

DISS. ETH No. 18818

**Selectivity of Subunit Association in the Assembly of  
Adhesive Type 1 Pili from *Escherichia coli***

A dissertation submitted to

**ETH ZURICH**

for the degree of

**Doctor of Sciences**

presented by

**Florian Sven Nico Dworkowski**

M.Sc., The University of Oklahoma

Born on March 31<sup>st</sup>, 1978

citizen of Germany

accepted on the recommendation of

Prof. Dr. Rudi Glockshuber

Prof. Dr. Wolf-Dietrich Hardt

2010



# Table of Content

<b>TABLE OF CONTENT</b>	<b>I</b>
<b>FIGURES</b>	<b>IV</b>
<b>ABBREVIATIONS</b>	<b>VI</b>
<b>1 ABSTRACT</b>	<b>1</b>
<b>2 ZUSAMMENFASSUNG</b>	<b>3</b>
<b>3 INTRODUCTION</b>	<b>5</b>
<b>3.1 Uropathogenic bacteria</b>	<b>5</b>
3.1.1 Bacterial surface organelles	6
3.1.2 Type 1 and P-type pili	7
3.1.3 The Chaperone/Usher pathway	8
3.1.4 Donor-strand complementation	12
<b>3.2 Protein folding</b>	<b>17</b>
3.2.1 Thermodynamic and kinetic stability of proteins	19
3.2.2 Two-state model in protein folding	21
<b>3.3 Protein-protein interactions</b>	<b>25</b>
<b>3.4 Macromolecular protein assemblies</b>	<b>26</b>
<b>3.5 Aim of this study</b>	<b>27</b>
<b>4 RESULTS</b>	<b>29</b>
<b>4.1 Overview of proteins and protein variants used throughout this study</b>	<b>29</b>
<b>4.2 Complex formation between subunit and ds peptide</b>	<b>31</b>
4.2.1 Formation of subunit:donor-strand complexes depends on peptide concentration	31
4.2.2 Fim subunit proteins form complexes with designated and non-designated ds peptides	32
<b>4.3 Kinetic stability of subunit:ds complexes against unfolding</b>	<b>33</b>

<b>4.4</b>	<b>Kinetics of subunit:ds peptide complex formation</b>	<b>37</b>
4.4.1	Interrupted refolding experiments	37
4.4.2	Formation of Subunit:ds complexes during refolding depends on the concentration of the ds peptide present	39
<b>4.5</b>	<b>Qualitative folding of FimG<sub>t</sub> in the presence of ds peptides</b>	<b>40</b>
4.5.1	Folding traces of FimG <sub>t</sub> :ds complexes	41
<b>4.6</b>	<b>Displacement of chaperone from chaperone:subunit complexes by ds peptides</b>	<b>42</b>
4.6.1	Donor-strand exchange	42
<b>4.7</b>	<b>Influence of subunit surface on DSE reaction</b>	<b>50</b>
<b>4.8</b>	<b>Competitive ds peptide binding to subunit proteins</b>	<b>52</b>
<b>4.9</b>	<b>Structural approach to determine the specificity of the usher for different chaperone:subunit complexes</b>	<b>54</b>
4.9.1	Crystallization of binary subunit:ds complexes	54
4.9.2	Crystallization of ternary complexes	54
<b>5</b>	<b>SUMMARY AND DISCUSSION</b>	<b>63</b>
<b>5.1</b>	<b>Natural subunit:ds peptide complexes are kinetically most stable against dissociation</b>	<b>63</b>
<b>5.2</b>	<b>Natural subunit:ds peptide complexes are formed fastest</b>	<b>63</b>
<b>5.3</b>	<b>Binding of ds peptides induces structural changes in subunit</b>	<b>65</b>
<b>5.4</b>	<b>Surface of trailing subunit does not influence binding of ds</b>	<b>65</b>
<b>5.5</b>	<b>Donor-strand complementation is not solely responsible for correct pilus formation</b>	<b>66</b>
<b>5.6</b>	<b>Subunits bind ds peptides with different attractive forces</b>	<b>67</b>
<b>5.7</b>	<b>Implications of the structure of the ternary complex between FimC:FimF<sub>t</sub>:FimD<sub>N</sub></b>	<b>70</b>
<b>6</b>	<b>OUTLOOK</b>	<b>72</b>
<b>7</b>	<b>MATERIALS AND METHODS</b>	<b>73</b>
<b>7.1</b>	<b>Materials</b>	<b>73</b>
7.1.1	Chromatographic columns and media	73
7.1.2	Chemicals	73
7.1.3	Oligonucleotides	74

---

7.1.4	Peptides	75
7.1.5	Bacterial strains	75
<b>7.2</b>	<b>Methods</b>	<b>76</b>
7.2.1	Properties of proteins used in this study	76
7.2.2	Design of expression vectors	76
7.2.3	Expression and purification of proteins	80
7.2.4	Generation of FimC:FimA <sub>t</sub> complex	87
7.2.5	Protein and peptide concentrations	88
7.2.6	Crystallization of FimC:FimF <sub>t</sub> :FimD <sub>N</sub> <sup>(1-125)</sup> [CFD] complex	88
7.2.7	Production of subunit:ds complexes	89
7.2.8	Unfolding kinetics	89
7.2.9	Quantitative unfolding of subunit:ds peptide complexes	90
7.2.10	ds peptide competition experiments	91
7.2.11	Interrupted refolding experiments	91
7.2.12	Chaperone displacement from chaperone:subunit complexes by ds peptides (DSE)	92
7.2.13	Mass spectroscopy of subunit:ds peptide complexes	93
7.2.14	Qualitative refolding of FimG <sub>t</sub>	94
7.2.15	pH dependence of subunit:ds peptide complex stability	94
<b>8</b>	<b>APPENDIX</b>	<b>95</b>
<b>8.1</b>	<b>Additional results</b>	<b>95</b>
8.1.1	Comparison of IPTG-induced and auto-induced protein expression	95
8.1.2	Comparison of cytosolic vs periplasmic expression of FimG <sub>t</sub>	96
8.1.3	MALDI-TOF MS of subunit:ds complex mixtures	96
8.1.4	ESI-TOF MS of subunit:ds mixtures	98
8.1.5	Dependence of unfolding rates of FimG <sub>t</sub> :dsF and FimG <sub>t</sub> :dsG on the pH	99
	<b>REFERENCES</b>	<b>101</b>
	<b>DANKSAGUNG</b>	<b>111</b>
	<b>CURRICULUM VITAE</b>	<b>112</b>

## Figures

Figure 3-1: Electron micrograph of <i>E. coli</i> strain W3110 showing extracellular appendices. ....	5
Figure 3-2: Schematics of type 1 pilus formation through the chaperone/usher pathway in <i>Escherichia coli</i> . ....	9
Figure 3-3: Structure of the outer membrane domain of the usher protein PapC, PapC <sub>130-640</sub> . ....	11
Figure 3-4: Schematics of the donor-strand complementation mechanism. ....	13
Figure 3-5: Minimal mechanism of subunit incorporation at the outer-membrane assembly platform. ....	14
Figure 3-6: The <i>fim</i> -Operon. ....	15
Figure 3-7: Representation of the protein folding energy funnel theory. ....	18
Figure 3-8: Example of a Chevron plot. ....	23
Figure 3-9: Schematic representation of the interrupted refolding assay. ....	24
Figure 4-1: Concentration dependence of the formation of FimG <sub>t</sub> :dsG and FimG <sub>t</sub> :dsF complexes during refolding. ....	32
Figure 4-2: Chevron plot of GdmCl dependent unfolding of FimG <sub>t</sub> :ds complexes. ....	34
Figure 4-3: Chevron plot of GdmCl dependent unfolding of FimF <sub>t</sub> :ds complexes. ....	36
Figure 4-4: Interrupted Refolding Experiments of FimG <sub>t</sub> :ds and FimF <sub>t</sub> :ds Complexes. ....	38
Figure 4-5: Concentration dependence of FimG <sub>t</sub> :dsF refolding kinetics. ....	40
Figure 4-6: Tryptophan fluorescence refolding traces of various FimG <sub>t</sub> :ds species. ....	41
Figure 4-7: Proposed mechanism of subunit:ds complex formation during refolding. ....	42
Figure 4-8: Time-dependence of DSE as resolved by analytical ion exchange chromatography. ....	43
Figure 4-9: DSE for FimC <sub>his</sub> :FimG <sub>t</sub> complex versus ds peptides as measured by analytical ion exchange chromatography. ....	44
Figure 4-10: DSE for FimC <sub>his</sub> :FimF <sub>t</sub> complex versus ds peptides as measured by analytical ion exchange chromatography. ....	45
Figure 4-11: Model of FimF aligned with dsA peptide in different orientations. ....	47
Figure 4-12: Purification of FimF <sub>t</sub> :dsA by anion exchange chromatography. ....	48
Figure 4-13: DSE reactions of FimC <sub>his</sub> :FimH <sub>p</sub> and FimC <sub>his</sub> :FimA <sub>t</sub> complexes vs. ds peptides. ....	49
Figure 4-14: Schematic representation of the surface influence assay. ....	50
Figure 4-15: Influence of subunit surface on the donor-strand exchange reaction. ....	51
Figure 4-16: Unfolding trace of FimG <sub>t</sub> :ds <sub>x</sub> complexes. ....	53
Figure 4-17: Crystallization and crystal formation of the FimC:FimF <sub>t</sub> :FimD <sub>N</sub> complex. ....	55
Figure 4-18: Crystal structure of CFD complex. ....	57
Figure 4-19: Stereo picture of the superimposition of FimC:FimF <sub>t</sub> :FimD <sub>N</sub> complex (CFD) with FimC:FimH <sub>p</sub> :FimD <sub>N</sub> complex (CHD). ....	58
Figure 4-20: Analytical size-exclusion run of FimC:FimA <sub>t</sub> :FimD <sub>N</sub> . ....	60
Figure 4-21: Model of possible artificial disulphide bonds in FimC:FimA <sub>t</sub> :FimD <sub>N</sub> (1-125). ....	61
Figure 5-1: Interactions between ds peptide and subunit residues in FimG <sub>t</sub> :dsF and FimG <sub>t</sub> :dsG complexes. ....	69
Figure 5-2: Interaction of small molecular inhibitors with chaperone proteins. ....	71
Figure 7-1: Principle of the PHUSION™ Site-Directed Mutagenesis (© Finnzymes). ....	78
Figure 8-1: Comparison of manual- and auto-induction. ....	95

---

<i>Figure 8-2: Comparison of FimG<sub>T</sub> purified from cytosol and periplasm: Kinetics of formation of the FimG<sub>T</sub>:dsF complex.</i>	96
<i>Figure 8-3: MALDI-TOF MS of FimG<sub>T</sub>:ds and FimF<sub>T</sub>:ds mixture.</i>	97
<i>Figure 8-4: ESI-MS spectra of FimG<sub>T</sub>:ds complexes.</i>	98
<i>Figure 8-5: Figure pH dependence of the unfolding rates of the FimG<sub>T</sub>:dsF and FimG<sub>T</sub>:dsG complexes.</i>	100
<i>Table 4-1: Overview over the Fim protein wild type subunits and their derivatives.</i>	30
<i>Table 4-2: Kinetic parameters of subunit:ds complex unfolding.</i>	37
<i>Table 4-3: Parameters of the subunit vs. ds peptide DSE reactions.</i>	46
<i>Table 4-4: Data collection statistics for FimC:FimF<sub>T</sub>:FimD<sub>N</sub> protein complex.</i>	55
<i>Table 4-5: Data refinement statistics for FimC:FimF<sub>T</sub>:FimD<sub>N</sub> protein complex.</i>	56
<i>Table 7-1: Oligonucleotide primers used in this study.</i>	74
<i>Table 7-2: Properties of donor-strand peptides.</i>	75
<i>Table 7-3: Properties of proteins and protein complexes used in this study.</i>	76

## Abbreviations

The following abbreviations are used throughout this study:

6xHis	Hexa Histidine-Tag	IPTG	Isopropyl- $\beta$ -D-thiogalactopyranosid
ASA	Accessible Surface Area	ITC	Isothermal Titration Calorimetry
CD	Circular Dichroism Spectroscopy	MAC	Metal Affinity Chromatography
CFD	FimC:FimF <sub>t</sub> :FimD <sub>N</sub> (1-125)	MALDI	Matrix Assisted Laser Desorption/Ionization
CHD	FimC:FimH <sub>p</sub> :FimD <sub>N</sub> (1-125)	MES	Morpholinoethane-sulfonic acid
CV	Column Volume	MS	Mass-Spectroscopy
DTT	Dithiothreitol	MOPS	3-morpholino-propanesulfonic acid
ds	Donor-Strand	Ni-NTA	(Ni <sup>2+</sup> )-charged nitrilotriacetic acid
DSC	Donor-Strand Complementation	NMR	Nuclear Magnetic Resonance Spectroscopy
DSE	Donor-Strand Exchange	o/n	over night
ddH <sub>2</sub> O	Double-distilled/deionized water	PAGE	Polyacrylamide Gel Electrophoresis
<i>E. coli</i>	Escherichia coli	Pap/pap	Pyelonephritis-associated pili / P-pili
EDTA	Ethylene diamine tetra-acetic acid	PBS	Phosphate Buffered Saline
FimA <sub>t</sub>	N-terminally truncated FimA	PDB	Protein database
FimD <sub>N</sub>	N-terminal domain of FimD	PEG	Polyethylene glycol
FimF <sub>t</sub>	N-terminally truncated FimF	SDS	Sodium Dodecyl Sulfate
FimG <sub>t</sub>	N-terminally truncated FimG	SEC	Size Exclusion Chromatography (see also <i>GF</i> )
FimH <sub>p</sub>	Pilin-domain of FimH	Tris	Tris-hydroxymethyl-aminomethane
FPLC	Fast Protein Liquid Chromatography	TOF	Time-of-Flight
$\Delta G^{0\ddagger}$	Gibbs Free Energy	TS	Transition State
$\Delta\Delta G^{0\ddagger}$	Change in Gibbs Free Energy	<i>wt</i>	Wild Type
GF	Gel-Filtration (see also <i>SEC</i> )		
GdmCl	Guanidin Hydrochlorid		
HIC	Hydrophobic Interaction Chromatography		
HPLC	High Performance Liquid Chromatography		
HT	High throughput		
IEX	Ion Exchange Chromatography		
Ig	Immunglobulin		



# 1 Abstract

Many pathogenic, gram-negative bacterial cells exhibit outer membrane attached macromolecular, filamentous surface organelles, called pili or fimbriae. Pili are crucial for cell adhesion to the host during infection and thus are considered a veritable drug target. Pili are highly oligomeric protein structures assembled via a common mechanism called the chaperon/usher pathway. Unfolded subunit proteins are captured by a folding chaperone which also mediates transfer of the subunit to an outer membrane assembly platform, the usher, where the subunits are incorporated into the growing fimbriae and translocated through a pore in the usher into the extra-cellular space where it adopts its final quaternary structure. The interaction between the subunits in the mature pilus is based on an N-terminal extension of the adjacent subunit acting in completing an incomplete Ig-like fold of the preceding subunit. This “donor-strand complementation” provides some of the most stable protein:protein interactions known to date. Another striking feature of the pili is that the various subunit proteins are incorporated into the organelle in a defined, strict order.

This thesis concerns the question of how this stringent order of pilus formation is achieved. As a model system the type 1 pilus system, involved in the infection of the human urinary tract by uropathogenic *Escherichia coli* strains, was used. Type 1 pili consist of up to three thousand main structural subunits, called FimA, comprising the helical pilus rod. The adhesin FimH at the distal end of the fibrillum is linked to the rod via the linker proteins FimG and FimF. Except for the adhesin, which has a receptor binding and pilin domain, all subunits are single-domain proteins exhibiting an incomplete Ig-like fold. The chaperone protein in the type 1 pilus system is FimC and the outer-membrane usher is called FimD. Recently another subunit, FimI, responsible for termination of pilus growth, was identified.

By investigation of the kinetic stability of various subunit:donor-strand (ds) complexes we could show that the natural subunit:ds combinations were most stable, albeit the subunits also formed stable complexes with other ds peptides. The stability against unfolding of the natural subunit:ds complexes was so high that it can be assumed these complexes do not dissociate once formed. The assessment of the rates of complex formation yielded similar results, the natural complexes were formed fastest, but other complexes could form as well. From the kinetic data obtained on the different subunit:ds interactions, the probability of spontaneous formation of a mature type 1 pilus, consisting of 500 FimA subunits and one copy of FimH,

FimG, FimF and FimI, with the correct subunit order was calculated to be  $1 : 5.2 \cdot 10^{292}$ , assuming identical initial subunit concentrations. Consequently, a very efficient proofreading system exists that governs quaternary structure formation in type 1 pili.

Indeed this task is fulfilled by the outer-membrane usher. The N-terminal, periplasmic domain of the usher, FimD<sub>N</sub>, was previously shown to bind to the FimC:subunit complexes and the ternary complex FimC:FimH<sub>p</sub>:FimD<sub>N</sub> was crystallized. It was postulated that FimD<sub>N</sub> distinguishes between chaperone bound subunits by binding each FimC:subunit complex with different affinity. To investigate this further, the ternary complex FimC:FimF<sub>t</sub>:FimD<sub>N</sub> was crystallized. The X-ray structure of the complex confirmed that FimD<sub>N</sub> discriminates between different FimC-bound subunits through specific contacts via its N-terminal tail (residues 1-24 of FimD), that adopts a defined structure in the ternary complexes, but is flexibly disordered in free FimD<sub>N</sub>. However, the affinities of FimD<sub>N</sub> for the different chaperone:subunit complexes do not differ by more than 40-fold. Consequently the formation of correct subunit:subunit contacts is mainly determined by the ability of the usher FimD to selectively catalyze formation of correct subunit:donor-strand pairs.

## 2 Zusammenfassung

Viele krankheitserregende, gram-negative Bakterien besitzen an der Oberfläche ihrer äusseren Membran fadenförmige, makromolekulare Oberflächenstrukturen, so genannte Pili oder Fimbriae. Diese Pili sind essentiell für die Zell-Adhäsion an das Gewebe des Wirts bei Infektionen und gelten daher als vielversprechendes Ziel im drug-targeting. Pili sind hetero-oligomere Proteinstrukturen, welche mittels des so genannten Chaperone/Usher-Pathway gebildet werden. Ungefaltete Untereinheitenproteine werden von einem Faltungshelferprotein, dem Chaperone, gebunden und zur Assemblierungsplattform in der äusseren Membran transportiert, dem so genannten Usher, wo diese in den wachsenden Pilus integriert und aus der Zelle transportiert werden, wo sie ihre endgültige Struktur annehmen. Die Interaktion zwischen den Untereinheitenproteinen im ausgewachsenen Pilus wird dabei durch die Aufnahme eines verlängerten Aminosäurenstranges am N-terminalen Ende der einzuarbeitenden Untereinheit in die vorhergehende Untereinheit erreicht. Dabei wird die unvollständige, Ig-ähnliche Struktur des Vorgängermoleküls von der nachfolgenden Untereinheit vervollständigt. Diese „donor-strand complementation“ gilt als eine der stärksten bekannten Protein-Protein Wechselwirkung.

Pili besitzen ausserdem die bemerkenswerte Eigenschaft, dass die verschiedenen Untereinheitenmoleküle in einer ganz bestimmten, strikten Reihenfolge in die Struktur eingearbeitet werden.

In dieser Dissertation wird die Frage adressiert, wie diese strikte Ordnung erlangt wird. Als Modellsystem dient das Type 1 Pilus System, welches eine wichtige Rolle in der Infektion von menschlichen Harnwegen durch uropathogene *Escherichia coli*-Stämme spielt. Type 1 Pili bestehen aus bis zu dreitausend Einheiten der hauptstrukturellen Untereinheit FimA, welche den helixförmigen Pilus-Stamm bilden. Das Adhäsion FimH ist über die Verbindungproteine FimG und FimA mit dem Stamm verbunden. Ausser dem Adhäsion, welches sowohl eine Rezeptorbindungs- als auch eine Pilin-Domäne besitzt, bestehen alle Untereinheitenproteine aus einer einzelnen Domäne und besitzen eine unvollständige Ig-ähnliche Faltung. Im Type 1 System wirkt das Protein FimC als Chaperone und das Protein FimD fungiert als Usher in der äusseren Membran. Kürzlich wurde auch ein weiteres Untereinheitenprotein, FimI, beschrieben, welches das Piluswachstum terminiert.

Durch die Untersuchung der kinetischen Stabilität verschiedener Untereinheiten:Donor-strand (ds)-Komplexe waren wir in der Lage zu zeigen, dass die intrinsischen Komplexe die höchste Stabilität aufwiesen. Allerdings bildeten die Untereinheitenproteine auch Komplexe mit nicht vorgesehenen Bindungspartnern. Die Entfaltungsresistenz der natürlichen Komplexe ist dabei so hoch, dass diese nicht mehr dissoziieren wenn sie sich einmal gebildet haben.

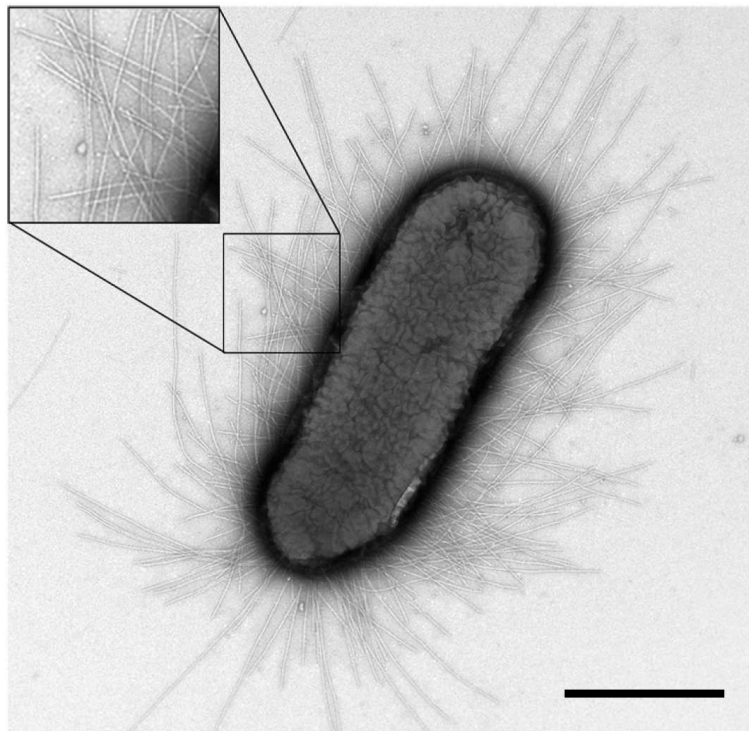
Die Erhebung der Raten der Bildung der Komplexe ergab ähnliche Ergebnisse. Die natürlichen Komplexe bildeten sich am schnellsten, allerdings wurden auch die nicht-natürlichen Komplexe gebildet. Mit Hilfe der kinetischen Daten, die für verschiedenen Untereinheiten:Donor-strand Komplexe bestimmt wurden, ist es möglich die Wahrscheinlichkeit für die spontane Ausbildung eines vollständigen Pilus mit der korrekten Sequenz an Untereinheiten, bestehend aus 500 FimA-Untereinheiten und je einer Untereinheit FimH, FimG, FimF and FimI, zu berechnen. Unter der Annahme, dass kein bisher noch unbekannter, zusätzlicher Selektionsmechanismus vorhanden ist, wurde diese Wahrscheinlichkeit, unter der Voraussetzung, dass alle Untereinheitenproteine in gleicher Ausgangskonzentration vorliegen, als  $1:5.2 \cdot 10^{292}$  bestimmt. Dies bedeutet im Wesentlichen, dass ohne ein weiteres Kontrollsystem keine Pili gebildet werden.

Diese Aufgabe erfüllt das Usherprotein in der äusseren Membran. Es wurde zuvor gezeigt, dass die N-terminale, periplasmatische Domäne des Ushers in der Lage ist die Chaperone:Untereinheiten-Komplexe zu binden und der ternäre Komplex FimC:FimH<sub>p</sub>:FimD<sub>N</sub> wurde kristallisiert. Es wurde postuliert, dass FimD<sub>N</sub> gezielt Chaperone:Untereinheiten-Komplexe unterscheidet, indem es die Komplexe mit unterschiedlicher Affinität bindet. Um dies zu verifizieren wurde der ternäre Komplex FimC:FimF<sub>t</sub>:FimD<sub>N</sub> kristallisiert. Die Röntgenstruktur dieses Komplexes bestätigte das FimD<sub>N</sub> die verschiedenen FimC-gebundenen Untereinheiten durch spezifische Kontakte seiner 24 N-terminalen Aminosäuren, welche nur im ternären Komplex eine definierte Struktur aufweisen und in freiem FimD<sub>N</sub> ungeordnet sind, auswählen kann. Allerdings weichen die Affinitäten von FimD<sub>N</sub> für verschiedene Chaperone:Untereinheiten-Komplexe nur um maximal das 40-fache voneinander ab. Hieraus folgt, dass die korrekten Untereinheit:Untereinheit-Kontakte hauptsächlich durch die Fähigkeit des Ushers, selektiv die Bildung von korrekten Untereinheiten:Donor-strand-Paaren zu katalysieren, bestimmt wird.

### 3 Introduction

#### 3.1 Uropathogenic bacteria

Urinary tract infections (UTI) are one of the major infections in humans caused by bacterial pathogens. An estimated 50% of women are believed to develop at least one urinary tract infection during their life span; 20-40% will suffer from reoccurring infections (Stanton and Dwyer, 2000). In 1999, uncomplicated UTIs were accounted for 46 million prescriptions in Western Europe and Canada alone, making this disease a major financial burden for the public healthcare systems (Naber, 2000).



**Figure 3-1: Electron micrograph of *E. coli* strain W3110 showing extracellular appendices.**

The cell was negatively stained with sodium phosphotungstate and imaged by TEM. The scale bar in the lower right corner represents a length of 1 $\mu$ m (from Hahn *et al.*, 2002).

The major causes of those infections are uropathogenic *Escherichia coli* (UPEC) strains (Svanborg and Godaly, 1997). These bacteria infect the urinary tract via outer-membrane attached adhesive surface organelles, so called pili or fimbriae, which allow the bacteria to selectively adhere to the respective host tissue (*Figure 3-1*, Brinton, 1959; Reid and Sobel, 1987). There is no conclusive explanation on how the UPEC enter the urinary tract, but it is assumed that sexual contact and transfer from fecal flora, especially into the female urethra,

are major causes for initial infection. Once the bacteria enter the host's urinary tract they use, among others, so called type 1 pili to attach to the Uroplakin Ia receptor on the urinary tract epithelium cells, forming fairly large biofilms in the process (Connell *et al.*, 1996; Zhou *et al.*, 2001). Using exocytic vesicles, they invade host cells, thus shielding themselves from the rather strong immune responses of the host, including strong sheer forces from urination, exfoliation of the epithelium layer and inflammatory immune responses (Bishop *et al.*, 2007; Mulvey *et al.*, 2001). Recently it was also shown that the UPEC can form large intracellular bacterial communities, so called pods, shielding themselves to the outside and escaping the immune response. It is believed that this might be a reason for the establishment of chronic UTIs and reoccurring infections even when no UPEC could be detected any more after treatment (Anderson *et al.*, 2003).

Understanding function and biogenesis of the adhesive surface organelles of the UPEC strains, such as type 1 and P type pili, might therefore improve the possibility for development of new treatments of UTI.

### **3.1.1 Bacterial surface organelles**

Gram-negative bacteria exhibit a vast array of different surface organelles on their outer membrane, most of which are hetero-multimeric structures (*Figure 3-1*). Their functions vary from facilitating motility and chemotaxis (e.g. type IV pili, flagella), secretion of virulence factors (e.g. flagella) over bacterial conjugation (F pili) to adherence to host-tissue (adhesive pili) (Fernandez and Berenguer, 2000; Ippen-Ihler and Minkley, 1986; Young *et al.*, 1999).

Adhesive pili, or *Fimbriae*, play a crucial role during infection by pathogenic bacterial strains. They determine specificity for a certain host-tissue and facilitate biofilm formation. *Fimbriae* can be subdivided into four major classes, depending on the mechanism utilized for their assembly. The first class consists of pili assembled by the so called Chaperone/Usher pathway, e.g. type 1 pili. Here, the unfolded structural protein is secreted into the cells periplasm where it is captured by a periplasmic chaperone protein, which facilitates folding and prevents premature aggregation of pilus components while transporting the newly folded pilus protein to an outer membrane assembly platform, termed the Usher. The protein is secreted into the extracellular space and incorporated in the growing pilus (Sauer *et al.*, 2000). The Chaperone/Usher pathway is very well investigated and utilized by a majority of pathogenic bacteria. It will be discussed in detail later on (see 3.1.3).

The second class of pili is assembled via a very similar mechanism also depending on a bi-molecular assembly platform/assembly factor system, but the components do not share any sequence similarity to the Chaperon/Usher pathway. CS1, CS2 and colonization factor antigen I (CFA/I) fimbriae, involved in enterotoxigenicity of enterotoxigenic *Escherichia coli* strains (ETEC), belong to this category (Smyth *et al.*, 1996).

In 1989, Olsén *et al.* described a new type of adhesive surface structure. Those organelles were termed *curli* because of their thin, coiled appearance. Curli are able to form lateral aggregates, thus forming significantly thicker structures. They are composed of a single type of subunit protein encoded by the gene *crl* which is found in most *E.coli* isolates but only certain strains were found to be able to process the gene-product into curli. (Olsen *et al.*, 1989). It was later proposed that the curli monomers are secreted to the cell surface where they are assembled into fibers by assembly proteins (Chapman *et al.*, 2002).

The bundle-forming type IV pili are formed via a fourth pathway. This pathway utilizes some 14 different assembly proteins, many of which are homologues to components of the Type II secretion machinery. Assembly of pili happens through a multistep mechanism, with the general steps of inter-membrane core-complex formation, recruitment of pilus-associated proteins in a trans-envelope manner, followed by recruitment of an ATPase and finally a potential bifurcation of the pathway, supported by the finding that some type IV pili like *A. tumefaciens* VirB/D4 pili can not only assemble as a classical pilus but also form a secretion channel (Christie *et al.*, 2005; Peabody *et al.*, 2003).

### 3.1.2 Type 1 and P-type pili

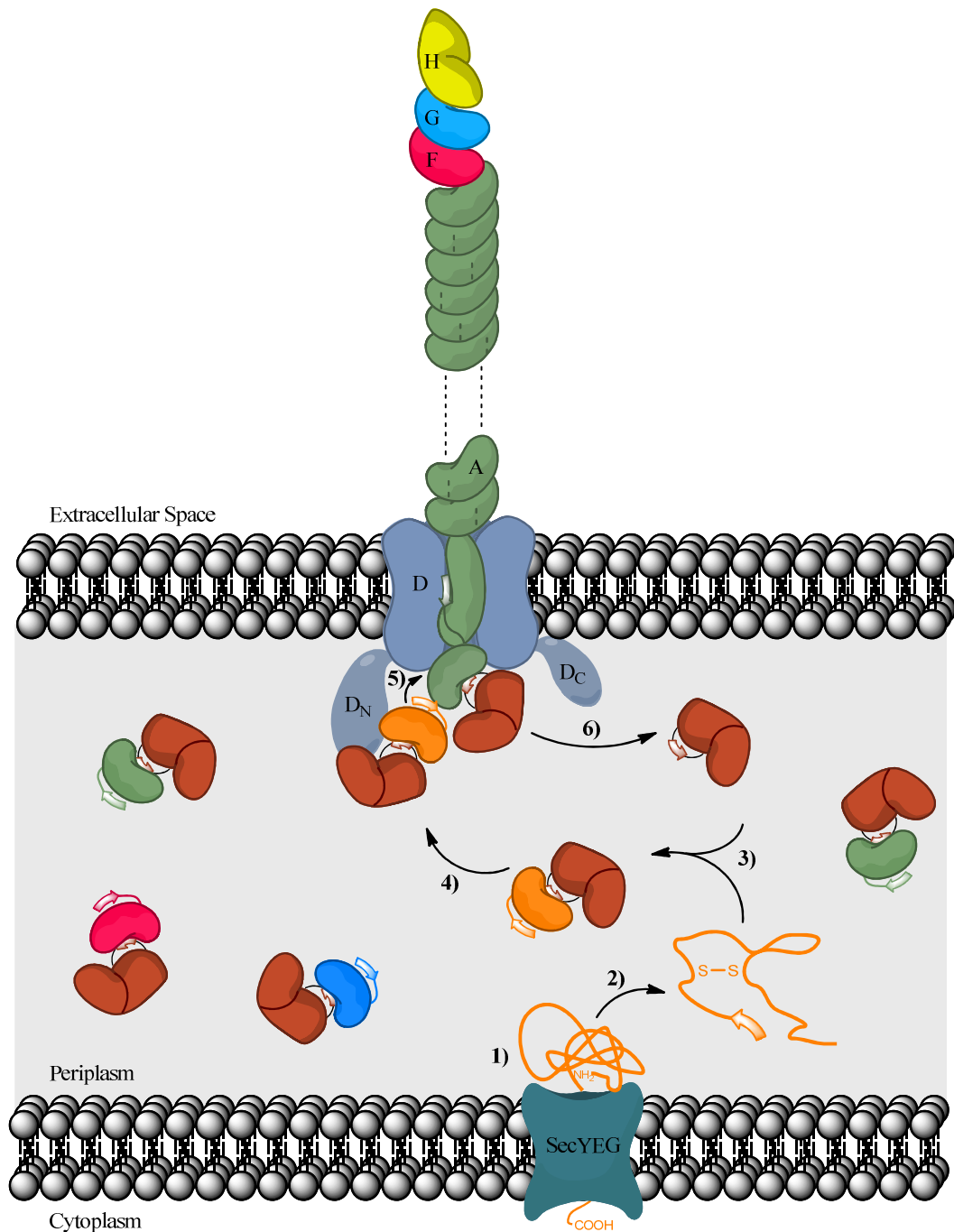
Even though UPEC strains can express some 30+ different surface organelles, the most common ones associated with UTIs are P type pili (encoded by the *pap* gene cluster) and type 1 pili (encoded by the *fim* gene cluster, *Figure 3-2*). P type pili are responsible for adherence of the bacteria to the gal $\alpha$ 1-4gal units of human kidney glycolipids and type 1 pili bind to the mannose moieties of the human bladder protein uroplakin Ia (Lund *et al.*, 1987; Zhou *et al.*, 2001). Both kinds of pili are very similar in terms of structure, consisting of a homomultimeric rod, forming a right handed helix with a diameter of about 6.9Å and a length of up to 2µm and a fibrillar tip consisting of an adhesin, which is linked to the rod by a linker formed by a varying number of adaptor subunit proteins. In case of type 1 pili the rod is formed by 500-3000 units of the major subunit FimA. The tip fibrillum is formed by one or several copies of the adaptor subunits FimF and FimG and a single copy of the adhesin FimH

at the very tip of the assembly (see *Figure 3-2*). In the case of P type pili the major subunit is PapA. The P pilus tip fibrillum is formed predominantly by several copies of the adaptor subunit PapE, which are linked to the pilus rod by the minor subunit PapK and to the adhesin PapG by the adaptor subunit PapF (reviewed in Vetsch and Glockshuber, 2005).

### **3.1.3 The Chaperone/Usher pathway**

Most pili are secreted and assembled via the chaperone-usher pathway, a common system in gram-negative bacteria, as for example in *Salmonella enterica*, *Klebsiella pneumoniae*, *Haemophilus influenzae*, or *Pseudomonas aeruginosa* (*Figure 3-2*, reviewed in Sauer *et al.*, 2004; Thanassi *et al.*, 2005). In this pathway, the pilus subunit entering the bacterial periplasm through the SEC pathway is captured by a periplasmic chaperone, which serves several purposes. It not only facilitates the folding of the subunits (Barnhart *et al.*, 2000), but also caps the subunit and thus prevents it from degradation and premature aggregation in the periplasm (Bullitt *et al.*, 1996; Kuehn *et al.*, 1991; Soto *et al.*, 1998). It was also shown that the chaperone can act as a folding catalyst, significantly accelerating the folding of the oxidized subunit in the periplasm (Bann *et al.*, 2004; Vetsch *et al.*, 2004). Furthermore, it primes the bound subunit for incorporation into the assembling pilus (Sauer *et al.*, 2002) and mediates the transfer of the subunit to the outer-membrane assembly platform, the usher (Lindberg *et al.*, 1989).





**Figure 3-2: Schematics of type 1 pilus formation through the chaperone/usher pathway in *Escherichia coli*.**

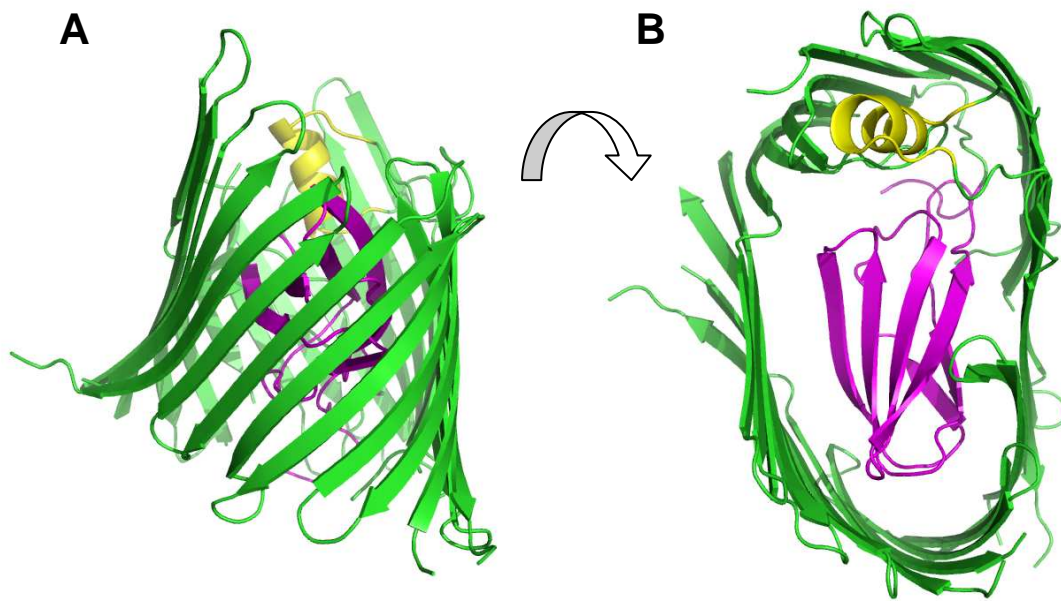
(A) The unfolded subunit proteins are secreted into the periplasm by the SecYEG secretory pathway (1). The disulphide bond present in all subunits is formed by cysteine oxidation (2) and the protein is captured by the periplasmic chaperone FimC (3). The chaperone:subunit complex is transported through the periplasm to the outer membrane assembly platform FimD (4) where the complex is detected and selected by the N-terminal extension of FimD, FimD<sub>N</sub>, and the correct subunit is incorporated into the growing pilus by donor-strand exchange (5), replacing the preceding chaperone from the accepting subunit in the process (6). The incorporated subunits are

translocated through the outer-membrane usher in a linear manner and form a helical structure on the outside of the cell.

The usher protein forms a pore of about 2 nm in diameter in the outer membrane, through which the subunits are translocated into the extracellular space (Li *et al.*, 2004; Saulino *et al.*, 1998). Since this pore is too small to accommodate a whole pilus rod (~7 nm in diameter), the subunits have to be translocated in a linear chain, adopting their final quaternary structure only in the extracellular space. The usher is likely to be also involved in the anchoring of the assembled pilus to the outer membrane (also see 3.1.4.3 below). The chaperones in the P type and type 1 pilus systems are PapD and FimC, respectively, and the usher is termed PapC for P pili and FimD for type 1 pili.

By *in vitro* reconstitution of type 1 pilus assembly it was shown that the usher also significantly accelerates the incorporation of the main structural subunit FimA into the growing pilus, revealing the usher to be a catalyst for pilus formation (Nishiyama *et al.*, 2008). A recent study utilizing the same approach in the P type system came to the same conclusion (Huang *et al.*, 2009).

For the type 1 pili usher FimD, Capitani *et al.* predicted four different domains using bioinformatic analysis (Capitani *et al.*, 2006b). Besides the already known N-terminal periplasmic domain (FimD<sub>N</sub>, residues 1-137) and the 22 strand  $\beta$ -barrel transmembrane domain they proposed the existence of two additional periplasmic domains, one at the C-terminus (FimD<sub>C</sub>, residues 658-833) and a third one interrupting the trans-membrane domain (FimD<sub>M</sub>, residues 222-327). At that time structural information was only available on the FimD<sub>N</sub> domain (Nishiyama *et al.*, 2005). In 2008, the first structural data on the transmembrane usher domain of PapC was published by Remaut *et al.* (Remaut *et al.*, 2008). In this study, a truncated version of the P type usher protein PapC, lacking the first 129 and last 169 amino acids, comprising a total of 511 amino acid residues, was structurally determined. They found the domain to be a 24 strand  $\beta$ -barrel with outer- and inner dimensions of 64Å by 45Å and 45Å by 25Å, respectively, with a height of approximately 45Å. The presence of a middle, periplasmic domain as proposed by Capitani *et al.* was also confirmed and was found to function as a plug for the trans-membrane pore (*Figure 3-3*). A single  $\alpha$ -helix at the extracellular end of the usher pore was described to act as a lateral anchor for the plug domain in its closed form.



**Figure 3-3: Structure of the outer membrane domain of the usher protein PapC, PapC<sub>130-640</sub>.**

(A) Lateral view on the outer membrane pore of PapC. The 24 stranded  $\beta$ -barrel pore is shown in green, the plug domain is magenta and the only  $\alpha$ -helix, believed to be involved in anchoring of the plug in its closed state, is colored yellow. (B) shows the same molecule turned by 90° and allowing a view into the pore from the extracellular side. (PDB Code 2vqi; adapted from Remaut *et al.*, 2008)

The structure of the transmembrane domain of PapC was modeled into the cryo-electron microscopy density map of FimD in complex with assembled Fim tip fibrillum together with the structures of all other complexes of Fim proteins known to be present at the usher during tip fibrillum assembly (FimC:FimH<sub>p</sub>:FimD<sub>N</sub>, PDB code 1ZE3, Nishiyama *et al.*, 2005; FimC:FimH, PDB code 1QUN, Choudhury *et al.*, 1999). Instead of a FimG structure the analogue structure from the Pap system, PapD:PapA:PapA (PDB code 2UY6, Verger *et al.*, 2007) was used as a model for FimC:FimF:FimG. Interestingly, this model was made using a twin-pore usher complex, because densitometric measurements of the purified Fim tip-fibrillum yielded molar ratios of 2:2:1:1:1 for its components FimD, FimC, FimH, FimG and FimF, respectively. Even though the authors concluded that FimC is present at only half the concentration determined due to different susceptibility to coomassie blue-staining, it is still assumed that the usher is present at a two-fold excess over the other subunits. In two different crystal forms of the protein a dimer along two-fold crystallographic axis was observed. Although the two protomers did not form any contact, it was assumed this was the

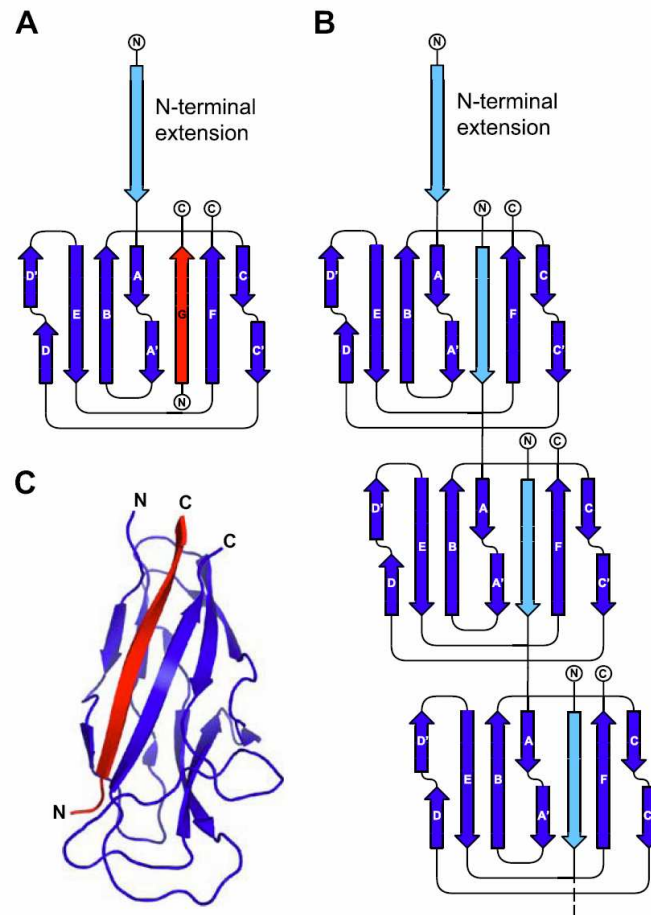
biologically functional assembly, despite the presence of another dimer interface in a head-to-head manner, which is not discussed in any detail.

Whether this twin-pore is really a biological relevant form or an experimental artifact has to be investigated in more detail, especially since the above described model reduces one of the two usher molecules to a very minor role in the assembly located at the second pore, which seems very unfavorable on the energy budget of the bacterial cell. Additionally, the authors agree on the fact that the resolution of the cryo-electron microscopy map was not high enough to place the N-terminal domain of the usher with certainty, rendering the other assignments somewhat questionable. Furthermore, functional studies on the co-expression of PapC and PapD by Thanassi and co-workers also showed that usher monomers seemed sufficient for pilus assembly *in vitro* (So and Thanassi, 2006; Thanassi *et al.*, 2002).

In 2009 another x-ray structure of a slightly larger PapC pore construct (residues 126-650) was published by Huang *et al.* (Huang *et al.*, 2009). This structure was derived from a different crystal form and did not exhibit lateral dimer formation of the usher along a 2-fold crystallographic axis, but rather showed a head-to-head dimerization of two usher molecules. This might indicate that, if the usher is present as a dimer in solution, it rather forms a head-to-head than a twin-pore. This interaction, if present *in vivo*, would only be possible to form in a inter-cellular manner and most likely would be intermitted through structural changes upon priming of the usher by the adhesin:chaperone complex.

### 3.1.4 Donor-strand complementation

In pilus systems, the subunit:subunit complexes form unusual strong interactions, withstanding even highly denaturing conditions. This strong interaction is achieved by an interaction termed donor-strand complementation (DSC). All pilus subunits are single-domain proteins, except for the adhesins, which consist of a pilin and a lectin domain, and display considerable structural homology. The pilus domains exhibit an incomplete immunoglobulin-like (Ig-like) fold of six  $\beta$ -strands, lacking the seventh  $\beta$ -strand that would complete the fold. During subunit capture, the periplasmic chaperone inserts an extended segment (termed donor-strand, ds) into the hydrophobic cleft caused by the missing  $\beta$ -strand in the subunit, and thus transiently becomes part of the subunits' tertiary structure (*Figure 3-4A,C*). It is noteworthy that the chaperones donor-strand always is incorporated in a parallel manner in respect to the F-strand of the subunit, which is in disagreement with an Ig-like fold.



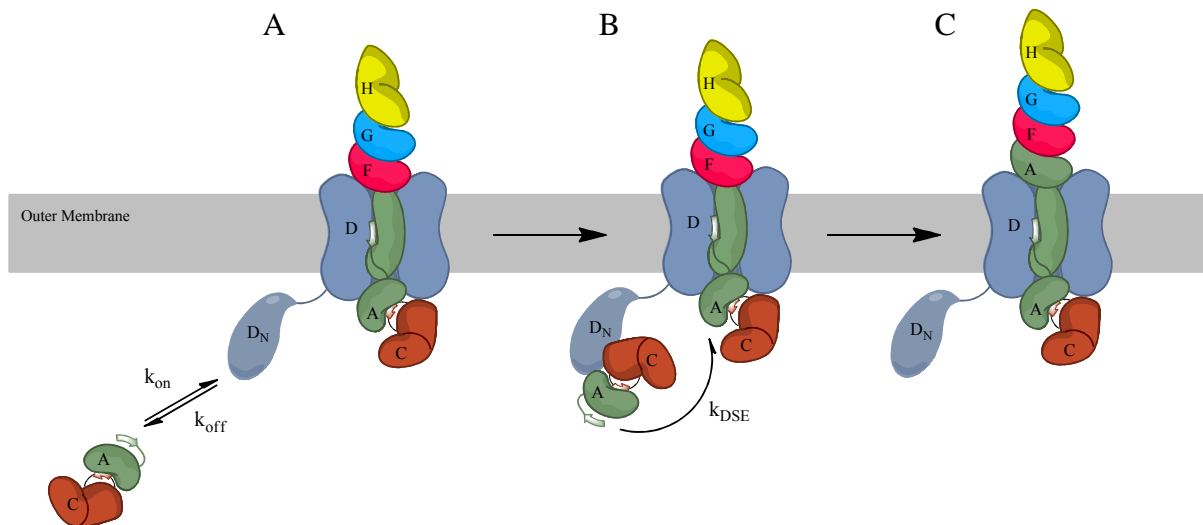
**Figure 3-4: Schematics of the donor-strand complementation mechanism.**

(A) shows the topological organization of a pilus subunit upon binding of the chaperone donor-strand. The incomplete immunoglobulin like fold ( $\beta$ -strands A to F) is completed by residues 101-110 from FimC and this donor strand becomes the missing G strand, which is oriented parallel to the F strand and anti-parallel to the A and A' strands. (B) When a subunit's donor-strand replaces the chaperone donor-strand during pilus biogenesis, the incorporated strand is oriented in the opposite orientation, parallel to strands A and A' and anti-parallel to the F strand. A linear chain of subunits is formed by this complementation mechanism. (C) shows the pilin domain of the adhesin FimH (blue, PDB Code 1QUN) in complex with the peptide corresponding to the donor-strand sequence of the chaperone FimC (red). (From Capitani *et al.*, 2006a)

Every subunit, except for the adhesin, contains a similar amino acid extension of about 15-20 residues at their N-terminal end. Upon assembly of a new subunit into the growing pilus at the outer membrane assembly platform, this donor-strand replaces the chaperone's donor strand in the preceding subunit in the quaternary structure of the pilus. In contrast to the interaction with the chaperone, the subunit's donor-strand is incorporated anti-parallel in respect to the F-strand, completing the Ig-like fold (*Figure 3-4B*). This mechanism is termed donor strand exchange (DSE) (Sauer *et al.*, 1999, reviewed in Vetsch and Glockshuber, 2005).

### 3.1.4.1 Pilus assembly

One striking feature in pilus assembly is the fact that the respective pilus subunits seem to be incorporated only in a defined order *in vivo*. In the case of type 1 pili this means that the adhesin FimH is always incorporated first, followed by the linker subunits FimG and FimF. Only then the major structural subunit FimA is integrated into the developing pilus (Saulino *et al.*, 2000). Since subunit incorporation occurs at the usher, the usher must be able to discriminate between the various chaperone:subunit complexes. Indeed, the various chaperone:subunit complexes were found to bind to the usher with different affinities, with the chaperone:adhesin complex exhibiting the strongest affinity towards the usher (Dodson *et al.*, 1993; Saulino *et al.*, 1998). Later it was shown that the N-terminal, periplasmic domain of the Usher, FimD<sub>N</sub>, selectively binds the chaperone:subunit complexes with  $K_D$  values in the range of 0.5 to 25  $\mu$ M (Nishiyama and Glockshuber, 2010; Nishiyama *et al.*, 2003). This suggests a minimal mechanism of subunit incorporation at the usher as depicted in *Figure 3-5*.



**Figure 3-5: Minimal mechanism of subunit incorporation at the outer-membrane assembly platform.**

(A) The chaperone-bound subunit transiently binds to the N-terminal domain of the usher with specific affinity. (B) The complex is brought into close proximity to the growing pilus where the virtually irreversible donor-strand exchange takes place. (C) One chaperone molecule is released and the subunit-chain is further translocated into the extra-cellular space.

Upon binding of the chaperone:subunit complex, it is brought into close proximity to the usher pore where the donor-strand exchange takes place. Since the interaction of the donor-

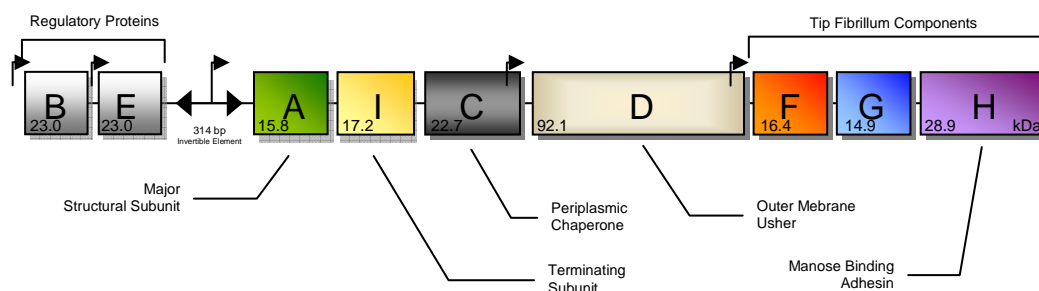
strand complemented subunits is so strong, this reaction can be assumed to be irreversible (Puorger *et al.*, 2008). Whether the Usher actively participates in this step, or if the DSE alone is able to account for the correct subunit incorporation, remains to be shown.

### 3.1.4.2 Regulation of type 1 pilus expression

The *fim* operon (Figure 3-6) contains nine genes. These are the genes *fimB* and *fimE*, encoding for the regulatory proteins FimB and FimE, followed by a 314bp invertible DNA fragment upstream of the *fimA* gene, containing the promoter for expression of the gene of the major structural subunit, as well as the gene *fimI*, encoding the molecular anchor protein, FimI, and *fimC* for the chaperone, FimC.

The gene encoding for the chaperone also contains the promoter for the expression of the gene *fimD*, encoding for the outer membrane usher, FimD, which, in turn, contains the promoter for the genes for the minor subunit proteins forming the pilus tip-fibrillum, *fimF*, *fimG* and *fimH*.

The regulatory proteins FimB and FimE are responsible for the phase-variable expression of the *fim* genes by switching the invertible element, and thus the promoter it contains, allowing expression of the main structural subunits (phase ON), resulting in piliation, or inhibiting the expression of those subunits (phase OFF), resulting in non-piliated cells (Blomfield, 2001). Here FimB mediates the phase variation both from phase-ON to phase-OFF and phase-OFF to phase-ON, while FimE predominantly changes the invertible element to the phase-OFF orientation (Abraham *et al.*, 1985; Klemm, 1986).



**Figure 3-6: The *fim*-Operon.**

The *fim*-operon consists of nine genes and one invertible element. Located on the upstream end of the invertible element, used in activity switching of the expression of the structural subunits, are the genes *fimB* and *fimE*, encoding for the regulatory subunits FimB and FimE. Downstream of the invertible element the genes *fimA*, *fimI*, *fimC* and *fimD* encoding for the main structural

subunit FimA, the putative molecular anchor FimI, the periplasmic chaperone FimC and the outer membrane assembly platform FimD. Finally, the genes *fimF*, *fimG* and *fimH* are encoding the tip fibrillum proteins FimF and FimG as well as the adhesin FimH, respectively. The numbers indicate the molecular mass of the respective mature proteins.

A typical type 1 fimbriated *E.coli* cell exhibits around 500 pili on its surface, each of which is composed of an average of 1000 subunit proteins. These roughly 500'000 protein units amount to about 8% of the cells total protein content.

Microarray gene expression studies by Schembri *et al.*, performed in *E.coli* K12 strain MG1655, showed that the genes of the *fim* cluster are expressed in a ratio roughly corresponding to the proposed composition of the mature pilus. The genes *fimA*, *fimI*, *fimC*, *fimD*, *fimF*, *fimG* and *fimH* were found to be expressed in ratios of 100, 28, 8, 9, 4, 10 and 5, respectively (Schembri *et al.*, 2002). Deletion of the *fimH* gene did not affect those ratios significantly, except for the lack of FimH. In a variant of the strain where the whole *fim*-operon was deleted, expression of Antigen43 (AG43), an outer membrane protein involved in inter-bacterial aggregation and together with type 1 pili accountable for biofilm formation *in vivo* (Hasman *et al.*, 2000; Roche *et al.*, 2001), was found to be up-regulated, whereas its expression was suppressed when *fim* was expressed.

Another study by Snyder *et al.* (Snyder *et al.*, 2004) compared the transcriptome of *E.coli* strain CFT073 grown in rich media to the same strain derived from the urine of infected mice. The *fimC* gene was found to be one of the 50 most up-regulated (~4.5-fold) genes in the cells taken from the UTI affected mice compared to the transcriptome of the cells derived from the *in vitro* culture. Also, *fimA* was found to be the fourth most expressed gene in the *in vivo* cells, only outdone by genes related to translation.

### 3.1.4.3 Termination of pilus biogenesis

Until recently it was not known how pilus growth is terminated in the type 1 system. In the P-pilus it was known that one of the minor subunits, PapH, is responsible for the termination of pilus growth by binding to the last unit of the main structural subunit, PapA, making up the helical pilus rod (Båga *et al.*, 1987; Verger *et al.*, 2006).

Recent work by Oleksander Igantov in our laboratory now showed that a similar mechanism is present in type 1 pili. Here the terminator is the minor subunit FimI. It terminates the growth of the pilus rod by irreversibly binding to the last incorporated main structural subunit



FimA on the periplasmic side of the usher pore thus preventing further incorporation of other subunits. It also functions as a molecular anchor, securing the mature pilus in the usher pore and thus in the outer membrane of the cell, as was shown by the accumulation of pili in the medium rather than at the bacterial surface when expression of FimI was suppressed (personal communication).

### **3.2 Protein folding**

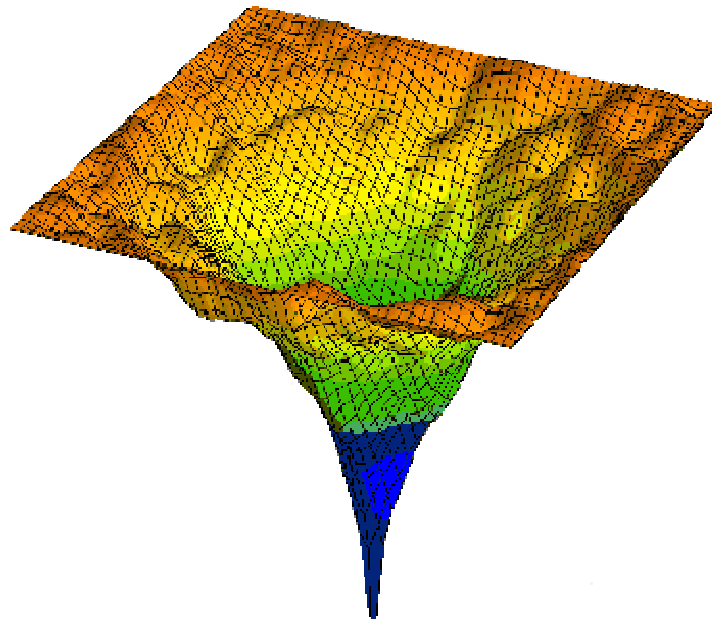
Correct protein folding is not only necessary for most proteins to be able to fulfill their biological functions but protein misfolding is also cause for numerous diseases in mammals and humans, including amyloid formation in Alzheimers disease (Finder and Glockshuber, 2007; Glenner, 1980).

Until the late sixties it was unknown how exactly proteins fold into their correct functional conformation. Only the experiments of Anfinsen and coworkers on the folding of Ribonuclease A *in vitro* led to the assumption that all information necessary for a protein to adopt its native and biological active tertiary structure is provided solely by its amino-acid sequence, since no cofactors or folding chaperones were present during the successful refolding experiments. Anfinsen postulated that the native state of a protein is identical with it's thermodynamically most stable conformation (Anfinsen, 1967; Anfinsen, 1973; Anfinsen and Scheraga, 1975; Haber and Anfinsen, 1962).

Around the same time, Cyrus Levinthal discussed the now famous "Levinthal Paradox", namely that due to the large number of degrees of freedom, spontaneous folding of a polypeptide chain toward a single specific conformation seemed highly unlikely. Would an average protein with a length of 101 amino acid sample all of its  $3^{300}$  (or  $10^{143}$ ) possible confirmations to obtain its native structure, protein folding would occur on a timescale longer than the age of the known universe and not on the milliseconds-to-minutes timescale observed *in vitro* and *in vivo* (Levinthal, 1968).

The apparent solution to this problem originates from the theory of the energy funnel in protein folding. This hypothesis suggests that proteins fold along a rough energy surface with local minima toward an energy minimized native state at the very bottom of the funnel-shaped energy landscape (*Figure 3-7*). Thus the protein can populate diverse states without having to adopt every structure possible, but still will fold towards a common native state. Since this energy funnel can have local minima from which the protein might not escape by itself, the necessity of some proteins for folding-helper proteins (chaperones and folding catalysts) to

adopt its native state can also be explained by this model (Dill and Chan, 1997; Wolynes *et al.*, 1995).



**Figure 3-7: Representation of the protein folding energy funnel theory.**

Shown is the rough energy landscape illustrating the solution to the Levinthal paradox. Proteins can fold using different pathways with local energy minima without the need to populate every possible conformation.

Major advances in the field of protein folding have been made since Anfinsen's famous postulate in 1975. Today, investigation of the "protein folding problem" is focused on answering the three predominant questions: (i) what is the mechanism of protein folding, (ii) can the native structure of a protein be predicted from its amino acid sequence and (iii) is there a "code" for protein folding (Dill *et al.*, 2008).

While there is still debate if there is a global mechanism for protein folding, the energy funnel theory is commonly accepted. However, the problem still remains of how a protein can find a route down this funnel in a quick and precise manner. Along diverse other models, most of which only focus either on the interpretation of experimental data or purely theoretical models, the so called "zipping and assembly" (ZA) theory was able to explain both theoretical as well as experimental protein folding data (Dill *et al.*, 2008). In this mechanism, it is assumed that small regions on the proteins amino acid chain fold independently toward a local structure, which in turn induces folding in previously unfolded regions of the chain. So the global folding problem is broken down into many local processes, thus raising the probability of

correct folding as well as the velocity of the reaction while minimizing the possible folding routes down the energy funnel (Dill *et al.*, 1993; Fiebig and Dill, 1993; Huang *et al.*, 2007).

The classical view on protein folding coined by Anfinsen was the assumption that there are many small, independent driving forces or interactions, such as hydrophobic interactions, hydrogen bond formation, or Van-der-Waals forces and that the amino-acid code encoded solely for the proteins secondary structure, which then in turn encoded for the ternary structure (Anfinsen and Scheraga, 1975). In the past 20 years it became more prominent that this is not always applicable. Instead of a strict sequential encoding of secondary and ternary structure it is now assumed that the ternary structure influences the secondary structure in the same way as the secondary structure determines the ternary structure and that there is a predominant driving force in the folding code.

Not denying the other forces their role in folding, the predominant driving force is assumed to be the hydrophobic interaction between amino acid side chains during protein folding (Dill, 1990a; Dill, 1990b; Dill, 1999; Yang *et al.*, 2007). The main reasons for this assumption are the facts that most proteins have a hydrophobic core, thus shielding hydrophobic amino acids from water. Also most proteins are more readily denatured in non-polar solvents. Finally, studies in which a proteins amino acid sequence was altered while keeping the hydrophobic properties of the residues intact allowed those modified proteins to adopt similar ternary structures as their native counterpart (Cordes *et al.*, 1996; Kamtekar *et al.*, 1993; Kim *et al.*, 1998).

### 3.2.1 Thermodynamic and kinetic stability of proteins

A protein's "stability" can relate to its thermodynamic or kinetic stability. The thermodynamic stability refers to the difference in the Gibbs free energy of the folded versus the unfolded state of a protein, thus stabilizing one state compared to the other. Typical values for the free energy of folding,  $\Delta G_{U-N}$ , range from -20kJ/mol to -60kJ/mol (Fersht, 1999). Even though these values represent only marginable stabilization, the negative nature of the values indicates the preferred state of the polypeptide chain to be the folded, native state. This is true for virtually all proteins, with rare exceptions including some cycline-dependent kinase inhibitors, the steroidogenic acute regulatory protein (StAR) or clusterin (Bose *et al.*, 1999; Dunker *et al.*, 2001; Wright and Dyson, 1999). The Gibbs free energy is dependent on the enthalpy and the entropy of a molecule. Upon folding of a polypeptide chain the degrees of freedom of this chain drastically reduces and so does the entropy. Consequently, for the

folded state to be thermodynamically more favorable, the enthalpy has to rise upon folding. The gain in enthalpy by the tight packing of amino acid side chains in the folded core of a native protein just barely balances the loss in entropy and there must be another participator in this equation. If the water molecules solvating the protein in aqueous solutions are taken into account, the equation is balanced out more easily, since the water molecules solvating the long, randomly coiled polypeptide chain are released upon protein folding and enhance the overall entropy (Fersht, 1999).

The kinetic stability refers to the fact that a certain molecular state of a polypeptide chain represents a minimum in the energy landscape, but not necessarily the energetically most favorable state. The energy barrier the protein would have to pass over to reach a possibly more favorable state defines its kinetic stability. The higher this barrier, or transition state (TS), is, the more stable the current state will be. When the transition state theory was developed in the 1930s, it was applied to the simple formation or breakage of a chemical bond in small molecules. The free energy of activation of a reaction,  $\Delta G^{0\ddagger}$ , is thus defined by the ratio of educt  $[N]$  to transition state  $[T^\ddagger]$  as shown in equations (4.1) and (4.2), where  $R$  refers to the universal gas constant and  $T$  is the absolute temperature.

$$\Delta G^{0\ddagger} = -RT \ln \frac{[N]}{[T^\ddagger]} = -RT \ln K \quad (4.1)$$

$$[T^\ddagger] = [N] \cdot e^{\frac{\Delta G^{0\ddagger}}{RT}} \quad (4.2)$$

Since the original theory was dealing with single bond breakages and formation in simple molecules, the decay of the TS toward the educt could be attributed to the vibrational frequency ( $\nu$ ) necessary to break the respective bond. This can be expressed by equation (4.3) with  $k_B$  being the Boltzman constant and  $h$  the Planck constant.

$$\begin{aligned} -\frac{d[N]}{dt} &= \nu [T^\ddagger] & \text{with} & \quad \nu = \frac{k_B T}{h} \\ \rightarrow -\frac{d[N]}{dt} &= [N] \frac{k_B T}{h} \cdot e^{\frac{-\Delta G^{0\ddagger}}{RT}} & & (4.3) \\ \rightarrow k &= \frac{k_B T}{h} \cdot e^{\frac{-\Delta G^{0\ddagger}}{RT}} \end{aligned}$$

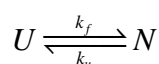
To account for the higher viscosity and friction and thus smaller probability of encounter of reaction partners necessary for bond formation, a pre-exponential factor, the so called transmission coefficient,  $\kappa$ , was introduced, but for most simple chemical reactions assumed to be equal 1. In protein folding, however, this factor becomes significant, due to the vast amount of bonds formed and the big influence of the solvent on the overall folding reaction (Kramers, 1940). Many studies on the identity of such a pre-exponential factor in protein folding have been performed, giving very different results in the range of  $10^6$  to  $10^8$  s<sup>-1</sup> and thus  $\kappa$  seems to be different for any given protein (Bieri *et al.*, 1999; Eaton *et al.*, 1997; Jacob *et al.*, 1997).

Another problem is the fact that the factor for protein folding differs from the one for unfolding of the same protein since in the folding reaction the viscosity of the medium is more significant than friction, while in the unfolding reaction, due to the tight packaging of the protein core, friction becomes more important. Since determination of such factors is tedious and time consuming, it is not practicable for most applications. To be able to give a reliable value of protein stability it is common to give the change in the free energy of activation,  $\Delta\Delta G^{0\ddagger}$ , compared to a reference instead of an absolute value, since the pre-exponential factors cancel out in such an equation (see equation (4.4)).

$$\Delta\Delta G^{0\ddagger} = \Delta G_1^{0\ddagger} - \Delta G_2^{0\ddagger} = -RT \ln\left(\frac{K_1}{K_2}\right) \quad (4.4)$$

### 3.2.2 Two-state model in protein folding

The two-state model of protein folding is the simplest possible case. It describes the assumption that there are only two states, the completely unfolded ( $U$ ) and the completely folded state ( $N$ ), separated by a transition state ( $TS$ ,  $\ddagger$ ) and without the formation of any Intermediates ( $I$ ). In this case, the rate of formation of the native molecule is  $k_f$ , and the rate of unfolding of the protein is  $k_u$  (Privalov and Khechinashvili, 1974; Tanford, 1970).



Small, single-domain proteins predominantly fold by a two-state pathway into the native state. This behavior is induced by the cooperative interaction of locally induced micro-structures in

the poly-peptide chain which then induce further interaction in previously remote areas of the chain. This results in the “all-or-nothing” behavior exhibited by these molecules (Dill *et al.*, 1993; Fersht, 1999).

### 3.2.2.1 Denaturant Dependence of Protein Folding and Unfolding

The rates of denaturant induced unfolding of a native protein depend on the concentration of denaturant present in the reaction (Greene and Pace, 1974; Santoro and Bolen, 1988; Tanford, 1970). In case of a two-state folder, when the rate of unfolding,  $k_u$ , is plotted logarithmically over the denaturant concentration,  $[D]$ , a linear dependence is observed according to equation (4.5), with  $k_u^{H_2O}$  the rate of unfolding at zero molar denaturant and  $m_u$  a factor for the dependence of the unfolding rate on the denaturant concentration, the kinetic m-value.

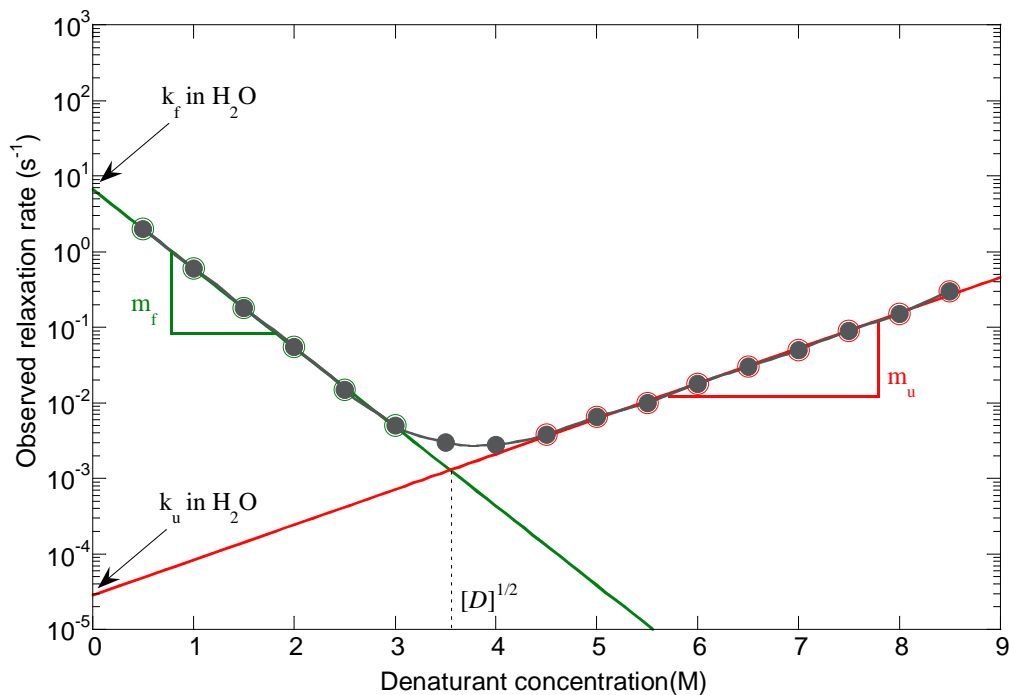
$$\ln k_u = \ln k_u^{H_2O} - m_u [D] \quad (4.5)$$

The rate of folding of a protein also depends on the denaturant concentration in and analogous way according to equation (4.6).

$$\ln k_f = \ln k_f^{H_2O} - m_f [D] \quad (4.6)$$

The combination of both, the logarithmic rates of folding and unfolding in dependence on denaturant concentration can be represented as a chevron plot (*Figure 3-8*). The name stems from the typical shape of the two arms, forming a V.

The observed apparent rate constant,  $k_{app}$ , is the sum of the single rates of folding,  $k_f$ , and unfolding,  $k_u$ :  $k_{app} = k_f + k_u$ . At the denaturation midpoint  $[D]^{1/2}$  both rates are equal and thus the reaction is in equilibrium



**Figure 3-8: Example of a Chevron plot.**

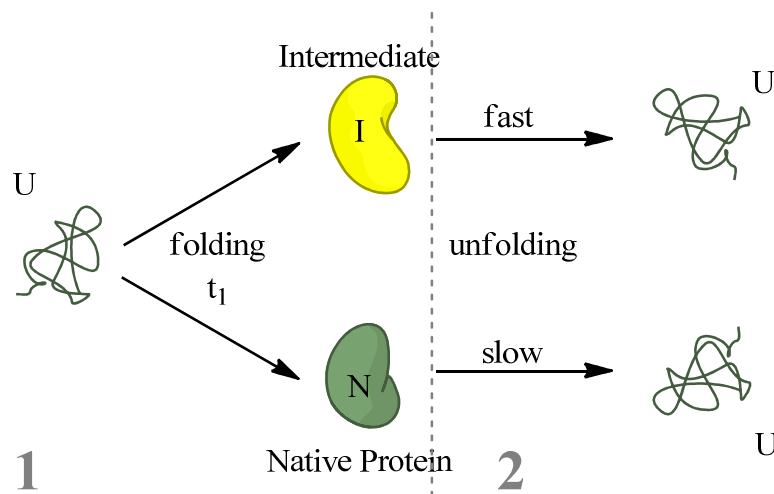
Chevron plots are a convenient way to represent kinetic protein folding data. The relaxation rates of folding and unfolding reactions are plotted over the respective concentration of denaturant (usually GdmCl or Urea) present in the reaction. By extrapolation of the linear parts of the folding and unfolding arm it is possible to obtain the rates of folding,  $k_f$ , and unfolding,  $k_u$ , at zero molar denaturant as well as the kinetic m-values, a measure for the surface accessibility of the TS.

The left, decreasing part of the Chevron plot is called the refolding arm, describing the dependence of the refolding relaxation rates on denaturant concentration, while the right, increasing part is the unfolding arm. When both linear arms are fitted linearly they intersect at the denaturation midpoint,  $[D]^{1/2}$ . At denaturant concentrations lower than  $[D]^{1/2}$  the rate of folding is the dominating factor in the reaction, at concentrations above  $[D]^{1/2}$  the unfolding rate is dominating.

Also, the kinetic m-values allow estimation on the denaturant accessible surface area (ASA) of the transition state relative to the native and unfolded state. The smaller the m-value is, the smaller the difference in ASA between the TS and the starting species is. Consequently, the solvent-accessibility of the TS is more similar to that of the native state if  $m_f > m_u$ . In consequence of this relation, kinetic unfolding data is a powerful tool to compare behavior of different mutants of a protein or the interaction of a protein with different binding partners, since a deviation of the kinetic m-values from the *wt* values is a strong indicator of changes in the ASA of a species and thus can help elucidation of complex structural functions.

### 3.2.2.2 Test for Native Molecules

To avoid that possible intermediates influence the measurement of formation of native molecules, interrupted refolding measurements are performed. Here, the protein is refolded for variable times and the refolding reaction is stopped by transferring the sample to unfolding conditions again. The spectroscopic change of the unfolding reaction is recorded, e.g. using tryptophan fluorescence. Using completely unfolded and folded protein as a spectroscopic reference, the amplitude of the spectroscopically measured unfolding trace can directly be converted into amount of folded protein at each time point. To avoid detecting possible folding intermediates, the unfolding conditions are chosen to only unfold the native molecules in the instruments time frame, per definition the most stable, species in the process, while all possible intermediates are already unfolded in the instruments dead-time or are at least well distinguishable from the native molecules (*Figure 3-9*, Schmid, 1983).



**Figure 3-9: Schematic representation of the interrupted refolding assay.**

The unfolded protein is refolded for varying times  $t_1$  (1). At that point the reaction is stopped by transfer to unfolding conditions (2). Unfolding conditions are chosen so that the fast unfolding of possible intermediates can well be distinguished from the slower unfolding of the native molecules. The amplitude of the slowest unfolding phase is directly proportional to the concentration of native molecules after refolding for the time  $t_1$ .



### 3.3 Protein-protein interactions

Protein-protein interactions are determined by hydrogen-bonds, Vand-der-Waals-, hydrophobic- and electrostatic forces. There is, however, not a single concept common to all known protein-protein interactions since the kind of interaction differs with the biological function of the complexes. The main differentiation is based on the identity of the interacting protomers, namely homo- or hetero-oligomeric complexes, the ability of the protomers to exist as monomeric entities *in vivo* (obligate vs. non-obligate complexes) and the fact that protein-protein interactions can be of transient or permanent nature.

Prominent examples of transient complexes are enzyme:inhibitor complexes. Albeit a fairly strong interaction between the catalyst and the inhibitor is necessary to avoid the respective reaction to take place, the inhibitor has to be readily released from the enzyme protein when the need arises. Permanent complexes on the other hand only dissociate if a strong, external trigger is present. This can be a denaturing condition (heat, solvent) or a cellular trigger.

Examples for homocomplexes are in most cases obligate (i.e. the monomers are not biological active or do not exist in monomeric form) and their formation is often considered as an extension of protein folding (quaternary structure formation). This can be true for heterocomplexes, too, but considered these are seldom co-localized upon expression they have to be at least meta-stable during the translocation to the respective binding partner. Heterocomplexes can also form transient or permanent complexes comprised of proteins which exist, and even might exhibit biological function, separately.

The interfaces in protein-protein interaction formation differ between permanent and transient complexes. The interfaces of permanent compounds are predominantly formed by hydrophobic patches whose shape exhibits a high degree of complementary towards the shape of the cognate binding partner. Those patches differ in size (200-400 Å<sup>2</sup>) and number (1-15) (Jones and Thornton, 1996; Lijnzaad and Argos, 1997; Stites, 1997). Statistical analysis of known protein-protein interfaces gave estimates on the size of the interface area of 550-4900 Å<sup>2</sup> with an average of about 800 Å<sup>2</sup> and a fraction on the monomers total surface of about 6-30%. It was also found that the presence of the amino acid residues arginine, histidine, tyrosine, tryptophan, serine, asparagine and hydrophobic residues in general was increased compared to the overall distribution in proteins (Davies and Cohen, 1996; Jones and Thornton, 1996; Stites, 1997). Next to the hydrophobicity of the interface region, the electrostatic forces form the second most important power in protein-protein interaction (Sheinerman *et al.*, 2000). It is believed that this is due to the fact that removal of solvent from charged groups is

facilitated since they more readily form interactions with other charged groups or even hydrophilic residues (Xu *et al.*, 1997a).

Another force involved in complex formation is the establishment of hydrogen bonds (H-bonds) between amino acids. Even though both the amino acid side-chains as well as the peptide backbone can be involved in H-bond formation, the amino acid side-chains form about  $\frac{3}{4}$  of all H-bonds in proteins. It was noted that, since the H-bonds are hardly ever oriented in an optimal manner, intra-proteinaceous H-bonds are only of weak or average strength, while H-bonds on the surface between protein and solvating water molecules are considered of strong character (Xu *et al.*, 1997b).

An additional factor in formation of protein-protein complexes is the induction of conformational changes in the binding partner upon binding. Especially obligate, permanent complexes which form during protein folding can induce strong structural changes in one another and thus form very tight bonds, up to the point where the single protein subunits complement each others fold. In transient complexes this effect is not as pronounced, since the already folded binding partners are fairly limited in the degree of structural change they can induce in each other.

### **3.4 Macromolecular protein assemblies**

A proteins biological relevance is in many cases not in a single molecular state but is defined by interaction with other proteins. Some proteins form di-, tri-, or tetramers, but many form macromolecular assemblies of higher multimeric states. This is true for all kingdoms of life. Prominent examples of macromolecular assemblies are the ribosomes, fatty acid synthases, proteosomes, or the nuclear pore complexes in eukaryotes (Cheng, 2009; Elad *et al.*, 2009; Maier *et al.*, 2008; Steitz, 2008). In the prokaryotic kingdom, assemblies on the outer surface play an especially vital role. So are the flagella and type IV pili important for chemotaxis and motility, the F-pilus is involved in bacterial conjugation and P type and type 1 pili are, among others, involved in bacterial anchoring and biofilm formation (Hall-Stoodley *et al.*, 2004; Van Houdt and Michiels, 2005).

Investigation of macromolecular complexes proved to be challenging due to their size and complexity. Dissection of assemblies into their respective subunits was as tedious as assembling them from the single proteins. Especially structural investigations are of great interest and at the same time the most challenging task in the field. Protein crystallization of

the complexes, sometimes ranging up to the mega-Dalton range in size, is very seldom successful but necessary to elucidate the function of the complex in its natural context.

Despite the interest in those assemblies, very little is known how the significant degree of specificity during their biogenesis is achieved. The most promising target of such investigations are the fimbriae in the bacterial kingdom, due to their relative simplicity compared to other structures and their comparatively easy handling. It is known that those macromolecular structures are assembled in a linear fashion, forming long chains of protein which then assemble into a final quaternary structure, in most cases a helical form. The linear assembly makes it quite easy to break the assembly steps down to single protein-protein interaction steps and by building up on the single steps it might prove possible to determine the mechanism of assembly and specify of the whole macromolecular complex.

### ***3.5 Aim of this study***

The aim of this study is to gain further insight into the mechanism by which the prokaryotic cell controls the correct incorporation of the various subunits into the growing type 1 pilus on the cell surface. This question will be assessed both on a biophysical approach, focusing on the interaction of the subunits themselves, as well as a structural biology approach looking into the interaction between the outer membrane usher and the respective chaperon:subunit complexes.



## 4 Results

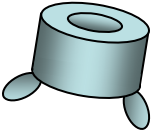
















Type 1 pili are adhesive macromolecular protein complexes on the surface of bacterial cells. They exhibit extreme resistance towards denaturing conditions. In fact, type 1 pilus subunit:subunit interactions are one of the most stable protein:protein interactions known to date (Eshdat *et al.*, 1981; Krogfelt and Klemm, 1988; Puorger *et al.*, 2008). The linear chain of proteins forming the type 1 pilus also exhibits a stringent sequence order of incorporated subunits. So is the adhesin FimH always incorporated as the initial subunit into the growing pilus, followed by a not exactly determined number of the linker subunits FimG and FimF before connecting to several hundred copies of the main structural subunit FimA, which forms the helical pilus rod (Hahn *et al.*, 2002; Vetsch and Glockshuber, 2005). This interaction of the subunits proceeds through a mechanism termed donor strand complementation (DSC), in which an N-terminal extension, the donor-strand (ds), of the subunit to be incorporated into the growing pilus becomes part of the ternary structure of the preceding subunit to complete the incomplete Ig-like fold of this protein. This interaction was shown to contribute significantly to the high degree of stability of the whole macromolecular assembly (Puorger *et al.*, 2008; Sauer *et al.*, 2002; Zavialov *et al.*, 2005).

One open question in the investigation of fimbriae assembled via the chaperon/usher pathway is, how the order of the subunits is determined. Is this achieved only at the outer membrane-usher during assembly or are the subunits themselves sensible for the correct binding partner, i.e. only prone to bind the designated donor-strand, even in absence of a controlling entity?

### **4.1 Overview of proteins and protein variants used throughout this study**

A tabular overview over the *wild type* Fim proteins and their derivatives used in this study is given in *Table 4-1* as a reference for the reader.

**Table 4-1: Overview over the Fim protein *wild type* subunits and their derivatives.**

Fim Subunit <i>wild type</i>	Derivatives	Schematics / Color Code	# aa	M <sub>r</sub> (Da)	Purification Tag	Expression	Description
<b>FimD</b> 	<b>FimD<sub>N</sub></b>		125	13'655		Periplasmatic	N-terminal domain of the outer membrane usher FimD. Expressed as soluble protein.
<b>FimC</b> 	<b>FimC<sub>yy</sub></b>		205	22'730		Periplasmatic	FimC variant in which the intrinsic Trp residues were replaced by Tyr residues to minimize fluorescence
	<b>FimC<sub>his</sub></b>		211	23'553	C-term. (6xhis)	Periplasmatic	FimC with C-terminal 6xhistidine tag
<b>FimH</b> 	<b>FimH</b>		279	29'068		Periplasmatic	
	<b>FimH<sub>p</sub></b>		122	12'266		Periplasmatic	C-terminal, pilin domain of adhesin protein FimH
<b>FimA</b> 	<b>FimA</b>		159	15'827		Whole Pili	
	<b>FimA<sub>t</sub></b>		140	13'973		Cytosolic, auto-induction	N-terminally truncated FimA
	<b>FimA<sub>A</sub></b>		184	18'024		Cytosolic, auto-induction	Self-complemented FimA. FimA ds-sequence is added to C-terminus via a 6 aa linker sequence
<b>FimG</b> 	<b>FimG<sub>t</sub></b>		132	13'659		Cytosolic, auto-induction	N-terminally truncated FimG
<b>FimF</b> 	<b>FimF<sub>t</sub></b>		142	14'835		Periplasmatic	N-terminally truncated FimF
	<b>FimF<sub>AF4</sub></b>		176	18'451		Cytosolic, auto-induction	Self-complemented FimF. FimF ds-sequence is added to C-terminus via a 6 aa linker sequence. N-terminal extension exchanged for ds-sequence of FimA

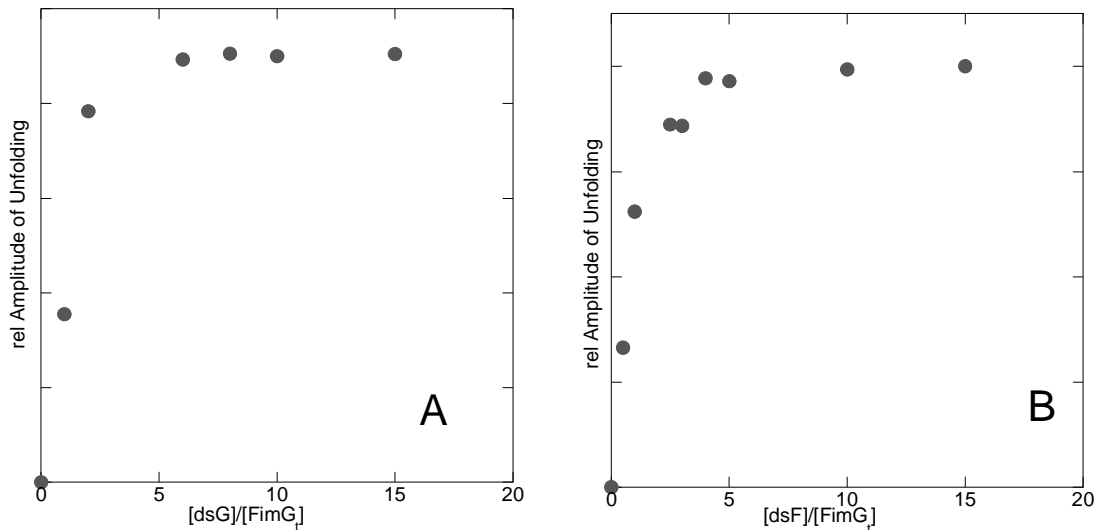
## **4.2 Complex formation between subunit and ds peptide**

To examine the hypothesis that subunit:donor-strand interaction is sufficient to regulate the correct order of incorporation of subunit proteins into the growing pilus, the interactions between single subunits and all possible donor- strands were investigated.

A natural starting point in the investigation of protein complexes is the establishment of complex formation and its parameters. Since this study will focus on the interaction of Fim subunit proteins with peptides corresponding to the N-terminal extensions of the Fim subunits, so called donor-strands (ds), formation of complexes between those proteins and peptides was investigated first. To avoid auto-aggregation between subunits, N-terminally truncated variants were used. Those protein variants are denoted by the suffix t, e.g. FimG<sub>t</sub>.

### **4.2.1 Formation of subunit:donor-strand complexes depends on peptide concentration**

In refolding reactions of unfolded FimG<sub>t</sub> protein in the presence of varying amounts of the donor-strand peptides dsF or dsG the dependence of complete refolding on the presence of an excess of peptide was assessed. For this experiment the protein was refolded out of storage condition (20mM Tris/HCl, pH8.0; 2.75M GdmCl) by dilution 1:11 in 20mM Tris/HCl, pH8.0 containing the respective amount of ds peptide. Final concentration of the protein was 3 $\mu$ M during refolding. After complete refolding the amount of formed complex was determined by transferring the reaction to unfolding conditions and following the decrease in tryptophan fluorescence emitted at 330nm. The amplitude is directly related to the fraction of complex molecules present in the reaction, since free protein would unfold in the dead time of the measurement due to the high concentration of denaturant used for unfolding (FimG<sub>t</sub>:dsG : [GdmCl] = 6.5M; FimG<sub>t</sub>:dsF : [GdmCl] = 7.0M). For both complexes a dependence on the concentration of ds peptide was observed (*Figure 4-1*).



**Figure 4-1: Concentration dependence of the formation of FimG<sub>t</sub>:dsG and FimG<sub>t</sub>:dsF complexes during refolding.**

(A) 1 μM FimG<sub>t</sub> refolded in presence of 0 to 15 μM dsG. (B) 1 μM FimG<sub>t</sub> refolded in presence of 0 to 15 μM dsF. The amplitude of the fluorescence decrease upon transfer to unfolding conditions was recorded at 330nm with an excitation wavelength of 295nm. Amplitudes are directly proportional to fractions of native complex.

#### 4.2.2 Fim subunit proteins form complexes with designated and non-designated ds peptides

For subsequent experiments pure complexes between a subunit and a given ds peptide had to be obtained. For that, unfolded subunit proteins were refolded in the presence of a 5-fold excess of the respective ds peptide by 1:11 dilution from storage buffer (pH8.0, containing 2.75M GdmCl) to a final denaturant concentration of 0.25M at pH8.0. After complete refolding, unbound protein and peptide were removed by anion exchange chromatography. Using this approach, complexes between the subunit FimG<sub>t</sub> and the ds peptides dsA, dsF, dsG, and dsI could be obtained in high purity. For complexes between FimF<sub>t</sub> and dsA or dsF the efficiency was slightly diminished compared to the FimG<sub>t</sub> complexes, FimF<sub>t</sub>:dsI yield was even lower due to high degree of aggregation, most probably due to the extended length of the dsI peptide (21 amino acids compared to the average 14-15), and the complex between FimF<sub>t</sub> and dsG could not be obtained at all. This shows that incorporation of FimG after FimF in the tip fibrillum is already inhibited on the level of the donor-strand complementation (DSC).



### 4.3 Kinetic stability of subunit:ds complexes against unfolding

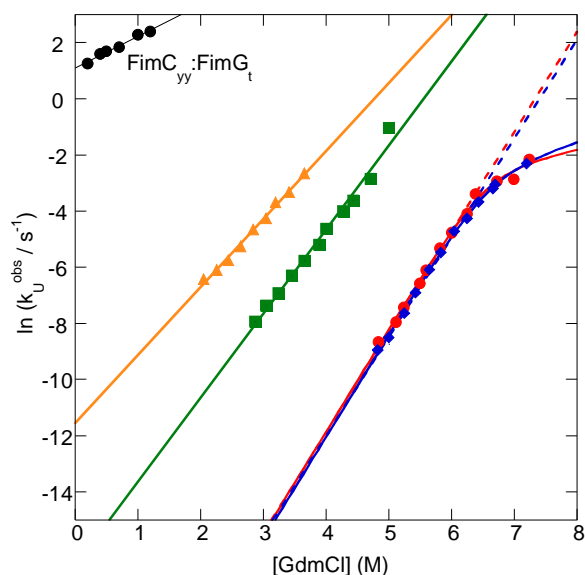
The purified subunit:ds complexes were consequently subject of investigation of complex stability. In particular the kinetic stability of the complexes against unfolding was of special interest since it was known that pili are extremely resistant against denaturation. To answer the question if this effect is due to the global ternary structure of the pilus or the same effect can also be observed on the single subunit level, the kinetic stability against unfolding of the purified subunit:ds complexes was assayed.

To probe the kinetic stability against unfolding of the subunit:ds complexes, the individual complexes were formed and purified as described. The purified complexes were transferred to unfolding conditions and the decrease in tryptophan fluorescence was recorded in a far-UV fluorescence spectrophotometer ( $\lambda_{\text{Ex}}=295\text{nm}$ ,  $\lambda_{\text{Em}}=330\text{nm}$ ). By repeating the measurements at different denaturant concentrations, it is possible to determine the dependence of the rate of unfolding on the denaturant concentration. Since the Fim proteins form very stable complexes with ds peptides, even high concentrations of GdmCl were not sufficient for effective unfolding, i.e. measurements in a laboratory timeframe. As a consequence, to be able to perform measurements in a reasonable timeframe, in addition to the use of very high concentrations of denaturant, the pH was also lowered to pH 2.0 to facilitate unfolding.

The unfolding traces of the decreasing fluorescence intensity recorded at 330nm during excitation of the sample at 295nm, were fitted to a single-exponential equation and the resulting apparent rate-constants of unfolding ( $k_u^{\text{app}}$ ) were logarithmically plotted against the GdmCl concentration (*Figure 4-2*). This kind of plot, also known as Chevron plot, yields valuable information on the unfolding reaction. Interpolation of the linear fit through the denaturant dependant rates of unfolding gives the relative rate of unfolding at 0M denaturant concentration ( $k_u^{\text{H}_2\text{O}}$ ) and the slope of the linear fit ( $m_u$ ) is an indicator for the accessible surface area of the transition state of the unfolding reaction (see 3.2.2.1).

As a reference for the calculation of increase in free energy of activation of unfolding ( $\Delta\Delta G_U^{0\ddagger}$ ) upon ds peptide binding relative to the FimC:subunit complex, the respective chaperone-bound subunit was also assayed. Since unfolding of the chaperone in this bimolecular complex might mask the unfolding signal of the subunit, unfolding was compared to that of free chaperone. In case of the FimC:FimG<sub>t</sub> complex this was found to be true and so a variant of the chaperone lacking its two tryptophan residues, FimC<sub>yy</sub>, was used. The complex FimC<sub>yy</sub>:FimG<sub>t</sub> complex was unfolded and due to the lack of tryptophan the unfolding of the chaperone was spectroscopically silent at the conditions used in the assay,

hence only the unfolding of FimG<sub>t</sub> was measured. For the FimC:FimF<sub>t</sub> complex a significantly slower unfolding was observed compared to free chaperone and thus the use of the modified chaperone was not necessary (data not shown).



**Figure 4-2: Chevron plot of GdmCl dependent unfolding of FimG<sub>t</sub>:ds complexes.**

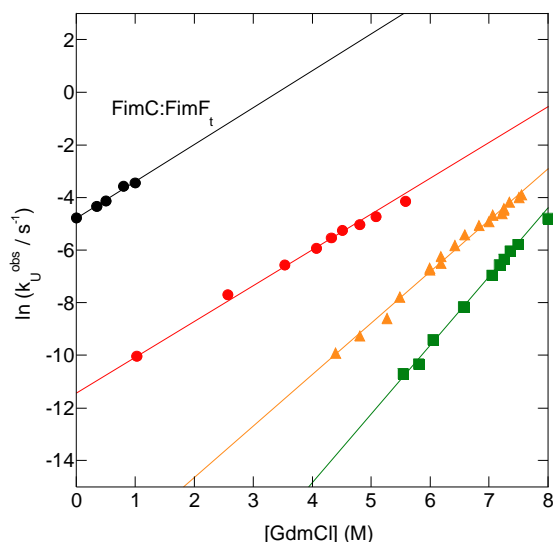
Natural logarithms of the observed rates of unfolding are plotted against the denaturant concentration. FimC<sub>yy</sub>:FimG<sub>t</sub> complex is depicted in black circles (●). Complex of FimG<sub>t</sub> with dsA (■), dsF (●), dsG (◆) and dsI (▲) are shown. For FimG<sub>t</sub>:dsA and FimG<sub>t</sub>:dsI linear dependence was observed and the curve fitted with a linear function, yielding kinetic m-values for unfolding ( $m_U$ ) of 2.99 and 2.42, respectively. For FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG a deviation from linearity was observed at high denaturant concentrations. This can be explained with a structural change in the transition state at high denaturant concentrations and was fitted according to a three-state model (Szabo, 1969) and modified for the missing refolding data (Puorger *et al.*, 2008). This yielded the individual values for the kinetic rate constants at zero molar denaturant concentration,  $k_{NI}^0$ ,  $k_{IN}^0$  and  $k_{IU}^0$ , as well as the shared kinetic m-values  $m_{NI}^0 = 2.41 \pm 0.07$ ,  $m_{IN}^0 = -0.76 \pm 0.28$  and  $m_{IU}^0 = 0.49 \pm 0.31$ . Comparison to the extrapolation to 0M GdmCl using a linear fit did show that the result was comparable, thus the simple fit was used for further analysis.

Extrapolation of linear fits for FimG<sub>t</sub>:dsA and FimG<sub>t</sub>:dsI complexes to zero molar GdmCl concentration yielded apparent rates of unfolding of  $k_{app}^{H_2O} = (1.409 \pm 0.039) \cdot 10^{-6} s^{-1}$  and  $k_{app}^{H_2O} = (6.531 \pm 0.131) \cdot 10^{-6} s^{-1}$ , respectively (see also *Table 4-2*). The Chevron plot of unfolding of FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG complexes showed deviation from linearity at high denaturant concentrations, indicative of a shift in the rate-limiting transition state at high

GdmCl concentrations (Bachmann and Kiefhaber, 2001). The data was subsequently fitted to three-state model according to equation (8.2), (Bachmann and Kiefhaber, 2001; Szabo, 1969) and modified for the lack of information on the folding arm of the Chevron plot as described by Puorger *et al.* (Puorger *et al.*, 2008). Direct comparison of the linear part of this fit to a simple linear fit omitting the data points at high denaturant concentrations showed that this deviation is minimal and the final rates of unfolding at 0M GdmCl differ only minimal compared to the high degree of error such an analysis implicates ( $k_{appFimG_t:dsG}^{H_2O, 3-State} = 2.77 \cdot 10^{-12}$  vs.  $k_{appFimG_t:dsG}^{H_2O, 2-State} = 4.42 \cdot 10^{-12}$  and  $k_{appFimG_t:dsF}^{H_2O, 3-State} = 3.36 \cdot 10^{-12}$  vs.  $k_{appFimG_t:dsF}^{H_2O, 2-State} = 4.43 \cdot 10^{-12}$ ) and thus the effect of the altered transition-state is not important for the extrapolation of the relative rates of unfolding at 0M denaturant concentration and consequently the simpler linear fit was used for further analysis.

Similar experiments were performed with the truncated subunit protein FimF<sub>t</sub> (Figure 4-3), which did only form complexes with the dsA, dsF and dsI peptides, but not with dsG. This finding indicates that a reverse order of subunits in the pilus tip fibrillum is already prohibited by the DSC mechanism, since FimF apparently is not able to bind the N-terminal extension of FimG, assumed to be its predecessor in the type 1 tip-fibrillum. The rates of unfolding determined were  $k_{FimF_t:dsA}^{H_2O} = 1.06 \cdot 10^{-11}$ ,  $k_{FimF_t:dsF}^{H_2O} = 1.08 \cdot 10^{-5}$  and  $k_{FimF_t:dsI}^{H_2O} = 8.92 \cdot 10^{-9}$ .

The complex between FimF<sub>t</sub> and dsA is by far the most stable one with a half-life of unfolding of around 2000 years at pH 2.0, again in good agreement with expectations since this is the designated combination *in vivo*. However, the subunit also forms fairly stable complexes ( $t_{FimF_t:dsF}^{1/2} = 15h$ ) with its own ds, indicative of the possibility of more than one FimF subunit in the mature type 1 pilus tip fibrillum.



**Figure 4-3: Chevron plot of GdmCl dependent unfolding of FimF<sub>t</sub>:ds complexes.**

Natural logarithms of the observed rates of unfolding are plotted against the denaturant concentration. Complex of FimF<sub>t</sub> with dsA (■), dsF (●) and dsI (▲) are shown. All traces exhibited linear behaviour and could be fitted with a linear function, yielding kinetic  $m$ -values for unfolding ( $m_U$ ) of 2.61, 1.36 and 1.95 for FimF<sub>t</sub>:dsA, FimF<sub>t</sub>:dsF and FimF<sub>t</sub>:dsI, respectively. FimF<sub>t</sub> did not form complexes with the dsG peptide and thus this complex could not be analyzed. Depicted is also the unfolding trace of FimC<sub>his</sub>:FimF<sub>t</sub> complex (●,  $m_U=1.41$ ).

The increase in kinetic stability between chaperone:subunit and subunit:ds complexes was calculated as change in free energy of activation of unfolding,  $\Delta\Delta G_U^{0\dagger}$ , relative to the chaperone bound subunit protein (*Table 4-2*).

FimC<sub>yy</sub>:FimG<sub>t</sub> complex was found to be extremely unstable at pH 2.0, so that measurements of unfolding kinetics were only possible in a stopped-flow apparatus on a millisecond timescale. The half-life of unfolding for this complex was found to be about 250 milliseconds at pH 2.0. FimC<sub>his</sub>:FimF<sub>t</sub>, on the other hand, was considerably less stable than the FimF<sub>t</sub>:ds complexes, but was well measurable at pH 2.0 with a half-life of unfolding at around 80 seconds, in comparison about 320-times slower than unfolding of the FimC<sub>yy</sub>:FimG<sub>t</sub> complex. The results of the investigation of the kinetic stability of the possible complexes formed by the pilus tip proteins indicated a pattern of stability which is in agreement with the assumed order of subunits in the mature type 1 pilus. However, since the formation of complexes with non-designated ds peptides was also possible, the selectivity can not be controlled by the DSC mechanism alone, the designated complexes must form significantly faster than the non-designated ones to avoid misincorporation.

**Table 4-2: Kinetic parameters of subunit:ds complex unfolding.**

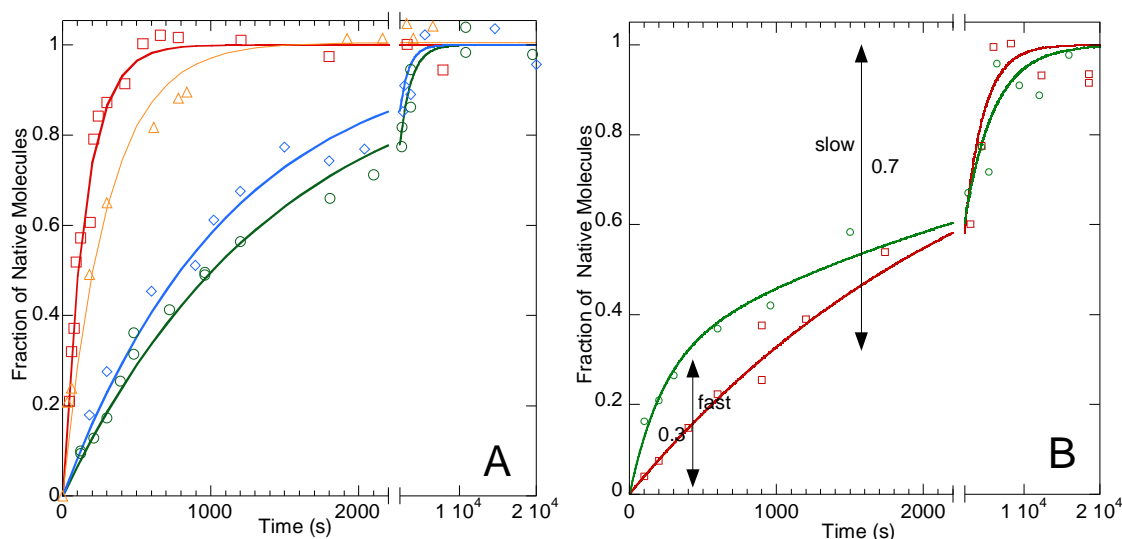
Complex	$m_U$	$k_U^{H_2O}$ (s <sup>-1</sup> )*	$t_{1/2}^{H_2O}$	$\Delta\Delta G_U^{0\dagger}$ (kJ·mol <sup>-1</sup> )
<b>FimC<sub>yy</sub>:FimG<sub>t</sub></b> <sup>a</sup>	1.126	(2.71±0.02)	(0.256±0.002) s	0
<b>FimG<sub>t</sub>:dsI</b>	2.421	(6.53±0.54) • 10 <sup>-6</sup>	(29.5±2.0) h	-32.07
<b>FimG<sub>t</sub>:dsA</b>	2.992	(1.41±0.21) • 10 <sup>-7</sup>	(57±9.7) d	-41.58
<b>FimG<sub>t</sub>:dsG</b>	3.519	(5.47±0.12) • 10 <sup>-12</sup>	(4020±88) y	-66.76
<b>FimG<sub>t</sub>:dsF</b>	3.566	(4.25±0.13) • 10 <sup>-12</sup>	(5170±152) y	-67.38
<b>FimC<sub>his</sub>:FimF<sub>t</sub></b>	1.405	(8.25±0.05) • 10 <sup>-3</sup>	(84±1) s	0
<b>FimF<sub>t</sub>:dsI</b>	1.953	(1.74±0.01) • 10 <sup>-8</sup>	(462±11) days	-32.41
<b>FimF<sub>t</sub>:dsA</b>	2.609	(1.11±0.12) • 10 <sup>-11</sup>	(1998±200)y	-50.64
<b>FimF<sub>t</sub>:dsG</b>	n.a.	n.a.	n.a.	n.a.
<b>FimF<sub>t</sub>:dsF</b>	1.362	(1.287±0.13) • 10 <sup>-5</sup>	(15±1.7) h	-16.02

\* Interpolated to 0M GdmCl, pH2.0  
<sup>a</sup> from Puorger *et al.*, 2008

## 4.4 Kinetics of subunit:ds peptide complex formation

### 4.4.1 Interrupted refolding experiments

To assess the rate of formation of the complexes, interrupted refolding experiments were performed. The unfolded subunit proteins FimG<sub>t</sub> or FimF<sub>t</sub> were refolded for various time spans in the presence of a 10-fold excess of the respective ds peptide at pH8.0 as described before. The reaction was stopped by transfer of the sample to unfolding conditions and the amount of accumulated complex was determined by recording the decrease of tryptophan fluorescence emitted at 330nm during excitation at 295nm upon unfolding. The unfolding conditions were chosen in a way that only the stable subunit:ds complexes would unfold in a measurable timeframe and all possible, but unstable or less stable intermediates would be already unfolded in the dead time of the experiment. For the respective complexes these conditions were at pH1.0 and denaturant concentrations as follows: [GdmCl]<sub>FimG<sub>t</sub>:dsA</sub>= 5.0M; [GdmCl]<sub>FimG<sub>t</sub>:dsF</sub>= 8.0M; [GdmCl]<sub>FimG<sub>t</sub>:dsG</sub>= 7.0M; [GdmCl]<sub>FimG<sub>t</sub>:dsI</sub>= 4.0M; [GdmCl]<sub>FimF<sub>t</sub>:dsA</sub>= 7.55M; [GdmCl]<sub>FimF<sub>t</sub>:dsF</sub>= 6.44M; [GdmCl]<sub>FimF<sub>t</sub>:dsI</sub>= 6.0M. The unfolding curves for the subunit:ds complexes were fitted to a single exponential and the resulting amplitudes plotted against incubation time (*Figure 4-4*).



DS	FimG <sub>t</sub>		FimF <sub>t</sub>		
	$k_{app}(s^{-1})$	$t_{1/2}(s)$	$k_{fast}(s^{-1})$	$k_{slow}(s^{-1})$	$t_{1/2}(s)$
dsA	$(6.84 \pm 0.31) \cdot 10^{-4}$	$1013 \pm 44$	$(4.50 \pm 2.56) \cdot 10^{-3}$	$(2.57 \pm 0.53) \cdot 10^{-4}$	$154 \pm 88$ $2697 \pm 461$
dsF	$(6.69 \pm 0.39) \cdot 10^{-3}$	$104 \pm 6$	n/a	$(3.96 \pm 0.35) \cdot 10^{-4}$	$1750 \pm 155$
dsG	$(8.70 \pm 0.53) \cdot 10^{-4}$	$796 \pm 48$	n/a	n/a	n/a
dsl	$(3.29 \pm 0.51) \cdot 10^{-3}$	$210 \pm 13$	-	-	-

**Figure 4-4: Interrupted Refolding Experiments of FimG<sub>t</sub>:ds and FimF<sub>t</sub>:ds Complexes.**

Unfolded FimG<sub>t</sub> (A) and FimF<sub>t</sub> (B) were refolded in the presence of a 10-fold excess of ds-peptides at pH8.0, 25°C. The reactions were stopped by transferring the sample to unfolding conditions and the fluorescence decrease upon unfolding was recorded. The resulting amplitudes were plotted over time and the data fitted to a single-exponential to yield  $k_f^{app}$ , except for the experiment with FimF<sub>t</sub>:dsA, which was fitted with a double-exponential. For FimG<sub>t</sub> the reaction with dsF was by far the fastest with a half-life of about 100s, followed by the reactions with dsG and dsA with half-lives of 800s and 1000s, respectively. Colors of the curves are as denoted in the table.

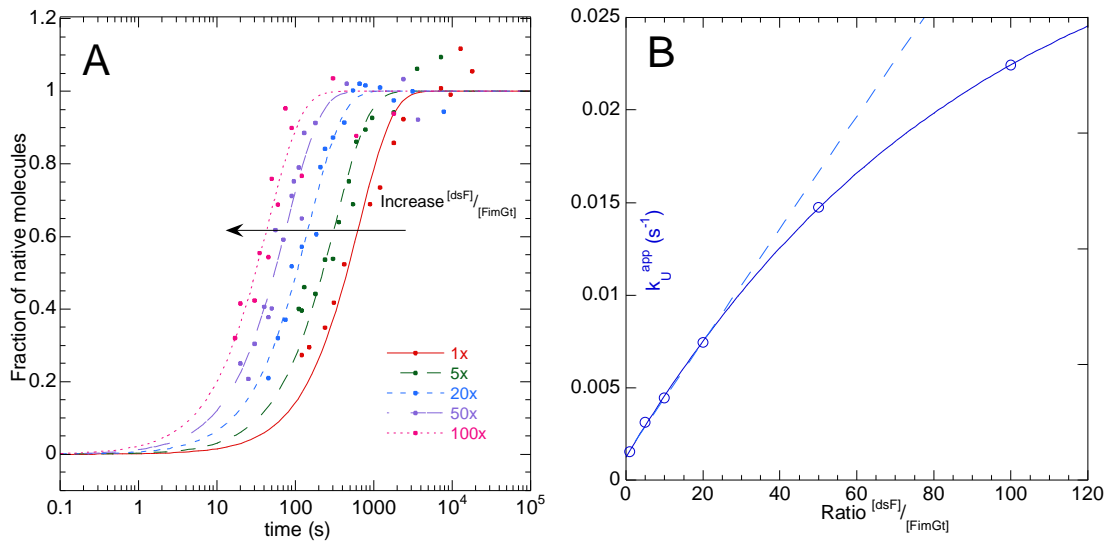
The data for the FimG<sub>t</sub>:ds complexes clearly shows that the most native like complex, FimG<sub>t</sub>:dsF ( $k_{app} = 6.69 \pm 0.39 \cdot 10^{-3} s^{-1}$ ), was formed about one order of magnitude faster than the FimG<sub>t</sub>:dsA ( $k_{app} = 6.84 \pm 0.31 \cdot 10^{-4} s^{-1}$ ) and FimG<sub>t</sub>:dsG ( $k_{app} = 8.70 \pm 0.53 \cdot 10^{-4} s^{-1}$ ) complexes, with FimG<sub>t</sub>:dsG being formed the slowest.

In case of the FimF<sub>t</sub>:ds complexes there was a different picture. For the reaction with dsA two phases could be observed in the interrupted refolding experiment. One of those phases was a fast phase with an apparent rate of  $k_1 = 4.50 \pm 2.56 \cdot 10^{-3} s^{-1}$  and a slower phase with an apparent

rate of  $k_2 = 2.57 \pm 0.53 \cdot 10^{-3} \text{ s}^{-1}$ . The ratio of fast-to-slow phase was about 3:7. This effect can most likely be attributed to the incorporation of the dsA peptide in a non-desired conformation, which is then displaced by the stable, native-like one. The fast rate of the FimF<sub>t</sub>:dsA formation is in the same dimension as the rate for formation of FimG<sub>t</sub>:dsF, while the complex FimF<sub>t</sub>:dsF is formed very slowly. As was shown before, FimF<sub>t</sub>:dsG is not formed at all.

#### 4.4.2 Formation of Subunit:ds complexes during refolding depends on the concentration of the ds peptide present

The previously described interrupted refolding experiment was not only performed with a 10-fold excess of peptides over subunits, but also with other ratios of ds peptide over protein. In *Figure 4-5A* the results for FimG<sub>t</sub> versus dsF are shown. The increase of dsF concentration while keeping FimG<sub>t</sub> constant accelerated the overall reaction. Plotting the apparent rates of the mono-exponential fits over the respective ratios of  $[\text{dsF}]/[\text{FimG}_t]$  would yield a linear dependence in case of a 1<sup>st</sup>-order reaction in regard to the peptide and in case of a 2<sup>nd</sup>-order reaction in regard to the peptide the plot would form part of a parabola. As seen in *Figure 4-5B*, the ratios show linearity at low excess of peptide over protein and a deviation of linearity at high excess of peptide with an decreasing slope toward a peptide concentration independent, constant rate of about  $0.034 \text{ s}^{-1}$ . This can be interpreted as the change of the reaction order from a rate-limiting 2<sup>nd</sup>-order to a 1<sup>st</sup>-order reaction at high concentrations of peptide present, indicative of the folding of the protein becoming the rate-limiting step over the binding of the ds peptide. The apparent rate of the 1<sup>st</sup>-order reaction ( $k_{0rxn}^{app} = 0.034 \text{ s}^{-1}$ ) is in good agreement with the rate of folding of free FimG<sub>t</sub>,  $k_f^{\text{FimG}_t} = 0.045 \text{ s}^{-1}$ .



**Figure 4-5: Concentration dependence of FimG<sub>t</sub>:dsF refolding kinetics.**

The plot (A) shows multiple interrupted refolding experiments, performed with various ratios of [dsF] over [FimG<sub>t</sub>]. It shows that the concentration of ds peptide has an influence on the overall reaction velocity. In (B) the apparent rates are plotted over the reactant ratios. The shapes of the plots indicate a pseudo-1<sup>st</sup>-order reaction. The deviation from linearity at high excess of ds peptide over FimG<sub>t</sub> indicates that folding becomes eventually the rate-limiting step.

This result clearly shows that complex formation between the subunit and the ds peptide during refolding is a multi-step process. First the unfolded protein adopts at least a partially folded state to which the free peptide then binds and forms the considerably more stable complex.

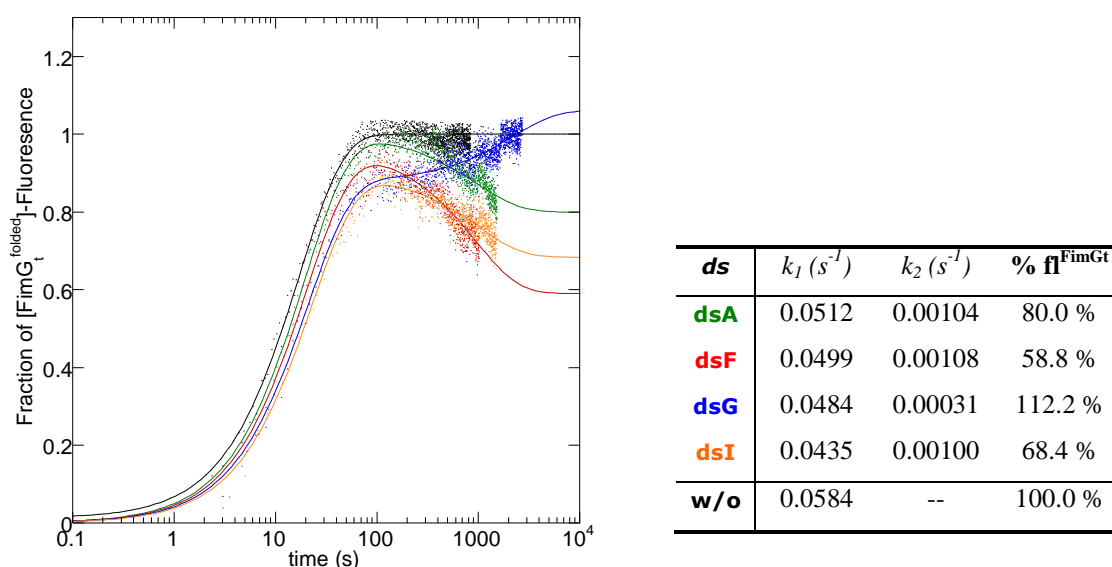
#### **4.5 Qualitative folding of FimG<sub>t</sub> in the presence of ds peptides**

From the previous experiments it seemed obvious that protein folding and ds peptide binding is not a concomitant process. It is conceivable that there are at least two different possible mechanisms in the complex formation: (i) the protein adopts its final tertiary structure before the peptide binds to the protein or (ii) the protein folds into an intermediate to which the peptide binds and induces complete ternary structure adoption upon binding. To elucidate the two possible hypotheses, the complete folding process in presence of the ds peptides was followed by change in the fluorescence intensity due to the change in the microenvironment of the respective fluorophores.



### 4.5.1 Folding traces of FimG<sub>t</sub>:ds complexes

The intrinsic fluorescence intensity of the various complexes was assayed by following the folding of FimG<sub>t</sub> in presence of dsA, dsF, dsG and dsI peptides via recording the change of tryptophan fluorescence emitted at 330nm with an excitation wavelength of 295nm. 33μM of unfolded FimG<sub>t</sub> protein in 2.75M GdmCl was refolded and the measurement started by 1:11 dilution with refolding buffer (pH8.0) containing a 5-fold excess of the respective ds peptide. The recorded folding traces showed significant differences in intrinsic fluorescence of completely refolded FimG<sub>t</sub>:ds complexes (*Figure 4-6*).

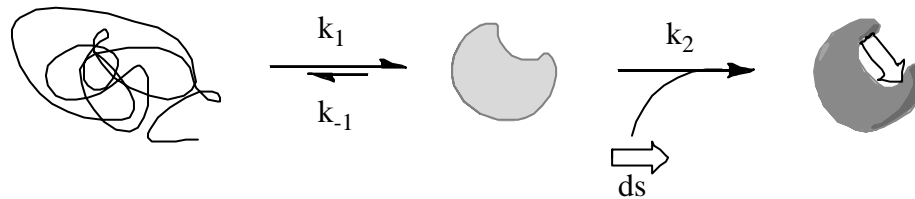


**Figure 4-6: Tryptophan fluorescence refolding traces of various FimG<sub>t</sub>:ds species.**

3μM FimG<sub>t</sub> were refolded in the presence of 5-fold excess of the single ds peptides dsA (green), dsF (red), dsG (blue) and dsI (orange). Folding of free FimG<sub>t</sub> is also shown (black). Final fluorescence intensity was normalized to that of FimG<sub>t</sub> without ds peptide and rate constants as well as relative final fluorescence intensity is summarized in the depicted table.

The traces also showed at least two steps in the folding reaction. By comparison with the folding traces of FimG<sub>t</sub> in the absence of ds peptides it was clear that the first step with a rate of about  $(0.058 \pm 0.004)s^{-1}$  could be attributed to folding of FimG<sub>t</sub> to a intermediate state with the rate  $k_1$ , to which then the respective ds peptide could bind with the rate  $k_2$  (*Figure 4-7*). Upon binding of the peptide the fluorescence signal decreased for dsA, dsF and dsI, with binding of dsF having the strongest influence on the signal, likely due to a very tight interaction, in consensus with the kinetic stability of this complex. The FimG<sub>t</sub>:dsG complex, on the other hand, showed a significant increase of the fluorescence signal upon binding of

the peptide. This indicates that this complex attains a somehow different conformation than the other subunit:ds complexes, although the kinetic stability was shown to be virtually identical to that of the FimG<sub>i</sub>:dsF complex.



**Figure 4-7: Proposed mechanism of subunit:ds complex formation during refolding.**

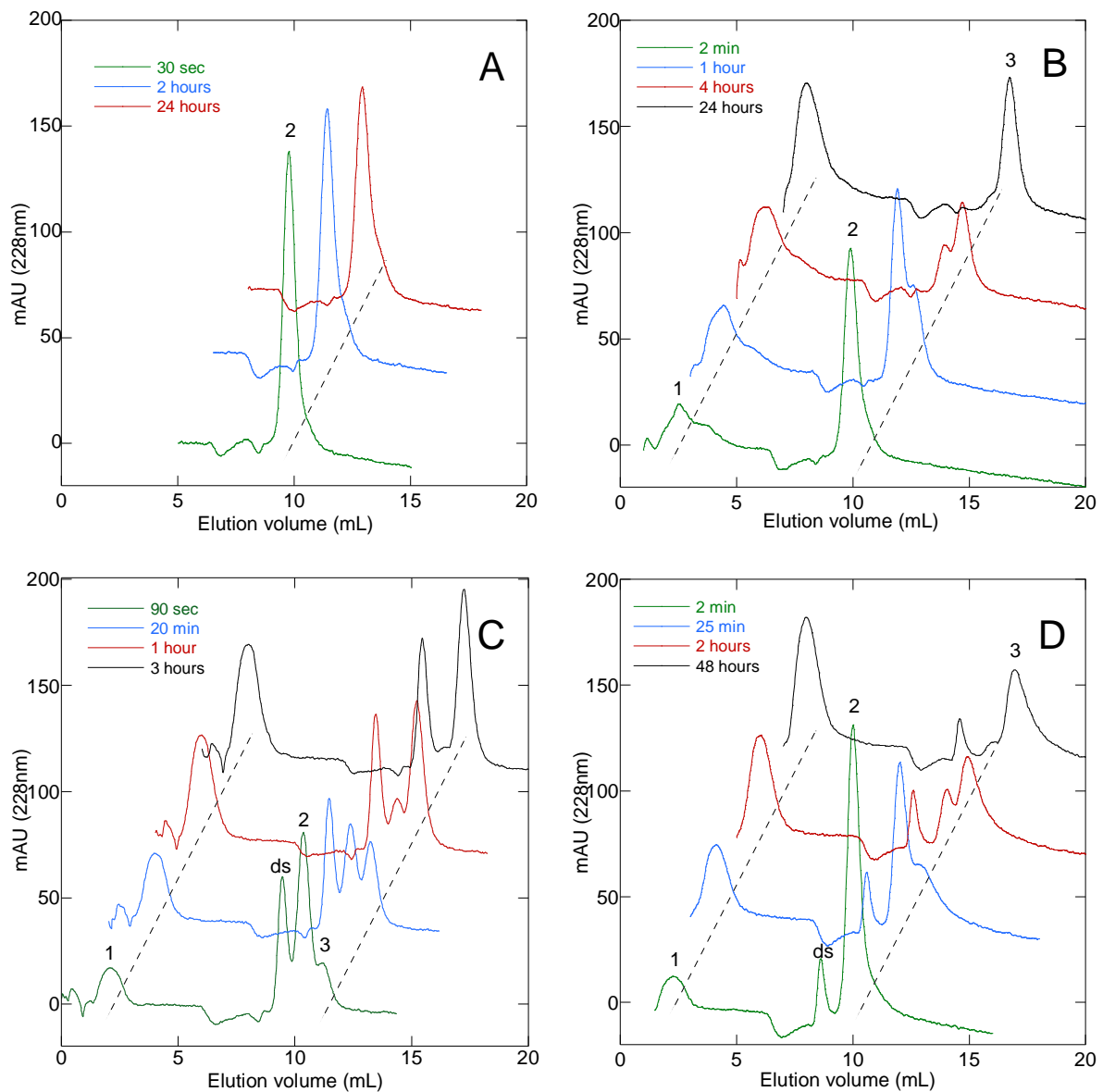
The unfolded protein adopts significant tertiary structure with a rate of  $k_1$  before binding of the ds takes place and the final conformation is attained with the rate  $k_2$ .

## 4.6 Displacement of chaperone from chaperone:subunit complexes by ds peptides

Since ds binding does not happen during subunit folding *in vivo*, but instead the ds is interacting with the completely folded, chaperone bound protein, the interaction between FimC:subunit complexes and the ds peptides was investigated

### 4.6.1 Donor-strand exchange

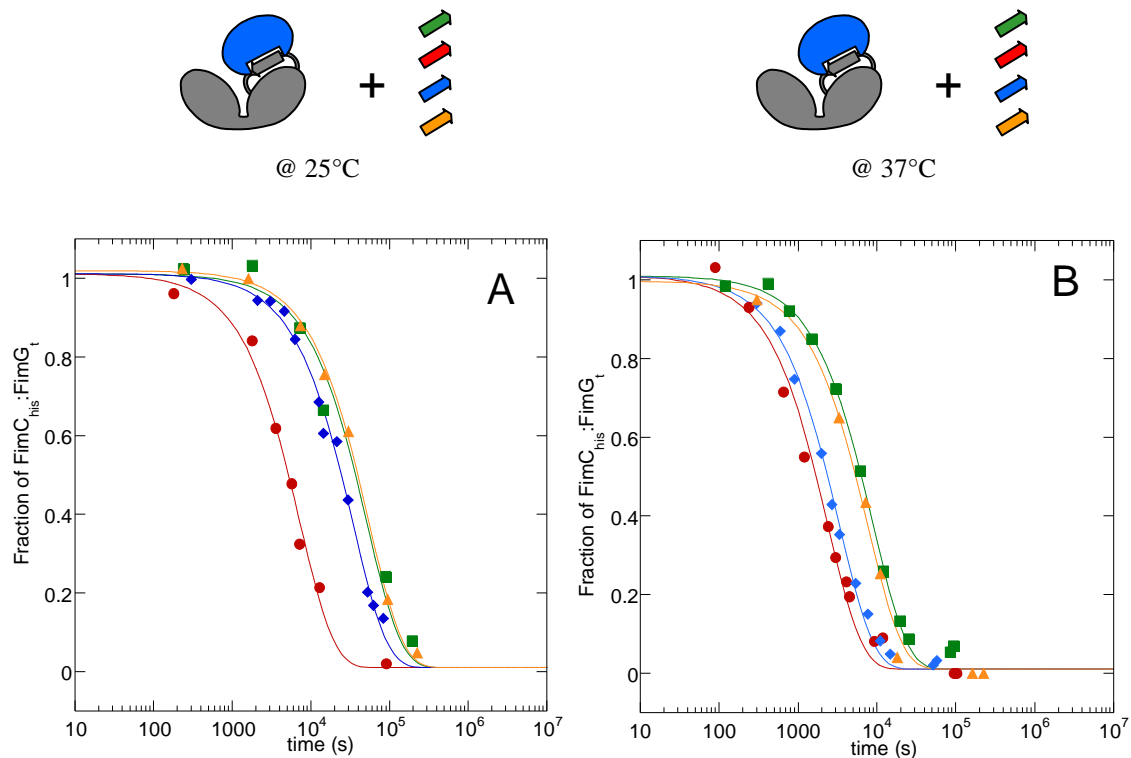
The ability of the donor-strand peptides to displace the chaperone FimC from chaperone:subunit complexes was assayed. To provide fast and accurate quantitation of species present at a given time point in the course of the reaction, analytical cation exchange chromatography was utilized. Conditions were chosen to minimize the peak overlay while still allowing normalization over total protein concentration for at least one component. Unless otherwise noted, disappearance of chaperone:subunit complex and appearance of free FimC was measured and normalized over the total chaperone concentration. Most subunit:ds species and ds peptides did not bind at those conditions, thus not interfering with peak integration and resolution. Typical elution profiles for all FimG<sub>i</sub>:ds complexes analyzed are shown in *Figure 4-8*.



**Figure 4-8: Time-dependence of DSE as resolved by analytical ion exchange chromatography.**

Elution profiles for  $2\mu\text{M}$   $\text{FimC}_{\text{his}}:\text{FimG}_{\text{t}}$  complex incubated at  $37^{\circ}\text{C}$  (A) without donor-strand peptide and with (B)  $20\mu\text{M}$  dsA, (C)  $20\mu\text{M}$  dsF and (D)  $20\mu\text{M}$  dsG. At each time point  $200\mu\text{L}$  of reaction were stopped by addition of  $800\mu\text{L}$  ice-cold running buffer and the mixture applied to a  $1\text{mL}$  Resource<sup>TM</sup> ion exchange column. Elution was done with a gradient of  $0\text{-}300\text{mM}$  NaCl over 20 column volumes. The peaks refer to the  $\text{FimG}_{\text{t}}:\text{ds}$  complex formed (1), the  $\text{FimC}_{\text{his}}:\text{FimG}_{\text{t}}$  complex (2) and the free chaperone  $\text{FimC}_{\text{his}}$  (3).

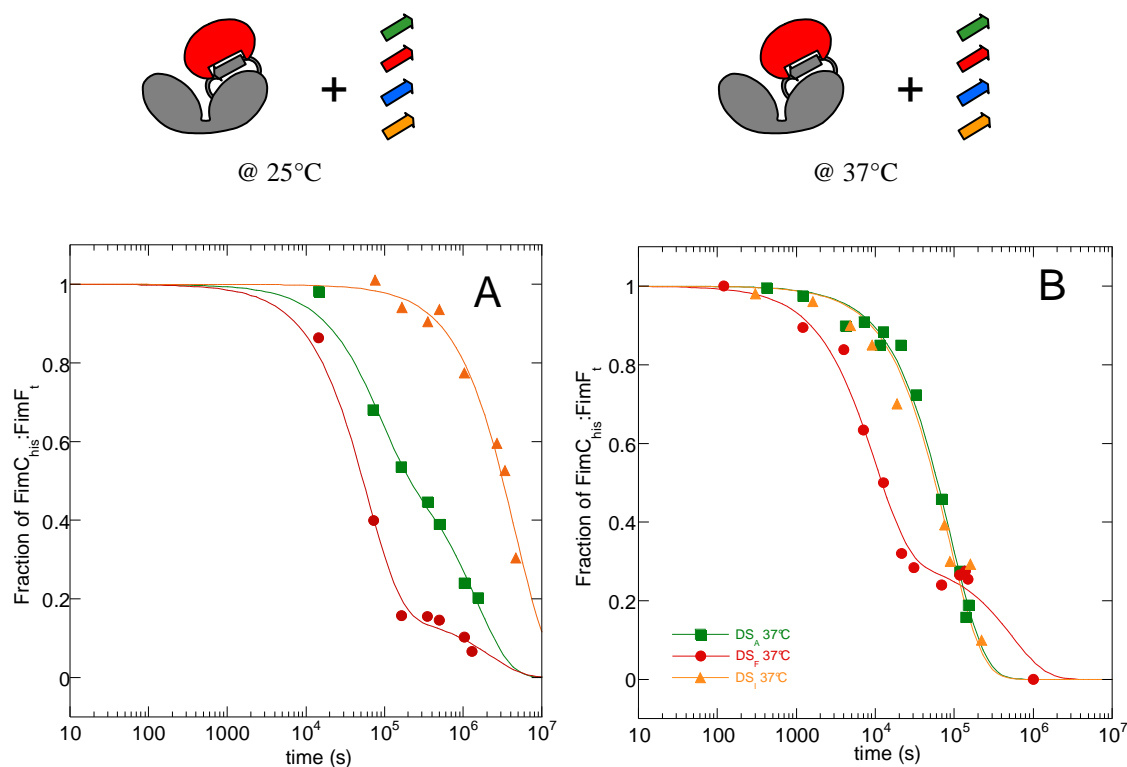
Integrated peak areas were normalized as described and plotted against incubation time. The resulting curves were then fitted to a single-exponential.



**Figure 4-9: DSE for FimC<sub>his</sub>:FimG<sub>t</sub> complex versus ds peptides as measured by analytical ion exchange chromatography.**

(A) DSE of FimC<sub>his</sub>:FimG<sub>t</sub> against dsF (●), dsA (■), dsG (◆) and dsI (▲) at 25°C. (B) DSE of FimC<sub>his</sub>:FimG<sub>t</sub> against dsF (●), dsA (■), dsG (◆) and dsI (▲) at 37°C.

The experiment using the chaperone:subunit complex FimC<sub>his</sub>:FimG<sub>t</sub> demonstrated that in fact the most native-like displacement reaction of FimC<sub>his</sub> by the peptide dsF happened the fastest at 25°C with a half-life of about 84 minutes, followed by the reaction with dsG and dsA with half-lives of about 6.5 and 10 hours, respectively (*Figure 4-9*). To ensure that this is also the case under more physiological conditions, the same experiment was performed at 37°C. Increase in the overall temperature during the DSE reaction from 25°C to 37°C did not result in an alteration of the succession pattern but in a general increase of reaction rates for formation of FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsA and FimG<sub>t</sub>:dsG complexes of 4-, 6- and 10-fold, respectively (Table 4-3).



**Figure 4-10: DSE for  $\text{FimC}_{\text{his}}:\text{FimF}_t$  complex versus ds peptides as measured by analytical ion exchange chromatography.**

(A) DSE of  $\text{FimC}_{\text{his}}:\text{FimF}_t$  against dsF (●), dsA (■) and dsI (▲) at 25°C. (B) DSE of  $\text{FimC}_{\text{his}}:\text{FimF}_t$  against dsF (●), dsA (■) and dsI (▲) at 37°C.

The results for the DSE measurements on the  $\text{FimC}_{\text{his}}:\text{FimF}_t$  complex showed a surprising phenomena. For the reactions of the complex with the donor-strand peptides dsF or dsA, more than one amplitude was observed (*Figure 4-10*). This feature was very dominant for the slow measurements at 25°C, giving ratios of about 80% to 20% of the fast versus the slow phase with rates of  $(1.659 \pm 0.180) \cdot 10^{-5} \text{ s}^{-1}$  and  $(4.598 \pm 3.185) \cdot 10^{-7} \text{ s}^{-1}$ , respectively, for the reaction of  $\text{FimC}_{\text{his}}:\text{FimF}_t$  with dsF and for the reaction of  $\text{FimC}_{\text{his}}:\text{FimF}_t$  with dsA ratios of about 45% to 55% of the fast versus the slow phase were observed with rates of  $(1.241 \pm 0.292) \cdot 10^{-5} \text{ s}^{-1}$  and  $(6.845 \pm 0.133) \cdot 10^{-7} \text{ s}^{-1}$ , respectively. This was also observed for the reaction of  $\text{FimC}_{\text{his}}:\text{FimF}_t$  with dsF at 37°C with ratios of about 70% to 30% of the fast versus the slow phase with rates of  $(1.002 \pm 0.133) \cdot 10^{-4} \text{ s}^{-1}$  and  $(1.871 \pm 0.879) \cdot 10^{-6} \text{ s}^{-1}$ , respectively.

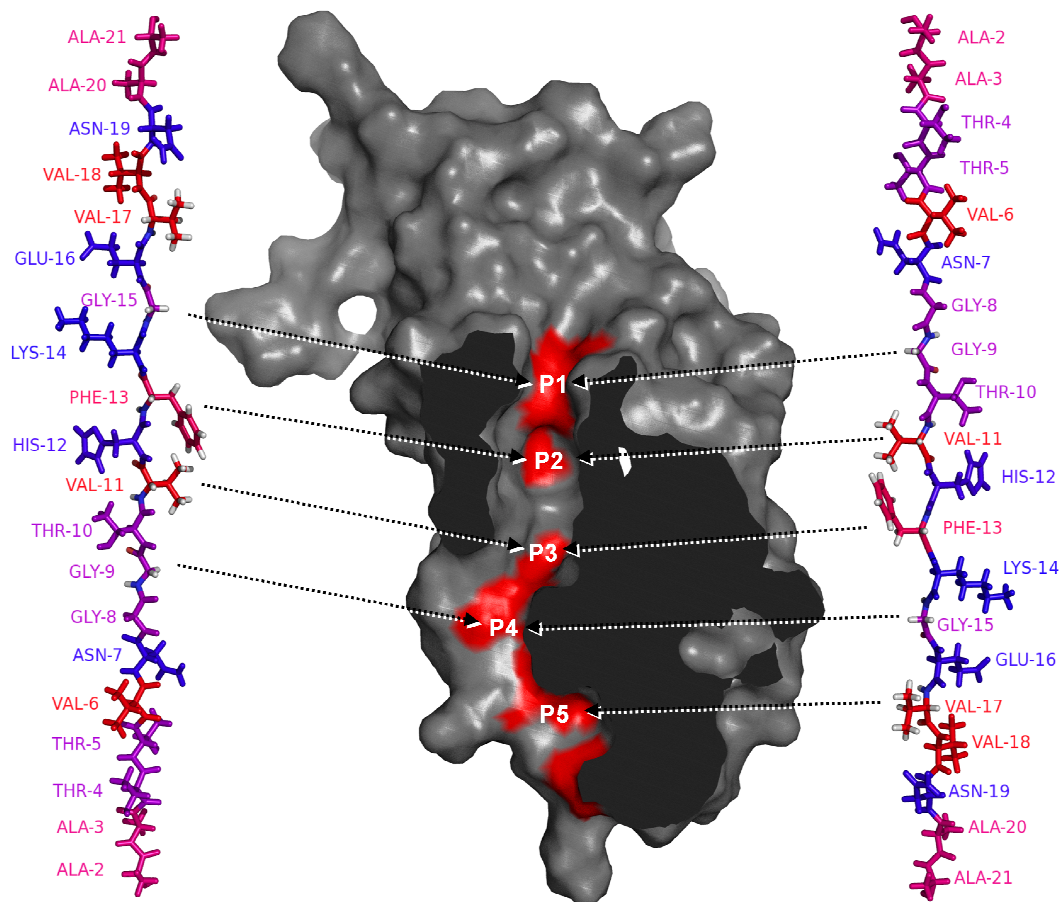
The temperature independent two phases during refolding with dsF are very likely due to precipitation, which could be observed for this species after prolonged incubation.

In case of dsA this effect is most likely due to the possibility of alternative binding modes of the donor-strand peptides to the subunit. Especially dsA is more likely than the other ds peptides to be incorporated in a parallel manner relative to the subunits F-strand instead of the expected anti-parallel orientation (*Figure 4-11*).

**Table 4-3: Parameters of the subunit vs. ds peptide DSE reactions.**

	$k_{obs}^{25^{\circ}C}$ (s <sup>-1</sup> )	$k_{obs}^{37^{\circ}C}$ (s <sup>-1</sup> )	$t_{1/2}^{25^{\circ}C}$	$t_{1/2}^{37^{\circ}C}$
<b>FimC<sub>his</sub>:FimG<sub>t</sub></b>				
<b>dsI</b>	(1.90±0.41)·10 <sup>-5</sup>	(1.32±0.69)·10 <sup>-4</sup>	(10.1±2.2) hr	(87.8±46.2) min
<b>dsA</b>	(1.94±0.38)·10 <sup>-5</sup>	(1.12±0.06)·10 <sup>-4</sup>	(9.94±1.94) hr	(103.4±5.5) min
<b>dsF</b>	(1.37±0.11)·10 <sup>-4</sup>	(4.19±0.32)·10 <sup>-4</sup>	(84.1±6.5) min	(27.6±2.1) min
<b>dsG</b>	(2.94±0.13)·10 <sup>-5</sup>	(2.99±0.14)·10 <sup>-4</sup>	(6.56±0.30) hr	(38.6±1.8) min
<b>FimC<sub>his</sub>:FimF<sub>t</sub></b>				
<b>dsI</b>	(2.16±0.13)·10 <sup>-7</sup>	(1.24±0.38)·10 <sup>-5</sup>	(893.5±54.3) hr	(99.4±19.4) hr
<b>dsA</b>	(1.24±0.29)·10 <sup>-5</sup> <sup>a</sup> (6.85±0.13)·10 <sup>-7</sup> <sup>b</sup>	(1.12±0.04)·10 <sup>-5</sup>	(15.5±3.7) hr <sup>a</sup> (281.3±5.5) hr <sup>b</sup>	(17.2±0.7) hr
<b>dsF</b>	(1.659±0.18)·10 <sup>-5</sup> <sup>a</sup> (4.60±3.19)·10 <sup>-7</sup> <sup>b</sup>	(1.00±0.13)·10 <sup>-4</sup> <sup>a</sup> (1.87±0.88)·10 <sup>-6</sup> <sup>b</sup>	(11.6±1.3) hr <sup>a</sup> (418.7±290.1) hr <sup>b</sup>	(115.3±15.3) min <sup>a</sup> (102.9±48.3) hr <sup>b</sup>
<b>dsG</b>	n/a	n/a	n/a	n/a
<b>FimC<sub>his</sub>:FimA<sub>t</sub></b>				
<b>dsI</b>	-	(8.28±0.38)·10 <sup>-6</sup>	-	(23.2±1.1) hr
<b>dsA</b>	-	(4.23±0.39)·10 <sup>-6</sup>	-	(45.4±4.2) hr
<b>dsF</b>	-	(1.93±0.32)·10 <sup>-6</sup>	-	(99.9±16.6) hr
<b>dsG</b>	-	(2.00±0.48)·10 <sup>-6</sup>	-	(96.1±22.8) hr
<b>FimC<sub>his</sub>:FimH<sub>p</sub></b>				
<b>dsI</b>	-	(1.94±0.41)·10 <sup>-5</sup>	-	(10.1±2.1) hr
<b>dsA</b>	-	<10 <sup>-8</sup>	-	
<b>dsF</b>	-	(2.19±0.16)·10 <sup>-4</sup>	-	(53±3.9) min
<b>dsG</b>	-	(7.24±0.62)·10 <sup>-5</sup>	-	(159±14) min
<sup>a</sup> rate of fast phase FimC <sub>his</sub> :FimF <sub>t</sub> DSE				
<sup>b</sup> rate of slow phase in FimC <sub>his</sub> :FimF <sub>t</sub> DSE				

The one residue which is common to all ds peptides is the glycine residue at the P4 position, which is necessary for the peptide to wrap around the hydrophobic bulge which was shown to be present at this location in all Fim subunit proteins for which structural data is available. So the alignment was based on the presence of a glycine at this position. When this is done with the dsA peptide in the reverse orientation, the residues at the remaining pockets, P1 to P4 and P5 are very similar to the residues in these locations when the dsA peptide is located in the correct orientation, making a different binding pattern to FimF very probable.

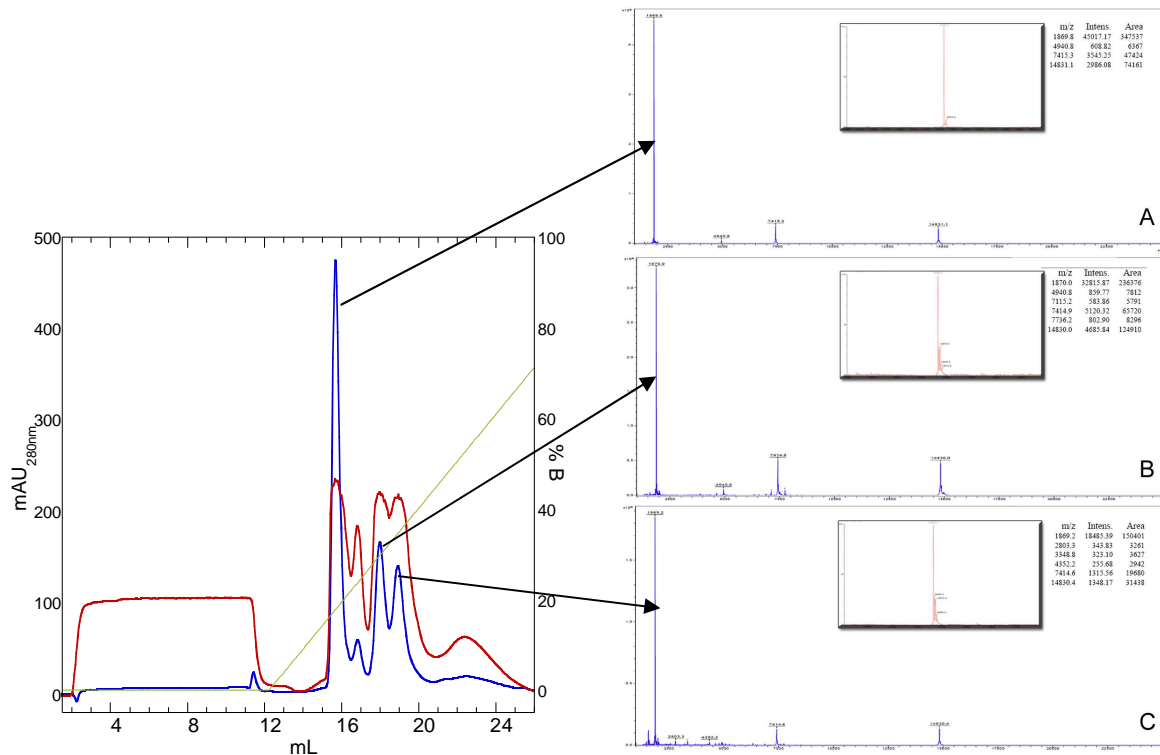


**Figure 4-11: Model of FimF aligned with dsA peptide in different orientations.**

The picture was generated using the FimF<sub>F</sub> NMR-structure (PDB 2JMR, Gossert *et al.*, 2008) and PyMol (DeLano, 2002). On the left side dsA is shown oriented in the non-natural orientation, parallel to FimF's F-strand, on the right side the native-like orientation is depicted. As one can see there is a very good likeliness for dsA to be incorporated in reverse. Residues for the peptides are colored due to their hydrophaticity. Red is hydrophobic, blue is hydrophilic and magenta is amphipathic.

The formation of different FimF<sub>t</sub>:dsA complexes with different elution behavior could be shown by a combination of ion exchange chromatography and mass spectroscopy. When the products from refolding FimF<sub>t</sub> in presence of dsA were purified by an anion exchange resin, multiple species were observed (*Figure 4-12*). To verify that those species were in fact different forms of the same subunit:ds complex samples were analyzed by mass spectroscopy. ESI-MS of all samples showed a single peak in the range from 10-20kDa at 14835Da, consistent with free FimF<sub>t</sub>. MALDI-MS in the low-range showed identical peaks at 1872Da in all samples, indicative of dsA, so there might in fact be different possibilities for dsA to

bind to FimF<sub>t</sub>. This is not relevant *in vivo*, since here the N-terminal extension is fixed and oriented correctly by attachment to its subunit protein.

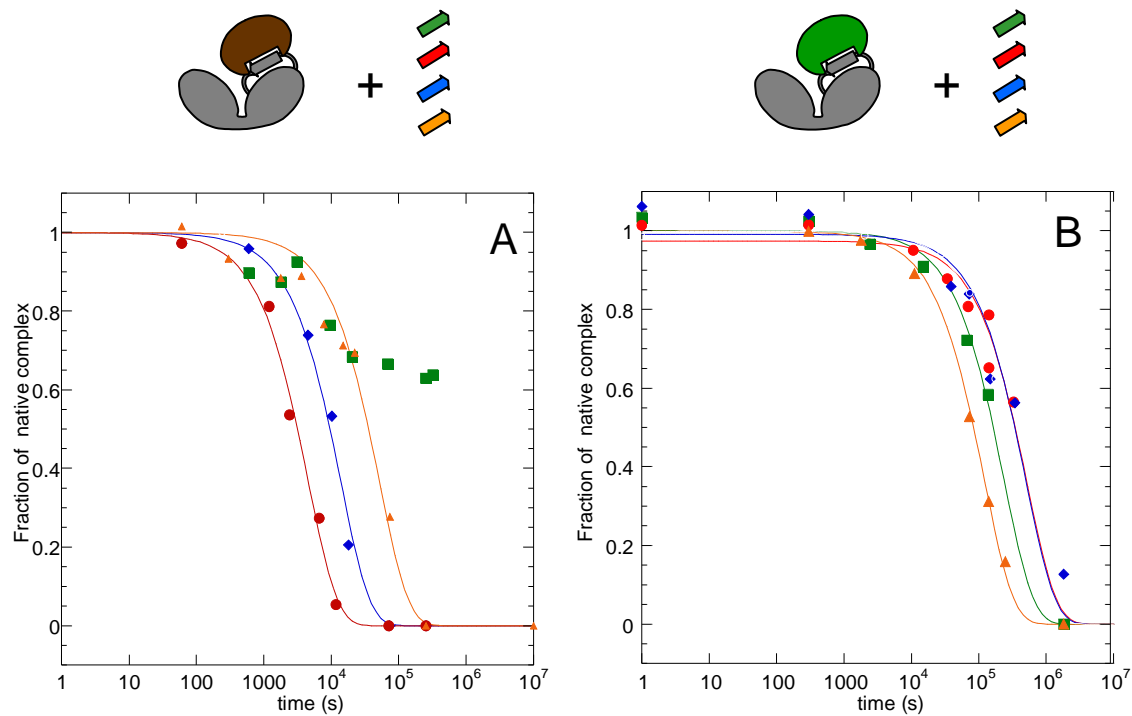


**Figure 4-12: Purification of FimF<sub>t</sub>:dsA by anion exchange chromatography.**

Unfolded FimF<sub>t</sub> was refolded in the presence of a 3-fold excess of dsA in 20mM Tris/HCl, pH8.0, for 20h at RT. Reaction mixture was applied to a ResourceQ™ anion exchange column and bound protein eluted by a gradient of 0-100% of 500mM NaCl. The chromatograph shows aromatic amino acid absorption at 280nm (blue) and peptide bond absorption at 210nm (red) as well as concentration of NaCl (green). Unbound peptide shows in the flowthrough (2-12mL). A, B and C refer to the MALDI-MS and ESI-MS spectra (inset) of the species eluted at 15.5, 18 and 19mL elution volume.

To obtain a full model of the DSE reactions during pilus assembly, the chaperone bound subunits FimH<sub>p</sub> and FimA<sub>t</sub> were also analyzed. FimH<sub>p</sub> is the C-terminal, pilin domain of the adhesin FimH. The complexes FimC:FimH<sub>p</sub> and FimC:FimA<sub>t</sub> were again incubated at 37°C with the same excess of ds peptides as for FimC<sub>his</sub>:FimF<sub>t</sub> and the resulting fractions of chaperon:subunit and free chaperone were determined by analytical ion exchange chromatography as described before. The data was plotted against reaction time (Figure 4-13) and the resulting rates of the time resolved formation of FimH<sub>p</sub>:ds and FimA<sub>t</sub>:ds complexes are summarized in Table 4-3.





**Figure 4-13: DSE reactions of  $\text{FimC}_{\text{his}}:\text{FimH}_p$  and  $\text{FimC}_{\text{his}}:\text{FimA}_t$  complexes vs. ds peptides.**

(A) is logarithmic representation of the time-dependance of the formation of the complexes between  $\text{FimC}_{\text{his}}:\text{FimH}_p$  and the ds peptides dsA (■), dsF (●), dsG (◆), and dsI (▲). (B) is an identical depiction of the reactions between ds peptides and the  $\text{FimC}_{\text{his}}:\text{FimA}_t$  complex.

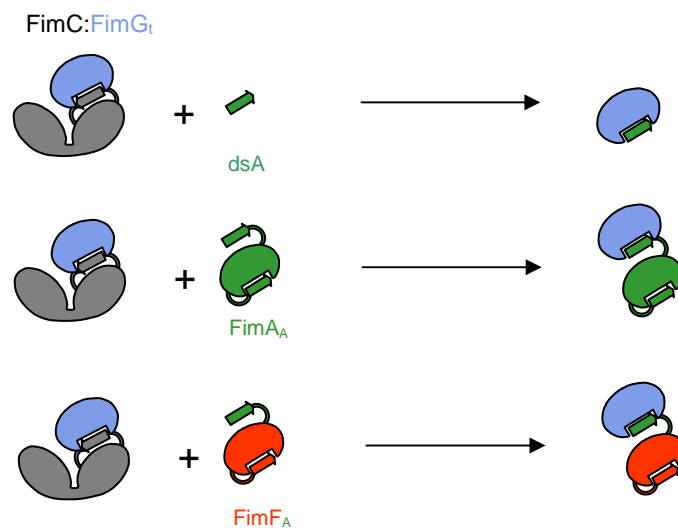
Interestingly, the pilin domain of FimH is forming complexes with the dsF peptide most readily, followed by dsG and dsI. Reaction with dsA was so slow that no rate for the reaction could be determined at physiological conditions and in the absence of accelerating agents. This could be a mechanism to avoid generation of pili lacking the linker proteins between the adhesin and the pilus rod.

The  $\text{FimC}_{\text{his}}:\text{FimA}_t$  complex only underwent very slow donor-strand exchange. On average, all rates were about one order of magnitude slower compared to the other chaperone:subunit complexes. This is not surprising, because even though FimA is the most abundant subunit in the mature pilus, the other subunits have to be incorporated before the structural subunit FimA. Also, the expression level of FimA is noticeable higher compared to the other subunits, increasing the probability of premature FimA incorporation. Additionally it was shown that the usher catalyzes the incorporation of FimA in the pilus significantly (Nishiyama *et al.*, 2008). From all ds peptides dsI is incorporated into  $\text{FimA}_t$  with the highest rate. This is in agreement with the role of FimI as a pilus growth terminator, in which it eventually must compete with other FimA for incorporation at the usher.

#### 4.7 Influence of subunit surface on DSE reaction

With regards to the previously obtained results, the next question that emerged was if the surface of the donor-strand harboring subunit played any role in the DSE reaction.

That is if the presence of protein surface in immediate proximity to the donor-strand either accelerates or inhibits the DSE reaction compared to the reaction with just the donor-strand peptide of a similar sequence. This theory was tested by performing the DSE experiment with both donor-strand peptide as well as a self-complemented, non truncated subunit protein harboring an N-terminal extension of the same sequence as the peptide (*Figure 4-14*). All experimental conditions were kept identical.

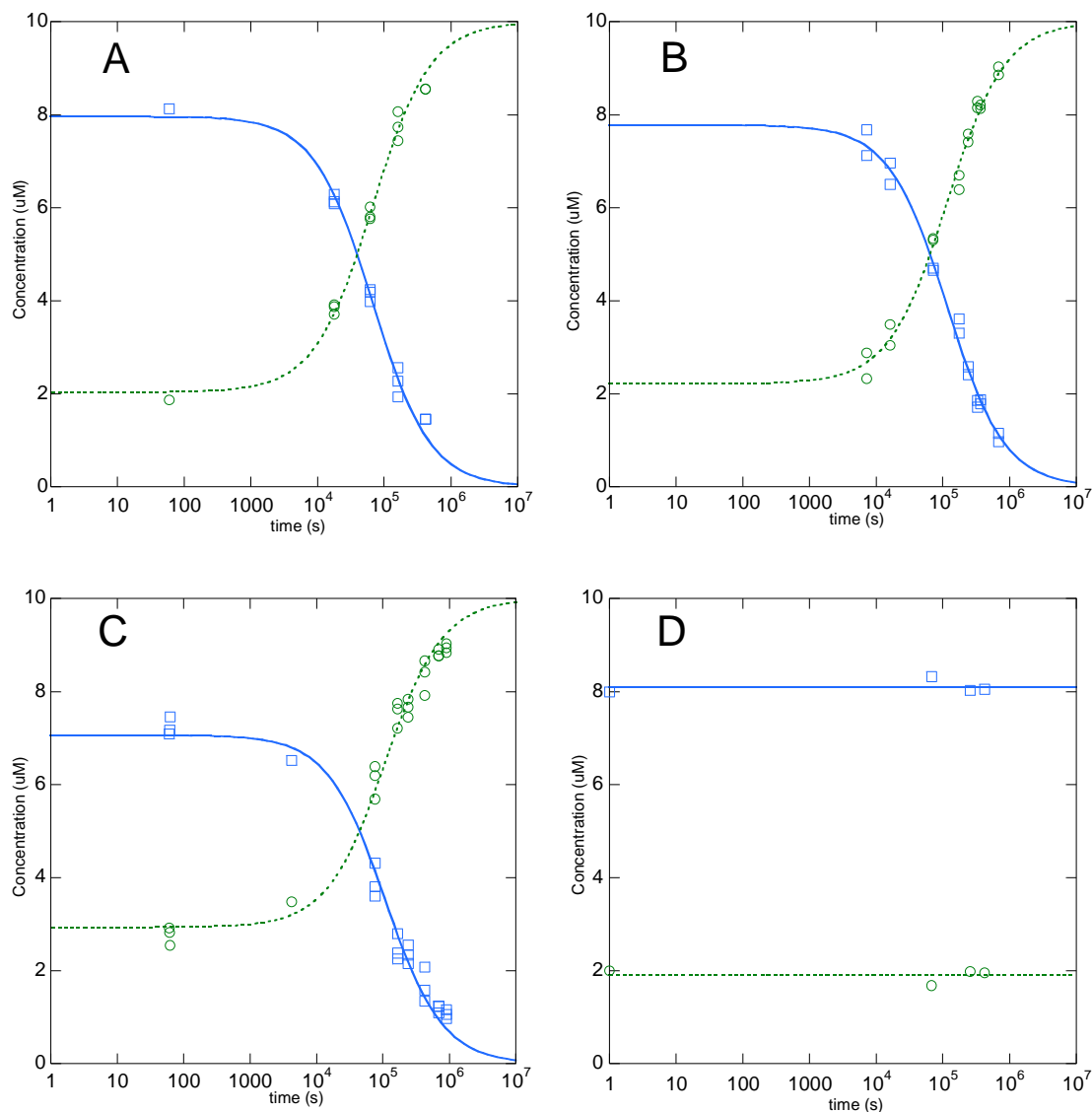


**Figure 4-14: Schematic representation of the surface influence assay.**

The chaperone bound FimG<sub>t</sub> subunit is incubated with equimolar amount of the free ds peptide dsA, the self-complemented FimA variant FimA<sub>A</sub> or the self-complemented FimF variant harbouring the N-terminal extension of FimA, FimF<sub>AF</sub>. When the surface of the subunit to which the ds is attached does not influence the exchange reaction, all constructs should behave the same.

In a first instance, the displacement reaction of the N-terminally his-tagged chaperone FimC<sub>his</sub> from the FimC<sub>his</sub>:FimG<sub>t</sub> complex by either the donor-strand peptide dsA or the self-complemented subunit protein FimA<sub>A</sub> was measured. 10 μM complex were mixed with a 3-fold excess of either dsA or FimA<sub>A</sub> at pH8.0 and samples were incubated at 25°C and the progress of the exchange reaction was measured using analytical cation exchange measurements. Peaks in the elution profiles were integrated and normalized to protein concentration and the results plotted against reaction time (see *Figure 4-15A/B*). The data was

fitted according to a 2<sup>nd</sup>-order reaction, yielding rate constants of  $(1.892 \pm 0.107) \cdot 10^{-6} \text{ M}^{-1} \text{ s}^{-1}$  and  $(1.113 \pm 0.054) \cdot 10^{-6} \text{ M}^{-1} \text{ s}^{-1}$  for the reactions with dsA and FimA<sub>A</sub>, respectively. Although this might indicate a minor acceleration of the process by the presence of the subunit surface, the difference in the rates is not conclusive enough to be certain, hence this minor difference might be due to the fact that the N-terminal extension attached to the FimA subunit is sterically slightly more restricted compared to the free peptide.



**Figure 4-15: Influence of subunit surface on the donor-strand exchange reaction.**

10 μM FimC<sub>his</sub>:FimG<sub>t</sub> complex were incubated with (A) 30 μM dsA peptide, (B) 30 μM FimA<sub>A</sub> and (C) 30 μM FimF<sub>AF4</sub>. Disappearance of FimC<sub>his</sub>:FimG<sub>t</sub> complex (□) and appearance of free FimC<sub>his</sub> (○) are plotted over time. Lines are fitted to the data according to a second-order reaction using KaleidaGraph. (D) is a control of 10 μM FimC<sub>his</sub>:FimG<sub>t</sub> complex incubated without binding partner.

To further investigate the surface influence on the DSE reaction, we turned to the identity of the subunit present, namely to determine if there is a difference in the behavior of the process when a donor-strand extension is attached to a non-natural subunit.

This was tested by generating a variant of the self-complemented subunit FimF<sub>F4</sub> in which the sequence of the N-terminal extension was changed from the natural FimF donor-strand to the sequence of the N-terminal extension of FimA. The resulting construct, FimF<sub>AF4</sub>, was measured toward the FimC<sub>his</sub>:FimG<sub>t</sub> complex in the same manner as before, yielding a 2<sup>nd</sup>-order rate-constant of  $(1.31 \pm 0.12) \cdot 10^{-6} \text{ M}^{-1} \text{ s}^{-1}$ . This compares very close to the rate measured for the more natural FimA<sub>a</sub> protein. In fact, within the error range, both rates appear to be identical, so that we conclude that the presence of subunit protein surface might have a negligible effect on the DSE reaction, most probably due to steric influence on the interaction with the preceding subunit, but is independent of the actual identity of said protein surface.

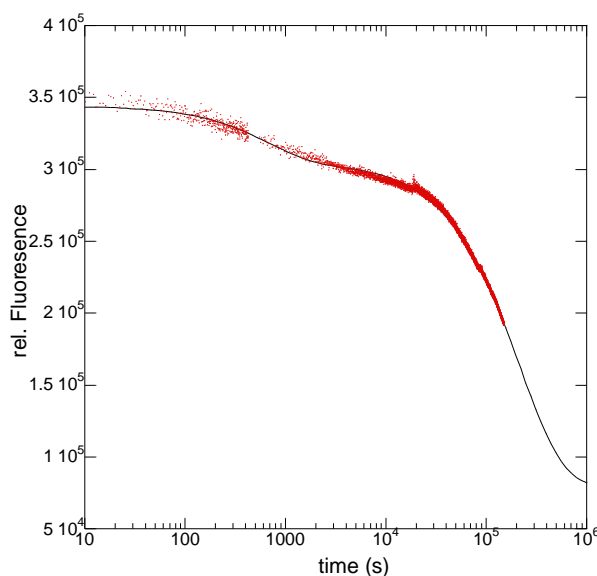
#### **4.8 Competitive ds peptide binding to subunit proteins**

Of special interest is the question, what happens when the subunit is offered more than one binding partner, i.e. ds peptide, at the same time during refolding. Will it select for a specific binding partner, more explicit, will it select for the “correct” binding partner? Or will there just be a statistical distribution of complexes according to the kinetics of the formation of the single complexes? In other words, is ds selectivity controlled by kinetic or thermodynamic means?

Testing this proved to be difficult, since some of the possible complexes, in particular FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG, exhibited very similar biophysical parameters like stability, size and charge. Marker-molecules like fluorescent probes would prove very useful but would also interfere with biological activity in such small systems. Initial tries to separate the possible components on conventional chromatographic media, such as ion exchange or size exclusion columns, proved ineffective due to the above mentioned similarities. Even by reverse phase-HPLC there were difficulties to distinguish between two of the standard complexes used to establish the method.

Another approach was to use unfolding measurements to distinguish between formed complexes, choosing the conditions in a way that it was possible to unfold the various complexes with different rates in a stirred fluorescence cuvette while following the unfolding reaction by tryptophan fluorescence.

The result of this experiment using FimG<sub>t</sub> and a 5-fold excess of each of the ds-peptides dsA, dsF, dsG and dsI is shown in *Figure 4-16*. The unfolding reaction, induced by using 3M GdmCl at pH2.0 shows two phases with apparent rates of  $(1.49 \pm 0.2) \cdot 10^{-3} \text{ s}^{-1}$  and  $(1.49 \pm 0.2) \cdot 10^{-3} \text{ s}^{-1}$  for the fast and slow phase, respectively. The rates correspond to the FimG<sub>t</sub>:dsI and FimG<sub>t</sub>:dsF/dsG complexes (compare 4.3). The amplitudes were corrected for the different intrinsic fluorescence using the factors determined in 4.5.1 and yielded a ratio of 1:7 of FimG<sub>t</sub>:dsI to FimG<sub>t</sub>:dsF/dsG complexes.



**Figure 4-16: Unfolding trace of FimG<sub>t</sub>:ds<sub>x</sub> complexes.**

FimG<sub>t</sub> was refolded in presence of a 5-fold excess of the ds peptides dsA, dsF, dsG and dsI for 48hours. The reaction mixture was transferred into a fluorescence spectrophotometer and unfolding of formed FimG<sub>t</sub>:ds complex induced by addition of 3M GdmCl, pH2.0. Unfolding curves were fitted using rates of unfolding determined by unfolding kinetic measurements. Two amplitudes with rates corresponding to FimG<sub>t</sub>:dsI and FimG<sub>t</sub>:dsF could be observed with values of 38'900cps and 226499cps, respectively. Using the correction factors from *Figure 4-6*, the ratios of FimG<sub>t</sub>:dsI to FimG<sub>t</sub>:dsF/dsG could be calculated to be 1:7.

## **4.9 Structural approach to determine the specificity of the usher for different chaperone:subunit complexes**

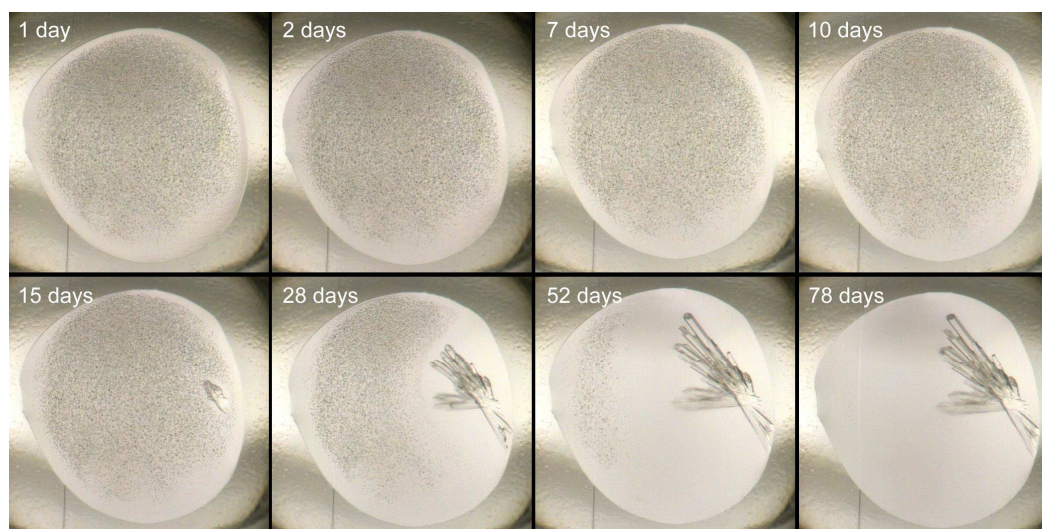
### **4.9.1 Crystallization of binary subunit:ds complexes**

To further investigate the binding of the ds peptides to the subunit and specifically the occupation of the groves in the subunits binding pockets P1 to P5 by different ds peptides, trials to crystallize FimG<sub>t</sub>:dsG and FimG<sub>t</sub>:dsA complexes as well as FimF<sub>t</sub>:dsG and FimF<sub>t</sub>:dsA complexes were performed. The complexes were purified by anion exchange chromatography and used in protein sitting drop vapor diffusion crystallization trials. Unfortunately, all the mentioned complexes exhibited a strong tendency to form aggregates to the point of total precipitation at higher concentrations. FimG<sub>t</sub>:dsG and FimG<sub>t</sub>:dsA complexes could be concentrated to about 8mg/ml and 6mg/ml before complete precipitation, but neither of the complexes was concentrated high enough to be crystallized as shown by HT concentration tests. For the FimF<sub>t</sub> complexes the achieved concentrations were even lower, so that again no crystallization trials were successful.

### **4.9.2 Crystallization of ternary complexes**

#### **4.9.2.1 Crystallization of the FimC:FimF<sub>t</sub>:FimD<sub>N</sub><sup>(1-125)</sup> (CFD) complex**

Initial trials to crystallize the ternary complex between chaperone FimC, subunit FimF and the N-terminal domain of the usher, FimD<sub>N</sub>, were performed with the 6xhis-tagged FimC<sub>his</sub>:FimF<sub>t</sub> complex due to its easy purification using nickel affinity chromatography. However, even though the ternary complex was readily formed, actual protein crystal formation proved unsuccessful, despite some 2000+ screened conditions, including testing various precipitants, protein-to-mother liquor ratios and temperatures. Subsequent trials with the non-tagged FimC:FimF<sub>t</sub> complex proved to be successful, yielding protein crystals at two different conditions using sitting drop vapor-diffusion, one in trials with a reservoir solution of 20% (w/v) PEG8'000 in 50mM succinic acid, pH4.0 and the second utilizing a reservoir solution of unbuffered 20% (w/v) PEG20'000 (see *Figure 4-17*). Crystals grown in the later one were used for X-ray data collection.



**Figure 4-17: Crystallization and crystal formation of the FimC:FimF<sub>7</sub>:FimD<sub>N</sub> complex.**

Sitting drop vapor-diffusion setup in 96well HT screening plates at 4°C. Drops consisted of 100nL reservoir solution and 100nL of protein solution (40mg/ml protein in 20mM Tris/HCl, pH8.0). Reservoir solution was unbuffered 20% (w/v) PEG20'000.

The successfully crystallized protein complex FimC:FimF<sub>7</sub>:FimD<sub>N</sub> (CFD) was subjected to X-ray diffraction at the beam line X06SA at the Swiss Light Source (SLS, Paul Scherrer Institute, Villigen) and the diffraction pattern was recorded using MAR CCD imaging plate. Crystals diffracted to a resolution limit of 1.76Å and in the space group *P3(1)21*. The summary of crystallographic data can be found in *Table 4-4*. The dataset was processed initially with the software HKL2000 (Minor *et al.*, 2006), yielding an electron density map of a resolution up to about 1.76Å.

**Table 4-4: Data collection statistics for FimC:FimF<sub>7</sub>:FimD<sub>N</sub> protein complex**

<i>Data collection parameters</i>	
Space group	<i>P3(1)21</i>
Cell dimensions (Å)	a = b = 126.948, c = 68.755
Wavelength (Å)	0.885616
Resolution range (Å)	33.0–1.76
Total reflections	344,406
Unique reflections (test reflections)	63,367 (1285)
Redundancy	5.4 (4.4) <sup>a</sup>
<i>R</i> <sub>meas</sub> (%)	8.1 (35.4) <sup>a</sup>
Completeness (%)	99.9 (99.3) <sup>a</sup>
<I/σ(I)>	19.6 (3.6) <sup>a</sup>

<sup>a</sup> Outermost resolution shell 1.82–1.76 Å

The structure of the ternary protein complex was then derived from this electron density map using molecular replacement with the software PHASER (McCoy *et al.*, 2005). Here the previously reported ternary protein complex FimC:FimH<sub>p</sub>:FimD<sub>N</sub> (CHD, PDB code 1ZE3, Nishiyama *et al.*, 2005) served as the search model, yielding initial phases for FimC and FimD<sub>N</sub>. Using this information, the software ARP/wARP (Perrakis *et al.*, 1999) was able to assign 96 out of 142 residues of the FimF<sub>i</sub> sequence to the density map. Further refinement was done by cycling the data several times through the model building software COOT (Emsley and Cowtan, 2004) and structural refinement software CNS (Brünger *et al.*, 1998), yielding a nearly complete model. Final refinements were done with the software PHENIX (Adams *et al.*, 2002) and validated with the software MolProbity (Davis *et al.*, 2007) and WHAT\_CHECK (Hooft *et al.*, 1996) (See Table 4-5). The final structure was submitted to the Protein Database (PDB) and assigned the PDB code 3BWU.

**Table 4-5: Data refinement statistics for FimC:FimF<sub>i</sub>:FimD<sub>N</sub> protein complex**

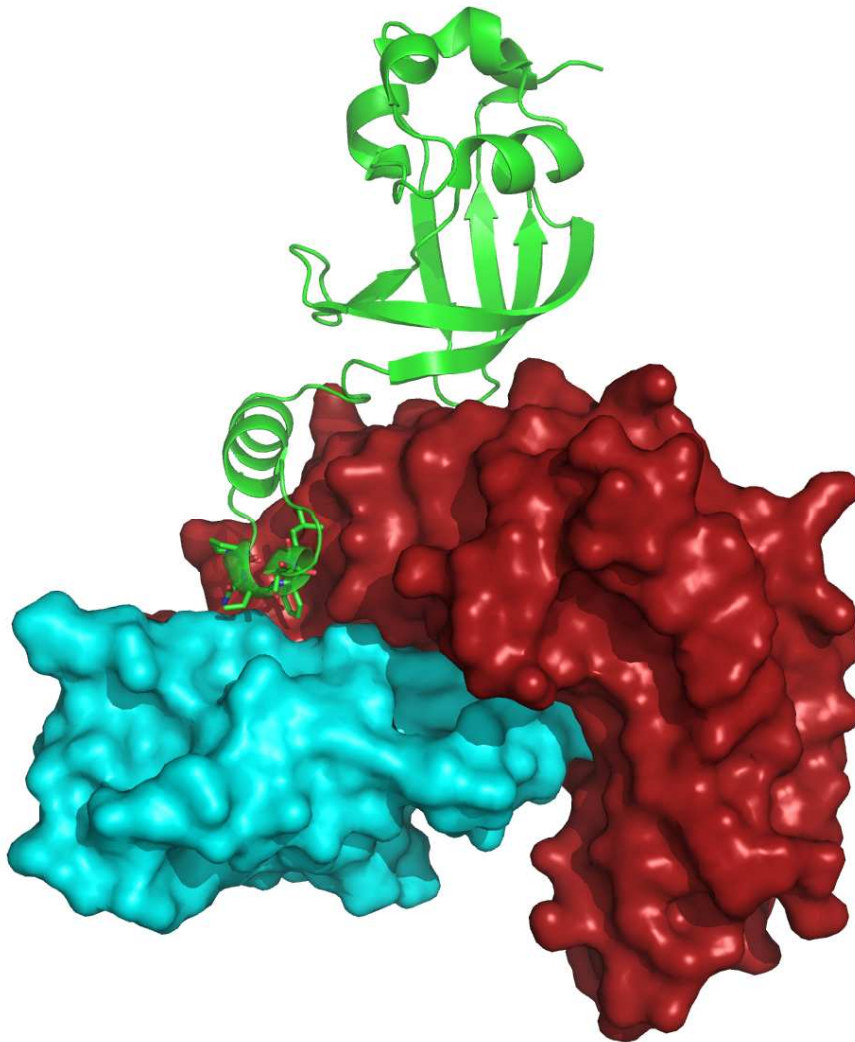
<i>Refinement statistics</i>	
<i>R</i> (%)	17.5
<i>R</i> <sub>free</sub> (%)	21.2
<i>Ramachandran plot</i>	
Favored region (%)	99.1
Number of residues in disallowed regions	0
Overall B-factor (protein + solvent)	23.1
Number of protein atoms	3511
Number of water molecules	641
Number of ligand atoms	59
R.M.S. bond lengths (Å)	0.003
R.M.S. bond angles (°)	0.635

The so obtained protein structure of CFD comprised almost the entire amino-acid sequence of the ternary protein complex. Only residues 26-39 of FimF<sub>i</sub>, 95-99 of FimC and the very terminal residues Gln1 (N-terminus) as well as 123-125 (C-terminus) of FimD<sub>N</sub> could not be accounted for due to the lack of electron density.

As expected, the donor strand of FimC, comprised of the residues 101-107, is bound to the hydrophobic groove of FimF<sub>i</sub>, thus completing a seven-stranded immunoglobulin-like fold of



FimF<sub>t</sub>. Thereby the donor-strand is inserted parallel to the F-strand of FimF<sub>t</sub>, common for all chaperone:subunit complexes, in contrast to subunit:subunit complexes where the ds is incorporated in an anti-parallel orientation relative to the F-strand.

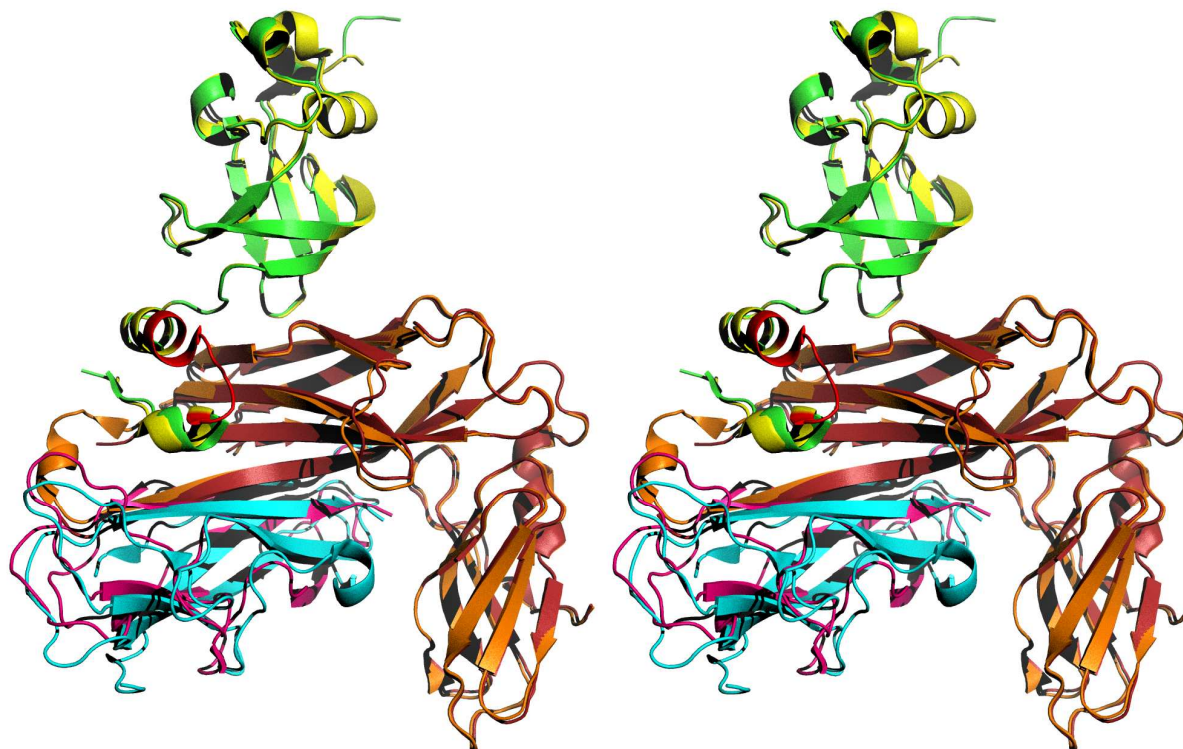


**Figure 4-18: Crystal structure of CFD complex.**

FimC is depicted in red, FimD<sub>N</sub> is green and FimF<sub>t</sub> is shown in cyan. On the N-terminal tail of FimD<sub>N</sub> the residues interacting with the FimC:FimF<sub>t</sub> complex are shown in as stick type. Pictures were made using PyMol (DeLano, 2002).

FimF<sub>t</sub> itself exhibits high similarity to the recently published NMR-structure of the self-complemented variant FimF<sub>F</sub> (PDB code 2JMR, Gossert *et al.*, 2008), with an R.M.S.D. of 1.8Å over the 128 aligned Cα-atoms of both structures.

Compared to the previously reported structure of the ternary protein complex FimC:FimH<sub>p</sub>:FimD<sub>N</sub> (CHD, PDB code 1ZE3, Nishiyama *et al.*, 2005), a high degree of similarity was observed, too. The structures of the proteins FimC and FimD<sub>N</sub>, present in both complexes, exhibit R.M.S.D. values of only 0.7Å and 0.5Å, respectively, while FimF<sub>t</sub> and FimH<sub>p</sub> show a conserved fold with an R.M.S.D. of 2.1Å over 117 aligned Cα-atoms, despite only 20% sequence identity.



**Figure 4-19: Stereo picture of the superimposition of FimC:FimF<sub>t</sub>:FimD<sub>N</sub> complex (CFD) with FimC:FimH<sub>p</sub>:FimD<sub>N</sub> complex (CHD).**

In CFD FimC is depicted in brown, FimF<sub>t</sub> in cyan, and FimD<sub>N</sub> in yellow. In CHD FimC is orange, FimH<sub>p</sub> is magenta, FimD<sub>N</sub> is green. The picture shows that the structures of FimC and FimD<sub>N</sub> in both complexes are very similar. Residues 10–18 of FimD<sub>N</sub> in CFD are depicted in red. All figures were prepared with PyMOL (DeLano, 2002)

The total area of interaction between FimD<sub>N</sub> and the chaperone:subunit complexes is comparable in both the CFD and CHD structures with values of 1070Å<sup>2</sup> and 1290Å<sup>2</sup>, respectively, as calculated using the software PISA (Krissinel and Henrick, 2007). The FimD<sub>N</sub> protein binds to the N-terminal domain of the chaperone FimC, opposite to the chaperones subunit-binding site, similar to what was seen before in the CHD complex. In the CFD complex, the interaction is of mostly polar nature, consisting of a specific salt bridge between residues Asp36 of FimD<sub>N</sub> and Arg66 of FimC, along with five more hydrogen-bonds, which

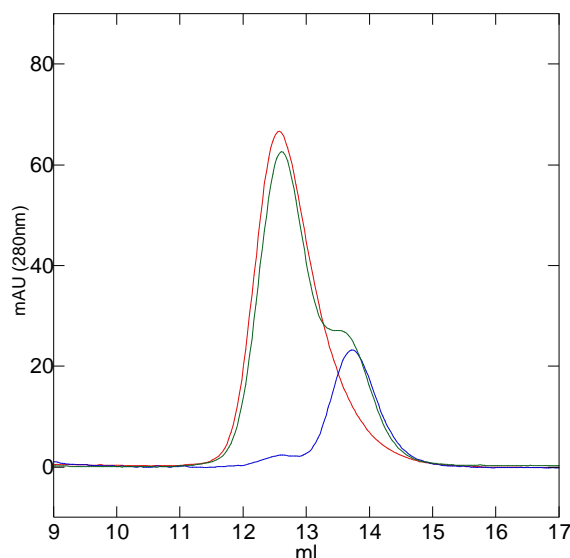
is the same as found before in the CHD structure. Among the interacting residues between FimD<sub>N</sub> and FimC only the residues Arg34, Gln108 and Gln109 of FimD<sub>N</sub> show alternative conformations compared to the CHD complex. While the body of FimD<sub>N</sub> binds to FimC opposite the subunit binding site, the N-terminal tail of FimD<sub>N</sub> (*Figure 4-19*, shown in red) reaches around the chaperone N-terminal domain and makes contact to the chaperone as well as the subunit FimF<sub>t</sub> in the cleft between the two latter proteins. Here the interactions between FimD<sub>N</sub> and FimC are predominantly hydrophobic, with the three aromatic residues Phe4, Phe8 and Phe22 of FimD<sub>N</sub> forming interactions with Leu32 and Ile90 of FimC. Again, the same was observed in the CHD complex. It was also shown before, that mutation of any of these residues resulted in the lack of pilus formation *in vivo* (Nishiyama *et al.*, 2005).

In the CHD structure it was shown that the N-terminal tail of the FimD<sub>N</sub> also was working as a detector for the chaperone-bound subunit, being the only part of FimD<sub>N</sub> actually contacting the subunit in the ternary complex. The same was found to be true for the new CFD complex, although the interface area between usher-domain and subunit is significantly smaller (120 Å<sup>2</sup>) compared to that in the CHD complex (210 Å<sup>2</sup>). Additionally, it appears to be less ordered, with no electron density for the residue Arg7 of FimD<sub>N</sub>, which formed hydrogen bonds to Thr212 and Thr200 of FimH<sub>p</sub> in the CHD structure. Also the CFD structure only shows one hydrogen bond between the residues Asn5 of FimD<sub>N</sub> and Glu82 of FimF<sub>t</sub>, compared to three hydrogen bonds between FimD<sub>N</sub> and FimH<sub>p</sub> in the CHD complex. This is in excellent agreement with the fact that the CFD complex is thermodynamically less stable than the CHD complex, indicating a lower affinity of FimD<sub>N</sub> for the FimC:FimF<sub>t</sub> complex than for the FimC:FimH<sub>p</sub> complex. This difference in affinity is probably the main driving force for the selection mechanism of the usher for chaperone:subunit complexes during pilus assembly.

#### **4.9.2.2 Crystallization of FimC:FimA:FimD<sub>N</sub> (CAD) and FimC:FimG:FimD<sub>N</sub> (CGD) complexes**

Crystallization of other chaperone:subunit complexes with FimD<sub>N</sub> were unsuccessful due to the low affinity of FimD<sub>N</sub> for FimC:FimA and FimC:FimG complexes. The inability to form stable ternary complexes was shown by analytical size-exclusion measurements (*Figure 4-20*). This was also reported earlier using Isothermal Titration Calorimetry (ITC) measurements by Nishiyama *et al.*, 2003.

To overcome the problem of ternary-complex dissociation during crystallization trials, rapid mixture of the components FimC:FimA<sub>t</sub> and FimD<sub>N</sub> directly in the HT 96well plates used for crystallization condition screening was utilized, unfortunately with limited success.



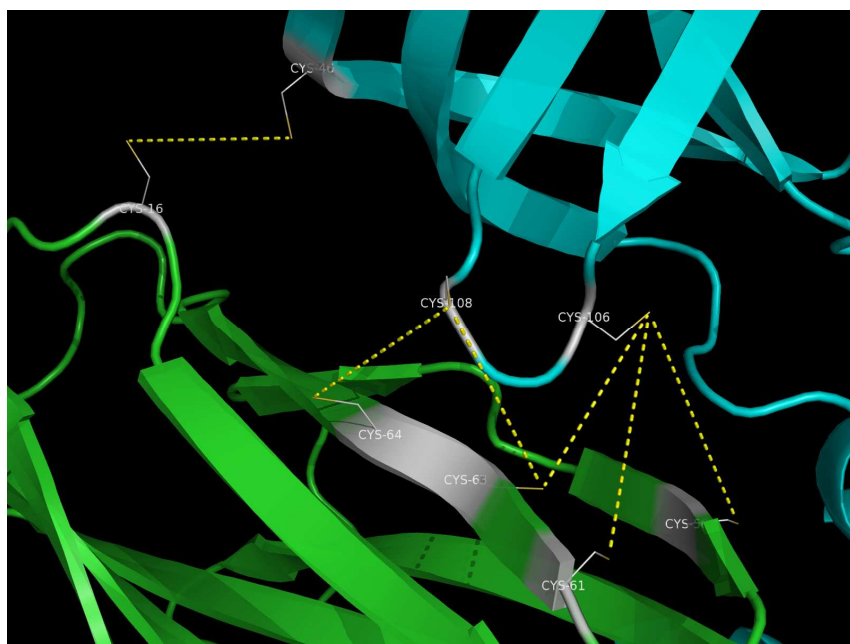
**Figure 4-20: Analytical size-exclusion run of FimC:FimA<sub>t</sub>:FimD<sub>N</sub>.**

Equimolar amounts of FimC:FimA<sub>t</sub> and FimD<sub>N</sub> were incubated o/n and applied to a Superdex75 10/300 SEC column in PBS, pH7.4. As references the single components are also shown. FimC:FimA<sub>t</sub> is red, FimD<sub>N</sub> is blue and the reaction mixture is green. As one can see, no ternary complex was formed.

To stabilize the interaction between the binary complex FimC:FimA<sub>t</sub> and the N-terminal domain of the usher FimD<sub>N</sub> mutants of FimC:FimA<sub>t</sub> and FimD<sub>N</sub> were designed in which the chaperone in the binary complex as well as the ushers periplasmic domain contains an artificial cysteine residue on the interface between chaperone and FimD<sub>N</sub> as modeled by alignment using the already solved structures of the ternary complexes CFD and CHD. By selective oxidation during ternary complex formation a covalent disulfide bond between the components would stabilize the weak interaction and thus allow for crystal formation.

The residues to be mutated to cysteines were chosen based on proximity to the respective other protein. All possible mutations are depicted in *Figure 4-21*. In the usher domain FimD<sub>N</sub> Val106 and Gln108 were chosen and separately mutated to Cys. In the chaperone protein co-expressed with the subunit FimA<sub>t</sub> in the plasmid pfimCfimAt the residues Thr64 and Asn63 were also changed to Cys one at a time by the same method. Although construction and expression of the constructs was successful, no conditions were found to allow selective

oxidation of disulphide bonds at the desired location, resulting in unspecific aggregation and cross-linking and very low yields of complex, unsuitable for crystallization.



**Figure 4-21: Model of possible artificial disulphide bonds in FimC:FimA<sub>t</sub>:FimD<sub>N</sub>(1-125).**

Close-up of the interaction region between FimD<sub>N</sub> (cyan) and FimC (green) in CFD structure. Possible artificial disulphide bonds and necessary cysteine residues to be introduced are depicted. Subsequently, the mutations T64C and N63C were chosen for FimC and V106C and Q108C form FimD<sub>N</sub>(1-125).



## 5 Summary and Discussion

### ***5.1 Natural subunit:ds peptide complexes are kinetically most stable against dissociation***

Unfolding kinetic measurements performed on the complexes between Fim subunit proteins and the ds peptides dsA, dsF, dsG and dsI revealed that the inherent subunit:ds combinations *in vivo* were the most stable species observed. For FimGt the complexes with dsF and dsG showed similar kinetic stabilities with half-lives of unfolding above 4000 years, indicative of the fact that those complexes will not dissociate once formed under physiological conditions during the lifetime of the cell. The same holds true for the FimF<sub>t</sub>:dsA complex.

Even though the other possible complexes, with the exception of FimF<sub>t</sub>:dsG, could also be obtained and were considerably more stable than the chaperone bound subunit proteins, the rate of unfolding of the innate complexes was several orders of magnitude higher. If the selectivity was determined at the subunit level this result can point toward a kinetically controlled mechanism in which the subunit can bind any ds and the non-inherent combinations will be displaced with the inherent one over time. However, since the dissociation of even the least stable complexes would take several hours or even days, this process would be way too slow for the fast pilus assembly observed *in vivo*.

### ***5.2 Natural subunit:ds peptide complexes are formed fastest***

The unfolding kinetics indicated that the selectivity during pilus assembly might be controlled by kinetic stability. Since the dissociation of non-inherent complexes is too slow to account for the fast formation of pili *in vivo*, the designated complex would have to form significantly faster than the non-inherent combinations. This was assessed by interrupted refolding measurement of the complex formation during refolding of the unfolded subunit proteins FimG or FimF<sub>t</sub>. The formation of the inherent complexes FimG<sub>t</sub>:dsF and FimF<sub>t</sub>:dsA was found to be about one order of magnitude faster than the formation of non-inherent complexes, but this difference in velocity is not sufficient to account for correct pilus tip formation since still a considerable amount of unnatural complexes would form. So the specificity for a ds peptide alone is not sufficient to explain the high magnitude of order in the pilus array.

Even though the subunit protein folds only in the periplasmic space where other subunits and thus donor-strands are present, the unfolded protein is captured by the chaperone FimC that

not only assists in folding but also inhibits interaction with other subunits by capping the ds binding groove of the subunit. Therefore, investigation of the binding of ds peptides to chaperone-bound subunit proteins will yield information on the formation in better agreement to the situation *in vivo*. Consequently, the chaperone bound subunit proteins FimG<sub>t</sub>, FimF<sub>t</sub>, FimA<sub>t</sub> and the pilin domain of the adhesin FimH, FimH<sub>p</sub> were incubated with the ds peptides dsA, dsF, dsG and dsI and the formation of subunit:ds complex as well as the release of free chaperone was assayed. A similar pattern to the refolding experiments was observed, alas with less pronounced differences in the rates of formation between desired and non-inherent combinations. Again, at 25°C the FimC<sub>his</sub>:FimG<sub>t</sub> complex underwent the fastest displacement reaction with the ds peptide dsF, though only about 5x faster than with the dsG peptide. At 37°C this difference was even less. For FimC<sub>his</sub>:FimF<sub>t</sub> a difference to the refolding reaction was observed. Instead of the peptide dsA the peptide dsF was bound more readily. This might be due to the fact that the transition state of the FimC<sub>his</sub>:FimF<sub>t</sub> complex and the FimF<sub>t</sub>:dsF complex is very similar, as can be derived from the very similar kinetic m-values determined during the measurement of denaturant dependent unfolding of the complexes. So when the chaperone dissociates from FimF<sub>t</sub>, the subunit is already in a conformation prone to bind the dsF peptide. The main structural subunit FimA<sub>t</sub> in complex with the chaperone showed the highest rate of chaperone displacement with the ds peptide corresponding to the N-terminal extension of the subunit FimI. This is in good agreement with the function of FimI as pilus assembly terminator, since the incorporation of the terminator has to compete with the incorporation of the next structural subunit to avoid infinite pilus growth. The affinity of FimC<sub>his</sub>:FimA<sub>t</sub> toward dsA is consequently found to be about half compared to dsI but higher than the affinity for the remaining ds peptides. Since up to 3000 FimA subunits can be incorporated in single pilus, the FimA subunit must be present in the periplasm in a much greater abundance than FimI, which is only incorporated once. Due to the high excess of FimA over FimI, even though FimI is incorporated more readily, FimA will be integrated at a much higher rate.

The pilin domain of the adhesin FimH also shows high affinity for the dsI peptide. This can be a means of the cell to down-regulate the amount of pili formed on the surface relative to available, primed usher pores. Even though an usher molecule might be activated due to the binding of the adhesin FimH, the binding of the terminator protein directly to the adhesin can suppress pilus growth at this pore and allow other pili to incorporate more FimA subunits and thus forming longer pili.



### **5.3 Binding of ds peptides induces structural changes in subunit**

A first indication that there might be structural differences between the various FimG<sub>t</sub>:ds complexes derived from the unfolding kinetics. The Chevron plots of the denaturant dependent unfolding reactions of the complexes did not show parallel curves, but variation in the slope of the linear fit. The slope of an arm of the Chevron plot, also known as the kinetic *m*-value, is a measure for the conformational similarity of the transition state to either the native or folded protein in a two-state model. A small kinetic *m*-value for the unfolding arm indicates the transition state more closely resembles the folded protein, while a shallower slope of the folding arm is indicative of the transition state being closer related to the unfolded state of the protein. Since the protein itself was always FimG<sub>t</sub> and only the bound peptide differed, there had to be a structural difference in the complex to account for the altered kinetic *m*-values.

Another strong indication of structural differences in the complexes came from the different intrinsic fluorescence intensity of the different complexes. The fluorescence change upon folding of FimG<sub>t</sub> with and without ds peptides present showed a two-step mechanism. The first step can be accounted as the folding of FimG<sub>t</sub> to at least a semi-folded state to which the ds peptides bind. Upon binding of the peptides the fluorescence intensity changed significantly, signifying a change in the micro-environment of the tryptophan residues.

### **5.4 Surface of trailing subunit does not influence binding of ds**

The complex formation of type 1 pilus subunit proteins with the peptides corresponding to the N-terminal extensions involved in the DSC mechanism was investigated fairly extensive. One fact important to consider is the fact that those peptides are taken out of their natural context by not being attached to the harboring subunit any more. To elucidate if this fact has any influence on the overall DSE reaction the reaction of FimC:FimG<sub>t</sub> with the free peptide dsA and the self-complemented FimA<sub>A</sub> variant was directly compared. In experiments where every parameter except the binding partner was kept constant, just a very minor difference in the rates of the DSE reaction were observed. The dsA peptide bound the subunit protein slightly faster than FimA<sub>A</sub>. This might be due to the fact that the subunit-bound ds is sterically slightly more limited than the free peptide. Otherwise the fact that a globular protein was attached to ds did not seem to make a difference. To verify this effect is not an artifact only present when the FimA variant was used, a FimF variant with an N-terminal extension altered to the

sequence of dsA, FimF<sub>AF4</sub>, was used. The rate of the DSE reaction was identical to that of the experiment using FimA<sub>A</sub>. This clearly shows that the identity of the ds bound subunit does not influence the DSE reaction.

### **5.5 Donor-strand complementation is not solely responsible for correct pilus formation**

The question set at the beginning of this dissertation was whether or not the subunits themselves are able to select for their inherent interaction partner. By forming complexes between the subunits FimG<sub>t</sub> and FimF<sub>t</sub> with all possible ds peptides and subjecting the purified complexes to unfolding kinetic measurements, we could show that the innate subunit:ds combinations were the kinetically most stable species. This result is in accordance with the hypothesis, because the kinetic stability against unfolding was so high that the complexes would not dissociate once formed. Since the non-innate subunit:ds combinations were stable enough to inhibit correct pilus formation considerably *in vivo*, the innate complexes would have to form faster than the non-inherent ones to avoid misassembly. The rate of formation of the complexes was assessed by interrupted refolding experiments. Although the formation of the inherent combinations between subunit and ds were found to be faster than the undesired combinations, the differences were not significant enough to avoid wrong subunit incorporation. Since the subunits were only offered one ds peptide at the time during those measurements, the possibility that the subunit would select for the correct binding partner through an unknown mechanism when offered more than one ds peptide at the time had to be investigated. The subunit was refolded in the presence of all ds peptides and the resulting mixture of complexes was purified and qualitatively analyzed by mass spectroscopy. All ds peptides could be detected by the mass spectroscopic measurements, proving that the subunit is not able to actively select for only one ds peptide.

Up to this point the isolated subunit was probed. Nevertheless, in the bacterial cell the subunits are not present as isolated proteins but are bound to the periplasmic chaperone. To assess if the chaperone-bound subunit behaves differently toward the ds peptides, donor-strand exchange (DSE) measurements were performed, in which the FimC:subunit complexes were incubated with the ds peptides and the disappearance of the chaperone:subunit complex was evaluated. The DSE reactions again were in agreement with a preference of the subunit for the inherent ds identity, but were extremely slow, with half-lives in the range of hours to days. Since pilus assembly happens in a timeframe of only minutes there must be an accelerating

factor *in vivo*. This factor was recently revealed to be the assembly platform FimD (Nishiyama *et al.*, 2008). The usher was shown to be able to significantly catalyze the incorporation of FimA subunits into the growing pilus. It is conceivable that the incorporation of the other subunits while also be accelerated at the usher. It has yet to be proven that the usher also actively selects for the correct subunit to be incorporated.

With the rates of formation of the various subunit:ds complexes during DSE as determined above, it is possible to calculate the probability of the formation of mature pilus with the correct order of subunits in absence of any other controlling mechanisms. With the assumption that all chaperone:subunit complexes are abundant in equimolar amounts in the periplasm and by using the rates of formation summarized in *Table 4-3*, the probability of the correct formation of a pilus with the stoichiometry of one FimH, one FimG, one FimF, 500 FimA and one FimI subunit was calculated to be  $1:5.2 \cdot 10^{-292}$ . Since a pilus containing only 500 FimA subunits is considered short, the probability for the average pilus will be even lower. Even if different expression levels of the proteins are taken into account, which were determined to be 5:10:4:100:28 for FimH, FimG, FimF, FimA and FimI, respectively (Schembri *et al.*, 2002), the probability is only marginally higher with a ratio of  $1:2.3 \cdot 10^{-283}$ .

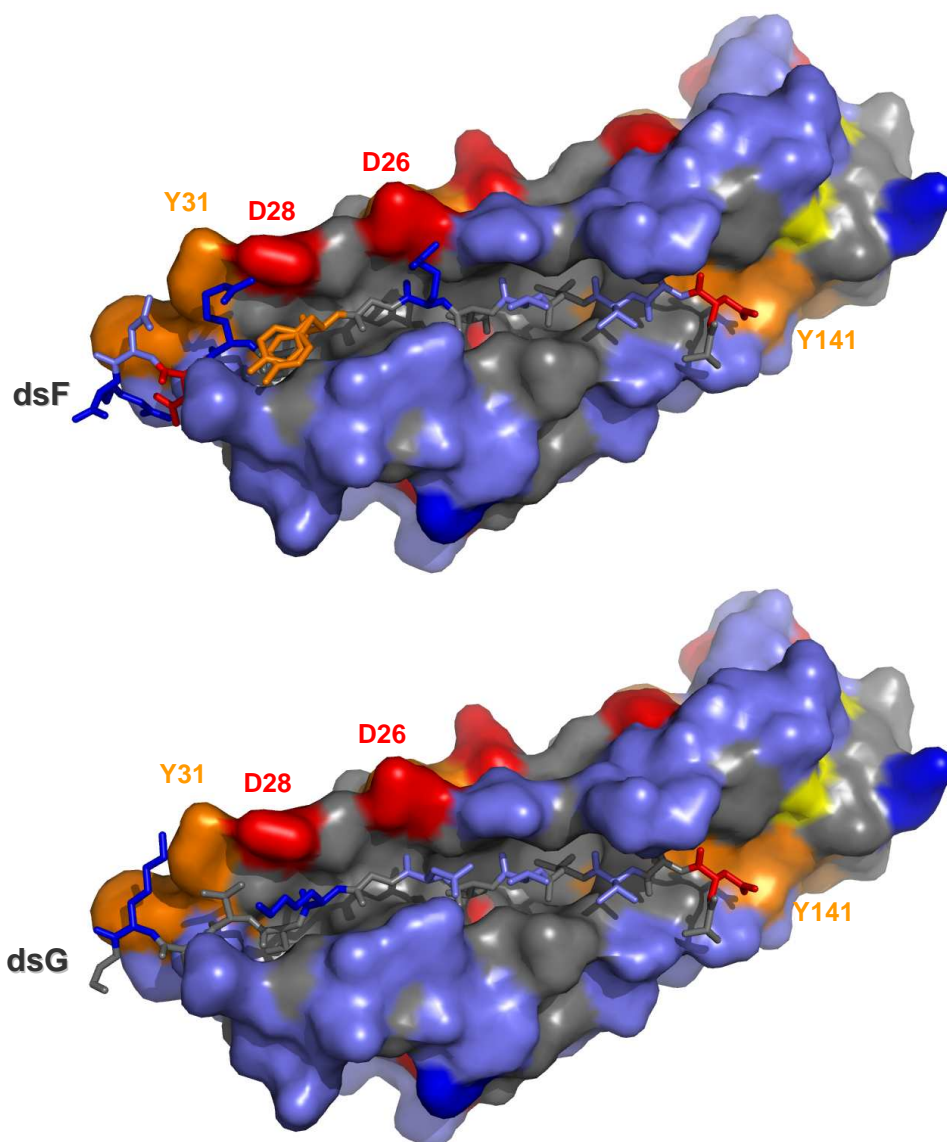
In a related study performed on the Pap pilus system, similar results were obtained (Rose *et al.*, 2008). Even though the authors reason that the DSE process is at least partially responsible for subunit ordering, when the probability for correct formation of a mature P type pilus is calculated as described above, with a stoichiometry of one adhesin PapG, one subunit PapF, five subunits PapE, one linker PapK, 1000 structural subunits PapA and one terminator protein PapH, the probability comes to  $1:3.8 \cdot 10^{-267}$ . From these results it is reasonable to conclude that the minimal mechanism of subunit incorporation at the usher as described in *Figure 3-5* must be expanded by an usher controlled step during the DSE process to assure correct pilus biogenesis, in which the usher must be able to selectively and significantly catalyze only the formation of the respective natural subunit:subunit interaction, similar to what was shown for the incorporation of the first FimA subunit (Nishiyama *et al.*, 2008).

## **5.6 Subunits bind ds peptides with different attractive forces**

During electrospray mass spectroscopy of subunit:ds complexes an unexpected behavior was observed (see Appendix). The complexes FimG<sub>i</sub>:dsF and FimG<sub>i</sub>:dsG, which showed virtual identical stabilities during denaturant induced unfolding measurements, behaved very differently in the mass spectrometer. The FimG<sub>i</sub>:dsF complex was basically undissociable and

peaks in the mass spectrograph corresponded nearly exclusively to the undissociated complex, not to the free peptide or protein. In the case of FimG<sub>t</sub>:dsG, however, no undissociated complex but only the free peptide and protein were observed. Even extensive fine-tuning of the conditions to suppress complex dissociation, like lowering cone voltage, injection gas pressure and even the vacuum did not enable us to observe undissociated FimG<sub>t</sub>:dsG complex. This led to the presumption that the different ds peptides are actually utilizing different modes of binding to the target subunit protein. The main cause of complex dissociation in mass spectroscopy is the high vacuum necessary for unhindered ion transport through the machine. This vacuum lowers the effect of hydrophobic interactions by removing the solvent shell and strengthens the impact of electrostatic interactions by eliminating charged molecules in the environment of the protein molecules. Thus it was speculated that binding of the dsF peptide is more dependent on electrostatic interactions compared to the dsG peptide, whose interaction with the subunit protein is more coined by hydrophobic interactions.

First indications that there is a difference in binding of the different ds peptides were given by observation of folding of the subunit FimG<sub>t</sub> in presence of the different ds peptides (*Figure 4-6*). The fluorescence traces of refolding were recorded at 330nm with an excitation wavelength of 280nm and compared to the refolding trace of FimG<sub>t</sub> in absence of any ds peptide. The first half of the reaction was mostly identical, raising the fluorescence intensity to a common level, indicative of a pathway were the free protein folds to a distinct intermediate and it is not until then that the ds peptide binds to the protein. Upon binding of the peptide, however, the traces showed distinct deviation from one another. With exception of dsG the fluorescence signal decreased upon incorporation of the ds peptide with dsF showing the biggest decline. In the case of dsG the fluorescence signal actually rises considerably. These differences in final fluorescence intensity are strongly indicative of differences in the structural conformation of fully folded and assembled FimG<sub>t</sub>:ds complexes. By looking at the structure of the FimG<sub>t</sub>:dsF complex (PDB Code 3BFQ, Puorger *et al.*, 2008) several charged residues around the ds-binding groove of FimG<sub>t</sub>, wrapping around the bound dsF peptide, could be identified (*Figure 5-1*). In the FimG<sub>t</sub>:dsF complex, the proteins Asp26 and Asp28 residues interact with the Arg8 and Arg12 residues of the peptide, respectively. The peptide residue Tyr10 probably also interacts with Asp28. Additionally there are interactions between the two Tyrosine residues Tyr 141 and Tyr31 with the peptides Asp2 and Asp13 residues formed, resulting in five electrostatic interactions, additionally to the binding force provided by the hydrophobic pockets in the subunit being occupied by the peptides hydrophobic residues.



**Figure 5-1: Interactions between ds peptide and subunit residues in FimG<sub>i</sub>:dsF and FimG<sub>i</sub>:dsG complexes.**

Top row shows the x-ray structure of FimG<sub>i</sub>:dsF (PDB code 3BFQ), bottom row shows the same protein with dsG modeled in, replacing dsF. The amino acid residues are colored by type (red are acidic residues, dark blue are basic residues, light blue are polar residues, gray are non-polar residues, yellow are cysteines and orange are aromatic residues). Pictures were made using PyMol (DeLano, 2002).

In case of the FimG<sub>i</sub>:dsG complex only three of the electrostatic interactions remain: the two tyrosine residues on the opposite ends of the binding pockets can interact with the peptide's Asp2 and Lys14 residues. In the actual groove, only the interaction between Asp28 and the peptide's Lys10 residue can be seen. This structural analysis strongly indicates a difference in

binding of the peptides. This hypothesis is supported by the pH dependence of unfolding of FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG complexes, which yielded transition midpoints of the titration curves which were in agreement with the aspartic acid residues at the binding pocket losing their charge and allowing for easier release of the peptide.

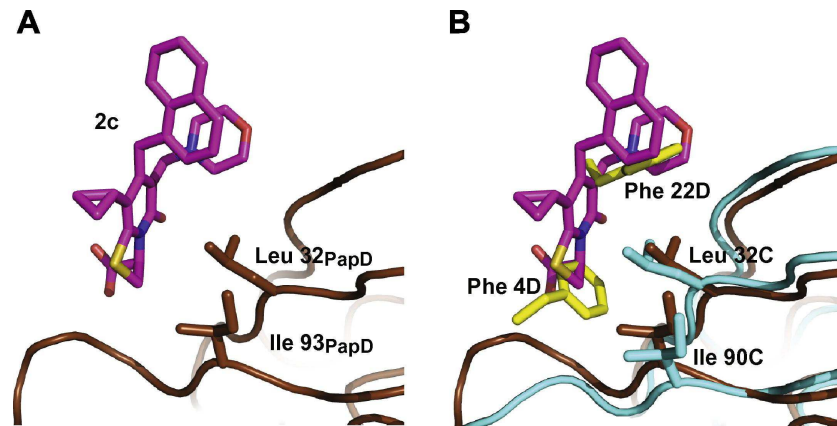
### **5.7 Implications of the structure of the ternary complex between FimC:FimF<sub>t</sub>:FimD<sub>N</sub>**

In both available ternary complex structures the chaperone FimC is the dominant interaction partner of the N-terminal usher domain (contributing 84–88% to the interaction interface), and the structures of FimC and FimD<sub>N</sub> and their relative orientations are almost the same. It is therefore well possible that the N-terminal tail of FimD<sub>N</sub> features an essentially conserved conformation when binding to other chaperone:subunit complexes. Still, the subtle differences observed between CFD and CHD are all found in the small FimD<sub>N</sub>:subunit interface and its influence on binding is not negligible since FimD<sub>N</sub> binds the other two chaperone:subunit complexes, FimC:FimG<sub>t</sub> and FimC:FimA, with at least 10–100 times lower affinity (Nishiyama *et al.*, 2003). Therefore, the conclusion was drawn that the small interface is critical for the affinity of FimD<sub>N</sub> towards chaperone–subunit complexes. Whether the affinity of FimD<sub>N</sub> towards chaperone:subunit complexes determines the order in which pilus subunits are assembled in the pilus fiber is still an open question. However, the low affinity of FimD<sub>N</sub> towards FimC:FimG<sub>t</sub> speaks against such a hypothesis: FimG is in the tip fibrillum and one would expect FimD<sub>N</sub> to show a similar affinity towards FimC:FimG<sub>t</sub> compared to that towards FimC:FimF<sub>t</sub>. In addition, a previous study demonstrated that polymerization of FimA subunits *in vitro* is strongly suppressed in the presence of FimD<sub>N</sub> (Vetsch *et al.*, 2006).

In respect to the above mentioned considerations, it was confirmed that FimD<sub>N</sub> acts as initial binding site of chaperone:subunit complexes at the usher. FimD<sub>N</sub> does however not appear to be involved in promoting donor strand exchange.

The hydrophobic residues Leu32 and Ile90 of FimC appear to be a “hot spot” (Bogan and Thorn, 1998) for the interface to the N-terminal tail of FimD<sub>N</sub>, promoting the conserved binding mode of the flexible tail. These residues are strongly conserved in pilus chaperones. Importantly, a recent study on small molecule drug candidates showed that one compound (2c) bound to the corresponding residues (Leu32 and Ile93) in the homologous chaperone PapD from P pili (*Figure 5-2A*, Pinkner *et al.*, 2006), preventing pilus assembly *in vivo*. This region is structurally equivalent to the one where the FimD<sub>N</sub> tail contacts FimC (*Figure 5-2B*).

The binding interface between N-terminal usher domains and chaperone:subunit complexes is thus likely to be conserved in different pilus systems assembled via the chaperone/usher pathway, and a promising target in the search for pilus assembly inhibitors.



**Figure 5-2: Interaction of small molecular inhibitors with chaperone proteins.**

(A) In the PapD-2c complex (PDB code 2J7L, Pinkner *et al.*, 2006), the pilicide 2c (magenta) interacts predominantly with highly conserved PapD residues Leu32 (PapD) and Ile93 (PapD) (brown). (B) Superposition of the PapD-2c complex (brown and magenta) on CFDPilD (cyan and yellow) shows that the pilicide 2c binds exactly in the FimD<sub>N</sub>:chaperone interface at the position occupied by Phe4D and Phe22D (in yellow) in CFDPilD. Pilicide 2c probably prevents binding of the usher to chaperone:subunit complexes, thereby inhibiting pilus assembly *in vivo*.

## 6 Outlook

We could show that the high degree of order in mature type 1 pili is not an effect of the DSC mechanism alone. The rates of formation of the subunit:subunit interactions indicate that without another control mechanism no correctly ordered pili would form.

Even though we could rule out a control on the level of subunit interaction, the question remains how exactly the bacterial cell controls correct pilus assembly. In light of recent results seems obvious that the usher is actively controlling the order of subunit protein to be incorporated. Naturally, to elucidate the exact mechanism of pilus ordering, one must turn to the usher. By using the new possibilities opened by the *in vitro* reconstitution of the functional assembly platform (Nishiyama *et al.*, 2008) new functions of the usher will be revealed. Here special attention should be focused on the C-terminal domain of the usher, a part of the protein whose function is totally unknown except for the fact that it is essential for the functional usher.

To further explain why the different ds peptides, despite their overall similarities, induce so big differences in the subunit stability, the newly discovered electrostatic locking of some subunits should be further investigated. Site directed mutagenesis of the relevant residues to change those properties would prove very useful.

Finally, more information on the structure of various compounds of the type 1 system, preferably complexes of naturally inherent interaction partners or assembly intermediates at the usher would be a great help in the unraveling of the pilus assembly machinery.



## 7 Materials and Methods

### 7.1 Materials

#### 7.1.1 Chromatographic columns and media

Resource<sup>™</sup> S and Resource<sup>™</sup> Q IEX columns, Supderdex SEC columns and HisTrap<sup>™</sup> MAC columns phenyl sepharose hydrophobic interaction and SepharoseQ FF anion exchange resin were from GE Healthcare (Upsala, Sweden). Ni-NTA medium for histidine selective metal affinity-chromatography was from Qiagen (Hilden, Germany). QA52 anion exchange resin was from Whatman (Kent, UK).

#### 7.1.2 Chemicals

Acrylamide, Bis-Tris, Dithiothreitol (DTT), D(-)- $\alpha$ -Aminobenzylpenicillin sodium salt (Ampicillin), Ethylenediaminetetraacetic acid (EDTA), kanamycine and N,N,N',N'-Tetramethylethylenediamine (TEMED) were from AppliChem (Dortmund, Germany).

Merck KGaA (Darmstadt, Germany) provided ammoniumpersulfate, sodium chloride (NaCl), sodium phosphate and Urea.

3-(N-morpholino)propanesulfonic acid (MOPS), Polyethyleneglycol 8'000 (PEG8'000), Tris(hydroxymethyl)aminomethane (Tris, Trizma<sup>™</sup>) and Polymyxin B sulfate were purchased from Sigma-Aldrich (St. Louis, USA).

Fluka (Buchs, Switzerland) supplied Ammonium acetate, Ammonium sulfate, Polyethyleneglycol 20'000 solution (PEG20'000) and Triton X-100.

Components for bacterial growth-media, namely Bacto<sup>™</sup> Yeast extract, Bacto<sup>™</sup> Tryptone and Agar, were supplied by BD Becton, Dickinson and Company (Franklin Lakes, USA).

And finally the large quantities of guanidine hydrochloride (GdmCl) needed for this study were acquired at NiGu Chemie GmbH (Waldkraiburg, Germany).

Sodium dodecyl sulfate (SDS) was purchased at Carl Roth KG (Karlsruhe, Germany).

### 7.1.3 Oligonucleotides

Oligonucleotide primers were synthesized and HPLC purified by Microsynth (Balgach, Switzerland) in a scale of 0.04 $\mu$ M. Lyophilized oligonucleotides were redissolved in ddH<sub>2</sub>O to a final concentration of about 100 $\mu$ M prior to use.

**Table 7-1: Oligonucleotide primers used in this study**

#	Name	Sequence (5'-3')
1	FD_Ft-HisDel_FW	CGG GCG TAA TGG AAT AAG GAT CAG CTT GGC
2	FD_Ft-HisDel_RW	GCC AAG CTG ATC CTT ATT CCA TTA CGC CCG
3	FD_FtC-seq_340FW	GCC TTG GCA GGT TTA GTT TTG G
4	FD_FtC-seq_770FW	CAA TGC CAC GGC TAC CTT CAC TCT T
5	FD_FtC-seq_795RW	AAG AGT GAA GGT AGC CGT GGC ATT G
6	FD_FtC-seq_1604RW	GTA TCA GGC TGA AAA TCT TCT CTC
7	FD_Ft-MutDel_FW	CCA ACC TGC TTG CAC TTG AAA ATA CGG TGT CAG C
8	FD_Ft-MutDel_RW	GCT GAC ACC GTA TTT TCA AGT GCA AGC AGG TTG G
9	pr_for_fimGt-cyt	CCG GTA CCG GTA GCA TAT GGC CAA ACC GTG TAC GGT TTC CAC C
10	pr_rev_fimGt-cyt	GCA TCA CCT ATA CCT ACA GCT AAG GAT CCG GCC ATC CGG TCG
11	pr_for_fimGtvvhis-cyt	CCG GTA CCG GTA GCA TAT GCA CCA CCA TCA CCA CCA TGT GGT TGC CAA AAC CGT GTA CGG TTT CCA CC
12	pr_rev_fimGtvvhis-cyt	GCA TCA CCT ATA CCT ACA GCT AAG GAT CCG GCC ATC CGG TCG
13	FD_Dn-V106C_FW	CTG CGC ATC TGG ATT GTG GTC AGC AGC GAC
14	FD_Dn-V106C_RW	GTC GCT GCT GAC CAC AAT CCA GAT GCG CAG
15	FD_Dn-Q108C_FW	GCA TCT GGA TGT TGG TTG CCA GCG ACT GAA CCT G
16	FD_Dn-Q108C_RW	CAG GTT CAG TCG CTG GCA ACC AAC ATC CAG ATG C
17	FD_CAT-N63C_FW	GGG AAA AAA AGA GAA TTG CTT ACG TAT TCT TGA TGC
18	FD_CAT-N63C_RW	GCA TCA AGA ATA CGT AAG CAA TTC TCT TTT TTT CCC
19	FD_CAT-T64C_FW	GGG AAA AAA AGA GAA TTG CTT ACG TAT TCT TGA TGC
20	FD_CAT-T64C_RW	GCA TCA AGA ATA CGT AAG CAA TTC TCT TTT TTT CCC
21	FimFAF4_fwd	[Phos]-CGT TCA CTT TAA AGG GGA AGT TGT TAA CGG CTG TAG TGT GGC CGC TGA ATC AA
22	FimFAF4_rev	[Phos]-GTC CCA CCA TTA ACC GTC GTG GCA GCC ATA TGT ATA TCT CCT TCT TAA AGT T

### 7.1.4 Peptides

Peptides were produced by either GeneScript (USA), Thermo Scientific (Germany) or JPT (Germany) and delivered lyophilized in a purity of at least 95%. Prior to use they were redissolved in ddH<sub>2</sub>O or 5% acetic acid (dsI<sup>6</sup>).

**Table 7-2: Properties of donor-strand peptides**

dsA:	AATTVNGGTVHFKEVVNA	pI: 6.79	Mw: 1872.00
dsF:	ADSTITIRGYVRDNG	pI: 6.00	Mw: 1637.70
dsG:	ADVTITVNGKVVAKR	pI: 9.99	Mw: 1570.80
dsI:	GNKWNTTLPGGNMQFQGVIIAE	pI: 6.22	Mw: 2262.50
dsI <sup>6</sup> :	TLPGGNMQFQGVIIAE	pI: 4.59	Mw 1674.90

### 7.1.5 Bacterial strains

For recombinant expression of periplasmic proteins the protease deficient *E.coli* strain HM125 was used Meerman and Georgiou, 1994.

For expression of proteins as inclusion bodies in the cytoplasm the *E.coli* strain BL21(DE3) from Novagen (Germany) was used (Studier and Moffatt, 1986).

BL21(DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm</i> (DE3)
HM125	KS272 <i>degP eda rpoH15</i>
W3110	F <sup>-</sup> <i>lambda<sup>-</sup> IN(rrn<sub>D</sub><sup>-</sup> rrn<sub>E</sub>)I, rph<sup>-</sup>1</i>
NovaBlue <sup>™</sup>	K12 <i>Tet<sup>+</sup> recA<sup>-</sup> endA<sup>-</sup> lacI<sup>q</sup></i>

## 7.2 Methods

### 7.2.1 Properties of proteins used in this study

An overview about proteins and their derivatives discussed in this study is given in *Table 4-1* in *Section 4.1*. The most important properties of the proteins used are summarized in *Table 7-3*.

**Table 7-3: Properties of proteins and protein complexes used in this study**

Protein	M <sub>r</sub> (Da)	# aa	pI <sub>calc</sub>	ε <sub>280nm</sub> (M <sup>-1</sup> cm <sup>-1</sup> )
FimC:FimF <sub>t</sub>	37'490	346	7.90	29'990
FimC <sub>his</sub> :FimF <sub>t</sub>	38'370	353	7.92	31'525
FimC <sub>his</sub> :FimG <sub>t</sub>	37'212	343	6.78	35'995
FimC:FimF <sub>t</sub> :FimD <sub>N</sub> <sup>(1-125)</sup>	51'184	472	5.44	37'610
FimC <sub>his</sub>	23'553	211	10.10	21'620
FimF <sub>t</sub>	14'548	139	5.00	8'370
FimG <sub>t</sub>	13'659	132	4.40	12'210
FimD <sub>N</sub> <sup>(1-125)</sup>	13'700	125	4.41	5'240
FimF <sub>F4</sub>	18'216	172	5.57	11'585
FimF <sub>AF4</sub>	18'451	176	5.86	10'095
FimA <sub>A</sub>	18'024	184	4.87	3'105

### 7.2.2 Design of expression vectors

#### 7.2.2.1 pfimG<sub>t</sub><sup>cyt</sup>

Starting with the template plasmid *pfimGtfimChis*, encoding the co-expression of the 6xHis tagged chaperone FimC<sub>his</sub> with the N-terminally truncated FimG<sub>t</sub>, the nucleotide sequence encoding for the truncated subunit was amplified by PCR using the oligonucleotide primers 9 and 10 (*Table 7-1*). N-terminally introduced *NdeI* restriction site and C-terminally introduced *BamHI* restriction site were used to generate sticky ends on the insert. The receptor plasmid pET11a was linearized with the same restriction enzymes and the insert ligated into the dephosphorylated plasmid. Successful clones were screened using LB agar amp<sup>+</sup> plates and the plasmid was prepared using a Wizard<sup>®</sup> Plus SV Miniprep kit (Promega, USA) and stored in ddH<sub>2</sub>O at -20°C. The correct nucleotide sequence was confirmed by commercial sequencing services (Syngene Biotech AG, Schlieren) using standard T7 primers provided by the service company. The resulting plasmid was termed *pfimGt-cyt*.

### 7.2.2.2 **pfimG<sub>t</sub>vvhis<sup>cyt</sup>**

This construct was generated for easier purification using metal affinity chromatography. The cloning procedure was performed analogous to the generation of *pfimGt-cyt* but using the oligonucleotide primers 11 and 12 instead of primers 9 and 10 for the insert generation. The oligonucleotide primer 11 introduces an N-terminal 6xHis tag linked to the N-terminus of FimG<sub>t</sub> by a ValVal linker. The resulting plasmid was termed *pfimGtvvhis-cyt*.

### 7.2.2.3 **pfimFt-fimC**

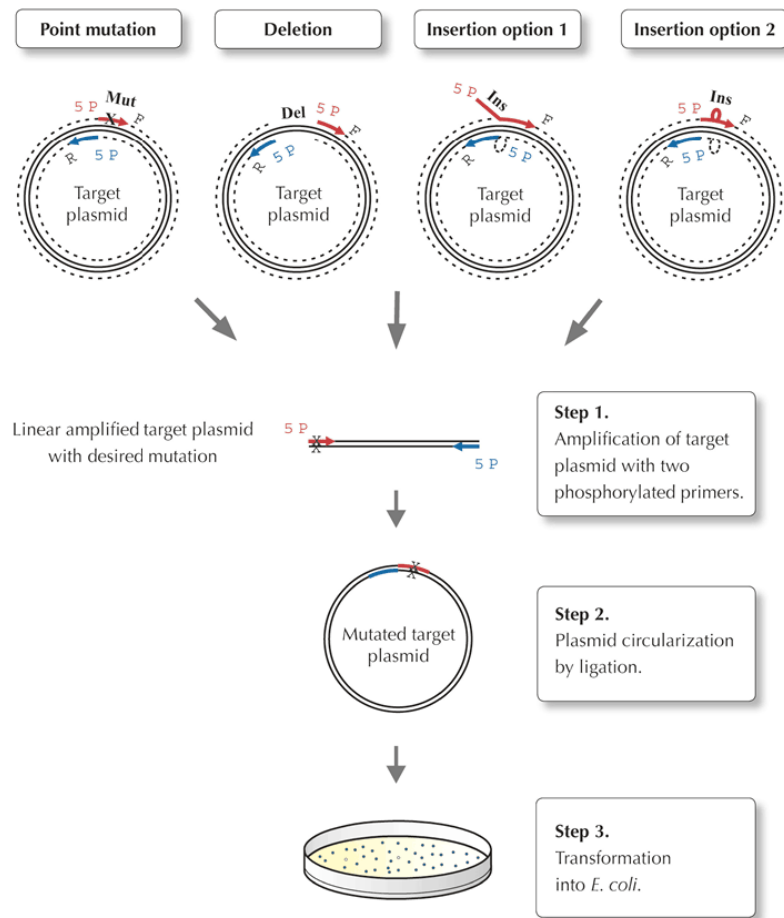
The expression plasmid *pfimFt-fimC* was generated by deletion of the C-terminal 6xHis tag on the chaperone from the template plasmid *pfimFt-FimChis* (Nishiyama *et al.*, 2003). This was done by QuikChange<sup>™</sup> following the standard protocol supplied and utilizing the oligonucleotide primers FD\_Ft-HisDel\_FW and FD\_Ft-HisDel\_RW (see Table 7-1). Successful clones were detected by Ampicillin resistance on LB-agar amp<sup>+</sup> plates and the plasmid was prepared using a Wizard<sup>®</sup> Plus SV Miniprep kit (Promega, USA) and stored in ddH<sub>2</sub>O at -20°C. Correct nucleotide sequence was confirmed by commercial sequencing services (Syngene Biotech AG, Schlieren) using primers 3-6 (Table 7-1).

### 7.2.2.4 **pfimFAF<sub>4</sub>**

#### 7.2.2.4.1 **PHUSION<sup>™</sup> Site-directed Mutagenesis**

Using the PHUSION<sup>™</sup> Site-directed Mutagenesis system, one is able to introduce and/or delete large arrays of nucleotide sequence in one single PCR-reaction. This is achieved by utilizing modified oligonucleotide primers, which exhibit an N-terminal phosphorylation and are designed to run away from each other, thus synthesizing the whole template plasmid plus/minus the desired modification. The resulting linear construct is then ligated using T4 ligase and the phosphorylated 5'-ends of both strands. The resulting linearized plasmid is then transformed in any suitable bacterial host and bacterial colonies are screened for the desired mutation.

Mutation efficiency is claimed to be around 80% and plasmids up to 10kbp can be used as templates.



**Figure 7-1: Principle of the PHUSION™ Site-Directed Mutagenesis (© Finnzymes)**

#### 7.2.2.4.2 Construction of *pFimFAF<sub>4</sub>*

Using the PHUSION™ Site-directed Mutagenesis kit (Finnzymes, Espoo, Finland) the N-terminal extension of FimFF<sub>4</sub>, a self-complemented variant of FimF carrying its own N-terminal extension as an artificial C-terminal extension connected to the protein by a 4 amino acid linker, encoded by the template plasmid *pfimFF4* (Gossert *et al.*, 2008) was changed to the N-terminal extension corresponding to that of native FimA. The artificial, C-terminal extension corresponding to the usually N-terminally located extension of FimF, introduced to inhibit polymerization and facilitate proper folding of the protein, was left unchanged.

Since the signaling sequence for the secretory pathway was already removed in the template plasmid, the protein is expressed in the hosts' cytoplasm.

To make the relatively extensive change of inserting 54 nucleotides into the plasmid, the oligonucleotide primers were designed in the way that the sequence which was to be inserted

---

was split roughly equally over both primers, the forward primer having a 30 nucleotide 5'-terminal extension and the reverse primer having a 24 nucleotide 5'-terminal extension. Both oligonucleotides were commercially phosphorylated at their respective 5'-ends. Actual mutagenesis was performed according to the protocol supplied with the PHUSION™ kit. The resulting vector was transformed chemically into NovaBlue™ competent cells (Novagen, Merck, Germany) and the plasmid was prepared using a Wizard® Plus SV Miniprep kit (Promega, USA) and stored in ddH<sub>2</sub>O at -20°C.

## 7.2.3 Expression and purification of proteins

### 7.2.3.1 Media:

Except for the 5052-Stock solution all media are autoclaved at 121°C for 20min and stored at room temperature (RT).

• 2YT (1L):	16g	Tryptone		5g	Yeast extract
	10g	Yeast extract		10ml	1M NaCl.
	5g	NaCl		2.5ml	1M KCl.
• LB (1L):	10g	Tryptone	10ml	1M MgCl <sub>2</sub>	
	5g	Yeast extract	10ml	1M MgSO <sub>4</sub>	
	10g	NaCl	20ml	1M Glucose	
• LB Agar (1L):	10g	Tryptone	• ZY (1L):	10g	Tryptone
	5g	Yeast extract		5g	Yeast extract
	10g	NaCl	• 50x 5052-Stock:	250g	Glycerol
	15g	Agar		25g	Glucose
• SOC (1L):	20g	Tryptone		100g	α-Lactose



### 7.2.3.2 Periplasmatic expression

All plasmids encoding for protein expression in the bacterial periplasm were transformed chemically into chemically competent cells of the protease-deficient *E.coli* strain HM125 (Meerman and Georgiou, 1994). Using the plasmids antibiotic resistance towards Ampicillin, successful transformation was verified by colony selection on LB-Agar plates containing 100µg/ml Ampicillin. Cultures containing the plasmid were successively grown at 30°C in 2YT medium containing the same amount of antibiotic. Large scale expression was done in 6x1.6L format until an optical density at 600nm of 0.7 was reached. Cultures were then induced by addition of 1mM IPTG. The cells were harvested by centrifugation after further incubation for 4h at unchanged conditions. The proteins were extracted from the periplasm by chemical lysis of the outer membrane by incubation for 2h at 4°C in 50mM Tris/HCl, pH7.5 containing 150mM NaCl, 5mM EDTA and 1mg/ml of polymixin-B-sulfate. Immediately afterwards the cell debris was removed by ultracentrifugation and the clear supernatant was dialyzed against the appropriate purification buffer to remove the antibiotic.

#### 7.2.3.2.1 *FimC:FimF<sub>t</sub>*

Buffer:

- Q-buffer (20mM Tris/HCl, pH8.0)
 

	A	B
Trizma	20mM	20mM
NaCl	---	500mM
  
- S-buffer (20mM Bis/Tris/HCl, pH5.9)
 

	A	B
Bis/Tris	20mM	20mM
NaCl	---	500mM
  
- PBS (pH 7.4)
 

	1L
NaH <sub>2</sub> PO <sub>4</sub> •H <sub>2</sub> O	(10mM)
Na <sub>2</sub> HPO <sub>4</sub> •7H <sub>2</sub> O	(10mM)
NaCl	140mM

After expression of the *FimC:FimF<sub>t</sub>* complex in the periplasm of *E.coli* HM125 (see above), the extracted proteins were dialyzed against 5L of Q-Buffer A at 4°C over night (o/n). Then the solution was applied to a QA52 anion-exchange resin to bind contaminating protein. Upon removal of the resin by centrifugation the supernatant was dialyzed against 5L S-buffer A at 4°C o/n and then applied to a ResourceS™ cation-exchange column. The protein was eluted

using a gradient of 0-100% S-Buffer B over 20 column volumes (CV). To even further enhance the purity of the sample for crystallization experiments, size-exclusion chromatography (SEC) was performed on a Superdex75 26/60 column, equilibrated in PBS. Purity of the sample was then verified by SDS-PAGE and the concentration determined by UV absorption at 280nm and subsequently adjusted to approximately 200 $\mu$ M for storage at -80°C. The final yield was 4mg of pure protein per Liter of bacterial culture.

#### 7.2.3.2.2 *FimC<sub>his</sub>:FimF<sub>t</sub>*

- Ni-NTA buffer (50mM NaP<sub>i</sub>, pH8.0)
 

	A	B
NaP <sub>i</sub>	50mM	50mM
NaCl	300mM	300mM
Imidazol	---	250mM
- S-buffer (20mM MOPS/NaOH, pH6.7)
 

	A	B
MOPS	20mM	20mM
NaCl	---	300mM

The periplasmic extract was dialyzed against 5L of Ni-NTA buffer A o/n at 4°C. Afterwards the extract was applied to a Ni-NTA column equilibrated in Ni-NTA buffer A and the column was washed until the UV absorption at 280nm showed a stable baseline again. After another washing step with 8% Ni-NTA buffer B (=20mM imidazole) the His-tagged protein complex was eluted by a step gradient using 20%, 40%, 80% Ni-NTA buffer b (= 50mM, 100mM, and 200mM imidazole, respectively). Again, each step was performed until a stable baseline was achieved again. Fractions containing the desired proteins were determined by SDS-PAGE and the pooled fractions dialyzed o/n against 5L of S-buffer A before applied to a ResourceS™ 6mL cation-exchange column, also equilibrated in S-buffer A. The final yield was 10mg of pure protein per Liter of bacterial culture.

#### 7.2.3.2.3 *FimC<sub>his</sub>:FimG<sub>t</sub>*

- Ni-NTA buffer (50mM NaP<sub>i</sub>, pH8.0)
 

	A	B
NaP <sub>i</sub>	50mM	50mM
NaCl	300mM	300mM
Imidazol	---	250mM
- S-buffer (20mM MOPS/NaOH, pH6.7)
 

	A	B
MOPS	20mM	20mM

NaCl	---	200mM
• PBS (pH 7.4)		
	1L	
NaH <sub>2</sub> PO <sub>4</sub> •H <sub>2</sub> O	(10mM)	
Na <sub>2</sub> HPO <sub>4</sub> •7H <sub>2</sub> O	(10mM)	
NaCl	140mM	

After dialyzing the periplasmic extract against 5L NiNTA-buffer A o/n, the solution was applied to a NiNTA column of an adequate size. Impurities were removed by washing the column in 8% NiNTA-buffer B (20mM imidazole) and the His-tagged protein complex was eluted via a gradient of 8-100% NiNTA-buffer B over 20 CV. After removal of imidazole by dialysis against 5L S-buffer A o/n, cation-exchange chromatography was performed on a ResourceS™ 6mL column. Proteins were separated by a gradient of 0-100% S-buffer B over 25CV and the fractions corresponding to FimC<sub>his</sub>:FimG<sub>t</sub> were pooled and concentrated to a volume of max. 10mL before being applied to a Superdex75 26/60 SEC column equilibrated in PBS. The final yield was 4.6mg of pure protein per Liter of bacterial culture.

#### 7.2.3.2.4 *FimD<sub>N</sub>*<sup>(1-125)</sup>

As described earlier (Nishiyama *et al.*, 2003)

• Q-buffer (20mM Bis/Tris/HCl, pH6.2)		
	A	B
Bis/Tris	20mM	20mM
NaCl	---	500mM
• PBS (pH 7.4)		
	1L	
NaH <sub>2</sub> PO <sub>4</sub> •H <sub>2</sub> O	(10mM)	
Na <sub>2</sub> HPO <sub>4</sub> •7H <sub>2</sub> O	(10mM)	
NaCl	140mM	

After dialyzing the periplasmic extract against 5L Q-buffer A o/n at 4°C, the solution was applied to a 30mL QA52 anion exchange column (Whatman, Germany) equilibrated in Q-buffer A. Protein was eluted with a gradient of 0-100% Q-buffer B over 20 CV. FimD<sub>N</sub><sup>(1-125)</sup> was detected by SDS-PAGE and fractions containing the protein were pooled. Ammonium sulfate was added to the pool to a final concentration of 20% (w/v) and applied to a 20 mL phenyl sepharose column (Amersham, Sweden) equilibrated in Q-buffer B. Elution s performed by a gradient of 100-20% Q-buffer B over 20CV. The pooled protein containing fractions were polished by application to a Superdex75 26/60 SEC column (Amersham, Sweden) equilibrated in PBS. Clean protein was pooled and analyzed by SDS-PAGE.

Concentration was adjusted to  $\geq 400\mu\text{M}$  using an Amicon ultracentrifugation device with an YM10 membrane (Millipore, USA) and frozen in  $\text{lqN}_2$  for storage at  $-80^\circ\text{C}$ . The final yield was 4.5mg of pure protein per Liter of bacterial culture.

### 7.2.3.3 Cytosolic expression and autoinduction

For cytosolic expression of proteins, the encoding plasmids were transformed into *E.coli* strain BL21(DE3) using electroporation. Successful transformation was ensured by plating the transformed cells on LB-Agar containing the appropriate antibiotic and successful transformants were subsequently grown in 2YT medium containing the appropriate amount of antibiotic.

#### 7.2.3.3.1 Autoinduction

For higher yields of proteins expressed into the cytosol, plasmids encoding for protein expression in the cytosol were expressed using the auto-induction method described by Studier *et al.* (Studier, 2005). Plasmids were transformed into *E.coli* cells of the strain BL21(DE3) by electroporation and successful transformants screened by antibiotic resistance on LB-Agar plates containing the respective antibiotic. Colonies were picked and transferred into 100mL of ZY5052-media (ZY media supplemented with 1x 5052 solution) containing the appropriate antibiotic and grown at  $37^\circ\text{C}$  o/n.

This o/n-culture was then used to inoculate the desired amount (usually 1L) of antibiotic enriched ZY5052 media in a 1:100 dilution. The culture then was incubated under vigorous agitation at  $37^\circ\text{C}$  for 24 hours. The bacterial cells then were harvested by centrifugation and stored at  $-20^\circ\text{C}$  until further processed.

#### 7.2.3.3.2 Isolation of Inclusion Bodies (IBs)

- Resuspension buffer (100mM Tris/HCl, pH8.0)

Trizma Base	100mM
EDTA	1mM

- Triton buffer (60mM EDTA/NaOH, pH7.0)

EDTA	60mM
NaCl	1500mM
TritonX 100	6% (w/v)

- Washing buffer (100mM Tris/HCl, pH8.0)

Trizma Base	100mM
EDTA	20mM

5g of cell pellets were thawed on ice in 25mL resuspension buffer and resuspended using an ultrathorax followed by sonication on ice. For solubilization of impurities 0.5 volumes of triton buffer were added and the solution was stirred for 30min at 4°C. The sample was centrifuged at 31'000g for 10min and the pellet resuspended in 25mL washing buffer using sonication. This washing step was repeated for a total of 4 washes. The now purified inclusion bodies were either stored at -20°C or further processed immediately.

### 7.2.3.3.3 *FimA<sub>t his</sub><sup>Cyt</sup>*

- Solubilization buffer (20mM MES/NaOH, pH6.0)

MES	20mM	
GdmCl	6M	
EDTA	1mM	
DTT	20mM	→ add immediately before use only!

- SEC buffer (20mM Tris/HCl, pH8.0)

Trizma Base	20mM	
GdmCl	6M	
DTT	1mM	→ add immediately before use only!

- NiNTA-buffer (50mM NaP<sub>i</sub>, pH8.0)

	A	B
NaP <sub>i</sub>	50mM	50mM
GdmCl	6M	6M
NaCl	300mM	300mM
Imidazole	---	250mM

1g of purified inclusion bodies were resuspended in 5mL solubilization buffer by sonication and stirred on ice for 30min. Impurities are removed by centrifugation at 31'000g for 20min and the supernatant is applied to a Superdex200 26/60 SEC column (Amersham, Sweden) equilibrated in SEC buffer. The pooled protein fractions were desalted and the buffer was exchanged to NiNTA-buffer A on HiPrep 16/20 desalting column (GE Healthcare, USA) before applied to a 10mL NiNTA column (Qiagen, Germany) equilibrated in the same conditions. Impurities are eluted by 8% NiNTA-buffer B (~20mM Imidazole) and the protein was eluted by a gradient of 8-100% NiNTA-buffer B. The pooled protein was diluted to <25µM and air oxidized by addition of 0.1µM CuCl<sub>2</sub> as a catalyzing agent. Oxydation was

stopped by addition of 0.2 $\mu$ M EDTA and the protein applied to another run of SEC as described above, lacking DTT in the running buffer. The eluted protein was analyzed by SDS-Page and stored in the unfolded state in GdmCl containing buffer at 4°C. The final yield was 7.6mg of pure protein per gram of wet bacterial biomass.

#### 7.2.3.3.4 *FimG<sub>t</sub><sup>Cyt</sup>*

- Solubilization buffer (20mM MES/NaOH, pH6.0)

MES	20mM		
Urea	8M		→ deionized before use
EDTA	1mM		
DTT	20mM		→ add immediately before use only!

- Q-buffer (20mM MES/NaOH, pH6.0)

	A	B	
MES	20mM	20mM	
Urea	8M	8M	→ deionized before use
EDTA	1mM	1mM	
NaCl	---	400mM	
DTT	20mM	20mM	→ add immediately
	before use only!		

- SEC buffer (20mM Tris/HCl, pH8.0)

Trizma Base	20mM
GdmCl	6M
NaCl	300mM

5g of purified inclusion bodies were resuspended in 20mL solubilization buffer by sonication and stirred on ice for 30min. Impurities are removed by centrifugation at 31'000g for 20min and the supernatant is applied to a 30mL Q-sepharose anion exchange column (Amersham, Sweden) equilibrated in Q-buffer A. The protein is eluted by a gradient of 0-60% Q-buffer B and the pooled fractions are applied directly to a Superdex75 26/60 SEC column (Amersham, Sweden) equilibrated in SEC buffer containing 1mM DTT. The eluted pool of protein was then diluted to a concentration of <20 $\mu$ M and 0.1  $\mu$ M of CuCl<sub>2</sub> was added for air oxidation of the disulfide bond contained in FimG. After complete oxidation the sample was again applied to the SEC column now using SEC buffer WITHOUT DTT and the eluted protein was analyzed by SDS-Page and stored in the unfolded state in GdmCl containing buffer at 4°C. The final yield was 4.5mg of pure protein per gram of wet bacterial biomass.

### 7.2.3.3.5 *FimA<sub>A</sub><sup>Cyt</sup>*

- Solubilization buffer (20mM Tris/HCl, pH8.0)

Trizma Base	20mM	
GdmCl	6M	
EDTA	1mM	
DTT	20mM	→ add immediately before use only!

- SEC buffer (20mM Tris/HCl, pH8.0)

Trizma Base	20mM
GdmCl	6M
EDTA	0.05mM

- Q-buffer (20mM Tris/HCl, pH8.0)

	A	B
Trizma Base	20mM	20mM
NaCl	---	250mM

1g of purified inclusion bodies were resuspended in 2mL solubilization buffer by sonication and stirred on ice o/n. Impurities are removed by centrifugation at 31'000g for 25min and the supernatant was applied to a Superdex200 26/60 SEC column (Amersham, Sweden) equilibrated in SEC buffer. The pooled protein fractions were diluted to <10μM total protein concentration and the protein was oxidized by addition of 1mM GSH and 0.5mM GSSG. The reaction was incubated for 2h at room temperature and the denaturant and oxidizing agents were removed by dialyzing against 2x5L 20mM Tris/HCl, pH8.0 over night. The dialyzed sample then was applied to a Resource<sup>TM</sup>Q 6mL anion exchange column (GE Healthcare, USA) equilibrated in Q-buffer A and the Protein was eluted by a gradient of 0-100% Q-buffer B. Purity was determined by SDS-PAGE and the appropriately concentrated protein was stored at -80°C. The final yield was 0.6mg of pure protein per gram of wet bacterial biomass.

### 7.2.3.3.6 *FimF<sub>AF4</sub><sup>Cyt</sup>*

The *FimF<sub>AF4</sub>* protein was expressed and purified analogous to *FimA<sub>A</sub>* (see above). The final yield was 0.8mg of pure protein per gram of wet bacterial biomass.

## 7.2.4 Generation of *FimC:FimA<sub>t</sub>* complex

To generate the chaperone:subunit complex *FimC:FimA<sub>t</sub>*, unfolded *FimA<sub>t</sub>* was refolded by rapid dilution from conditions containing 3M GdmCl to 0.25M GdmCl in the presence of an 1.5fold excess of native *FimC* protein. Buffer conditions were 20mM MOPS/NaOH, pH6.7 and 200mM NaCl. The reaction was incubated at RT for 2hours and dialyzed against 5L

20mM Tris/HCl, pH8.0 over night at 4°C. Unbound protein was removed by purification on a Superdex75 26/60 SEC column equilibrated in PBS, pH7.4.

### 7.2.5 Protein and peptide concentrations

Protein concentrations were determined via measurements of UV-light absorption of fluorophoric amino acids at 280nm using the following molar extinction coefficients:

FimC	24'000 M <sup>-1</sup> cm <sup>-1</sup> .
FimC <sub>his</sub>	21'620 M <sup>-1</sup> cm <sup>-1</sup> .
FimG <sub>t</sub>	12'210 M <sup>-1</sup> cm <sup>-1</sup> .
FimF <sub>t</sub>	8'370 M <sup>-1</sup> cm <sup>-1</sup> .
FimA <sub>this</sub>	3'105 M <sup>-1</sup> cm <sup>-1</sup> .
dsI	5'500 M <sup>-1</sup> cm <sup>-1</sup> .

Except for dsI, all used peptides lack fluorophoric amino-acids, thus peptide concentrations were determined via measurements of UV-light absorption of the peptide-bond at 205nm using the following molar extinction coefficients:

dsA	54'868 M <sup>-1</sup> cm <sup>-1</sup> .
dsG	40'569 M <sup>-1</sup> cm <sup>-1</sup> .
dsF	49'675 M <sup>-1</sup> cm <sup>-1</sup> .
dsI <sup>6</sup>	43'393 M <sup>-1</sup> cm <sup>-1</sup> .

### 7.2.6 Crystallization of FimC:FimF<sub>t</sub>:FimD<sub>N</sub><sup>(1-125)</sup> [CFD] complex

The ternary FimC-FimF<sub>t</sub>-FimD<sub>N</sub> complex was generated by incubation of equimolar amounts of FimD<sub>N</sub> and the FimC:FimF<sub>t</sub> complex in 20mM Tris/HCl, pH 8.0 for 2 hours at room temperature. To remove unbound components the reaction mixture was purified by gel filtration on a Superdex75 26/60 column in PBS, pH 7.4. The purified complex was dialyzed against 20mM Tris/HCl, pH8.0 to remove the excessive salt and afterwards concentrated to 40mg/ml using AMICON Ultra ultracentrifugation tubes.

The resulting protein sample was then used in sitting drop crystallization screenings using a Microsys4000 nanoliter dispenser. Drops consisting of 100nl protein and 100nl reservoir solution were pipetted in 96well scale. Crystals with the space group P3(2)2(1) were achieved during incubation at 4°C with reservoir solutions of a) 20% (w/v) PEG8'000 (Sigma) in



50mM succinic acid, pH4, and b) 20% (w/v) PEG20'000 (Fluka) after 3-4 weeks of incubation.

### 7.2.7 Production of subunit:ds complexes

Subunit:ds complexes were formed by refolding of unfolded subunit protein in the presence of a 3-5fold excess of ds peptides. Unfolded protein in 2.8M GdmCl was rapidly diluted in a 10fold volume of 20mM Tris/HCl, pH8.0 containing the respective ds peptide and incubated over night at room temperature. The reaction was consequently desalted on a HiTrap desalting 26/10 column (GE Healthcare) before unbound protein and peptides were removed by anion exchange chromatography on a ResourceS 1ml column (GE Healthcare) equilibrated in 20mM Tris/HCl, pH8.0 using a gradient of 0-500mM NaCl.

### 7.2.8 Unfolding kinetics

Unfolding kinetics for subunit:ds complexes were measured by unfolding of the complex in a fluorescence spectrophotometer (QM7/2003, PTI) and following the change of tryptophan fluorescence emitted at 330nm when excited at 295nm. Unfolding was initialized by transferring the complex to unfolding buffer containing various amounts of denaturant. All measurements were performed using 30mM glycine/HCl buffer, pH2.0 and varying amounts of GdmCl as denaturant, using purified complex at a final concentration of 3 $\mu$ M in a 1mL stirrable fluorescence cuvette. Raw unfolding signals were fitted to a single exponential as described by *equation (8.1)* with  $t$  the time,  $S_0$  the initial signal,  $A$  the amplitude of the signal change and  $k$  the apparent rate of the change. The so derived rates of unfolding were plotted logarithmical over the GdmCl concentration. The resulting unfolding arm of the chevron plot was used to determine the rate of unfolding of the complex at zero molar denaturant concentration ( $k_V^{H_2O}$ ) by linear extrapolation of the curve to zero molar denaturant.

$$S_t = S_0 + A \cdot e^{-kt} \quad (8.1)$$

For some complexes a deviation from linearity was observed in the chevron plot at high concentrations of denaturant. Those traces were fitted to equation (8.2), assuming a three-state model with a change in the transition state distribution at high denaturant concentrations (Bachmann and Kiefhaber, 2001; Puorger *et al.*, 2008).

For  $U \xrightleftharpoons[k_{IU}]{k_{UI}} I \xrightleftharpoons[k_{NI}]{k_{IN}} N$ :

$$rate_{app} = \frac{-B \pm \sqrt{B^2 - 4 \cdot C}}{2} \quad \text{with } B = -(\zeta_{UI} + \zeta_{IU} + \zeta_{IN} + \zeta_{NI}) \text{ and } C = \zeta_{UI} \cdot (\zeta_{IN} + \zeta_{NI}) + \zeta_{IU} \cdot \zeta_{NI}$$

$$rate_{app} = \frac{(\zeta_{UI} + \zeta_{IU} + \zeta_{IN} + \zeta_{NI}) \pm \sqrt{(\zeta_{UI} + \zeta_{IU} + \zeta_{IN} + \zeta_{NI})^2 - 4 \cdot (\zeta_{UI} \cdot (\zeta_{IN} + \zeta_{NI}) + \zeta_{IU} \cdot \zeta_{NI})}}{2}$$

Here the parameters stand for the rate of the folding from the unfolded protein toward the intermediate,  $\zeta_{UI}$ , the respective back-reaction,  $\zeta_{IU}$ , and the rate of formation of the native molecule from the intermediate,  $\zeta_{IN}$ , and the respective back-reaction,  $\zeta_{NI}$ . To allow for kinetic analysis of data, the rates are substituted for the denaturant dependant rate constants:  $\zeta = k \cdot e^{m \cdot x}$ . Here  $k$  is the rate constant and  $m$  is the kinetic m-value, a measure for the denaturant dependence of a rate constant.

Since the deviation from linearity appears at high concentrations of denaturant only, the change in transition state is relative to the folded protein, not the unfolded protein. To simplify the equation  $k_{UI}$  can thus be set to zero:

$$k_{app}^0 = \frac{(k_{IU}^0 \cdot e^{m_{IU} \cdot x} + k_{IN}^0 \cdot e^{m_{IN} \cdot x} + k_{NI}^0 \cdot e^{m_{NI} \cdot x}) \pm \sqrt{(k_{IU}^0 \cdot e^{m_{IU} \cdot x} + k_{IN}^0 \cdot e^{m_{IN} \cdot x} + k_{NI}^0 \cdot e^{m_{NI} \cdot x})^2 - 4 \cdot (k_{IU}^0 \cdot e^{m_{IU} \cdot x}) \cdot (k_{NI}^0 \cdot e^{m_{NI} \cdot x})}}{2} \quad (8.2)$$

The rate of unfolding at 0M denaturant concentration was subsequently calculated using the obtained parameters according to equation (8.3).

$$k_{app}^0 = \frac{k_{NI}^0}{k_{IN}^0} \cdot k_{IU}^0 \quad (8.3)$$

### 7.2.9 Quantitative unfolding of subunit:ds peptide complexes

To establish concentration dependence of complex formation between subunit and ds peptide during refolding, unfolded subunit protein was refolded in variable excess amounts of ds peptide. 1  $\mu$ M of protein in 20mM Tris/HCl, pH8.0, containing 2.8M GdmCl, was refolded by 1+10 dilution in 20mM Tris/HCl, pH8.0, containing the respective amount of ds peptide. To

ensure completeness of the reaction, the solutions were incubated at room temperature overnight. The amount of formed complex was determined by following the signal change upon unfolding in a Quantamaster7 fluorescence spectrophotometer (PTI) by transfer of 200uL of the sample into 800uL of 200mM Histidine/HCl, pH1.0, containing 7.0M, 6.2M or 5.0M GdmCl for FimG<sub>t</sub>:dsF, FimG<sub>t</sub>:dsG and FimG<sub>t</sub>:dsA complexes, respectively. The samples were excited at 295nm and emission recorded at 330nm.

Amplitudes of the unfolding curves were fitted to a single-exponential equation with a fixed k-value. The resulting amplitudes are directly proportional to the amount of complex formed in that reaction.

### **7.2.10 ds peptide competition experiments**

To evaluate how two or more ds peptides compete in binding the folding subunit, the unfolded subunit is refolded in excess of equimolar amounts of the respective ds peptides. After defined incubation times the reaction was cooled to 4°C to avoid any possible further reaction. The formed complex species are detected by measuring fluorescence change during unfolding. By carefully choosing unfolding conditions it is possible to predominantly unfold just one of the possible species at a time, thus one is able to quantify the present subunit:ds complexes.

Refolding was performed in the way that the final concentration of protein was be 5μM and the peptides was be in the respective excess. Refolding was performed by diluting one part protein solution contained in 2.8M GdmCl in 20mM Tris/HCl, pH8.0, with 10 parts ds solution in 20mM Tris/HCl, pH8.0.

Unfolding measurements were performed by diluting 100μL of the refolding mixture with 900μL unfolding buffer in a 1mL stirrable fluorescence cuvette (path length 10mm) in a PTI Quantamaster7 fluorescence spectrophotometer (PhotoMed, Germany). The excitation wavelength was 295nm and emission was recorded at 330nm. The resulting fluorescence traces were further analyzed and fitted to a multiple first-order model with KaleidaGraph.

### **7.2.11 Interrupted refolding experiments**

To asses the rate of formation of subunit:ds complexes by refolding of the subunit in the presence of ds peptide, interrupted refolding measurements, also known as “N-tests”, were performed. The unfolded protein was refolded by dilution 1:11 in presence of an excess of the

respective ds-peptide from conditions containing 2.75M GdmCl as denaturant to a final denaturant concentration of 0.25M. The refolding reaction is stopped at various time points by addition of a defined concentration of denaturant, thus initializing unfolding again. Unfolding is followed by fluorescence spectroscopy and the unfolding conditions are chosen in a way that only the native species can be observed and possible intermediates or badly folded artifacts are already denatured.

Refolding conditions were chosen in a way that the final protein concentration was 3 $\mu$ M. Unfolding was initialized by addition of 825 $\mu$ L unfolding premix to 175 $\mu$ L of the refolding reaction in a 1mL sitrable fluorescence cuvette (path length 1cm). Measurements were performed in a PTI Quantamaster7 fluorescence spectrophotometer (PhotoMed, Germany) with excitation an emission wavelength set to 295nm and 330nm, respectively.

### 7.2.12 Chaperone displacement from chaperone:subunit complexes by ds peptides (DSE)

To determine the rate of displacement of the chaperone from chaperone:subunit complexes by the donor-strand of a particular subunit, the subunit:chaperone complex was incubated with an defined excess of the respective ds peptide at 25°C or 37°C. At defined time points the exchange reaction was stopped by addition of 800 $\mu$ L ice-cold running buffer (20mM MOPS/NaOH, pH6.7) to 200 $\mu$ L of the reaction mixture. The sample was then applied to a 1mL ResourceS<sup>TM</sup> cation-exchange column equilibrated in 20mM MOPS/NaOH, pH6.7 and eluted using a gradient of 0 to 500mM NaCl in 20mM MOPS/NaOH, pH6.7 over 15 column volumes. For a better resolution UV absorbance was detected at 228nm instead of the 280nm. The peaks in the resulting chromatograph were fitted and integrated using the Windows<sup>TM</sup> software “PeakFit” (SeaSolve Software Inc.). Peak areas were successively normalized against total protein concentration using equations (8.4) to (8.6).

$$[FimC_{his}]_t = \frac{PA(FimC_{his})}{PA(FimC_{his}) + \frac{PA(FimC_{his} : SU)}{f}} \cdot [FimC_{his}]_{total} \quad (8.4)$$

$$[FimC_{his} : SU]_t = \frac{PA(FimC_{his} : SU)}{PA(FimC_{his} : SU) + PA(FimC_{his}) \cdot f} \cdot [FimC_{his} : SU]_{total} \quad (8.5)$$

$$f = \frac{\mathcal{E}_{228nm}(FimC_{his} : SU)}{\mathcal{E}_{228nm}(FimC_{his})} \quad (8.6)$$

Resulting protein concentrations were plotted against reaction time and fitted according to a pseudo 1<sup>st</sup>-order reaction (see equation (8.7) with  $S_t$ : signal at time t;  $S_0$ : signal at time 0;  $k^{pseudo}$ : apparent rate constant; A: amplitude of signal change) using the Windows<sup>TM</sup> software “KaleidaGraph” (Synergy Software) and consequentially normalized to fractions of total protein.

$$S_t = A \cdot e^{-k_1^{pseudo} \cdot t} + S_0 \quad (8.7)$$

## 7.2.13 Mass spectroscopy of subunit:ds peptide complexes

### 7.2.13.1 MALDI-TOF mass spectroscopy

Unfolded FimG<sub>t</sub> and FimF<sub>t</sub> were separately refolded in presence of a 3-fold excess of the donor-strand peptides dsA, dsF, dsG, and dsI by dilution with 20mM Tris, pH8.0, to a final denaturant concentration of 0.25M and incubation at room temperature o/n. For matrix assisted Laser desorption/ionistaion time-of-flight mass spectroscopy (MALDI-TOF MS), samples were applied to a Supdex75 10/300 SEC column equilibrated in 50mM ammonium acetate, pH5.0, to remove undesired components and exchange the buffer to a volatile solvent suitable for mass spectroscopy. Concentrations of samples were adjusted and send to the Functional Genomics Center Zürich (FGCZ) for MALDI-TOF MS measurements.

### 7.2.13.2 Electrospray mass spectroscopy

For quantitative measurements and a higher resolution of the MS spectra measurements were performed on an LTQ-Orbitrap XL ESI mass spectrometer (Thermo Fisher Scientific, USA). Samples were buffered in 50mM ammonium acetate, pH5.0, by buffer exchange and a TriVersa<sup>TM</sup> NanoMate chip-based infusion system (Advion BioSciences, Inc., USA) in combination with an ESI nozzle chip was used for injection. Parameters were chosen as mild as possible without losing signal to ensure as little complex dissociation as possible. Data was analyzed with the Xcalibur<sup>TM</sup> software package supplied with the instrument.

### 7.2.14 Qualitative refolding of FimG<sub>t</sub>

To assay the specific fluorescence of free FimG<sub>t</sub> and FimG<sub>t</sub>:ds complexes for normalization of quantitative fluorescence spectra, the unfolded protein was refolded from 33μM protein in 2.8M GdmCl by dilution of 100μL protein solution with 1mL of 20mM Tris/HCl, pH8.0 containing 33μM of the respective ds peptide in a 1cm stirrable fluorescence cuvette. Refolding was followed by change in tryptophan fluorescence emitted at 330nm with excitation at 295nm in a PTI Quantmaster7 fluorescence spectrophotometer. All measurements were carried out at 25°C. Refolding traces were plotted over incubation time and normalized using the refolding of FimG<sub>t</sub> in absence of any ds peptide as reference. Resulting curves were fitted to a double exponential according to equation (8.8) except for refolding of free FimG<sub>t</sub> which was fitted with a single exponential (equation (8.1)).

$$S_t = S_0 + A_1 \cdot e^{-k_1 t} + A_2 \cdot e^{-k_2 t} \quad (8.8)$$

### 7.2.15 pH dependence of subunit:ds peptide complex stability

To assess the influence of pH on the unfolding rates of subunit:ds complexes, 5μM of FimG<sub>t</sub>:ds<sub>G</sub> and FimG<sub>t</sub>:ds<sub>F</sub> were unfolded at constant concentrations of denaturant in buffers of varying acidity. The concentration of GdmCl as denaturant was kept constant as verified by refractive index determination at 7.73M and the unfolding reaction was initiated by transferring 100μL of sample to 900μL of the respective buffer containing denaturant in a stirrable fluorescence cuvette (1cm path length) and following the change in tryptophan fluorescence with excitation at 295nm and Emission recorded at 330nm. The buffers were 20mM Tris/HCl, pH8.0; 50mM MOPS/NaOH, pH7.5; 50mM MES/NaOH, pH6.0; 50mM Acetate, pH4.5; 30mM Glycine/HCl, pH3.0; 30mM Glycine/HCl, pH2.0; 200mM Histidine/HCl, pH1.0.

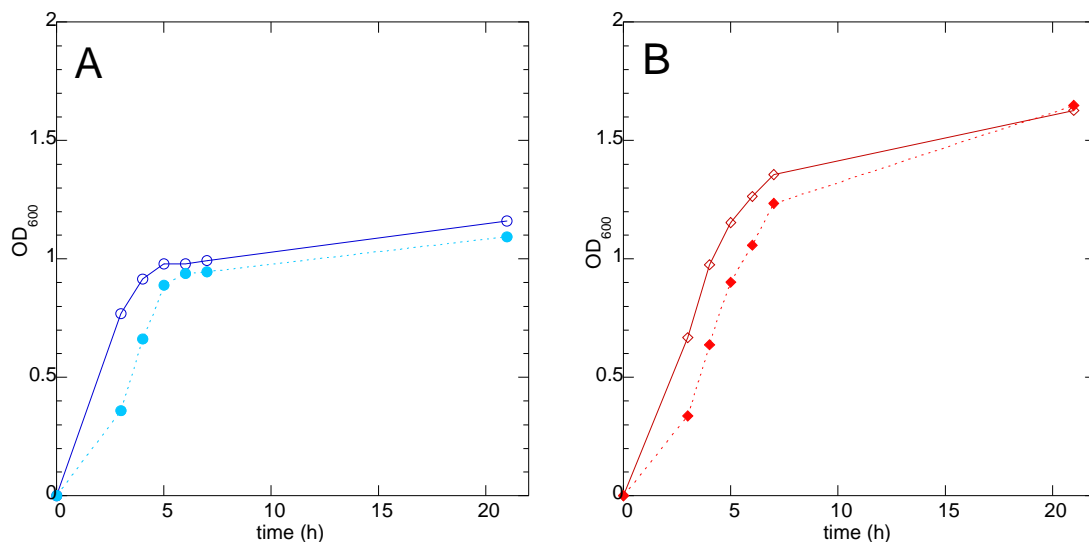
## 8 Appendix

### 8.1 Additional results

#### 8.1.1 Comparison of IPTG-induced and auto-induced protein expression

Since the expression of Fim-constructs lacking the signal-sequence for translocation into the periplasm happens as protein aggregates in the cytosol, the use of *E.coli* strain BL21 was proffer, allowing for the use of the auto-induction procedure described by Grabski *et al.*, 2005. To ensure the use of this method did not interfere with protein expression of our constructs, conventional IPTG-induction of the *lac* operon was used as a comparison.

Figure 8-1 shows cell growths for the constructs pfimFF4 and pfimFAF4, encoding for self-complemented FimF-variants with the natural N-terminal extension (FimFF<sub>4</sub>) and altered N-terminal extension corresponding to that of FimA (FimFAF<sub>4</sub>). Direct comparison of IPTG induction at OD<sub>600</sub>=0.7 and auto-induction show virtual identical cell-growth. Analysis of the protein expression of the cultures showed a about 30% higher yield of the target proteins when using auto-induction.

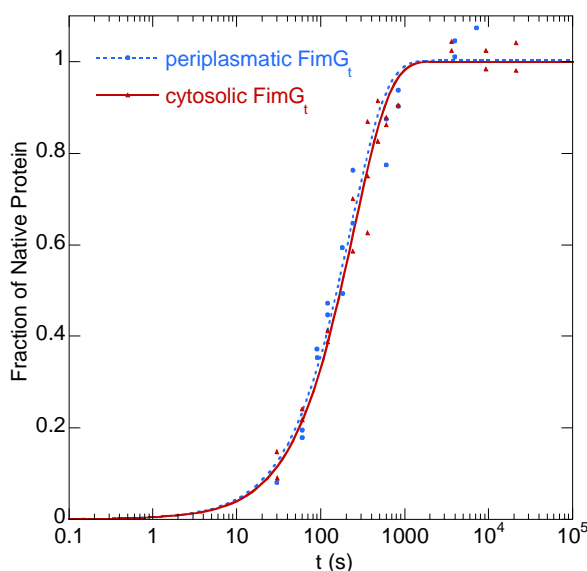


**Figure 8-1: Comparison of manual- and auto-induction.**

1L of 2YT media (open symbols) and 1L of auto-induction media (filled symbols) were inoculated with 10mL from the same pre-culture. (A) is FimFF<sub>4</sub>, (B) is FimFAF<sub>4</sub>. All cultures were incubated at 37°C. Cell growth was monitored by taking 1mL samples at various time-points. The 2YT culture was induced with 100mM IPTG upon reaching an OD<sub>600</sub> of 0.7 and the auto-induction culture was left alone. Both methods are comparable in respect to cell growth.

### 8.1.2 Comparison of cytosolic vs periplasmic expression of FimG<sub>t</sub>

To ensure the newly produced FimG<sub>t</sub> protein from the cytoplasm behaves the same way as the previously used variant, purified from the periplasm, we did a comparison using a double-jump experiment analyzing the refolding of the respective protein in presence of dsF as described in Chapter 7.2.11, was performed with identical concentrations of both variants. The resulting graphs are shown as overlay in Figure 8-2. Within error the both protein exhibited the same affinity and velocity for folding and binding of dsF, indicating that there is no problem using the way more fertile periplasmatic production method.



**Figure 8-2: Comparison of FimG<sub>t</sub> purified from cytosol and periplasm: Kinetics of formation of the FimG<sub>t</sub>:dsF complex.**

Interrupted refolding experiment with identical concentrations of both proteins and 10-fold excess of dsF, performed at pH7.0 and 25°C. Refolding was stopped by GdmCl induced unfolding of the formed native complex in a fluorescence spectrophotometer and amplitudes of unfolding curves were plotted over refolding time. Data was fitted to a single-exponential function and the rates were found to be within error identical. FimG<sub>t</sub> from expression as inclusion bodies in the cytosol is depicted by red, solid line, FimG<sub>t</sub> from the periplasm is shown in red, dashed line.

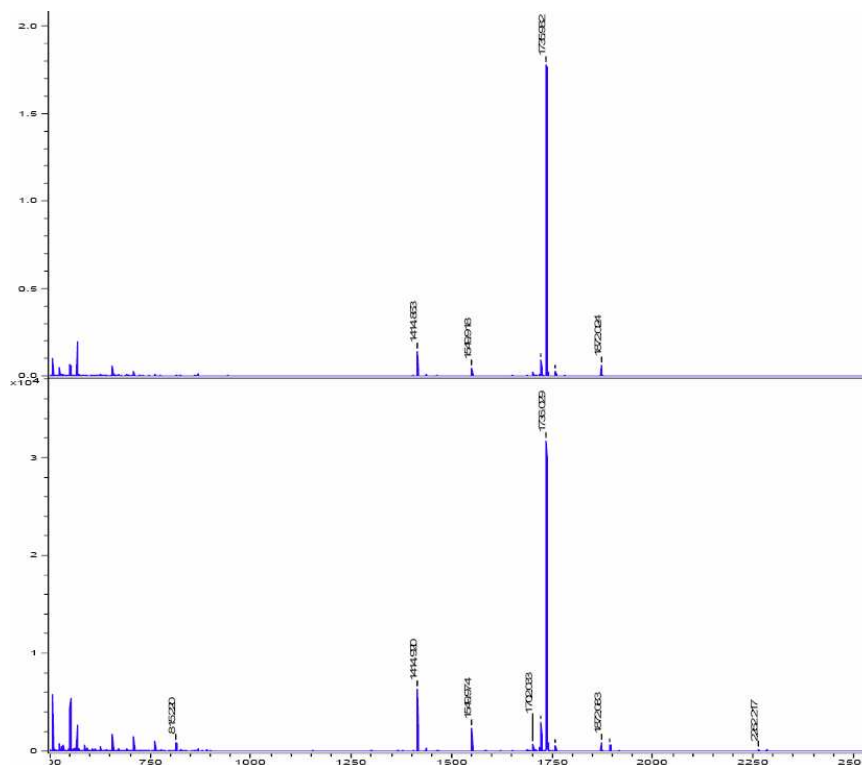
### 8.1.3 MALDI-TOF MS of subunit:ds complex mixtures

A qualitative approach to determine the identity of formed complex during FimG<sub>t</sub> or FimF<sub>t</sub> refolding in the presence of ds-peptides was mass spectroscopic analysis of the products. Unfolded FimG<sub>t</sub> or FimF<sub>t</sub> was refolded in the presence of a 5-fold excess of ds peptides (dsA,



dsF, dsG and dsI) and the reaction products were purified by separation on both size-exclusion and ion exchange chromatography. The pooled elution fractions corresponding to the peaks making up the subunit:ds complexes were subjected to MALD-TOF mass spectroscopy at the Functional Genomics Center Zürich (FGCZ, UZH, Switzerland).

A typical result is shown in Figure 8-3. The peaks at 1415m/z, 1737m/z, 1872m/z and 2262m/z correspond to the ds peptides dsG, dsF, dsA and dsI respectively. The presence of all ds peptides in the purified samples of subunit:ds complexes indicates that no discrimination mechanism is present at the subunit level, yielding a statistical distribution of subunit:ds complexes rather than only the thermodynamically favored one. Unfortunately it is not possible to use this data for quantification since the matrix desorbed complexes have significantly different ionization properties and desorption is not reproducible. Accordingly, the peak intensities shown here can not be an indication of the amount of respective peptide and thus subunit:ds complex, present.

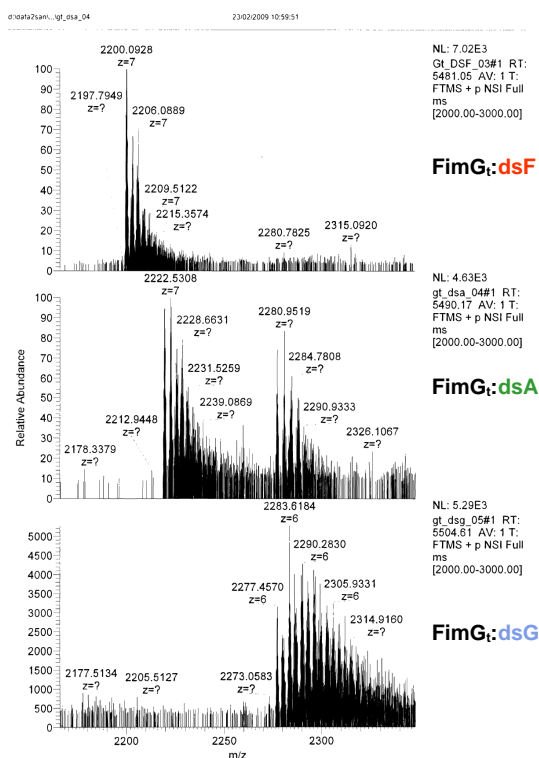


**Figure 8-3: MALDI-TOF MS of FimG<sub>t</sub>:ds and FimF<sub>t</sub>:ds mixture.**

(**Top**) FimG<sub>t</sub> was offered the possible binding partners dsA, dsF, dsG and dsI simultaneously. Sample was freed from unused reactants by size-exclusion chromatography and analyzed by MALDI-TOF mass spectroscopy. All ds peptides were detected in the purified complex fraction indicating a statistical distribution of subunit:ds complexes formed during refolding. (**Bottom**) Identical experiment with FimF<sub>t</sub>

### 8.1.4 ESI-TOF MS of subunit:ds mixtures

For quantitative assessment of the various complexes formed by simultaneous binding of various ds peptides to type 1 pilus subunits the method was changed from MALDI-TOF to the more gentle and reproducible ESI-TOF MS with the goal to measure intact complexes. Purified FimG<sub>t</sub>:ds complexes were analyzed using an OrbiTrap™ ESI-TOF mass spectrometer. Ionization energy, solvent pressure and nozzle temperatures were optimized in the attempt to keep the complexes stable during MS measurement. FimG<sub>t</sub>:dsF again stayed intact, while FimG<sub>t</sub>:dsA partially dissociated. FimG<sub>t</sub>:dsG, which was previously found to be kinetically and thermodynamically extremely stable in solution, totally dissociated in the process (Figure 8-4). This was very surprising and might indicate a significant difference in binding of different ds peptides to FimG<sub>t</sub>, resulting in different susceptibility to the denaturing conditions in the mass spectrometer.



**Figure 8-4: ESI-MS spectra of FimG<sub>t</sub>:ds complexes.**

10 μM of purified FimG<sub>t</sub>:dsF, FimG<sub>t</sub>:dsA and FimG<sub>t</sub>:dsG complexes were analyzed using an OrbiTrap™ ESI mass spectrometer. Conditions were optimized to minimize complex dissociation. Only FimG<sub>t</sub>:dsF complex stayed intact, while FimG<sub>t</sub>:dsA and FimG<sub>t</sub>:dsG at least partially dissociated.

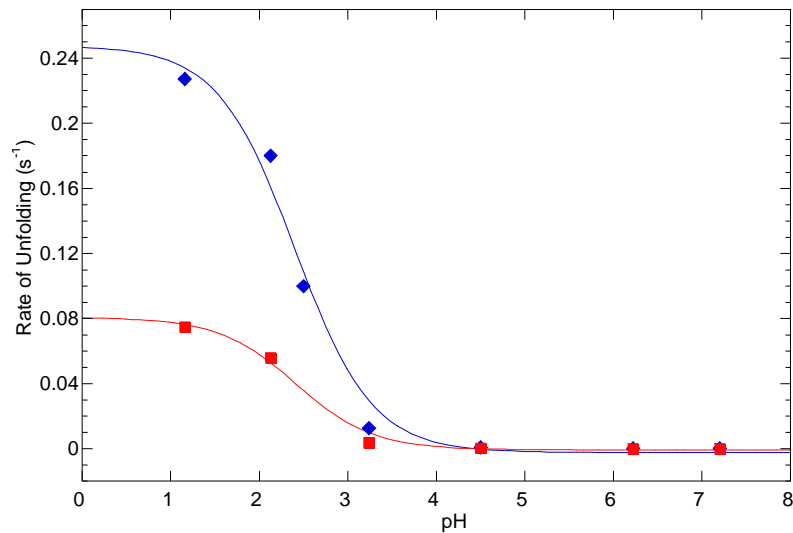
The main reason for complex dissociation in mass spectroscopy is the high vacuum in the detector which is weakening hydrophobic interactions and strengthening electrostatic interactions due to the removal of solvating water molecules. By switching to a qTOF<sup>TM</sup> mass spectroscope, which can operate at a significantly lower vacuum, we tried to reduce this effect. Again, the same pattern of dissociation of the complexes was found, making it impossible to distinguish between partially dissociated FimG<sub>t</sub>:dsA and full dissociated FimG<sub>t</sub>:dsG complexes.

This led to the hypothesis that in the FimG<sub>t</sub>:dsF complex the ds binds to the subunit protein *via* interactions governed more predominantly by electrostatic forces and the FimG<sub>t</sub>:dsG complex relies mostly on hydrophobic interactions.

### **8.1.5 Dependence of unfolding rates of FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG on the pH**

The poor stability of the FimG<sub>t</sub>:dsG complex in the vacuum during mass spectroscopy was surprising and led to the hypothesis that the binding of different ds peptides relies on different forces of interaction. So the FimG<sub>t</sub>:dsF complex, which was completely stable during mass spectroscopy, would rely more on electrostatic interactions, which are strengthened in the vacuum, than the FimG<sub>t</sub>:dsG complex.

To test this hypothesis, the pH susceptibility of the unfolding kinetics of the FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG complexes was assayed. This was done by unfolding equimolar amounts of the purified complexes in buffers of varying pH and constant denaturant concentration. 3 μM complexes were unfolded in 7.5M GdmCl contained in buffers with the pH values of 7.2, 4.5, 3.25, 6.2, 2.12 and 1.2, respectively. The rate of the unfolding reaction as recorded by following the decrease in tryptophan fluorescence was plotted against the pH (Figure 8-5). The transition mid-points of the titration curves were determined to be at pH 2.410 and pH 2.403 for the FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG complexes, respectively.



**Figure 8-5: Figure pH dependence of the unfolding rates of the FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG complexes.**

FimG<sub>t</sub>:dsF (■) and FimG<sub>t</sub>:dsG (◆) complexes were unfolded in 7.5M GdmCl at varying pH. The rate of unfolding, as determined by change in intrinsic fluorescence upon unfolding, was plotted against the pH. Transition midpoints were found to be at pH 2.410 and pH 2.4003 for FimG<sub>t</sub>:dsF and FimG<sub>t</sub>:dsG, respectively.

This is in good agreement with the isoelectric point of aspartic acid ( $pI = 2.7$ ) and indicates that the peptide is more readily dissociated from the protein once the charged residues on the rim of the peptide binding groove lose their charge.

As is seen in Figure 8-5, the complexes are not equally susceptible to the change in pH. The effect is much more pronounced for the FimG<sub>t</sub>:dsG complex. Since in this complex only one aspartic acid residue seems to be involved in anchoring the peptide, the stabilizing properties of the interaction is lost more readily at low pH than in the FimG<sub>t</sub>:dsF complex.

## References

- 1 Abraham, J.M., Freitag, C.S., Clements, J.R. and Eisenstein, B.I. (1985) An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc Natl Acad Sci U S A*, **82**, 5724-5727.
- 2 Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.-W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. Vol. 58, pp. 1948-1954.
- 3 Anderson, G.G., Palermo, J.J., Schilling, J.D., Roth, R., Heuser, J. and Hultgren, S.J. (2003) Intracellular bacterial biofilm-like pods in urinary tract infections. *Science*, **301**, 105-107.
- 4 Anfinsen, C.B. (1967) The formation of the tertiary structure of proteins. *Harvey Lect*, **61**, 95-116.
- 5 Anfinsen, C.B. (1973) Principles that govern the folding of protein chains. *Science*, **181**, 223-230.
- 6 Anfinsen, C.B. and Scheraga, H.A. (1975) Experimental and theoretical aspects of protein folding. *Adv Protein Chem*, **29**, 205-300.
- 7 Bachmann, A. and Kiefhaber, T. (2001) Apparent two-state tendamistat folding is a sequential process along a defined route. *J Mol Biol*, **306**, 375-386.
- 8 Båga, M., Norgren, M. and Normark, S. (1987) Biogenesis of *E. coli* Pap pili: PapH, a minor pilin subunit involved in cell anchoring and length modulation. *Cell*, **49**, 241-251.
- 9 Bann, J.G., Pinkner, J.S., Frieden, C. and Hultgren, S.J. (2004) Catalysis of protein folding by chaperones in pathogenic bacteria. *Proc Natl Acad Sci U S A*, **101**, 17389-17393.
- 10 Barnhart, M.M., Pinkner, J.S., Soto, G.E., Sauer, F.G., Langermann, S., Waksman, G., Frieden, C. and Hultgren, S.J. (2000) PapD-like chaperones provide the missing information for folding of pilin proteins. *Proc Natl Acad Sci U S A*, **97**, 7709-7714.
- 11 Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. and Kiefhaber, T. (1999) The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc Natl Acad Sci U S A*, **96**, 9597-9601.
- 12 Bishop, B.L., Duncan, M.J., Song, J., Li, G., Zaas, D. and Abraham, S.N. (2007) Cyclic AMP-regulated exocytosis of *Escherichia coli* from infected bladder epithelial cells. *Nat Med*, **13**, 625-630.
- 13 Blomfield, I.C. (2001) The regulation of pap and type 1 fimbriation in *Escherichia coli*. *Adv Microb Physiol*, **45**, 1-49.

- 14 Bogan, A.A. and Thorn, K.S. (1998) Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology*, **280**, 1-9.
- 15 Bose, H.S., Whittal, R.M., Baldwin, M.A. and Miller, W.L. (1999) The active form of the steroidogenic acute regulatory protein, StAR, appears to be a molten globule. Vol. 96, pp. 7250-7255.
- 16 Brinton, C.C. (1959) Non-Flagellar Appendages of Bacteria. *Nature*, **183**, 782-786.
- 17 Brünger, A.T., Adams, P.D., MariusClore, G., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. **54**, 905-921.
- 18 Bullitt, E., Jones, C.H., Striker, R., Soto, G., Jacob-Dubuisson, F., Pinkner, J., Wick, M.J., Makowski, L. and Hultgren, S.J. (1996) Development of pilus organelle subassemblies in vitro depends on chaperone uncapping of a beta zipper. *Proc Natl Acad Sci U S A*, **93**, 12890-12895.
- 19 Capitani, G., Eidam, O., Glockshuber, R. and Grutter, M.G. (2006a) Structural and functional insights into the assembly of type 1 pili from Escherichia coli. *Microbes Infect*.
- 20 Capitani, G., Eidam, O. and Grutter, M.G. (2006b) Evidence for a novel domain of bacterial outer membrane ushers. *Proteins*, **65**, 816-823.
- 21 Chapman, M.R., Robinson, L.S., Pinkner, J.S., Roth, R., Heuser, J., Hammar, M., Normark, S. and Hultgren, S.J. (2002) Role of Escherichia coli curli operons in directing amyloid fiber formation. *Science*, **295**, 851-855.
- 22 Cheng, Y. (2009) Toward an atomic model of the 26S proteasome. *Current Opinion in Structural Biology*, **19**, 203-208.
- 23 Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S.J. and Knight, S.D. (1999) X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. *Science*, **285**, 1061-1066.
- 24 Christie, P.J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S. and Cascales, E. (2005) BIOGENESIS, ARCHITECTURE, AND FUNCTION OF BACTERIAL TYPE IV SECRETION SYSTEMS. *Annual Review of Microbiology*, **59**, 451.
- 25 Connell, I., Agace, W., Klemm, P., Schembri, M., MÄfrild, S. and Svanborg, C. (1996) Type 1 fimbrial expression enhances Escherichia coli virulence for the urinary tract. Vol. 93, pp. 9827-9832.
- 26 Cordes, M.H., Davidson, A.R. and Sauer, R.T. (1996) Sequence space, folding and protein design. *Curr Opin Struct Biol*, **6**, 3-10.
- 27 Davies, D.R. and Cohen, G.H. (1996) Interactions of protein antigens with antibodies. *Proc Natl Acad Sci U S A*, **93**, 7-12.

- 28 Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S. and Richardson, D.C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res*, **35**, W375-383.
- 29 DeLano, W.L. (2002) The PyMOL User's Manual. DeLano Scientific LLC., Palo Alto, CA.
- 30 Dill, K.A. (1990a) Dominant forces in protein folding. *Biochemistry*, **29**, 7133-7155.
- 31 Dill, K.A. (1990b) The meaning of hydrophobicity. *Science*, **250**, 297-298.
- 32 Dill, K.A. (1999) Polymer principles and protein folding. *Protein Sci*, **8**, 1166-1180.
- 33 Dill, K.A. and Chan, H.S. (1997) From Levinthal to pathways to funnels. *Nat Struct Biol*, **4**, 10-19.
- 34 Dill, K.A., Fiebig, K.M. and Chan, H.S. (1993) Cooperativity in protein-folding kinetics. *Proc Natl Acad Sci U S A*, **90**, 1942-1946.
- 35 Dill, K.A., Ozkan, S.B., Shell, M.S. and Weikl, T.R. (2008) The protein folding problem. *Annu Rev Biophys*, **37**, 289-316.
- 36 Dodson, K.W., Jacob-Dubuisson, F., Striker, R.T. and Hultgren, S.J. (1993) Outer-membrane PapC molecular usher discriminately recognizes periplasmic chaperone-pilus subunit complexes. *Proc Natl Acad Sci U S A*, **90**, 3670-3674.
- 37 Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., Ausio, J., Nissen, M.S., Reeves, R., Kang, C., Kissinger, C.R., Bailey, R.W., Griswold, M.D., Chiu, W., Garner, E.C. and Obradovic, Z. (2001) Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling*, **19**, 26-59.
- 38 Eaton, W.A., Munoz, V., Thompson, P.A., Chan, C.K. and Hofrichter, J. (1997) Submillisecond kinetics of protein folding. *Curr Opin Struct Biol*, **7**, 10-14.
- 39 Elad, N., Maimon, T., Frenkiel-Krispin, D., Lim, R.Y.H. and Medalia, O. (2009) Structural analysis of the nuclear pore complex by integrated approaches. *Current Opinion in Structural Biology*, **19**, 226-232.
- 40 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.
- 41 Eshdat, Y., Silverblatt, F.J. and Sharon, N. (1981) Dissociation and reassembly of Escherichia coli type 1 pili. *J Bacteriol*, **148**, 308-314.
- 42 Fernandez, L.A. and Berenguer, J. (2000) Secretion and assembly of regular surface structures in Gram-negative bacteria. *FEMS Microbiol Rev*, **24**, 21-44.
- 43 Fersht, A.R. (1999) *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*. W.H. Freeman and Company, New York.

- 44 Fiebig, K.M. and Dill, K.A. (1993) Protein core assembly processes. *The Journal of Chemical Physics*, **98**, 3475-3487.
- 45 Finder, V.H. and Glockshuber, R. (2007) Amyloid-beta aggregation. *Neurodegener Dis*, **4**, 13-27.
- 46 Glenner, G.G. (1980) Amyloid deposits and amyloidosis: the beta-fibrilloses (second of two parts). *N Engl J Med*, **302**, 1333-1343.
- 47 Gossert, A.D., Bettendorff, P., Puorger, C., Vetsch, M., Herrmann, T., Glockshuber, R. and Wuthrich, K. (2008) NMR structure of the Escherichia coli type 1 pilus subunit FimF and its interactions with other pilus subunits. *J Mol Biol*, **375**, 752-763.
- 48 Grabski, A., Mehler, M. and Drott, D. (2005) The Overnight Express Autoinduction System: High-density cell growth and protein expression while you sleep. *Nat Meth*, **2**, 233-235.
- 49 Greene, R.F., Jr. and Pace, C.N. (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin. *J Biol Chem*, **249**, 5388-5393.
- 50 Haber, E. and Anfinsen, C.B. (1962) Side-chain interactions governing the pairing of half-cystine residues in ribonuclease. *J Biol Chem*, **237**, 1839-1844.
- 51 Hahn, E., Wild, P., Hermanns, U., Sebbel, P., Glockshuber, R., Haner, M., Taschner, N., Burkhard, P., Aebi, U. and Muller, S.A. (2002) Exploring the 3D molecular architecture of Escherichia coli type 1 pili. *J Mol Biol*, **323**, 845-857.
- 52 Hall-Stoodley, L., Costerton, J.W. and Stoodley, P. (2004) Bacterial biofilms: from the Natural environment to infectious diseases. *Nat Rev Micro*, **2**, 95-108.
- 53 Hasman, H., Schembri, M.A. and Klemm, P. (2000) Antigen 43 and type 1 fimbriae determine colony morphology of Escherichia coli K-12. *J Bacteriol*, **182**, 1089-1095.
- 54 Hooft, R.W., Vriend, G., Sander, C. and Abola, E.E. (1996) Errors in protein structures. *Nature*, **381**, 272.
- 55 Huang, F., Sato, S., Sharpe, T.D., Ying, L. and Fersht, A.R. (2007) Distinguishing between cooperative and unimodal downhill protein folding. *Proc Natl Acad Sci U S A*, **104**, 123-127.
- 56 Huang, Y., Smith, B.S., Chen, L.X., Baxter, R.H. and Deisenhofer, J. (2009) Insights into pilus assembly and secretion from the structure and functional characterization of usher PapC. *Proc Natl Acad Sci U S A*, **106**, 7403-7407.
- 57 Ippen-Ihler, K.A. and Minkley, E.G. (1986) The Conjugation System of F, The Fertility Factor of Escherichia Coli. *Annual Review of Genetics*, **20**, 593.
- 58 Jacob, M., Schindler, T., Balbach, J. and Schmid, F.X. (1997) Diffusion control in an elementary protein folding reaction. *Proc Natl Acad Sci U S A*, **94**, 5622-5627.



- 59 Jones, S. and Thornton, J.M. (1996) Principles of protein-protein interactions. *Proc Natl Acad Sci U S A*, **93**, 13-20.
- 60 Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M. and Hecht, M.H. (1993) Protein design by binary patterning of polar and nonpolar amino acids. *Science*, **262**, 1680-1685.
- 61 Kim, D.E., Gu, H. and Baker, D. (1998) The sequences of small proteins are not extensively optimized for rapid folding by natural selection. *Proc Natl Acad Sci U S A*, **95**, 4982-4986.
- 62 Klemm, P. (1986) Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in Escherichia coli. *Embo J*, **5**, 1389-1393.
- 63 Kramers, H.A. (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **7**, 284-304.
- 64 Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, **372**, 774-797.
- 65 Krogfelt, K.A. and Klemm, P. (1988) Investigation of minor components of Escherichia coli type 1 fimbriae: protein chemical and immunological aspects. *Microb Pathog*, **4**, 231-238.
- 66 Kuehn, M.J., Normark, S. and Hultgren, S.J. (1991) Immunoglobulin-like PapD chaperone caps and uncaps interactive surfaces of nascently translocated pilus subunits. *Proc Natl Acad Sci U S A*, **88**, 10586-10590.
- 67 Levinthal, C. (1968) Are there pathways for protein folding? *Extrait du Journal de Chimie Physique*, **65**, 2.
- 68 Li, H., Qian, L., Chen, Z., Thibault, D., Liu, G., Liu, T. and Thanassi, D.G. (2004) The outer membrane usher forms a twin-pore secretion complex. *J Mol Biol*, **344**, 1397-1407.
- 69 Lijnzaad, P. and Argos, P. (1997) Hydrophobic patches on protein subunit interfaces: characteristics and prediction. *Proteins*, **28**, 333-343.
- 70 Lindberg, F., Tennent, J.M., Hultgren, S.J., Lund, B. and Normark, S. (1989) PapD, a periplasmic transport protein in P-pilus biogenesis. *J Bacteriol*, **171**, 6052-6058.
- 71 Lund, B., Lindberg, F., Marklund, B.I. and Normark, S. (1987) The PapG protein is the alpha-D-galactopyranosyl-(1----4)-beta-D-galactopyranose-binding adhesin of uropathogenic Escherichia coli. *Proc Natl Acad Sci U S A*, **84**, 5898-5902.
- 72 Maier, T., Leibundgut, M. and Ban, N. (2008) The Crystal Structure of a Mammalian Fatty Acid Synthase. Vol. 321, pp. 1315-1322.
- 73 McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. and Read, R.J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr*, **61**, 458-464.

- 74 Meerman, H.J. and Georgiou, G. (1994) Construction and characterization of a set of *E. coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Biotechnology (N Y)*, **12**, 1107-1110.
- 75 Minor, W., Cymborowski, M., Otwinowski, Z. and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr*, **62**, 859-866.
- 76 Mulvey, M.A., Schilling, J.D. and Hultgren, S.J. (2001) Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect Immun*, **69**, 4572-4579.
- 77 Naber, K.G. (2000) Survey on antibiotic usage in the treatment of urinary tract infections. *J Antimicrob Chemother*, **46 Suppl 1**, 49-52; discussion 63-45.
- 78 Nishiyama, M. and Glockshuber, R. (2010) The outer membrane usher guarantees the formation of functional pili by selectively catalyzing donor-strand exchange between subunits that are adjacent in the mature pilus. *J Mol Biol*, **396**, 1-8.
- 79 Nishiyama, M., Horst, R., Eidam, O., Herrmann, T., Ignatov, O., Vetsch, M., Bettendorff, P., Jelesarov, I., Grutter, M.G., Wuthrich, K., Glockshuber, R. and Capitani, G. (2005) Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD. *Embo J*, **24**, 2075-2086.
- 80 Nishiyama, M., Ishikawa, T., Rechsteiner, H. and Glockshuber, R. (2008) Reconstitution of Pilus Assembly Reveals a Bacterial Outer Membrane Catalyst. Vol. 320, pp. 376-379.
- 81 Nishiyama, M., Vetsch, M., Puorger, C., Jelesarov, I. and Glockshuber, R. (2003) Identification and characterization of the chaperone-subunit complex-binding domain from the type 1 pilus assembly platform FimD. *J Mol Biol*, **330**, 513-525.
- 82 Olsen, A., Jonsson, A. and Normark, S. (1989) Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. *Nature*, **338**, 652-655.
- 83 Peabody, C.R., Chung, Y.J., Yen, M.R., Vidal-Ingigliardi, D., Pugsley, A.P. and Saier, M.H., Jr. (2003) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology*, **149**, 3051-3072.
- 84 Perrakis, A., Morris, R. and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol*, **6**, 458-463.
- 85 Pinkner, J.S., Remaut, H., Buelens, F., Miller, E., Åberg, V., Pemberton, N., Hedenström, M., Larsson, A., Seed, P., Waksman, G., Hultgren, S.J. and Almqvist, F. (2006) Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria. Vol. 103, pp. 17897-17902.
- 86 Privalov, P.L. and Khechinashvili, N.N. (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *Journal of Molecular Biology*, **86**, 665-684.

- 87 Puorger, C., Eidam, O., Capitani, G., Erilov, D., Grütter, M.G. and Glockshuber, R. (2008) Infinite Kinetic Stability against Dissociation of Supramolecular Protein Complexes through Donor Strand Complementation. *Structure*, **16**, 631-642.
- 88 Reid, G. and Sobel, J.D. (1987) Bacterial Adherence in the Pathogenesis of Urinary Tract Infection: A Review. *Reviews of Infectious Diseases*, **9**, 470-487.
- 89 Remaut, H., Tang, C., Henderson, N.S., Pinkner, J.S., Wang, T., Hultgren, S.J., Thanassi, D.G., Waksman, G. and Li, H. (2008) Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell*, **133**, 640-652.
- 90 Roche, A., McFadden, J. and Owen, P. (2001) Antigen 43, the major phase-variable protein of the Escherichia coli outer membrane, can exist as a family of proteins encoded by multiple alleles. *Microbiology*, **147**, 161-169.
- 91 Rose, R.J., Verger, D., Daviter, T., Remaut, H., Paci, E., Waksman, G., Ashcroft, A.E. and Radford, S.E. (2008) Unraveling the molecular basis of subunit specificity in P pilus assembly by mass spectrometry. pp. -.
- 92 Santoro, M.M. and Bolen, D.W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063-8068.
- 93 Sauer, F.G., Barnhart, M., Choudhury, D., Knight, S.D., Waksman, G. and Hultgren, S.J. (2000) Chaperone-assisted pilus assembly and bacterial attachment. *Curr Opin Struct Biol*, **10**, 548-556.
- 94 Sauer, F.G., Futterer, K., Pinkner, J.S., Dodson, K.W., Hultgren, S.J. and Waksman, G. (1999) Structural basis of chaperone function and pilus biogenesis. *Science*, **285**, 1058-1061.
- 95 Sauer, F.G., Pinkner, J.S., Waksman, G. and Hultgren, S.J. (2002) Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell*, **111**, 543-551.
- 96 Sauer, F.G., Remaut, H., Hultgren, S.J. and Waksman, G. (2004) Fiber assembly by the chaperone-usher pathway. *Biochim Biophys Acta*, **1694**, 259-267.
- 97 Saulino, E.T., Bullitt, E. and Hultgren, S.J. (2000) Snapshots of usher-mediated protein secretion and ordered pilus assembly. *Proc Natl Acad Sci U S A*, **97**, 9240-9245.
- 98 Saulino, E.T., Thanassi, D.G., Pinkner, J.S. and Hultgren, S.J. (1998) Ramifications of kinetic partitioning on usher-mediated pilus biogenesis. *Embo J*, **17**, 2177-2185.
- 99 Schembri, M.A., Ussery, D.W., Workman, C., Hasman, H. and Klemm, P. (2002) DNA microarray analysis of fim mutations in Escherichia coli. *Mol Genet Genomics*, **267**, 721-729.
- 100 Schmid, F.X. (1983) Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry*, **22**, 4690-4696.

- 101 Sheinerman, F.B., Norel, R. and Honig, B. (2000) Electrostatic aspects of protein-protein interactions. *Curr Opin Struct Biol*, **10**, 153-159.
- 102 Smyth, C.J., Marron, M.B., Twohig, J.M. and Smith, S.G. (1996) Fimbrial adhesins: similarities and variations in structure and biogenesis. *FEMS Immunol Med Microbiol*, **16**, 127-139.
- 103 Snyder, J.A., Haugen, B.J., Buckles, E.L., Lockatell, C.V., Johnson, D.E., Donnenberg, M.S., Welch, R.A. and Mobley, H.L.T. (2004) Transcriptome of Uropathogenic *Escherichia coli* during Urinary Tract Infection. Vol. 72, pp. 6373-6381.
- 104 So, S.S. and Thanassi, D.G. (2006) Analysis of the requirements for pilus biogenesis at the outer membrane usher and the function of the usher C-terminus. *Mol Microbiol*, **60**, 364-375.
- 105 Soto, G.E., Dodson, K.W., Ogg, D., Liu, C., Heuser, J., Knight, S., Kihlberg, J., Jones, C.H. and Hultgren, S.J. (1998) Periplasmic chaperone recognition motif of subunits mediates quaternary interactions in the pilus. *Embo J*, **17**, 6155-6167.
- 106 Stanton, S.L. and Dwyer, P.L. (2000) Preface. *Urinary Tract Infection in the Female*, xi-xii.
- 107 Steitz, T.A. (2008) A structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol*, **9**, 242-253.
- 108 Stites, W.E. (1997) Protein-protein interactions: Interface Structure, Binding Thermodynamics, and Mutational Analysis. *Chem Rev*, **97**, 1233-1250.
- 109 Studier, F.W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, **41**, 207.
- 110 Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, **189**, 113-130.
- 111 Svanborg, C. and Godaly, G. (1997) BACTERIAL VIRULENCE IN URINARY TRACT INFECTION. *Infectious Disease Clinics of North America*, **11**, 513-529.
- 112 Szabo, Z. (1969) Kinetic characterization of complex reaction systems. In Bamford, C. and Tipper, C. (eds.), *Comprehensive Chemical Kinetics*. Elsevier Publishing Company, Amsterdam, pp. 1-80.
- 113 Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv Protein Chem*, **24**, 1-95.
- 114 Thanassi, D.G., Stathopoulos, C., Dodson, K., Geiger, D. and Hultgren, S.J. (2002) Bacterial outer membrane ushers contain distinct targeting and assembly domains for pilus biogenesis. *J Bacteriol*, **184**, 6260-6269.
- 115 Thanassi, D.G., Stathopoulos, C., Karkal, A. and Li, H. (2005) Protein secretion in the absence of ATP: the autotransporter, two-partner secretion and chaperone/usher pathways of gram-negative bacteria (review). *Mol Membr Biol*, **22**, 63-72.

- 116 Van Houdt, R. and Michiels, C.W. (2005) Role of bacterial cell surface structures in *Escherichia coli* biofilm formation. *Research in Microbiology*, **156**, 626-633.
- 117 Verger, D., Bullitt, E., Hultgren, S.J. and Waksman, G. (2007) Crystal structure of the P pilus rod subunit PapA. *PLoS Pathog*, **3**, e73.
- 118 Verger, D., Miller, E., Remaut, H., Waksman, G. and Hultgren, S. (2006) Molecular mechanism of P pilus termination in uropathogenic *Escherichia coli*. *EMBO Rep*, **7**, 1228-1232.
- 119 Vetsch, M., Erilov, D., Moliere, N., Nishiyama, M., Ignatov, O. and Glockshuber, R. (2006) Mechanism of fibre assembly through the chaperone-usher pathway. *EMBO Rep*.
- 120 Vetsch, M. and Glockshuber, R. (2005) Formation of adhesive pili by the chaperone/usher pathway. *Protein Folding Handbook*, **II**, 965-986.
- 121 Vetsch, M., Puorger, C., Spirig, T., Grauschopf, U., Weber-Ban, E.U. and Glockshuber, R. (2004) Pilus chaperones represent a new type of protein-folding catalyst. *Nature*, **431**, 329-333.
- 122 Wolynes, P.G., Onuchic, J.N. and Thirumalai, D. (1995) Navigating the folding routes. *Science*, **267**, 1619-1620.
- 123 Wright, P.E. and Dyson, H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology*, **293**, 321-331.
- 124 Xu, D., Lin, S.L. and Nussinov, R. (1997a) Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. *J Mol Biol*, **265**, 68-84.
- 125 Xu, D., Tsai, C.J. and Nussinov, R. (1997b) Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Eng*, **10**, 999-1012.
- 126 Yang, J.S., Chen, W.W., Skolnick, J. and Shakhnovich, E.I. (2007) All-atom ab initio folding of a diverse set of proteins. *Structure*, **15**, 53-63.
- 127 Young, G.M., Schmiel, D.H. and Miller, V.L. (1999) A new pathway for the secretion of virulence factors by bacteria: The flagellar export apparatus functions as a protein-secretion system. Vol. 96, pp. 6456-6461.
- 128 Zavialov, A.V., Tischenko, V.M., Fooks, L.J., Brandsdal, B.O., Aqvist, J., Zav'yalov, V.P., Macintyre, S. and Knight, S.D. (2005) Resolving the energy paradox of chaperone/usher-mediated fibre assembly. *Biochem J*, **389**, 685-694.
- 129 Zhou, G., Mo, W.J., Sebbel, P., Min, G., Neubert, T.A., Glockshuber, R., Wu, X.R., Sun, T.T. and Kong, X.P. (2001) Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from in vitro FimH binding. *J Cell Sci*, **114**, 4095-4103.



## Danksagung

Mein Dank gilt zu allererst Rudi Glockshuber dafür das er mich in seiner Arbeitsgruppe aufgenommen und mir diese Arbeit ermöglicht hat. Ohne seine wertvolle Hilfe und Unterstützung wäre dieses Projekt sicher nicht möglich gewesen.

Prof. Wolf-Dietrich Hardt danke ich für seine Bereitschaft Teil meines PhD Komitees zu werden und später als Korreferent für meine Doktorarbeit zur Verfügung zu stehen.

Meinen Kollaborateuren an der Universität Zürich, Dr. Oliv Eidam und Dr. Guido Capitani, danke ich für die fruchtbare Zusammenarbeit an der Kristallstruktur des CFD-Komplexes und Dr. Bertran Gerrits am Functional Genomics Center Zürich danke ich für die geduldige Einweisung und Unterstützung an OrbiTrap, qTOF, NanoMate und co.

Dr. Chasper Puorger gilt mein ganz besonderer Dank für die Geduld mit der er sein immenses Wissen zu Proteinfaltung und Pilusbiogenese immer wieder und wieder in den Dienst meines Projektes gestellt hat sowie viele gesellige Abende in, um und um Zürich herum.

Patrick Kunz danke ich für seine Begeisterung für das Projekt und seine Hilfe bei der Untersuchung der FimF<sub>1</sub>-Komplexe während seines Forschungspraktikums.

Natürlich geht mein Dank auch an alle andere Mitstreiter auf dem E-Stock, Ehemalige genauso wie die Jetzigen, ganz besonders aber an die Besatzung des E2, auch die „Fim-Mafia“ genannt, Dr. Chasper Puorger, Christoph Giese, Dr. Oleksander Ignatov, Dr. Mireille Nishiyama, Dr. Michi Vetsch, Dr. Maria Crespo Solans und Dr. Denis Erilov.

Aber auch die „Grenzgänger“ Dr. Patrick Frei, Dr. Verena Finder, Elisabeth Mohorko und Sebastian Falk werde ich in sehr guter Erinnerung behalten, denn mit niemandem konnte man besser feiern wenn es nötig war.

Und nicht zu vergessen natürlich Monika Marti, ohne deren steten Einsatz und klare Worte nichts in unsrem Institut, aber auch nichts bei den HappyHours, funktioniert hätte.

Meinen Eltern danke ich, dass sie so viele Jahre Geduld mit mir hatten, mich immer unterstützt haben und mich niemals fragten, wann ihr Junge denn nun endlich mal einen „richtigen“ Beruf ergreift und von der Schule weg kommt. Nun ist es wohl endlich soweit.

Und nicht zuletzt möchte ich natürlich auch „kurz“ Dr. Mareike Kurz danken, die Sonne, Strand und Heimat aufgegeben hat damit ich hier in der kalten Schweiz meinen „Doktor machen“ konnte. Merci. Für alles. Und: Dito!

Florian Dworkowski,  
Zürich, Dezember 2009