorescence detection (Figure 8). The GdmCl dependent unfolding equilibrium shown in Figure 6 of the isolated lectin domain (black curve) and the lectin domain in the FimH-DsG complex (green curve) showed a stability increase by 10.3 kJ/mol relative to isolated FimHL. The blue curve shows that the domain interactions in the FimH-DsG A188D variant are weaker, rationalizing the population of a domain-separated state.

![Figure 6. Determination of the free energy of folding.](image)

(A) Fluorescence emission spectra (excitation: 280 nm) were recorded for the folded proteins (solid lines) and unfolded proteins in 5 M GdmCl (broken lines) (B) GdmCl induced equilibrium transitions were monitored by the fluorescence intensity increase at 350 nm (excitation: 280 nm) during unfolding.
In general, the interdomain interaction in two-domain proteins can be quantified through the gain in thermodynamic stability of the less stable domain in the context of the full-length protein relative to the isolated, less stable domain (Bhaskara et al. 2011). Previous data on the denaturant-dependent unfolding equilibrium of FimH revealed a three-state mechanism in which both domains unfold separately (Vetsch et al. 2002). But most importantly, experiments on the self-complemented FimH pilin domain with C-terminal FimG donor strand showed that no unfolding of the self-complemented construct occurs after one day of incubation at 5 M GdmCl (Erilov et al. 2007, Puorger et al. 2008). Based on this, the assumption was made that self-complemented FimH_P is folded at GdmCl concentrations below 3.5 M. This is also relevant for the experiments shown in Figure 8.

**Figure 7.** Three-state mechanism of unfolding/refolding equilibrium of full-length FimH. In the self-complemented FimH protein with the peptide DsG (blue) the less stable lectin domain (H_L, red) unfolds prior to the more stable pilin domain (H_P, yellow). N = native state, I = intermediate state, U = unfolded state.
2. RESULTS  Ligand binding is coupled to domain interaction in FimH

Figure 8. Unfolding kinetics of the lectin domain in the context of the pilin domain. Proteins were rapidly mixed with GdmCl to a final concentration of 3.5 M. The fluorescence change was monitored at 320 nm with selective excitation of the tryptophan residue in FimH_L at 297 nm. For comparison the data were normalized, since the amplitudes of the kinetic traces were differing. All kinetic traces were fitted with a single exponential function. (A) Comparison of different FimH·DsG interface variants with FimH_L. (B) Comparison of FimH·DsG interface variants with different linker lengths to FimH_L.

The observed unfolding rates were summarized and compared to FimH·DsG (WT) in Table 2. All together, the values follow the same trend as the binding affinities for ligands confirming the coupling between ligand binding and domain interaction energy in FimH.

Table 2. Summary and comparison of unfolding kinetics and dissociation constants of the FimH variants.

<table>
<thead>
<tr>
<th></th>
<th>$k_{\text{obs}}$ (unfolding) in 3.5 M GdmCl (s$^{-1}$)</th>
<th>$k_{\text{obs}}$ comparison (mutant/WT)</th>
<th>$K_d$ comparison (WT/mutant)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>HM</td>
</tr>
<tr>
<td>FimH·DsG (WT)</td>
<td>0.003 ± 0.001</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>FimH_L</td>
<td>0.222 ± 0.002</td>
<td>74 ± 25</td>
<td>3.3 ± 0.4 x 10$^3$</td>
</tr>
<tr>
<td>FimH·DsG A188D</td>
<td>0.022 ± 0.001</td>
<td>7 ± 2</td>
<td>2.3 ± 0.6 x 10$^1$</td>
</tr>
<tr>
<td>FimH·DsG A188W</td>
<td>0.024 ± 0.001</td>
<td>8 ± 3</td>
<td>1.5 ± 0.1 x 10$^1$</td>
</tr>
<tr>
<td>FimH·DsG 159-AAAAKA-160</td>
<td>0.197 ± 0.001</td>
<td>66 ± 22</td>
<td>2.6 ± 0.3 x 10$^3$</td>
</tr>
<tr>
<td>FimH·DsG 159-S-160</td>
<td>0.123 ± 0.001</td>
<td>41 ± 14</td>
<td>4.0 ± 0.4 x 10$^2$</td>
</tr>
<tr>
<td>FimH·DsG 159-SGS-160</td>
<td>0.172 ± 0.001</td>
<td>57 ± 19</td>
<td>7.8 ± 0.9 x 10$^2$</td>
</tr>
<tr>
<td>FimH·DsG 159-SGSGS-160</td>
<td>0.189 ± 0.001</td>
<td>63 ± 21</td>
<td>2.1 ± 0.2 x 10$^3$</td>
</tr>
</tbody>
</table>
2. RESULTS

Ligand binding is coupled to domain interaction in FimH

2.2.3 Discussion

The establishment of a model system to study the catch-bond mechanism of FimH immediately suggested the mutational analysis of FimH variants. Based on these results the prediction was that FimH variants with weakened interdomain interactions would show higher ligand binding affinity. The sensitivity for the analysis was high, since the isolated lectin domain showed a three orders of magnitude higher affinity for an established model ligand compared to the complemented FimH full-length protein. To test this idea experimentally, single amino acid replacements at a position (A188), which was already been identified to disturb the domain interactions in FimH (Anderson et al. 2007, Aprikian et al. 2007, Tchesnokova et al. 2008, Yakovenko et al. 2008) were introduced. Electrostatic repulsion by a charged residue (A188D) and steric repulsion with a bulky amino acid (A188W) had the same intermediate effect and resulted on average in 16-fold binding affinities gains for the two tested ligands. In a second set of FimH variants, the domain separated state was mimicked by inserting amino acids between the lectin and pilin domain of FimH. A construct with a stable model helix (Marqusee et al. 1989, Scholtz et al. 1992) between the two domains could bind an established model ligand with almost the same binding affinity as the isolated lectin domain. With introduction of a flexible linker from one over three to seven amino acids, the separation of the two FimH domains was varied stepwise. The construct with one amino acid showed two orders of magnitude higher affinity. Seven amino acids enforced three orders of magnitude affinity gain. To sum up, the affinity increase correlated with the physical separation of the FimH domains for both ligands and therefore confirmed the proposed catch-bond mechanism of FimH.

Additionally, experiments for comparing the free energy of domain interactions were performed. The assumption was made that the interdomain interaction in FimH can be defined as the gain in thermodynamic stability of the less stable domain. Important in this context was the fact that FimH_L is less stable and unfolds at lower GdmCl concentrations compared to complemented FimH_P (Erilov et al. 2007, Puorger et al. 2008, Vetsch et al. 2002). Moreover, fluorescence emission spectra of folded and unfolded FimH had previously demonstrated that the only two tryptophan residues located in the lectin domain of FimH are highly quenched in the folded state and show a huge signal increase upon unfolding (Vetsch et al. 2002). The
2. RESULTS  

Ligand binding is coupled to domain interaction in FimH

Explanation for this effect can be seen in the crystal structures of FimH, where the tryptophan residues are closely stacked. Guanidinium chloride induced equilibrium unfolding and refolding transitions of the lectin domain in the A188D mutant showed a stability decrease relative to the WT in FimH-DsG complexes. These results provoked the idea to measure unfolding kinetics and compare the observed rate constants for all variants. All unfolding traces were monophasic. The final comparison of unfolding rates and binding affinities confirmed the correlation between ligand binding properties and domain separation in FimH.

In previous studies mainly FimH variants with low affinity for ligands have been identified (Rodriguez et al. 2013, Schwartz et al. 2013). In several studies, almost exclusively the FimH A188D mutant was used as a model system for the high-affinity state of FimH. It was reported to display a 16-fold increase in affinity in bacterial binding assays with purified pili (Aprikian et al. 2007), which was confirmed in this study. Today, the highest reported affinity for mannose of a lectin is a construct in which the residues 186-201 in the pilin domain of FimH were replaced by residues from the FocH adhesin (Sokurenko et al. 2001). This construct showed almost 300-fold higher affinity to mannose compared to wild type cells (Aprikian et al. 2007).

Here, for the first time, full-length FimH proteins were constructed which showed three orders of magnitude higher affinities to ligands compared to the wild type protein. These constructs can potentially be transformed into bacterial cells producing plasmid encoded type 1 pili with high affinity for terminal mannoses thereby mimicking the domain separated state, which occurs under flow conditions. Moreover, the strength of shear force can be simulated with different flexible linker lengths. This model system can be used for different purposes. For example, it may serve as a reference, representing defined different states of FimH in assays using bacterial cells.
2. RESULTS

Ligand binding is coupled to domain interaction in FimH

2.2.4 Materials and Methods

Materials

Oligonucleotides were purchased from Microsynth (Balgach, Switzerland) and the QuikChange Site-Directed Mutagenesis Kit from Stratagene (La Jolla, CA, USA). Q-sepharose fast flow, SP-sepharose fast flow, the HiLoad 26/60 Superdex 75 column were purchased from GE Healthcare (Little Chalfont, UK). Chemicals of highest available purity were purchased from Sigma, Merck, or AppliChem. AA-grade Guanidinium Chloride (GdmCl) was obtained from NIGU Chemie (Waldkraiburg, Germany). The synthetic DsG peptide (sequence: ADVTITVNGKVVAKR) was purchased from JPT (Germany).

Mutagenesis of Expression plasmids

Six different FimH variants were created. Two with single amino acid replacements and four with amino acid insertions between the domains (between position 150-160). The plasmids for the periplasmic co-expression of full-length FimH with FimC was used as template in the PCR reaction. For single amino acid replacements, the following primers were used:

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>FimH_A188D_f</td>
<td>CTTACCGTTTATATGACAAAAAGCCAAAAACCTG</td>
</tr>
<tr>
<td>FimH_A188D_r</td>
<td>CAGGTTTTGGCTTTTTCAACAAATAACCGTAAG</td>
</tr>
<tr>
<td>FimH_A188W_f</td>
<td>CTTACCGTTTATTGGAAGCCAAAAACCTG</td>
</tr>
<tr>
<td>FimH_A188W_r</td>
<td>CAGGTTTTGGCTTTTCACAAATAACCGTAAG</td>
</tr>
</tbody>
</table>

The following primers were used for inserting amino acids between the FimHL and FimHP domain:

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>FimH_159-S-160_f</td>
<td>AGCGGCTGGGATTTCTGCTG</td>
</tr>
<tr>
<td>FimH_159-S-160_r</td>
<td>GCCAGTAGGCAACACCACATC</td>
</tr>
<tr>
<td>FimH_159-SGS-160_f</td>
<td>AGCGGCGCGGCTGCGATTTCTGCTG</td>
</tr>
<tr>
<td>FimH_159-SGS-160_r</td>
<td>GCCAGTAGGCAACACCACATC</td>
</tr>
<tr>
<td>FimH_159-SGSGS-160_f</td>
<td>GCAGCGGCGCAGCGGTAGCGGGCAAGCTGCGATTTCTGCTG</td>
</tr>
<tr>
<td>FimH_159-SGSGS-160_r</td>
<td>GCCAGTAGGCAACACCACATC</td>
</tr>
<tr>
<td>FimH_159-AAAAKA-160_f</td>
<td>GCGGCGGCGGCCGAAAGCGGCTGCGATTTCTGCTG</td>
</tr>
<tr>
<td>FimH_159-AAAAKA-160_r</td>
<td>GCCAGTAGGCAACACCACATC</td>
</tr>
</tbody>
</table>
2. RESULTS

Ligand binding is coupled to domain interaction in FimH

Production of FimH-DsG complexes

Experiments were performed as previously described (Chapter 2.1.5).

During ESI-MS spectrometry the FimH-DsG complexes partially dissociated. In the following table masses of the protein-peptide complexes are listed. In aqueous solutions the complexes were stable against dissociation and unspecific aggregation.

<table>
<thead>
<tr>
<th>FimH variant</th>
<th>yield (mg/1L)</th>
<th>theoretical mass (Da)</th>
<th>observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>1.5-2.5</td>
<td>30635.3</td>
<td>30636.0</td>
</tr>
<tr>
<td>A188D</td>
<td>0.5-1.0</td>
<td>30679.4</td>
<td>30679.5</td>
</tr>
<tr>
<td>A188W</td>
<td>8.8</td>
<td>30750.5</td>
<td>30751.0</td>
</tr>
<tr>
<td>159-S-160</td>
<td>3.3</td>
<td>30722.4</td>
<td>30723.0</td>
</tr>
<tr>
<td>159-SGS-160</td>
<td>3.3</td>
<td>30866.6</td>
<td>30868.0</td>
</tr>
<tr>
<td>159-SGSGSGS-160</td>
<td>6.9</td>
<td>31154.8</td>
<td>31155.5</td>
</tr>
<tr>
<td>159-AAAAAKA-160</td>
<td>1.3</td>
<td>31118.9</td>
<td>31119.0</td>
</tr>
</tbody>
</table>

Determination of protein concentrations

Protein concentrations were measured via the specific absorbance at 280 nm, using the following extinction coefficients: FimH-DsG (WT and all other mutants) (35’090 M⁻¹cm⁻¹), FimH-DsG A188W (40’590 M⁻¹cm⁻¹), FimH₅ (24’670 M⁻¹cm⁻¹). The concentration of the DsG peptide was determined using \( \varepsilon_{205\text{nm}} = 42’650 \) M⁻¹cm⁻¹.

Determination of HM and GN-FP-4 dissociation constants

Experiments were performed as previously described (Chapter 2.1.5).

Stopped–flow unfolding kinetics

Unfolding kinetics were measured on a SX-18MV stopped-flow spectrofluorimeter (Applied Photophysics) with two 2.5 mL syringes and final concentrations of 1 µM protein and 3.5 M GdmCl after mixing. Experiments were done at 25 °C in 20 mM MOPS-NaOH pH 7.4 by monitoring the change of fluorescence emission at 320 nm with selective excitation of the tryptophan residue of FimHL at 297 nm. All kinetic traces were fitted with a single exponential function and normalized.
2. RESULTS
Ligand binding is coupled to domain interaction in FimH

GdmCl-dependent folding transitions
For unfolding transitions, protein solutions were rapidly diluted with buffer containing different GdmCl concentrations. For refolding transitions, the protein was first incubated in 7 M GdmCl, 20 mM MOPS-NaOH, pH 7.4 during 20 hours at 25 °C and then rapidly diluted into buffer containing different GdmCl concentrations. Folding transitions were measured in a stirred 1x0.4 cm quartz cuvette at 25 °C by following the change in the fluorescence signal (excitation: 280 nm emission: 320 nm) on a QM-7/2003 spectrofluorimeter (PTI) with a peltier element. The volume for each sample was 1 mL and the final protein concentration 1 µM. All folding transitions were fitted according to the two-state model (Santoro et al. 1988).
2.3 A natural ligand of FimH – The glycoprotein uromodulin

J. Eras, G. Weiss, M. Pilhofer, R. Glockshuber

Contribution of the author: the idea and the experimental work, with the exception of the UMOD tomograms shown in Figure 7, the LS-MS/MS analysis in Table 1 and Supplementary Table 1, and the MALDI-MS analysis shown in Supplementary Figure 4.
2. RESULTS

2.3.1 Introduction

Pathogenic bacteria use adhesive organelles to mediate binding to host cell receptors. This is a crucial early event during the establishment of urinary tract infections (UTI). This process is controlled by three elements: E. coli adhesins, host receptors, and host defenses mechanisms. Type 1 pili are known to be highly mannose-sensitive. *In vitro* experiments have demonstrated that type 1 piliated *E. coli* bind to uroplakins Ia and Ib (Wu et al. 1996) in a FimH-dependent fashion (Zhou et al. 2001). It was also shown that uroplakins Ia and Ib carry a single high mannose glycan (Malagolini et al. 2000). In the recent years many data became available about *E. coli* adhesins and their urothelial receptors, but there is a lack of information about the role of host defense mechanisms. A potential candidate for this is the glycoprotein uromodulin (UMOD), the most abundant protein isolated from human urine. The fact that UMOD carries high-mannose glycans prompted the idea that UMOD may serve as a soluble host defense factor. By competing with urothelial receptors for type 1 piliated *E. coli*, UMOD may prevent the bacterial cells from binding to the urothelial surface. In several studies the interaction between type 1 pili bearing *E. coli* cells or purified type 1 pili and UMOD was investigated. One of the first studies showed that type I piliated *E. coli* are trapped by UMOD (Orskov et al. 1980). Later it was demonstrated that UMOD is the main urinary protein binding specifically to type 1 piliated *E. coli*, removal of glycans from UMOD prevents this binding, and that binding of *E. coli* cells to uroplakin receptors is blocked by UMOD (Pak et al. 2001). To get further insights into this mechanism, *in vitro* complex formation between the relevant type 1 pilus subunit FimH and UMOD was examined in this study.

UMOD is synthesized in the thick ascending limb of tubular epithelial cells in the kidney consisting of 640 amino acids. A 24 amino acid long signal peptide directs its insertion in the endoplasmic reticulum. In this cell compartment the GPI anchor is added, formation of 24 intramolecular disulfide bonds is catalyzed, and *N*-glycosylation takes place. The 26 amino acid long propeptide is removed and the protein is transferred to the Golgi, where the glycan chains are modified. It was recently shown that the serine protease hepsin cleaves the GPI anchor releasing the protein in the extracellular space (Brunati et al. 2015). In its mature form containing 563 amino acids it must reach the plasma membrane in a monomeric conformation.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

But in the urine it is found as a filament. Several models exist about the polymerization mechanism, for example derived from the recently solved crystal structure of the ZP domain of UMOD (Bokhove et al. 2016). Nevertheless, the orientation of UMOD monomers within a filament could not be explained so far. Therefore, the architecture of UMOD filaments was studied with electron cryotomography.

Figure 1. Crystal structure of homodimers of the ZP domain of human UMOD (Bokhove et al. 2016) (pdb-code: 4WRN). Homodimer with one ZP domain in cyan and the second ZP domain divided in N-terminal domain (ZP-N, red) and C-terminal domain (ZP-C, green). In the homodimer the ZP-N domains interact via a hydrophobic interface. ZP = zona pellucida.
In Figure 2 a schematic representation of the domain composition of UMOD is shown. It contains four epidermal growth factor domains (I-IV) made of about 40 amino acids each, a domain of unknown function (D8C) and a C-terminal ZP (zona pellucida) domain. The number of Cysteine residues, which are all involved in the formation of intramolecular disulfide bridges, is indicated in Figure 2.

**Figure 2. Schematic representation of the domain composition of UMOD.** Potential glycosylation sites (Asn-X-Ser/Thr motif X ≠ Pro) are highlighted with Asn. The number of Cysteine residues is indicated. Four epidermal growth factor (EGF) domains are labeled with I-IV. D8C: domain containing 8 Cysteine residues. ZP-N: N-terminal zona pellucida domain, ZP-C: C-terminal zona pellucida domain. The proteolytic cleavage site of hepsin for release of UMOD from the luminal face of kidney cells and the GPI anchor site are indicated.

On reducing SDS-PAGE UMOD runs as a 105 kDa protein, the glycans contributing to approximately 30 % of its mass. Detailed studies of released glycans from the protein using NMR approaches revealed a high heterogeneity and resulted in the assignment of 63 complex-type N-glycans also containing oligomannose-type carbohydrate trees ranging from MansGlcNAc2 to Man7GlcNAc2 (Hard et al. 1992, van Rooijen et al. 1998, 1998). It is currently reported, that out of eight potential N-glycosylation sites (Asn14, Asn52, Asn56, Asn208, Asn251, Asn298, Asn372, and Asn489) only the first site (Asn14) is not glycosylated and one site (Asn251) bears oligomannose-type carbohydrate chains (van Rooijen et al. 1999).
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

This page was intentionally left blank.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

2.3.2 Results

Purification of UMOD from its natural source and determination of the molar extinction coefficient

The purification of UMOD from urine was performed as described (Serafini-Cessi et al. 1989). Its procedure was fast and yielded high amounts of pure protein between 30 - 50 mg/L of urine. Figure 3A documents the purification of UMOD. The protein eluted from the additional size exclusion chromatography in one broad post-tailored peak (data not shown). Comparable amounts and purity were obtained from four additional male individuals (data not shown). All experiments in this study were performed with the preparation of the author’s protein.

Figure 3. Purification of UMOD from human urine and determination of the molar extinction coefficient. (A) Coomassie-stained 15 % SDS-polyacrylamide gel documenting the purification process. Lane M: molecular mass standard, lane 1: urine-sample, lane 2: flow-through fraction after the filtration, lane 3: washing step fraction after filtration and lane 4: supernatant after resuspension in water and centrifugation of the purified protein. (B) Determination of the molar extinction coefficient at 280 nm. Spectra of the native (black line) and unfolded protein in 8 M GdmCl (red line) were measured in 20 mM MOPS-NaOH pH 7.4. The absorbance spectrum of native UMOD was corrected for light scattering.
As a prerequisite for the quantitative analysis of the interaction between type 1 pili (FimH) and UMOD the amounts of monomeric UMOD had to be determined in solution. Size exclusion chromatography runs under denaturing conditions were performed (Supplementary Figure 1D) and proved that UMOD fibers could be dissociated into monomers and secondary structure (high content of β-sheet) was lost at high GdmCl concentrations (Supplementary Figure 1A). Furthermore, it could be shown that the filaments are stable at RT and secondary structure begins to be lost above 50 °C (Supplementray Figure 1B). The molar extinction coefficient was determined: \( \varepsilon_{280, \text{native}} = 117'107 \text{ M}^{-1}\text{cm}^{-1} \) (Figure 3B).
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

The binding stoichiometry of the FimH<sub>L</sub>-UMOD complex

To gain more insight into the complex formation between type 1 pili and UMOD the isolated lectin domain (FimH<sub>L</sub>) of FimH was used as a model. It was chosen due to its approximately 2500-fold higher affinity to Mannose compared to full length-FimH (Chapter 2.1.3.2, Figure 9C). Size exclusion chromatography could be used as a method, since the spontaneous dissociation of the complex was very slow (t<sub>1/2</sub> ≥ 1h) (Supplementary Figure 2). Two experiments were performed to estimate the binding stoichiometry. First, UMOD was mixed with excess FimH<sub>L</sub> (guaranteeing ≥ 98 % complex formation), the complexes were isolated and the individual protein amounts quantified with SDS-PAGE (Figure 4). For three different amounts of UMOD (1, 2 and 3 µM) and a constant concentration of FimH<sub>L</sub> (20 µM) applied to the column the complex ratios were estimated to be 3, 2.7, and 2.8 (Figure 4D).
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

Figure 4. UMOD-FimH\textsubscript{L} complex isolation and SDS-PAGE analysis. (A) Three different ratios of FimH\textsubscript{L}/UMOD were mixed and loaded to a Superose 6 10/300 column connected to an ÄKTA purifier system. Size exclusion profiles are shown and the corresponding collected fractions marked with dotted lines. (B) Fractions containing the UMOD-FimH\textsubscript{L} complex (fraction I) and unbound FimH\textsubscript{L} (fraction II) and calibration samples with known protein concentrations were analyzed on a Coomassie- stained 15 % SDS-polyacrylamide gel. M: molecular mass standard (n = 3, only one representative gel is shown). (C) Each band intensity (pixel) from FimH\textsubscript{L} in the complex with UMOD (lanes marked under I) and free FimH\textsubscript{L} (lanes marked under II) was quantified with the Software ImageJ. The average sum of band intensities from three gels was plotted against the FimH\textsubscript{L}/UMOD ratio. (D) The complex ratio from three individual gels was estimated with the calibration lanes (inlay) and was found to be between two and three FimH\textsubscript{L} molecules binding per UMOD monomer. Inlay: The averaged band intensities (n = 3) of the calibration lanes were plotted against the known protein concentration applied to the gel.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

Secondly, a titration experiment was performed and the FimH₉ peak areas used to quantify the protein amounts (Figure 5). The result of this experiment was a FimH₉/UMOD ratio of 2.3.

Figure 5. Determination of the binding stoichiometry of the UMOD-FimH₉ complex. (A) Samples containing 0 - 20 µM FimH₉ were loaded onto a Superose 6 Increase 10/300 GL column connected to a HPLC system and the absorbance at 280 nm was recorded. (B) The peak areas were plotted against the protein concentration to generate a straight calibration line. (C) A constant concentration of 20 µM FimH₉ was mixed with different UMOD concentrations (0.1 - 8 µM) and the peak areas of the FimH₉ size exclusion profiles quantified. (D) The calibration line was used to determine the amount of FimH₉ bound to UMOD and plotted against the UMOD concentration. The slope of the linear part corresponds to the number of FimH₉ molecules (here 2.3) bound per UMOD monomer.
In summary, the results show that two to three FimHL domains bind per UMOD monomer. It cannot be excluded that more than one FimHL binds per glycosylation site. In collaboration with Dr. Chia-wei Lin (Prof. Dr. Markus Aebi, ETH Zürich) UMOD filaments were digested with trypsin and the resulting peptides analyzed with LC-MS/MS methodology (Supplementary Table 1). Two out of eight potential glycosylation sites (Asn-X-Ser/Thr, X ≠ Pro) could be identified carrying high-mannose type glycans. Results are summarized in Table 1.

**Table 1. MS/MS-analysis of UMOD peptides.**

<table>
<thead>
<tr>
<th>Potential glycosylation site</th>
<th>Asn position (mature protein)</th>
<th>MS/MS result (glycosylation type)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>could not be identified</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>could not be identified</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>could not be identified</td>
</tr>
<tr>
<td>4</td>
<td>208</td>
<td>could not be identified</td>
</tr>
<tr>
<td>5</td>
<td>251</td>
<td>high-mannose</td>
</tr>
<tr>
<td>6</td>
<td>298</td>
<td>complex</td>
</tr>
<tr>
<td>7</td>
<td>372</td>
<td>complex</td>
</tr>
<tr>
<td>8</td>
<td>489</td>
<td>high-mannose</td>
</tr>
</tbody>
</table>

Additionally, intact UMOD filaments and the same sample treated with the deglycosidase PNGase F under native conditions was sent to the functional genomics center (University of Zurich) (Supplementary Figure 4). Very broad peaks show a heterogeneous distribution of species, probably due to different glycosylation patterns. The theoretical mass of the secreted, oxidized, and non-glycosylated UMOD polypeptide chain was calculated to be 61’356 Da. The result of the MALDI-MS analysis was 85’072 Da for the intact and 71’096 Da for the deglycosylated sample. Under these experimental conditions 16 % of the mass (13’976 Da) could be removed using PNGase F. Assuming an average mass of 180 Da for one monosaccharide (Morelle et al. 2007) and that all 8 potential glycosylation sites are occupied this would translate into the reasonable average number of approximately 10 sugars per N-glycosylation site. The additional, not removed mass (9’740 Da), could either be attributed to N-glycans not accessible for PNGase F in the native UMOD filament (together with the detected mass-shift translating into 16-17 sugars per N-glycosylation site) or partially come from potential O-glycosylation of UMOD (Easton et al. 2000, Jovine et al. 2002).
2. RESULTS

**FimHₗ binds to UMOD with moderate binding affinity**

The UMOD·FimHₗ complex was further analyzed to obtain the apparent dissociation constant. Therefore, a set of three titration experiments was performed (Figure 6A-C). Here, complexes were prepared with constant FimHₗ concentrations (0.5, 1 and 2.5 µM) and the UMOD concentration was varied from 0 to 3 µM. The samples were analyzed with analytical size exclusion chromatography. The peak areas of FimHₗ were determined, plotted against the UMOD concentration, globally fitted to the order of mass law with open parameters for the FimHₗ concentrations and normalized (Figure 6D). The $K_{d, \text{apparent}} = (4.0 \pm 1.0) \cdot 10^{-7}$ M.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

Figure 6. Determination of the apparent dissociation constant of the UMOD-FimH<sub>L</sub> complex. Three set of experiments were done. The concentration of FimH<sub>L</sub> was kept constant at 0.5 µM (A), 1 µM (B) and 2.5 µM (C) and the UMOD concentration was varied between zero and 3 µM. As a control, one sample containing the respective amount of FimH<sub>L</sub>, 3 µM UMOD, and 200 µM HM was injected (red line). Samples were loaded onto a Superose 6 10/300 GL column connected to a HPLC system and the absorbance at 280 nm was recorded. (D) The peak areas of FimH<sub>L</sub> were normalized and plotted against the UMOD concentration. Data points were fitted globally with the standard binding equation.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

Structural analysis of UMOD filaments with cryoelectron tomography

Today, structural information about UMOD consist mainly of studies with negative-stain electron microscopy of UMOD filaments and one X-ray crystal structure of a ZP-MBP fusion protein construct for the monomer (Bokhove et al. 2016). But no information is available about the molecular architecture of UMOD filaments. Therefore, cryo-electron tomograms were recorded by Gregor Weiss (Prof. Dr. Martin Pilhofer, ETH Zürich) of UMOD. UMOD filaments seemed to be a good specimen for this technique. Three representative results are shown in Figure 7. Figure 7A shows the native UMOD filaments with the main chain zigzag structure and side-arms in a regular pattern connected to the inner core. In Figure 7B, the UMOD-FimH₇ complex is shown. Extra-electron density can be seen in comparison to the free UMOD filaments (Figure 7A). The Elastase-resistant core (see also Supplementary Figure 5) of the protein is shown in Figure 7C. In summary, the results show that the fibers adopt a regular zigzag architecture in which successive monomers are rotated by 180° relative to each other and the C-terminal ZP domain of uromodulin forms the core of the fibril.

Figure 7. Sub-tomogram averages of UMOD filaments. (A) 10 μM UMOD (B) 10 μM UMOD + 100 μM FimH₇ (guaranteeing >99 % occupancy) and (C) 10 μM Elastase-treated UMOD fibrils were blotted on to copper grids and plunge-frozen in liquid ethane-propane. Cryoelectron tomograms were recorded and sub-tomogram averages created. Scale bar: 10 nm.
2. RESULTS

2.3.3 Discussion

There are several examples in nature of host defense mechanisms against infection. For example, in the upper respiratory and the digestive tract, a thick layer of mucus covering the epithelial cells prevents the adhesion of pathogens. Here, bacterial lectins usually bind to glycoproteins or glycolipids displayed by host cells causing infections (Cavallone et al. 2004). Another example is the protease lysozyme, which is secreted in the lacrimal fluid and splits the muramic acid linkage in bacterial cell walls to prevent eye infections. But there are also examples where pathogens bind to soluble host proteins, preventing their recognition as foreign by the host immune system. For example, group A streptococci can be coated with human fibrinogen, avoiding to be recognized by neutrophils (Whitnack et al. 1982). Or Schistosomes envelope themselves in host proteins to be protected from host defense mechanisms (Smithers et al. 1969). Here, a potential defense factor against urinary tract infections, the protein UMOD was investigated. It is the most abundant protein in human urine. It could be shown that at least two FimH molecules can bind to one UMOD monomer. Taking the approximately 2500-fold lower affinity to Mannose of full-length FimH compared to FimH into account the determined apparent $K_d$ of 0.4 μM would be translated to be 1 mM. Nevertheless, since UMOD is a huge polymeric filament in its native form it presents multiple binding sites for type 1 pili and the weak binding affinity would be still sufficient for complex formation in vivo. High UMOD concentrations could even form a three-dimensional net around the bacteria. This scenario may have two effects. Preventing binding of type 1 pili to the host or suppressing the division of bacterial cells which would reduce bacterial infection rates.

The structural analysis of UMOD filaments revealed a fishbone-like structure which was never been observed in this detail before. Additional electron density was detected when the lectin domain of FimH was added to native UMOD filaments, confirming complex formation. It was previously shown that the N-terminal region of UMOD (aa 1-291), mainly containing three EGF like domains, could be digested with pancreatic Elastase (Jovine et al. 2002). The stable, C-terminal fragment (running at approximately 43 kDa on reducing SDS-PAGE) could be obtained and was still in a multimeric state. Electron cryotomography of this sample explained that the previously reported zick-zack structure was caused by the central core of the
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

filament, consistent with studies demonstrating that the ZP domain is sufficient for UMOD polymerization (Jovine et al. 2006). To get deeper insight into the polymerization mechanism UMOD filaments were dissociated with GdmCl (Supplementary Figure 1 C,D). In 3 M GdmCl UMOD seemed to form a small fraction of higher oligomers and monomers, and above 4 M GdmCl only UMOD monomers were detected. Unfortunately this reaction was not reversible and attempts to refold or repolymerize UMOD were not successful (data not shown). Interestingly, UMOD filaments could also be dissociated into monomers at low pH (data not shown). Whether this process is reversible was not investigated further.

Enzymatic deglycosylation of UMOD confirmed that the main fraction of glycans contains complex-type oligosaccharides (Supplementary Figure 3). After PNGase F treatment, which cleaves high mannose and complex oligosaccharides from N-linked glycoproteins, the size was reduced from 105 kDa to approximately 80 kDa. Whereas Endo H treatment, which removes only high mannose glycans, did not result in a major change of the protein size. This confirmed the previously reported results that the carbohydrate moiety accounts for approximately 20-30 % of the THP molecule by weight (Fletcher et al. 1970, Muchmore et al. 1985). To get further insights in the site-specific glycosylation, trypsinated UMOD peptides were analyzed with LC-MS/MS approaches (Supplementary Table 1). Out of eight glycosylation sites four could be assigned, two carrying high-mannose type glycans. This would be in agreement with the complex stoichiometry of two to three binding sites for FimH₇ per UMOD monomer. It could either be that one out of the not identified glycosylation sites is a high-mannose type glycan or that more than one FimH₇ molecule can bind per glycosylation site.

In light of the increased emergence of antibiotic resistance, FimH has become an attractive target for developing an alternative strategy against UTI with anti-adhesive drugs blocking the binding site of FimH (Hartmann et al. 2011). In general, studying carbohydrate-lectin interactions might probably influence the development of high-affinity antiadhesive molecules targeting bacterial lectins. In this context it is worth to mention that many of the most promising anti-adhesive compounds are high-valency glycopolymers (Sattin et al. 2016). Since UMOD is also displaying multiple binding sites in its filamentous structure principles for the development of antiadhesive drugs could be transformed from the current study.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

2.3.4 Supplementary information

Supplementary Figure 1. Far-UV CD spectra, thermal, and thermodynamic stability of UMOD fibrils. (A) Spectra of native fibrils (solid line) and denatured monomers in 8 M GdmCl (broken line) were measured at pH 7.4 and 25 °C (protein concentration: 1 µM). (B) Thermal unfolding transition of UMOD fibrils at pH 7.4. Unfolding was followed by the CD signal change at 210 nm upon heating. The experimental data were tentatively fitted according to a two-state model and normalized (protein concentration: 1 µM). (C) GdmCl induced equilibrium transition was monitored by the CD signal change at 220 nm (protein concentration: 1 µM) and tentatively fitted according to a two-state model. (D) Size exclusion chromatography at variable GdmCl concentrations. Samples were incubated in the corresponding buffer o. n. and passed through a Superose 10/300 GL column equilibrated in 20 mM MOPS-NaOH pH 7.4 containing the respective GdmCl concentration at RT on an ÄKTA system. The size exclusion retention profiles were normalized.
Supplementary Figure 2. Dissociation kinetics of FimHL from UMOD. A solution with equimolar concentrations (3 µM each) of FimHL and UMOD or HM was incubated for 18 h. The complexes were mixed with 3 µM GN-FP-4, and the decrease in GN-FP-4 fluorescence at 520 nm as a consequence of UMOD or HM dissociation and GN-FP-4 binding was recorded. This confirms the high stability of the FimHL·UMOD complex and suitability for size exclusion chromatography experiments. Experiments were performed in 10 mM MOPS-NaOH pH 7.4, 30 mM NaCl at 25 °C.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

Supplementary Figure 3. Deglycosylation kinetics of UMOD fibrils with the enzymes Endo H and PNGase F. Samples were incubated at 37 °C and taken from one reaction tube after the indicated time points. The reaction was quenched with SDS-loading buffer and samples were analyzed on a Coomassie-stained 10 % polyacrylamide gel.

Supplementary Figure 4. MALDI-MS spectra of glycosylated and PNGase F digested UMOD fibrils. Native UMOD fibrils (black) or digested under native conditions with PNGase F (red) were analyzed at the functional genomics center (University of Zurich). The weighted average mass of all points in one peak (red peak: 67’500-74’000 m/z and black peak: 80’000-90’000 m/z) was calculated and is indicated. The calculated mass of the monomeric, deglycosylated UMOD peptide after secretion is also given.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

Supplementary Figure 5. Proteolytic digestion of UMOD fibrils with Elastase. Samples were incubated at 37 °C, pH 9 for 1h with different Elastase concentrations (indicated in µM) and analyzed on a Coomassie-stained 15 % polyacrylamide gel.
The following data presented in Supplementary Table 1 were measured and evaluated by Dr. Chia-wei Lin (Prof. Dr. Markus Aebi, ETH Zürich).

**Supplementary Table 1.** Mass spectrometry results of UMOD peptides after digestion with Trypsin and assignment to glycosylation types.

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>m/z</th>
<th>z</th>
<th>[M+H]+</th>
<th>[peptide+H]+</th>
<th>Glycan Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACAGGYVYN251</strong></td>
<td>1517.093</td>
<td>4</td>
<td>6065.373</td>
<td>4848.8167</td>
<td>Man5</td>
</tr>
<tr>
<td></td>
<td>1246.286</td>
<td>5</td>
<td>6227.431</td>
<td>4848.8167</td>
<td>Man6</td>
</tr>
<tr>
<td></td>
<td>1557.606</td>
<td>4</td>
<td>6227.426</td>
<td>4848.8167</td>
<td>Man6</td>
</tr>
<tr>
<td></td>
<td>1598.117</td>
<td>4</td>
<td>6389.468</td>
<td>4848.8167</td>
<td>Man7</td>
</tr>
<tr>
<td><strong>QDFN298</strong></td>
<td>1486.625</td>
<td>3</td>
<td>4457.874</td>
<td>1700.86081</td>
<td>(HexNac2Hex3)Fuc(HexNacHex)2NeuAc2HexNac2</td>
</tr>
<tr>
<td></td>
<td>1594.657</td>
<td>3</td>
<td>4781.97</td>
<td>1700.86081</td>
<td>(HexNac2Hex3)Fuc(HexNacHex)4NeuAc2</td>
</tr>
<tr>
<td></td>
<td>1196.244</td>
<td>4</td>
<td>4781.977</td>
<td>1700.86081</td>
<td>(HexNac2Hex3)Fuc(HexNacHex)4NeuAc2</td>
</tr>
<tr>
<td><strong>N372</strong></td>
<td>1063.464</td>
<td>4</td>
<td>4250.857</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)(HexNacHex)2NeuAc</td>
</tr>
<tr>
<td></td>
<td>1099.975</td>
<td>4</td>
<td>4396.898</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)Fuc(HexNacHex)2NeuAc</td>
</tr>
<tr>
<td></td>
<td>1154.747</td>
<td>4</td>
<td>4615.989</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)(HexNacHex)3NeuAc</td>
</tr>
<tr>
<td></td>
<td>1466.3</td>
<td>3</td>
<td>4396.899</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)Fuc(HexNacHex)2NeuAc</td>
</tr>
<tr>
<td></td>
<td>1417.614</td>
<td>3</td>
<td>4250.841</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)(HexNacHex)2NeuAc</td>
</tr>
<tr>
<td></td>
<td>1063.464</td>
<td>4</td>
<td>4250.857</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)(HexNacHex)2NeuAc</td>
</tr>
<tr>
<td></td>
<td>1136.237</td>
<td>4</td>
<td>4541.946</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)(HexNacHex)2NeuAc2</td>
</tr>
<tr>
<td></td>
<td>1227.521</td>
<td>4</td>
<td>4907.083</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)(HexNacHex)3NeuAc2</td>
</tr>
<tr>
<td></td>
<td>1318.803</td>
<td>4</td>
<td>5272.21</td>
<td>2337.1727</td>
<td>(HexNac2Hex3)(HexNacHex)4NeuAc2</td>
</tr>
<tr>
<td><strong>FALLMTN</strong></td>
<td>1875.762</td>
<td>2</td>
<td>3750.524</td>
<td>2372.0531</td>
<td>Man6</td>
</tr>
<tr>
<td><strong>CYATPSSN489</strong></td>
<td>1250.846</td>
<td>3</td>
<td>3750.537</td>
<td>2372.0531</td>
<td>Man6</td>
</tr>
</tbody>
</table>

2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

2.3.5 Materials and Methods

Materials

Chemicals were purchased from Sigma, Merck, and AppliChem. AA-grade Guanidinium Chloride (GdmCl) was obtained from NIGU Chemie (Waldkraiburg, Germany). Gel filtration columns were purchased from GE Healthcare.

Protein purification

1 L of urine from one healthy male (myself) was collected and stored in 0.1 % sodium azide at 4 °C. The urine was neutralized with 1 M NaOH and filtered through a layer of 20 g diatomaceous earth. The filter was washed with 1 L 25 mM NaH₂PO₄-NaOH pH 7.5, 140 mM NaCl. The matrix was stirred in 150 mL deionized water for 30 min and the solution centrifuged at 20000 g, 4 °C for 30 min. The purification steps were documented on a 15 % SDS-PAGE gel (Figure 3A). The supernatant was concentrated using 15 mL centrifugal filter devices (NMWL: 100 kDa) and passed over a Superdex 200 26/60 gel filtration column (GE Healthcare, USA) equilibrated with 10 mM NaH₂PO₄-NaOH pH 7, 30 mM NaCl. Fractions containing pure protein were pooled dialyzed against deionized water and stored at -80 °C.

Determination of protein concentration

The molar extinction coefficient of monomeric UMOD at 280 nm was determined by measuring UV spectra of native and unfolded (8 M GdmCl) protein in 20 mM MOPS-NaOH pH 7.4 (Gill et al. 1989) (Figure 3B). The spectra were corrected for light scattering as described (Colon 1999). ε₂₈₀ = 117’107 M⁻¹cm⁻¹ was used to determine the protein concentration.

CD spectroscopy and thermal stability

Far-UV CD spectra were recorded at 25 °C using a J-715 spectropolarimeter (Jasco, USA) and a 0.1 cm cuvette. Sensitivity, step size, scan speed and response were set to 100 mdeg/s, 0.5 mm, 20 mm/min and 2s, respectively. The protein concentration was 2.5 µM in 10 mM NaH₂PO₄-NaOH pH 7. The spectrum of the unfolded protein was measured in 8 M GdmCl. Eight spectra were accumulated, averaged, and baseline subtracted. Raw ellipticities were converted to mean residue ellipticities as described (Kelly et al. 2005).
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

The thermal unfolding transition of UMOD in 10 mM NaH$_2$PO$_4$-NaOH pH 7 was monitored by the change of the far-UV CD signal at 210 nm with a heating rate of 1 °C min$^{-1}$ on a J-715 spectropolarimeter (Jasco, USA).

**Analytical gel filtration under denaturating conditions**

The dissociation of UMOD fibers was analyzed with gel filtration. Therefore samples with a final volume of 100 µL, containing 1 µM UMOD and different GdmCl concentrations were mixed, incubated o.n. and loaded on a Superose 6 10/300 GL column (GE Healthcare, USA) connected to an ÄKTA system (GE Healthcare, USA). The samples were eluted at RT with 10 mM NaH$_2$PO$_4$-NaOH pH 7.0 and the corresponding GdmCl concentration (flow rate of 0.75 mL/min).

**Preparative gel filtration and isolation of the UMOD·FimH$_L$ complex**

To isolate the UMOD·FimH$_L$ complex, 20 µM of FimH$_L$ were pre-incubated overnight with different UMOD concentrations (1, 2 and 3 µM). The samples were passed over a Superose 6 10/300 GL column (GE Healthcare, USA) connected to an ÄKTA system (GE Healthcare, USA) and eluted at RT with 10 mM MOPS-NaOH pH 7.4, 30 mM NaCl and a flow rate of 1 mL/min. The peak fractions (indicated with dotted lines) of the complex (eluted at ca. 7.5 ml) and free FimH$_L$ (eluted at ca. 19.5 ml) were collected. The fractions and four additional calibration samples were loaded on 15 % SDS-PAGE gels in triplicates to eliminate gel artefacts. The band intensities were determined and background corrected with the software ImageJ to estimate the binding stoichiometry of the UMOD·FimH$_L$ complex.

**Analytical gel filtration and determination of UMOD·FimH$_L$ complex stoichiometry**

To determine the binding stoichiometry of the UMOD·FimH$_L$ complex samples were loaded onto a Superose 6 Increase 10/300 GL column (GE Healthcare, USA) connected to a HPLC system (Agilent Technologies, USA) and eluted at RT with 10 mM MOPS-NaOH pH 7.4, 30 mM NaCl, and a flow rate of 0.75 mL/min. Eluted proteins were detected by measuring the absorbance at 280 nm. Peak areas of the elution profiles were analyzed using the HPLC software package (Agilent Technologies, USA).
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

Analytical gel filtration and determination of apparent dissociation constant of the UMOD·FimH<sub>L</sub> complex

To determine the apparent dissociation constant samples were prepared with constant concentration of FimH<sub>L</sub> at 0.5, 1 and 2.5 µM and the UMOD concentration was varied between zero and 3 µM. As a control, one sample containing 3 µM and 200 µM HM was prepared. Samples were loaded onto a Superose 6 10/300 GL column connected to a HPLC system (Agilent Technologies, USA) and the absorbance at 280 nm was recorded. The peak areas of FimH<sub>L</sub> were normalized and plotted against the UMOD concentration. Data points were fitted globally with the standard binding equation.

Displacement of UMOD from the lectin domain of FimH by GN-4-FP

The rate constant of dissociation (k<sub>off</sub>) for UMOD from FimH<sub>L</sub> at 25 °C in 20 mM MOPS-NaOH pH 7.4 was measured indirectly by binding of excess GN-FP-4 to FimH<sub>L</sub> after dissociation of UMOD, recorded with the decrease in GN-FP-4 fluorescence at 520 nm (excitation at 497 nm). A mixture of FimH<sub>L</sub> and UMOD (3 µM each) was pre-incubated for at least 18 h. This complex was mixed with equimolar amounts of GN-FP-4 and the fluorescence was recorded in a 1.0 x 0.4 cm quartz cuvette on a QM 7/2003 spectrofluorimeter (PTI, USA) equipped with a magnetic stirrer every 10 min for 10 s, averaged and fitted according to first-order kinetics.

Deglycosylation and proteolytic digestion

According to the manufacturers protocol 10 µg of UMOD were incubated at 100 °C in 0.5 % SDS, 40 mM DTT for 10 min. The solution was cooled and the deglycosylation reaction started by adding Endo H or PNGase F in 50 mM Sodium acetate pH 6.0. After the indicated incubation time at 37 °C samples were taken and the reaction was quenched with 8 M urea. The samples were analyzed on a 10 % SDS-PAGE gel (Supplementary Figure 3).

The proteolytic digestion of UMOD was done with pancreatic Elastase (Promega, USA). The lyophilized enzyme was resuspended in water, centrifuged and the concentration determined using the molar extinction coefficient at 280 nm (55390 M<sup>-1</sup> cm<sup>-1</sup>). Ten samples, including controls and samples where the concentration of UMOD was kept constant at 13 µM and the enzyme concentration was varied
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

between 10 and 0.07 µM were incubated at 37 °C for one hour in 20 mM Tris-pH 9.0. Reactions were stopped by adding protease inhibitor (final concentration: 10 mM PMSF in DMSO). The samples were analyzed on a 15 % SDS-PAGE gel (Supplementary Figure 5).

Preparation of frozen-hydrated specimens for Electron Cryotomography (ECT)

Preparations of purified UMOD were mixed with 10 nm BSA-coated colloidal gold particles (1:4 v/v, Sigma). A 3.5 µL droplet of the mixture was applied to a glow-discharged copper EM grid (R2/1, Quantifoil, Germany). The grid was automatically blotted from both sides, plunge-frozen in liquid ethane-propane (37 % / 63 %) using a Vitrobot (FEI, USA) (Iancu et al. 2006, Tivol et al. 2008) and kept in liquid nitrogen.

Electron Cryotomogramm imaging

Purified UMOD was examined by electron cryotomography. Images were recorded on a Tecnai Polara TEM (FEI, USA) equipped with post-column GIF 2002 imaging filter and K2 Summit direct electron detector (Gatan, USA), or on a Titan Krios TEM (FEI, USA) equipped with a Quantum LS imaging filter and K2 Summit (Gatan, USA). Both microscopes were operated at 300kV and the imaging filters with a 20 eV slit width. The pixel size at the specimen level ranged from 3.45 Å to 4.05 Å. Tilt series covered an angular range from -60° to +60° with 3° increments and -4 to -8 µm defocus. The total dose of a tilt series was between 80-100 e-/Å². Tilt series and 2D projection images were acquired automatically using UCSF Tomo (Zheng et al. 2007) and SerialEM (Mastronarde 2005). Three-dimensional reconstructions were generated using the IMOD program suite (Kremer et al. 1996).

Sub-tomogram averaging

Tomograms used for subtomogram averaging were CTF-corrected in IMOD. UMOD filaments were identified visually in individual tomograms and their longitudinal axes were modeled with open contours in 3dmod. Model points, the initial motive list, and the particle rotation axes were generated using the stalkInit program from the PEET package (Heumann et al. 2011, Nicastro et al. 2006). This approach allowed the definition of each structures longitudinal axis as the particle y-axis. 614 to 1156 individual structures were averaged using PEET. A random particle was chosen as a first reference. Missing wedge compensation was activated.
2. RESULTS

A natural ligand of FimH – The glycoprotein uromodulin

This page was intentionally left blank.
3. GENERAL DISCUSSION AND OUTLOOK

Finding novel and innovative ways to target bacterial infections is a key challenge in chemical biology research. The focus of my work was on sugar-protein interactions involved in urinary tract infections. A detailed mechanistic insight into the interaction of the type 1 pilus adhesin FimH with a model ligand by biophysical and structural analysis shed light on the catch-bond mechanism of FimH. Furthermore, this mechanism was corroborated with a mutational analysis of the FimH protein. Finally, complex formation of FimH with one of its natural ligands, the glycoprotein uromodulin (UMOD), was studied.

The kinetic characterization of ligand binding, in the absence of mechanical force, of the physiologically relevant full-length FimH protein revealed an extremely dynamic lectin-sugar interaction. Most interestingly was that the interactions between the lectin and the pilin domain in full-length FimH accelerate ligand dissociation more than 100’000-fold. Especially this fast off-rate under static conditions makes one of the future goals regarding the therapeutic use of FimH antagonists very promising (Hartmann et al. 2011). A synthetic ligand would have a good chance to bind to FimH, thereby blocking bacterial attachment to the urinary tract, the first critical step in UTIs. Such a ligand should ideally have low dissociation rates in comparison to natural ligands, displayed on epithelial cells of the urinary tract. To screen for such a ligand, the fluorescein-labeled mannoside GN-FP-4 could be used. GN-FP-4 exhibits an about 7-fold increase in fluorescence polarization at 520 nm upon binding to FimH. Consequently, the physiologically relevant target FimH-DsG could be used very conveniently in 96-well format high-throughput screens. GN-FP-4 binds to FimH-DsG with a $K_d$ of $(1.3 \pm 0.1) \cdot 10^{-7}$ M. In analogy to the established assay with FimH, the dissociation rate of identified high-affinity FimH-DsG binders could be determined. In this inverse competition experiment complexes between FimH-DsG and promising ligands would be first formed and later GN-FP-4 added. Importantly, it would be necessary to show that the displacement rate of the new ligands by GN-FP-4 is first-order and independent of GN-FP-4 concentration. Thus, the observed rate in GN-FP-4 polarization increase would be identical to the rate of dissociation of the new ligand from FimH-DsG.
Already 30 years ago, the first antiadhesive small molecules, oligomannosides and aryl α-D-mannosides, binding to the lectin domain of FimH and therefore inhibiting bacterial adhesion, were reported (Aronson et al. 1979, Firon et al. 1982, 1983). These compounds were tested in mouse models and showed promising results, concerning the reduction of UTIs. This initiated the research on several monovalent mannose-based FimH antagonists. Here, the non-mannose component that remains after hydrolysis of a glycoside (aglycon) was varied (Kleeb et al. 2016). This was done due to the fact that in crystal structures of the lectin domain of FimH with several of those compounds, it could be seen, that the entrance to the mannose binding site is formed by two tyrosines and one isoleucine, supporting hydrophobic contacts with the aglycon (de Ruyck et al. 2016, Fiege et al. 2015, Wellens et al. 2012). Today, several mono- and multivalent FimH anti-adhesive molecules with nanomolar affinities were generated and tested in mouse models (Hartmann et al. 2011). But today, none of these compounds is in clinical phase trials (Jarvis et al. 2016, Mydock-McGrane L. et al. 2016, Mydock-McGrane L. K. et al. 2016). This primarily is due to the fact that these compounds show poor bioavailability due to their high hydrophobicity. The ideal compound should be administered orally. But for FimH-antagonists which are administered orally and active in the urinary tract, high solubility in aqueous solutions and fast renal excretion rates are needed (Ernst et al. 2009). Therefore the prodrug principle (Albert 1958), in which a drug is delivered and activated at the target site has been transferred from other drugs to antiadhesive-dugs (Stella et al. 2007, Stewart et al. 1986, Wire et al. 2006). The first example is a FimH-antagonist which was demonstrated to undergo intestinal absorption and fast renal clearance (Klein et al. 2010). In a second example the aqueous solubility could be improved up to 140-fold using a phosphate prodrug (Kleeb et al. 2016). As mentioned above the potency of these drugs has been verified in mouse models but today they are not commercially available.

The fact that FimH is highly specific for terminal mannoses, the simplest idea of consuming D-Mannose as a prophylaxis for recurrent UTIs was just recently tested in humans (Kranjcec et al. 2014). In this study around 300 women, with history of acute cystitis and initial antibiotic treatment were split into three groups: One group received 2 g of D-mannose powder daily, the second received 50 mg Nitrofurantoin daily, and the third did receive just water as a control. 15 % in the D-mannose group, 20 % in Nitrofurantoin group, and 61 % in the control group suffered under recurrent
UTIs. The authors summarize that there was no significant difference in the rate of recurrent UTIs between the intake of D-mannose and the antibiotic. Nevertheless, patients of the D-mannose group had a significantly lower risk of side effects compared to the antibiotic group. In this context it is interesting to mention that D-mannose powders are sold by companies for prevention of UTIs in humans (Falcento or Nutrin, Germany).

In the first part of the thesis, additional results clearly demonstrated that the three orders of magnitude higher affinity of the isolated lectin domain of FimH relative to the full-length protein FimH is conserved and independent of the ligand. Two conformational states were assigned to these proteins. The full-length FimH-DsG protein with associated lectin and pilin domain was defined as the associated state (A-state). The isolated lectin domain of FimH termed the separated state (S-state). The above mentioned result can be summarized and are consistent with a thermodynamic cycle (Figure 1) in which the A and S-states of FimH differ in ligand affinity by a factor of 1500-3000.

\[
\begin{align*}
A_{\text{free}} + \text{ligand} & \xrightarrow{K_d (\text{FimH-DsG})} A_{\text{bound}} \cdot \text{ligand} \\
S_{\text{free}} + \text{ligand} & \xrightarrow{K_d (\text{FimH})} S_{\text{bound}} \cdot \text{ligand}
\end{align*}
\]

Figure 1. Thermodynamic cycle of the coupling between domain separation and ligand (HM) binding in the absence of tensile mechanical force. As no experimental evidence for a significant population of \(S_{\text{bound}}\) could be obtained, the lower limit for the energy required for domain separation in the ligand-bound state was set to at least 10 kJ/mol. This translates into at least 30 kJ/mol for domain separation in the ligand free-state of FimH, corresponding to an \(A_{\text{free}}/S_{\text{free}}\) ratio of > 200000/1. \(A = \) associated state, \(S = \) separated state, \(K_d = \) dissociation constant, energy-values are given for the solid dashes.
To quantify the equilibrium constant in the absence of ligand, GdmCl dependent equilibrium folding transitions of FimH\(_L\) and FimH·DsG were measured. Here, the assumption was made that the interdomain interaction in FimH can be defined as the gain in thermodynamic stability of the less stable domain. Important in this context was the fact that FimH\(_L\) is less stable and unfolds at lower GdmCl concentrations compared to complemented FimH\(_P\) (Erilov et al. 2007, Puorger et al. 2008, Vetsch et al. 2002). Compared to FimH\(_L\), the transitions of FimH·DsG showed a stabilization of the lectin domain by -10.3 kJ/mol. The change in free energy of binding to HM between FimH·DsG and the isolated lectin domain is -20.1 kJ/mol. Therefore the binding energy exceeds the interdomain stabilization by 9.8 kJ/mol. Two possible explanations for this discrepancy will be shortly discussed in the following. Preliminary results (data not shown) demonstrated that FimH·DsG increases its ligand affinity at high ionic strength. The simplest explanation for this would be that the interaction between the pilin domain and the lectin domain in FimH·DsG can be disrupted at high ionic strength, leading to the high-affinity state of FimH (S-state). Since GdmCl is the hydrochloride salt of guanidine and therefore influences the ionic strength, the value for the domain interaction energy would be different. Another explanation for the discrepancy between the binding energy and the interdomain stabilization could be that FimH\(_L\) is not a two-state folder.

The results of my experiments, especially the introduction of peptide-complemented FimH as a model system, representing FimH in its natural form, provide the basis for analyzing FimH mutants found in clinical isolates from UTI patients (Chen et al. 2009, Iebba et al. 2012, Schwartz et al. 2013). To the best of my knowledge there is no comprehensive quantitative analysis in the literature, especially including \textit{in vitro} measurements of ligand binding kinetics. This would explain and reveal the relevance of FimH mutations in UTI infections and might have an influence in the development process of anti-adhesive drugs.

To further validate the hypothesis that FimH variants with weakened domain interactions bind tighter to mannose, their swimming behavior could be investigated on BSA-mannose coated surfaces by light microscopy. For this \textit{in vivo} approach plasmid encoded FimH variants expressed in a FimH deletion strain should be used. The prediction would be that strains expressing FimH variants with perturbed domain interactions should show longer average resting times on BSA-mannose coated
surfaces, and, due to the slower on-rates observed for the FimH lectin domain, longer average distances in the swimming phase. Furthermore, a synthetic FimH gene library with randomized residues in the domain interface could be generated. This would enable the selection of FimH variants with improved adhesion or improved motility. Experimentally this could be done on agar plates, where motility correlates with the velocity of bacterial lawn formation. To select for strains with low affinity (strong domain interactions) this assay could be modified by using agar plates displaying mannose on the surface. High affinity binders could be selected with microfluidic devices on mannose-coated surfaces (Swiecicki et al. 2013).

Another interesting future research aspect concerning FimH would be a molecular evolution approach. FimH could be used as a scaffold for generating lectins with high selectivity for different carbohydrate ligands. Research in this direction might be very promising because of two facts: First, my experiments have demonstrated the high selectivity of FimH for mannose and secondly a recent study has shown that nature has used the fold of the lectin domain of FimH already more than once. In the paper the authors demonstrate that other pilus systems, like the F9 pilus carry a lectin (FmlH) with high specificity for galactose (Conover et al. 2016). As mentioned above the overall structure was almost the same, with root-mean-squared deviation of 1.3 Å. By taking a closer look into the crystal structures, the specificity change from mannose to galactose was due to three point mutations in close proximity to the ligand binding site. This result emphasizes the high potential of using for example the isolated lectin domain of FimH, due to its high affinity to ligands, as a model protein for the development of screening assays in lectin design. The resulting protein variants, with high selectivity for different carbohydrates, could later be used as tools in a variety of applications. It is well known that changes in glycosylation patterns in plasma proteins are involved in a number of diseases such as cancer (Varki et al. 2009). Therefore, newly designed carbohydrate-specific lectins could be used in early-stage diagnosis and prognosis in clinical tests (Damborsky et al. 2016). Furthermore, these lectins could be valuable tools in deciphering the site-specific glycosylation profiles of glycoproteins.

Very promising are ongoing studies on the structure determination of the UMOD filaments mainly using electron cryotomography. Here the advantage can be used that UMOD produced from primary cultures of TAL cells from mouse kidneys shows
the same properties as UMOD prepared from its natural source (Glaudemans et al. 2014). Very convenient is that UMOD will be just secreted from the cells in the medium and can be harvested easily from the cell culture medium. Additionally, even UMOD knock out cell lines exist, allowing the plasmid encoded production of UMOD filaments (Brunati et al. 2015). This will offer the potential to answer many more interesting questions. Studying the influence of glycosylation on filament formation by removal of the glycosylation sites, expression of UMOD fragments, and investigating the effect of mutations found in patients for example with chronic kidney diseases (RampOLDi et al. 2011).
4. REFERENCES


4. REFERENCES


4. REFERENCES


4. REFERENCES


4. REFERENCES


4. REFERENCES


4. REFERENCES


4. REFERENCES


Orskov I, Ferencz A, Orskov F. Tamm-Horsfall Protein or Uromucoid Is the Normal Urinary Slime That Traps Type-1 Fimbriated Escherichia-Coli. Lancet, 1980, 8173:887-.


4. REFERENCES


Sato H, Frank DW. Multi-Functional Characteristics of the Pseudomonas aeruginosa Type III Needle-Tip Protein, PcrV; Comparison to Orthologs in other Gram-negative Bacteria. *Frontiers in microbiology*, 2011:142.


4. REFERENCES


4. REFERENCES


Acknowledgements

First of all, I would like to thank my PhD supervisor Prof. Dr. Rudi Glockshuber for giving me the opportunity to work on highly interesting PhD projects, for his great ideas concerning experiments, and for giving me the freedom to follow own ideas.

I am grateful to Prof. Dr. Eilika Weber-Ban for always valuable input during our joint seminars and for being co-referee of my thesis. I am also grateful to Prof Dr. Julia Vorholt and Prof. Dr. med. Olivier Devuyst for being co-referees of my thesis.

I am grateful to Dr. Chasper Puorger for guiding me during the start of my PhD.

Dr. Roman-Peter Jakob (Prof. Dr. Timm Meier, Biozentrum Basel) is acknowledged for work on the X-ray structure determination of FimH. I would like to thank Dr. Michal Walzcack (Prof. Dr. Gerhard Wider, ETH Zurich) for measuring NMR spectra of FimH. I thank Gregor Weiss (Prof. Dr. Martin Pilhofer, ETH Zurich) for doing electron tomography experiments with uromodulin.

I am grateful to Karin Buholzer (Prof. Dr. Ben Schuler, Universitity of Zurich) for initial single molecule FRET experiments with FimH.

For MS-analysis of uromodulin I want to thank Dr. Chia-wei Lin (Prof. Dr. Markus Aebi, ETH Zürich). I would like to thank Sonia Youhanna (Prof. Dr. med. Olivier Devuyst, University of Zurich) for delivery of samples for uromodulin preparations.

I thank all past and present members of the Glockshuber group for many interesting discussions and a great time: Carmen Palomino, Christoph Giese, Daniel Roderer, Dawid Zyla, Elisabeth Mohorko, Fabia Canonica, Helene Fäh-Rechsteiner, Helge Abicht, Hiang Dreher, Lena Bolten, Marcel Bolten, Martin Schärer, Maximilian Sauer, Rafal Zdanowicz, Silvia Napolitano, Thomas Spirig, Toni Vagt, Verena Eggli, and Zuzana Becarova-Zigova.

I would like to thank my friends, Dr. Malte Gersch, Dr. Melanie Weiss, and Stefan Gerhardy for being always available as constructive discussion partners and Dr. Chris Herbert Stainley for nice evening runs. I am grateful to Dr. Carsten Magnus for proof-reading parts of my thesis.

Last but not least, I am very grateful to my parents and sisters for the constant support over the last years.