Folding and stability of type 1 pilus subunits from *Escherichia coli*

Chasper Puorger

Institute for Molecular Biology and Biophysics
ETH Zürich

2006

Diss. ETH Nr. 16480

Diss. ETH Nr. 16480

Folding and stability of type 1 pilus subunits from *Escherichia coli*

Dissertation submitted to the

Swiss Federal Institute of Technology Zürich

for the degree of

Doctor of Natural Sciences

presented by

Chasper Puorger

Dipl. Natw. ETH

born on July 1. 1976

citizen of Ramosch, Graubünden, Switzerland

accepted on the recommendation of:

Prof. Dr. R. Glockshuber, examiner

Prof. Dr. E. Weber-Ban, coexaminer

2006

Table of contents

1. Abstract

Adhesive organelles on the surface of a variety of pathogenic bacteria mediate specific binding to host tissue. This is a crucial early event during the establishment of infections. In a large number of gram-negative bacteria a conserved assembly machinery, the “chaperone-usher pathway” is responsible for formation of such organelles that all consist of homologous protein subunits. The subunits form 1:1 complexes with a dedicated chaperone in the periplasm that targets the subunits to the usher in the outer membrane. The usher is an integral outer membrane protein that mediates subunit assembly and translocation to the cell surface. Structures of chaperone-subunit complexes have revealed that the subunits have an incomplete immunoglobulin (Ig)-like fold that is completed by a strand provided by the chaperone. This strand was termed “donor strand”. The donor strand of the chaperone runs parallel to the last strand of the subunit, which leads to a non-canonical Ig-like fold. Upon incorporation of a subunit into the pilus, the donor strand of the chaperone is replaced by an N-terminal segment of the adjacent subunit. This subunit donor strand is oriented anti-parallel relative to the last strand of the subunit, creating a canonical Ig-like fold.

Here, type 1 pili from uropathogenic *Escherichia coli* are used to study fundamental aspects of fiber formation by the chaperone-usher pathway. Each type 1 pilus consists of up to 3000 copies of the subunit FimA and a few copies of the subunits FimF, FimG and FimH at the distal end of the pilus. FimH, the adhesive subunit of type 1 pili has no donor strand and can therefore only locate to the tip of the pilus structure.

We first investigated the role of donor strand orientation and sequence on subunit stability. For this, different donor strands were fused to either the N- or the C-terminus of the pilin domain if FimH, $FimH_p$, which is shortened by the lectin domain. Denaturant-dependent folding transitions showed that such intramolecular donor strand complementation always leads to a marked increase in thermodynamic stability. In $FimH_p$ variants with N-terminally added donor strand, the orientation of the donor strand is as observed in chaperone-subunit complexes. These variants reached equilibrium within one day, demonstrating dynamic folding and unfolding. In contrast, variants with the donor strand of FimF or FimG fused to the C-terminus of

$FimH_p$ did not reach equilibrium. These variants exactly mimic the interaction between FimF or FimG and FimH within the pilus. The same was observed for FimG with the donor strand of FimF attached to the C-terminus ($FimG_{tcF9}$) and, even more pronounced, for FimG complemented with a peptide corresponding to the donor strand of FimF. Together, these data demonstrated that subunits possess an extreme kinetic stability against unfolding once incorporated into the pilus. To test this hypothesis, the unfolding kinetics of complemented subunits were investigated. We consistently found that unfolding of these subunits is extremely slow. For example, self-complemented FimG with a C-terminal FimF donor strand ($FimG_{tcF9}$) e.g. unfolds with a half-life of 100 years at pH 2. In contrast to the extraordinary kinetic stability, the thermodynamic stability of $FimG_{tcF9}$ is high (- 45 kJ mol^{-1}) but not exceptional. In addition, we discovered that dissociation of assembled subunits requires unfolding of the donor strand complemented subunit. Consequently, it is the enormous kinetic stability of the individual subunits in the pilus that efficiently prevents dissociation of fibers that are assembled by the chaperone-usher pathway.

Finally, the role of the periplasmic chaperone FimC in subunit folding was investigated. We found that the chaperone binds non-native subunits and accelerates folding of the bound subunit. Specifically, folding of FimG is accelerated more than 100-fold by FimC. As molecular chaperones in general only increase the yield but not the rate by which their substrates fold, FimC represents a previously unknown type of folding catalyst.

2. Zusammenfassung

Eine Vielzahl pathogener Bakterien besitzt adhesive Organellen auf der Oberfläche, die für das Anhaften an Zellen des Wirtsorganismus verantwortlich sind. Dies ist ein wichtiger erster Schritt im Infektionsprozess. In vielen gram-negativen Bakterien werden die Protein-Untereinheiten über einen Weg in die adhesiven Organellen eingebaut, für den der englische Begriff „chaperone-usher pathway“ geprägt wurde. Die Untereinheiten bilden dabei im Periplasma 1:1 Komplexe mit dem zugehörigen Faltungshelfer (chaperone), welcher die Untereinheiten dann an die Assemblierungsplattform (usher) in der äusseren Membran begleitet. An dieser Assemblierungsplattform, die einen Kanal in der äusseren Membran bildet, werden die Untereinheiten in den Pilus eingebaut und an die Zelloberfläche transferiert.
Aus den Strukturen von Faltungshelfern mit gebundener Untereinheit wurde ersichtlich, dass die Untereinheiten eine unvollständige, immunoglobulin (Ig)-ähnliche Faltung besitzen, die durch einen Strang des Faltungshelfers vervollständigt wird. Dieser Strang wurde als Donor-Strang bezeichnet. Der Donor-Strang vom Faltungshelfer verläuft parallel zum letzten Strang der Untereinheit, was zu einer atypischen Ig-ähnlichen Faltung führt. Nach dem Einbau in den Pilus wird der Donor-Strang des Faltungshelfers durch den N-terminalen Abschnitt der nachfolgenden Untereinheit ersetzt. Dieser Donor-Strang hat eine antiparallele Orientierung zum letzten Strang der Empfängeruntereinheit, was zu einer typischen Ig-ähnlichen Faltung führt.

In dieser Arbeit wurden Typ 1 Pili von uropathogenen *Escherichia coli* benutzt um grundlegende Einsichten in die Ausbildung von fibrillären Strukturen durch den chaperone-usher Weg zu erhalten. Jeder Typ 1 Pilus besteht aus bis zu 3000 Kopien der Untereinheit FimA und einigen wenigen Kopien der Untereinheiten FimF, FimG und FimH am äussersten Ende des Pilus. FimH, das für die Anhaftung an die Wirtszellen verantwortlich ist, besitzt keinen Donor-Strang und kann deshalb nur am äussersten Ende des Pilus eingebaut sein.
Als erstes haben wir die Rolle der Aminosäuresequenz und der Orientierung des Donor-Strangs auf die Stabilität der Untereinheit untersucht. Verschiedene Donor-Stränge wurden entweder an den N- oder C-Terminus der pilin Domäne von FimH, $FimH_p$, einer verkürzten Variante von FimH, gebunden. Denaturierungsmittel-

abhängige Faltungsübergänge zeigten, dass diese Ergänzung durch einen intramolekularen Donor-Strang in jedem Fall zu einer höheren thermodynamischen Stabilität führt. Die Orientierung des Donor-Strangs in den $FimH_p$ Varianten mit N-terminalem Donor-Strang ist dieselbe wie in Komplexen zwischen dem Faltungshelfer und den Untereinheiten. Diese Varianten erlangten schon nach einem Tag den Gleichgewichtszustand in Faltungsübergängen. Im Gegensatz dazu erlangten die Varianten mit den am C-Terminus gebunden Donor-Strängen von FimF oder FimG keinen Gleichgewichtszustand. Diese Varianten haben die gleiche Donor-Strang Orientierung wie bei Untereinheiten im Pilus. Dasselbe Verhalten wurde auch für FimG mit dem Donor-Strang von FimF am C-Terminus ($FimG_{tcF9}$) und noch stärker für den Komplex zwischen FimG und dem Peptid, das dem Donor-Strang von FimF entspricht, beobachtet. Zusammengefasst bedeuten diese Ergebnisse, dass eine extrem hohe kinetische Stabilisierung der Untereinheiten gegen Entfaltung vorhanden ist, wenn die Untereinheiten einmal im Pilus eingebaut sind. Um diese Hypothese zu testen, wurden Entfaltungskinetiken von durch einen Donor-Strang ergänzten Untereinheiten untersucht. Die Untereinheiten entfalten alle sehr langsam. $FimG_{tcF9}$, zum Beispiel, entfaltet mit einer Halbwertszeit von über 100 Jahren bei pH 2. Im Gegensatz zu der extrem hohen kinetischen Stabilität ist die thermodynamische Stabilität von $FimG_{tcF9}$ hoch ($-45\ kJ\ mol^{-1}$) jedoch nicht aussergewöhnlich. Wir haben ausserdem herausgefunden, dass die Dissoziation von assemblierten Untereinheiten, die Entfaltung derselben voraussetzt. Folglich wird die Dissoziation von Organellen, die durch den chaperone-usher Weg gebildet werden, durch die enorme kinetische Stabilität der einzelnen Untereinheiten im Pilus verhindert.

Schliesslich haben wir noch die Rolle des Faltungshelfers FimC bei der Faltung von Pilus Untereinheiten untersucht. Wir haben herausgefunden, dass der Faltungshelfer nicht-gefaltete Untereinheiten bindet und deren Faltung beschleunigt. Die Faltung von FimG wird durch FimC um einen Faktor von mehr als 100 beschleunigt. Da molekulare Faltungshelfer im Allgemeinen die Faltungsausbeute, jedoch nicht die Faltungsgeschwindigkeit erhöhen, stellt FimC einen bis anhin unbekannten Faltungsbeschleuniger dar.

3. Introduction

Type 1 pili are very stable organelles displayed on uropathogenic *Escherichia coli*. They are resistant to sodium dodecyl sulphate and harsh conditions are needed for unfolding. The extraordinary high stability of the pili is important for the persistence of the uropathogenic *E. coli* in the host. Knowledge of the determinants for this high stability are thus of general interest.

Protein stability is determined by thermodynamic and kinetic stability. The thermodynamic stability can be determined with GdmCl-induced equilibrium transitions and is determined as the Gibbs free energy difference between unfolded and folded protein as described below. The kinetic stability of a protein is defined by the height of the transition state compared to the folded state and can be determined with unfolding kinetics.

Type 1 pili are assembled via the chaperone-usher pathway. It involves transient complexes of the subunits with the periplasmic chaperone FimC. FimC prevents aggregation, a feature that is typical for molecular chaperones. In addition it accelerates the folding reaction of the subunits with which it interacts. Thus it belongs to the group of folding catalysts, which encompass among others prolyl peptidyl isomerases and oxido-reductases.

3.1. Protein folding

Folding of proteins has been extensively studied for years. A complete understanding of the complex mechanisms involved was, however, not obtained. It is known that proteins can only carry out their function in a properly folded conformation. Misfolded proteins, in contrast, are the cause of numerous diseases in humans (Dobson, 2003). Therefore an understanding of the mechanisms leading to properly folded and functional polypeptides is highly desirable.

In the sixties and early seventies Anfinsen described experiments, where he and his co-workers succeeded to refold unfolded Ribonuclease A *in vitro* to its native, catalytically active conformation (Anfinsen, 1973). This showed that the information for the tertiary structure is encoded in the primary structure and was taken as evidence that folding was a fast, spontaneous process. In contrast to this, Levinthal (Levinthal, 1968) considered that taking together all possible conformations the

polypeptide chain can adopt (e.g. 3^{100} different conformations for a 101 amino acid protein (Zwanzig et al., 1992)) and the maximal velocity for the conformational changes, extremely slow folding rates would be expected contradicting the actual situation *in vivo* and *in vitro*. The conclusion drawn from this so-called Levinthal paradox was that proteins fold via a well-defined pathway.

These apparently contradictive theories of Anfinsen and Levinthal were then merged to the rugged energy landscape theory. Folding follows a rough funnel with the energy minimised folded structure at its bottom (Chan and Dill, 1998; Dill and Chan, 1997; Wolynes et al., 1995) (Figure 1). The protein can thereby populate microscopic diverse conformations. The rugged energy landscape model gives a good explanation for how proteins can fold without having to explore every possible conformation but still can fold via different pathways.

Two-state folders adopt only two macroscopic states, the native and the unfolded state. But some proteins populate transient intermediates during folding. In the

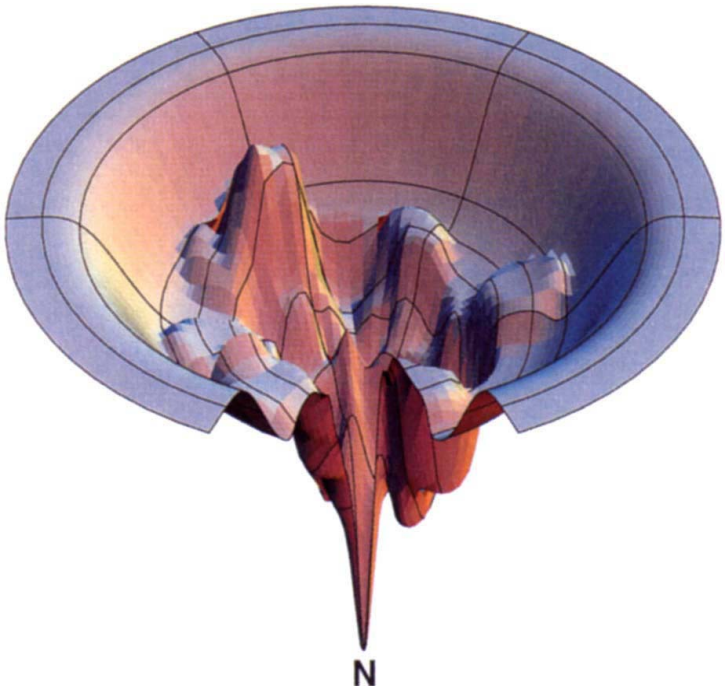


Figure 1: Rugged energy landscape with local minima and energy barriers. Derived from (Dill and Chan, 1997)

folding funnel these states are described as local energy minimum wells, where the proteins are temporarily trapped. Escape from such a minimum needs a certain amount of energy and intermediates are therefore thought to slow down the folding rate.

Intermediates in protein folding can be on- or off-pathway. An on-pathway intermediate is a necessary step in the folding path, whereas an off-pathway intermediate is an unproductive side product which has to be reversed to continue folding.

3.2. Thermodynamic stability in protein folding

The thermodynamic stability of a protein is defined as the Gibbs free energy ΔG^0_{stab} being the difference between the free energy of the unfolded and the free energy of the folded protein. The Gibbs free energy is also defined by the difference of free enthalpy and entropy.

$$\Delta G^0_{stab} = G^0_{folded} - G^0_{unfolded} = \Delta H^0_{stab} - T\Delta S^0_{stab} \qquad (1)$$

Proteins are only marginally stable, the free energy of folding (ΔG_{N-U}) typically ranges from - 20 to - 60 kJ mol^{-1} (Fersht, 2000). The negative values for the free energy of folding reflect the preference for the native state. The thermodynamic stability is the result of numerous weak attractive and repulsive interactions (Dill, 1990; Jaenicke, 1991). Not only interactions between the single atoms of the protein have to be taken into account, but also interactions with the solvent, in particular in the unfolded state, contribute to the thermodynamic stability. Important energetic contributions come from electrostatic interactions between charged residues or residues with a dipole, from nonpolar or van der Waals interactions, from intra- and intermolecular hydrogen bonds and from hydrophobic interactions. The entropy term is defined by the configurational freedom of the polypeptide, water release from polar and hydrophobic residues (see below) and change in chain flexibility. The thermodynamic stability is the result of a fine-tuned balance between all these contributions.

Many of the hydrophobic residues of a protein are shielded from water in the folded state because they are packed in its hydrophobic core, whereas they are solvent-accessible in the unfolded state. The water molecules arrange around accessible hydrophobic residues to form so-called “icebergs” as they make hydrogen bonds between each other but not with the hydrophobic residues (Lee, 1991; Symons, 1975). This leads to an ordered entropically unfavourable arrangement. Upon folding the water molecules are released and thus contribute positively to the entropy (Brady

and Sharp, 1997), partially compensating the entropy loss due to folding of the polypeptide chain.

Both enthalpy and entropy of folding depend on temperature. With increasing temperature the enthalpy of folding becomes more negative, favouring the native state and the entropy of folding becomes more negative, favouring the unfolded state (Figure 2). The temperature dependence of the free energy of folding is derived from the dependence of enthalpy and entropy from temperature (equation 2):

$$\Delta G_{N-U,T_2} = \Delta H_{N-U,T_1} + \Delta C_{N-U,p}\left(T_2 - T_1\right) - T_2\left[\Delta S_{N-U,T_1} + \Delta C_{N-U,p} \ln\left(\frac{T_2}{T_1}\right)\right] \quad (2)$$

With increasing temperature the term $T\Delta S_{N\text{-}U}$ becomes more negative and causes unfolding of the protein when it becomes more negative than $\Delta H_{N\text{-}U}$.
In contrast to chemical reactions, where reactants and products have similar specific

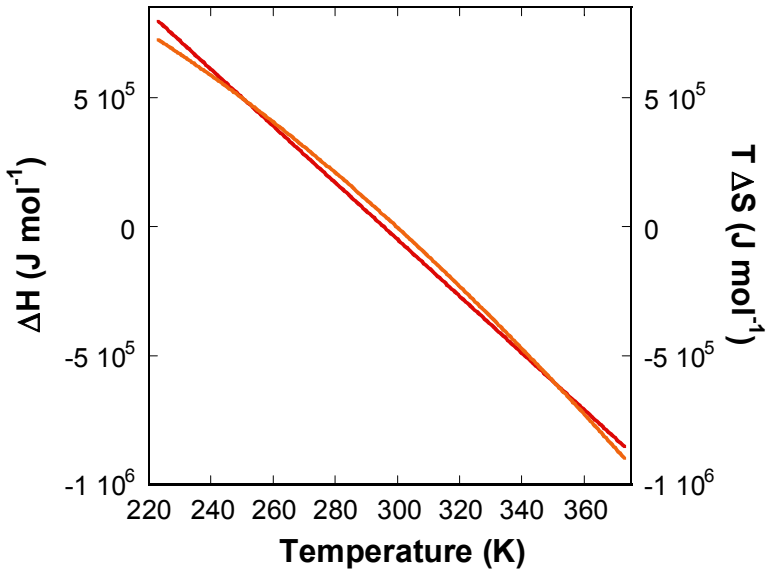


Figure 2: Temperature dependence of enthalpy and entropy. ΔH has a linear dependency on temperature (red) whereas T ΔS depends on the ln of temperature and thus shows a curvature (orange).

heat capacities (C_p), the difference in heat capacity between native and unfolded protein is quite large. For protein folding reactions, the heat capacity change is always negative. The high value for the heat capacity change implicates that there is in addition a cold denaturation point when the enthalpy of folding becomes less

negative with decreasing temperature and is again dominated by the entropy component.

3.3. Transition state theory and Kramers' theory in protein folding

Like in chemical reactions the two-state folding reactions of proteins can be treated with the transition state theory, albeit with some restrictions.

Transition state theory assumes two ground states separated by a distinct transition state. For chemical reaction in the gas phase equation (3) describes the correlation of the rate constant k and the free activation energy $\Delta G^{0\ddagger}$:

$$k = \kappa \frac{k_B T}{h} e^{-\frac{\Delta G^{0\ddagger}}{RT}} \qquad (3)$$

where k_B is the Boltzmann constant and h the Planck constant. The transmission coefficient κ describes the amount of back reaction and can ideally be set 1, leading to the conventional transition state theory.

The pre-exponential factor in equation (3) describes the vibrational motion of a chemical bond, which is about 6 10^{12} s^{-1} at room temperature. However, protein folding is much more complex than a simple chemical reaction. The polypeptide chain has to perform many motions and many non-covalent inter- and intramolecular interactions have to be formed and broken. According to Kramers' theory (Kramers, 1940) and its advancements (Hänggi et al., 1990), the pre-exponential factor is strongly dependent on the viscosity of the solvent and on internal friction in the unfolded and the native conformation. Several groups have estimated the maximal folding rate for proteins and thereby tried to determine the pre-exponential factor (Bieri et al., 1999); (Jacob et al., 1997); (Eaton et al., 1997). The resulting pre-exponential factor though lies in the range of 10^6 and 10^8 s^{-1}. Bieri *et al* have determined the fastest possible contact formation for short peptides by triplet-triplet energy transfers (Bieri et al., 1999). Their data suggest rate constants not faster than 5 10^7 s^{-1}, which can be seen as an absolute upper rate limit for protein folding.

While protein folding is strongly dependent on the viscosity of the solvent and less on the internal friction, unfolding of proteins is more dependent on internal friction than on the solvent viscosity (Ansari et al., 1992). Therefore the pre-exponential factor determined for folding has not the same value as the pre-exponential factor for unfolding. During protein unfolding, other factors like the influx of water into the

hydrophobic core might additionally influence the unfolding rates and it is therefore difficult to estimate the maximal rate constant for unfolding.

The structural parameters of a transition state can be determined with the Φ-value analysis. The Φ-value analysis compares the effect of an amino acid exchange in the polypeptide chain on the stability of the native state N and the transition state ‡ (equation (4)) (Fersht et al., 1992).

$$\Phi = \frac{\Delta\Delta G_{\ddagger-U}}{\Delta\Delta G_{N-U}} \qquad (4)$$

$\Delta\Delta G_{N-U}$ is the difference between the Gibbs free energies of folding of the protein variant and the wild type protein. $\Delta\Delta G_{\ddagger-U}$ is the free energy difference between the transition states of the protein variant and the wild type protein. A Φ-value of 1 denotes that the corresponding amino acid has the same environment in the folded protein and in the transition state. On the other hand a Φ-value of 0 means that a specific interaction found in the native state is not formed in the transition state.

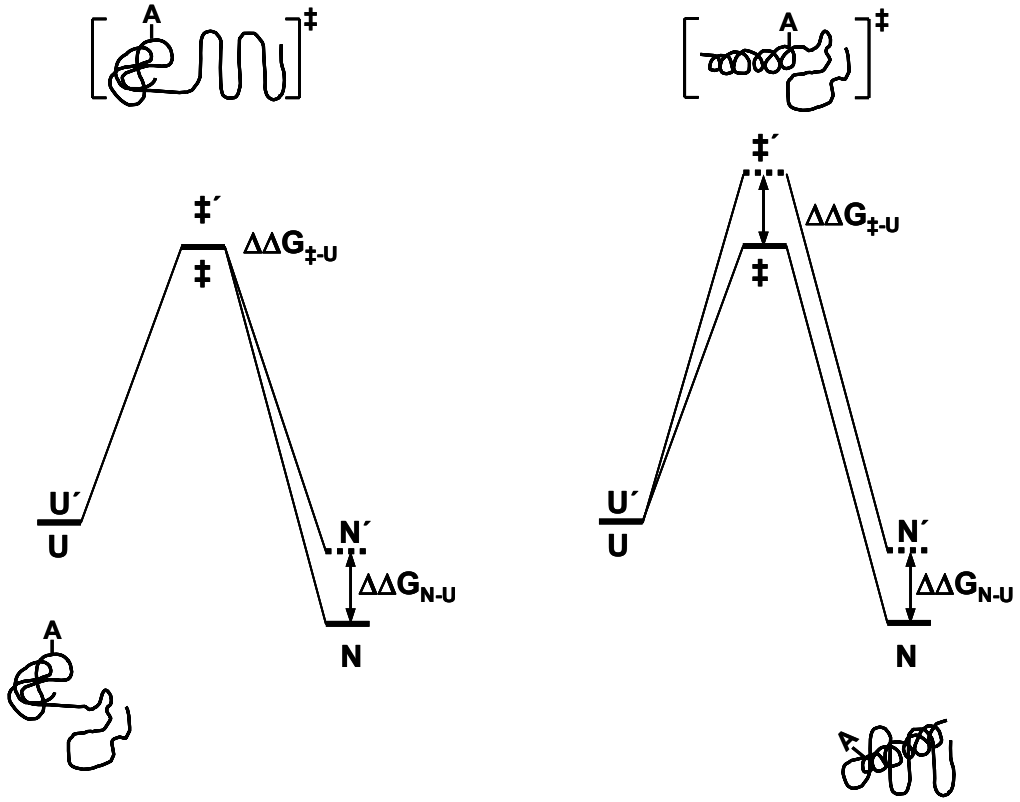


Figure 3: Free energy diagrams for protein folding along the two extreme pathways. Residue A in the helix is mutated giving the mutant which is denoted by ´. The free energies of U and U´ are superposed.

Ideally one can determine structural elements already present in the transition state (Figure 3). However, if the Φ-value is between 0 and 1 the situation becomes more difficult. In this case individual interactions have to be taken into account.

3.4. Molecular chaperones

Anfinsen has shown that proteins refold spontaneously *in vitro*. Information for the tertiary structure thus resides exclusively in the amino acid sequence (Anfinsen, 1973). But not all proteins fold efficiently. This has two main reasons. First, the polypeptide chains emerging from the ribosome do not have the complete primary sequence information until the synthesis is completed. During this period hydrophobic sequences are exposed and the chain misfolds or is degraded because it has not achieved its tertiary structure and is susceptible to proteases. A second important factor is the very crowded cell environment. The total protein concentration inside cells is in the order of 300 mg ml^{-1} (Ellis, 2001). In this environment aggregation of proteins exposing non-native hydrophobic patches is favoured. A large variety of factors is involved in the cellular folding process (Dobson and Karplus, 1999), including molecular chaperones. Molecular chaperones recognise extended hydrophobic patches. In general, molecular chaperones do not accelerate the folding reaction, but rather increase the efficiency of the folding reaction.

Molecular chaperones interacting with newly synthesized proteins are grouped into distinct classes, small chaperones with a size of up to 200 kDa and large chaperones of more than 800 kDa (Ellis and Hartl, 1999). According to their size the mode of action can be distinguished. Examples for small cytoplasmic chaperones in *Escherichia coli* are DnaK (hsp70 family) or DnaJ (hsp40 family). Small chaperones bind to small hydrophobic regions and thus prevent premature folding and aggregation during polypeptide synthesis. They release their substrates in an unfolded state. Function of small chaperones is often associated with ATP-hydrolysis and a variety of co-chaperones.

Large chaperones, such as the *E. coli* chaperonin family member GroEL-GroES, form a cage, where single, fully synthesized proteins are separated from other proteins and the cytoplasmic environment. In this cage the proteins are then allowed to complete folding. GroEL-GroES is the only chaperone essential for viability of *E. coli*.

Not surprising many chaperones are part of the heat shock or stress response. Many proteins tend to unfold and aggregate under stress conditions such as high temperature and need to be either rescued from this unproductive state or removed from the cellular space by proteolytic degradation.

Secretory proteins have to be at least in a partially unfolded conformation for the transfer across the membrane (Arkowitz et al., 1993; Randall and Hardy, 1986). Thus the possibility of unproductive interactions is given, invoking chaperone activity in the periplasm. A variety of periplasmic chaperones interact with proteins exported into the periplasm. Among the known periplasmic chaperones, many have additional functions like formation of disulfide bonds (Dsb-proteins (Missiakas and Raina, 1997)), peptidyl-prolyl *cis-trans* isomerisation (PPIases (Behrens et al., 2001)) or degradation of misfolded proteins (e.g. DegP (Spiess et al., 1999)). Periplasmic chaperones are important for export and incorporation of outer membrane proteins (e.g. SurA (Rouviere and Gross, 1996)) or of components of extra-cellular organelles (e.g. FimC-like chaperone family).
FimC-like chaperones belong to the family of steric chaperones that provide steric information for folding of their substrates (Ellis, 2000). FimC is essential for the assembly of type 1 pili. It interacts with the structural subunits via donor strand complementation, where one β-strand of FimC completes the fold of the subunit (Choudhury et al., 1999). Other members of the steric chaperone family are pro-peptides that accelerate folding of the protein and are subsequently cleaved off. A well investigated member is the pro-peptide of the α-lytic protease (Shinde and Inouye, 2000).

3.5. Folding Catalysts

Folding of proteins is usually fast. However, two steps required for protein folding are very slow: disulfide bond formation and *cis-trans* isomerisation of prolyl peptidyl bonds. *In vivo*, these steps are catalysed by specialised proteins.

Folding of many secreted proteins depends on efficient and correct formation of disulfide bonds that often stabilise the tertiary structure of proteins. Spontaneous disulfide bond formation occurs by air oxidation but this reaction is very slow. In *E.*

coli disulfide bond formation is catalysed by the Dsb proteins and thus folding is accelerated. DsbA is a very potent but unspecific oxidant with a CXXC amino acid motif typical for oxido-reductases. Oxidised DsbA transfers its disulfide-bond to the substrate protein and is thereby reduced. To regain its activity it has to be oxidised by DsbB (Bardwell et al., 1993; Dailey and Berg, 1993; Missiakas et al., 1993), an integral inner membrane protein that transfers the electron equivalents via quinone to molecular oxygen (Bader et al., 1999; Kobayashi and Ito, 1999). Many substrate proteins of DsbA have more than two cysteines. In such proteins, wrong disulfide bonds are likely to be introduced, as DsbA is an unspecific oxidant (Wunderlich and Glockshuber, 1993). These are reversed and corrected by the disulfide isomerase DsbC (Zapun et al., 1995). DsbC has to be kept in the reduced state to retain its isomerase activity, which is performed by the inner membrane protein DsbD. DsbD is in turn reduced by the cytoplasmic protein thioredoxin (Rietsch et al., 1997).

Cis-trans isomerisation of prolyl peptidyl bonds is another very slow step in protein folding. For all peptidyl bonds except prolyl-peptidyl, the *trans*-conformation is energetically strongly favoured. Only about 0.05 % of non-proline peptide bonds in native proteins are in *cis*-conformation (Herzberg and Moult, 1991; MacArthur and Thornton, 1991; Stewart et al., 1990). In Xaa-Pro peptide bonds the Gibbs free energy difference is much lower and the *cis*-conformation is significantly populated (10 – 40 %). Because of the high activation energy (about 90 kJ mol^{-1}) the *cis-trans* isomerisation of Xaa-Pro bond is slow with half-times between 10 - 100 s (Cheng and Bovey, 1977; Grathwohl and Wuthrich, 1976a; Grathwohl and Wuthrich, 1976b). Prolyl-peptidyl isomerase of the cyclophilin and FK binding protein families catalyse isomerisation of prolyl-peptide bonds *in vitro* (Schmid, 1993). Consequently the rate limiting step in folding of many proteins caused by *cis-trans* isomerisation is accelerated.

Pro-peptides form a third kind of folding catalyst (Peters et al., 1998). Various proteins are synthesized with a pro-region that assists folding (Baker et al., 1993; Gray and Mason, 1990). A variety of proteases belongs to this family, of which α-lytic protease is a well characterised example. Folding of mature α-lytic protease has a high activation barrier. A pro-region reduces the height of this kinetic barrier and

stabilises the folded α-lytic protease (Baker et al., 1992). The pro-region is subsequently auto-catalytically cleaved. After cleavage of the pro-peptide, the active protease is trapped in the folded state by a high activation barrier for unfolding.

3.6. Adhesive organelles

Bacterial adhesive organelles are involved in attachment of numerous pathogens to hosts, often causing severe diseases. Several different ways to assemble adhesive organelles on the surface of gram negative bacteria are known, including the chaperone-usher pathway (Sauer et al., 2000), the type II secreton (Peabody et al., 2003) and the extracellular nucleation precipitation pathway (Chapman et al., 2002). At least 37 different adhesive surface filaments are found to assemble via the

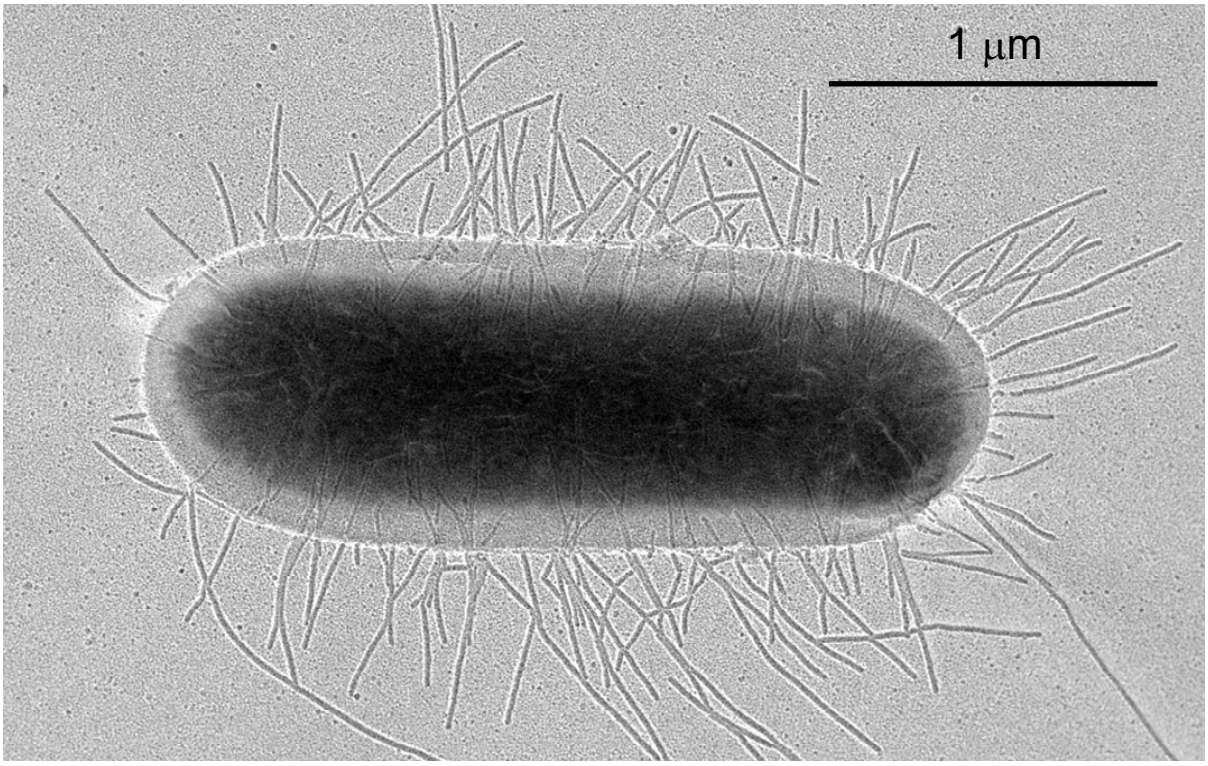


Figure 4: Electron micrograph of *Escherichia coli* bearing type 1 pili (E. Hahn & P. Wild, University of Zürich).

chaperone-usher pathway (Vetsch and Glockshuber, 2005). Prominent and well characterised representatives are type 1 pili on *E. coli* strains causing cystitis (Jones et al., 1993), P pili on *E. coli* strains involved in pyelonephritis and cystitis (Hultgren et al., 1991) and F1 antigen on *Yersinia pestis* causing plague (MacIntyre et al.,

2001). All filaments assembled by the chaperone-usher pathway are composed of multiple copies of one or more subunit proteins and can be divided in two distinct subfamilies according to their architecture: rigid rod like pili (e.g. type 1 pili and P pili) composed of several specialised subunits and flexible non-fimbrial fibres (e.g. F1 antigen) that are composed of only one or two different subunit species. Interestingly, the chaperones of the two subfamilies can be divided into two groups (Hung et al., 1996). FGS chaperones have a short loop between the F1 and the G1 β-strand and mediate the formation of rod like pili (FGS stands for F1-G1 Short). In the FGL (F1-G1 Long) chaperones the corresponding loop is longer. FGL chaperones are involved in formation of non-fimbrial fibres. The assembly mechanism and the interactions of subunits are governed by the same principles in all adhesive fibers assembled by the chaperone-usher pathway.

We used type 1 pili to study the chaperone-usher pathway. An electron micrograph of a type 1 piliated *E. coli* and a schematic illustration of a type 1 pilus are shown in Figures 4 and 5, respectively. The pilus is built up of a distal flexible tip fibrillum joint to the rigid pilus rod. The rod is formed by up to 3000 FimA subunits, assembled in a right handed helix (Hahn et al., 2002). The tip fibrillum is composed of the adhesin FimH followed by the linker proteins FimG and FimF. FimH has two domains, an N-terminal lectin domain and a C-terminal pilin domain with 25 – 30 % sequence identity to the other subunits. The lectin domain of FimH binds mannosylated glycoproteins that are present on the surface of various cell types. In the urinary tract FimH binds to uroplakin Ia and Ib, abundant proteins of the apical cell plasma membrane (Zhou et al., 2001). In addition FimH mediates internalisation of the bacterial cells and thus evasion from extracellular host defence (Mulvey et al., 1998); (Martinez et al., 2000).

For type 1 pilus biogenesis the subunits are transferred in an unfolded conformation across the inner membrane via the SecYEG secretion apparatus. In the periplasm the signal sequence for export is cleaved of by signal peptidase and a disulfide bond conserved in all subunits is introduced by the oxido-reductase DsbA. Introduction of this bond is not necessary for the biogenesis of type 1 pili (Jacob-Dubuisson et al., 1994a). However it was shown that biogenesis in other pilus systems is dependent of functional DsbA. In contrast to the type 1 pilus chaperone FimC, the P pilus chaperone PapD and the F1 antigen chaperone Caf1M have a disulfide bond which

is important for stability of the chaperone or interaction with the subunits (Jacob-Dubuisson et al., 1994a; Zav'yalov et al., 1997).

The periplasmic chaperone FimC assists folding of the subunits (Vetsch et al., 2004), and prevents aggregation (Jones et al., 1997) and premature oligomerisation in the periplasm (Soto et al., 1998). At the outer membrane usher FimD, the subunits are incorporated into the pilus which emerges to the extracellular space. FimC

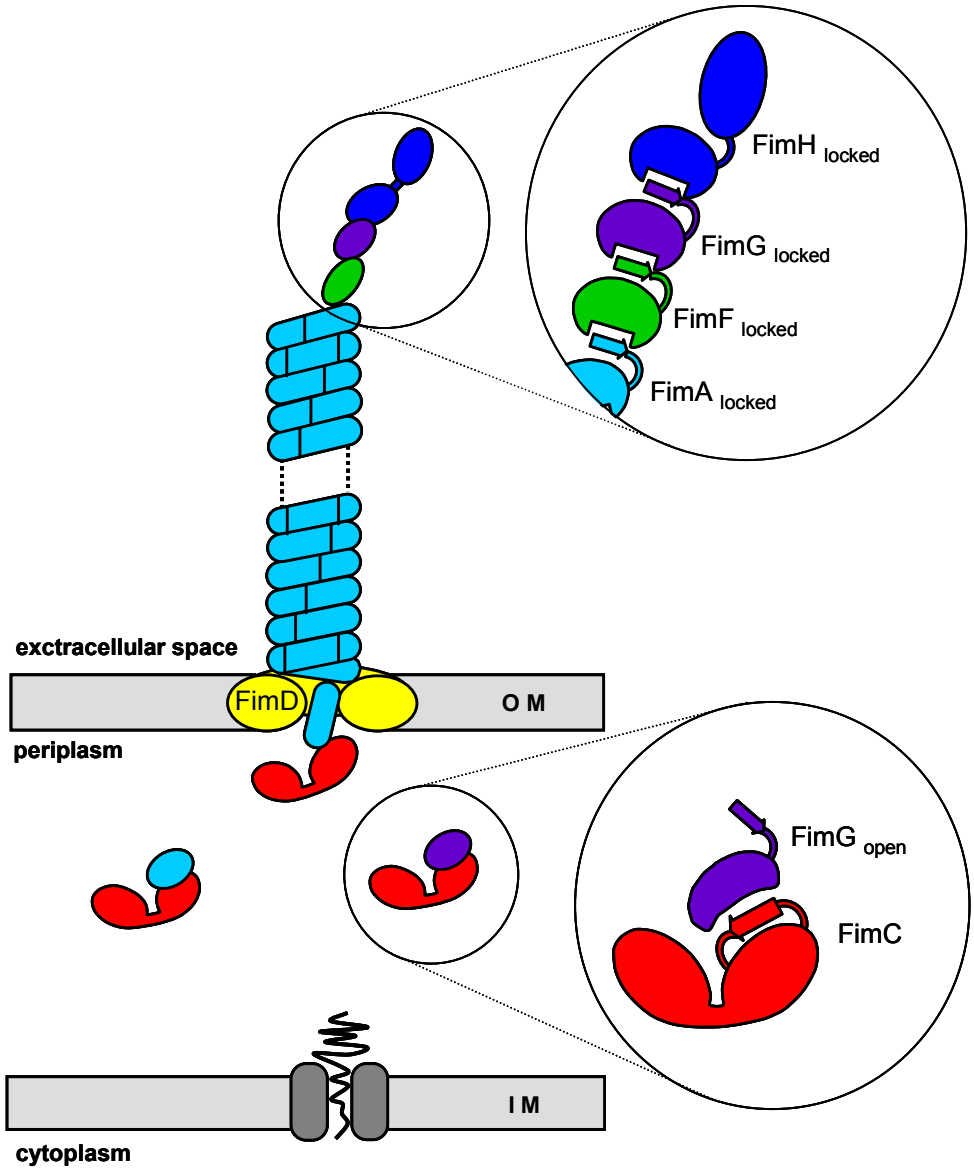


Figure 5: Schematic illustration of a type 1 pilus. Pilus subunits enter the periplasm via the Sec pathway. FimC assists folding of the subunits, prevents their aggregation and retains them in an open, assembly competent form. At the outer membrane usher FimD the subunits are assembled into the pilus as FimC is replaced by another subunit. The subunit thereby undergoes a transition to a more stable, locked conformation. IM: inner membrane, OM: outer membrane.

dissociates from the inserted subunit and is replaced by the next incorporated subunit.

The X-ray structure of the FimC-FimH complex (Choudhury et al., 1999) shows that FimC interacts solely with the pilin domain of FimH and not with the lectin domain. The pilin domain of FimH has an immunoglobulin-like fold but lacks the seventh β-strand. This results in a hydrophobic groove that is solvent-exposed in the monomeric state. In the heterodimeric complex with FimC (Figure 6) the G1 β-strand

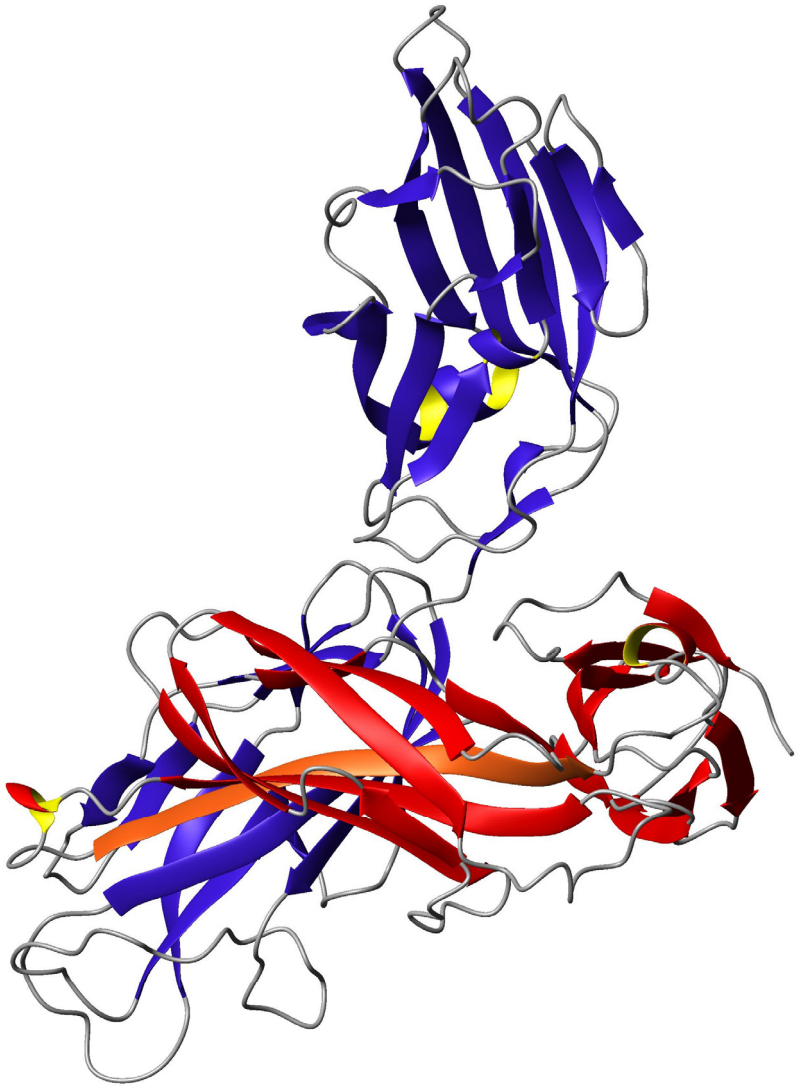


Figure 6: X-Ray structure of the FimC-FimH complex. Adapted from (Choudhury et al., 1999). FimC (red) interacts solely with the pilin domain of FimH (blue, lower part). The main interaction succeeds via the G1 β-strand from FimC (orange) that completes the Ig-like fold of the pilin domain of FimH, a phenomenon termed donor strand complementation.

of the N-terminal domain of the chaperone occupies this cleft and shields it from the solvent. This contact is called donor strand complementation, FimC is the donor and the pilin domain of FimH is the recipient. In the complex with the chaperone, the subunit is retained in an activated, assembly-competent conformation (Sauer et al., 2002; Zavialov et al., 2003).

After incorporation of the subunit into the pilus, the chaperone is replaced by the subsequent subunit. The N-terminal donor strand extension of the next subunit replaces the chaperone donor strand. Whereas the donor strand of the chaperone has a parallel orientation relative to the F-strand of the acceptor subunit (Choudhury et al., 1999; Sauer et al., 1999; Zavialov et al., 2003) the donor strand in the assembled pilus has an antiparallel orientation, resulting in a canonical immunoglobulin-like fold (Sauer et al., 2002; Zavialov et al., 2003) (Figure 7). In addition the subunit collapses to a thermodynamic more stable closed conformation (Zavialov et al., 2005). This transition to the more stable conformation is thought to

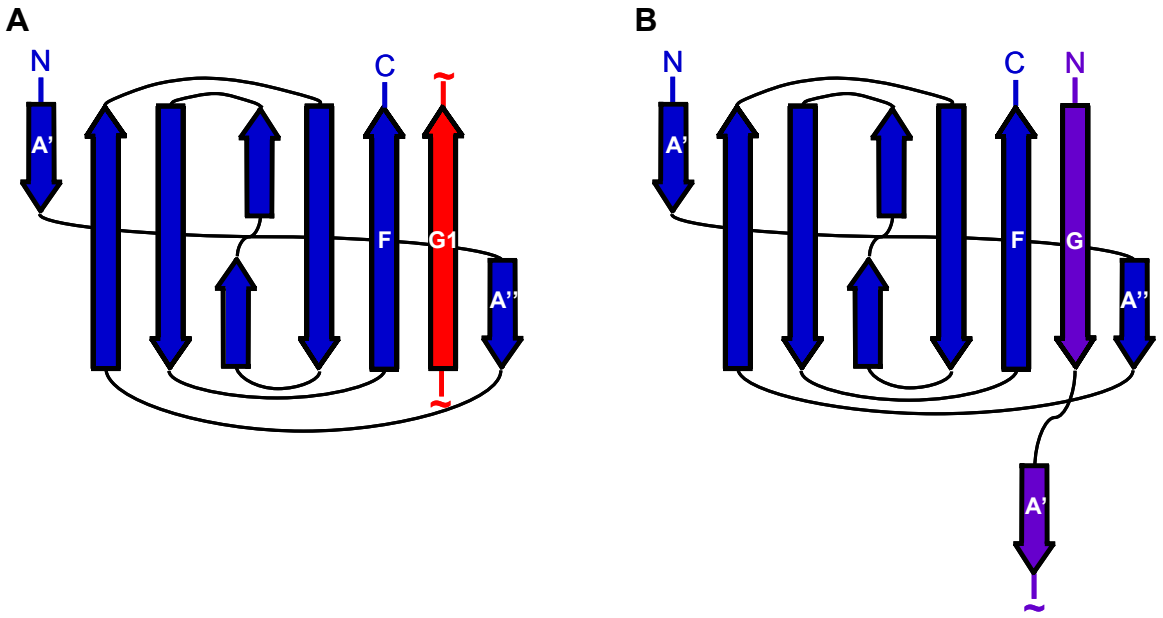


Figure 7: Topology models of pilus subunits with the donor strand of the chaperone **(A)** or the donor strand of another subunit **(B)** inserted.

be the driving force for pilus biogenesis that appears to be a spontaneous process, independent of any cellular energy source (Jacob-Dubuisson et al., 1994b).

The role of the outer membrane usher FimD is not fully understood. It serves as assembly platform and recognises the chaperone-subunit complexes (Saulino et al., 1998). FimD is a trans-membrane protein with at least one periplasmic domain

(Nishiyama et al., 2005; Nishiyama et al., 2003). This N-terminal periplasmic domain interacts with FimC-subunit complexes but neither with the chaperone nor the subunits alone. The first 25 amino acids of FimD are disordered in the monomer but show a finger-like arm structure in the ternary complex with the chaperone-subunit complex. The arm interacts with both FimC and the subunit. It is therefore most likely responsible for recognition of chaperone-subunit complexes and discrimination between the different subunits.

Oligomerisation of subunits from chaperone-subunit complexes *in vitro* is extremely slow (M. Vetsch, personal communication). For the cells it takes about three minutes to assemble the up to 3000 subunits that make an entire pilus (Dodd and Eisenstein, 1984). The assembly platform FimD is a very obvious candidate for accelerating the incorporation of the subunits. About the mechanism one can only speculate at the very moment.

4. Materials and methods

4.1. Materials

Oligonucleotides were purchased from Microsynth (Balgach, Switzerland), the QuikChange Site-Directed Mutagenesis Kit from Stratagene (La Jolla, CA, USA). QA52- and SE52-cellulose were obtained from Whatman (Maidstone, UK), Phenyl Sepharose, the Resource Q, Q-sepharose fast flow, the HiLoad 26/60 Superdex 75 and the Superdex 75 HR 10/30 columns were purchased from GE Healthcare (Little Chalfont, UK). Nickel-NTA was from Qiagen (Hilden, Germany). Bis(2-hydroxyethyl)iminotris(hydroxymethyl)methane (Bis-Tris), ethylenediaminetetraacetic acid (EDTA), ampicillin and isopropyl-β-D-thiogalactopyranoside (IPTG) were from AppliChem GmbH (Darmstadt, Germany). 3-N-morpholino propansulfonic acid (MOPS), 2-Amino-2-(hydroxymethyl)-1,3-propanediol (Tris) and polymyxin-B sulphate were purchased from Sigma (Deisenhofen, Germany). AA-grade Guanidinium Chloride (GdmCl) was obtained from NIGU Chemie (Waldkraiburg, Germany). All other chemicals were purchased from Fluka (Buchs, Switzerland).

Lyophilised peptides (>95 % pure) were purchased from Jerini (first set) and from Genscript (second set) and dissolved in H_2O.

Set 1:

DsFimA: AATTVNGGTVHFKGEVVNAR

DsFimC: TENTLQLAIISRIKR

DsFimG: DVTITVNGKVVAKR

DsFimF$_{G15R}$: ADSTITIRGYVRDNR

Set 2:

DsFimF$_{I5A}$: ADSTATIRGYVRDNG

DsFimF$_{I7A}$: ADSTITARGYVRDNG

DsFimF$_{I7R}$: ADSTITRRGYVRDNG

DsFimF$_{G9A}$: ADSTITIRAYVRDNG

DsFimF$_{V11A}$: ADSTITIRGYARDNG

DsFimF$_{wt}$: ADSTITIRGYVRDNG

DsFimF$_{triple}$: ADSTATARGYARDNG

In set 1 the peptides were designed with a C-terminal arginine to increase the solubility.

4.2. Design of self-complemented variants

Eight self-complemented variants of $FimH_p$ (pilin domain of FimH; residues 158-279 of FimH) were designed. Four with the donor strands of FimA (GGTVHFKGEVVNA) FimC (TENTLQLAIISRIK), FimF (DSTITIRGYVRDN) and FimG (DVTITVNGKVVAK) attached to the C-terminus via the linker DNKQ yielding $FimH_{pcA}$, $FimH_{pcC}$, $FimH_{pcF}$, and $FimH_{pcG}$, respectively, and four with the donor strands of FimA, FimC, FimF and FimG attached to the N-terminus via the linker GGSGG yielding $FimH_{pnA}$, $FimH_{pnC}$, $FimH_{pnF}$ and $FimH_{pnG}$, respectively. In addition self-complemented $FimG_t$ (FimG lacking the natural donor strand, i.e. residues 1-12) variants $FimG_{tcF4}$, $FimG_{tcF7}$ and $FimG_{tcF9}$ were obtained by fusing the FimF donor strand (ADSTITIRGYVRDN) to the C-terminus of $FimG_t$ via either a DNKQ amino acid linker ($FimG_{tcF4}$) or a Gly_7 or a Gly_9-linker ($FimG_{tcF7}$ and $FimG_{tcF9}$, respectively). An overview of the constructs is given in Figure 8.

For cloning of these variants standard PCR procedures and genomic DNA of *E. coli* strain W3110 as template were used. The DNA fragments encoding $FimG_{tcF4}$ and the different self-complemented variants of $FimH_p$ with the natural signal sequences for periplasmic expression were ligated into the pTrc99a vector (Pharmacia) under control of the IPTG inducible trc promoter.

Plasmids encoding for $FimG_{tcF7}$ and $FimG_{tcF9}$ were obtained by site-directed mutagenesis with pFimGtcF4, the plasmid encoding for $FimG_{tcF4}$, as template plasmid.

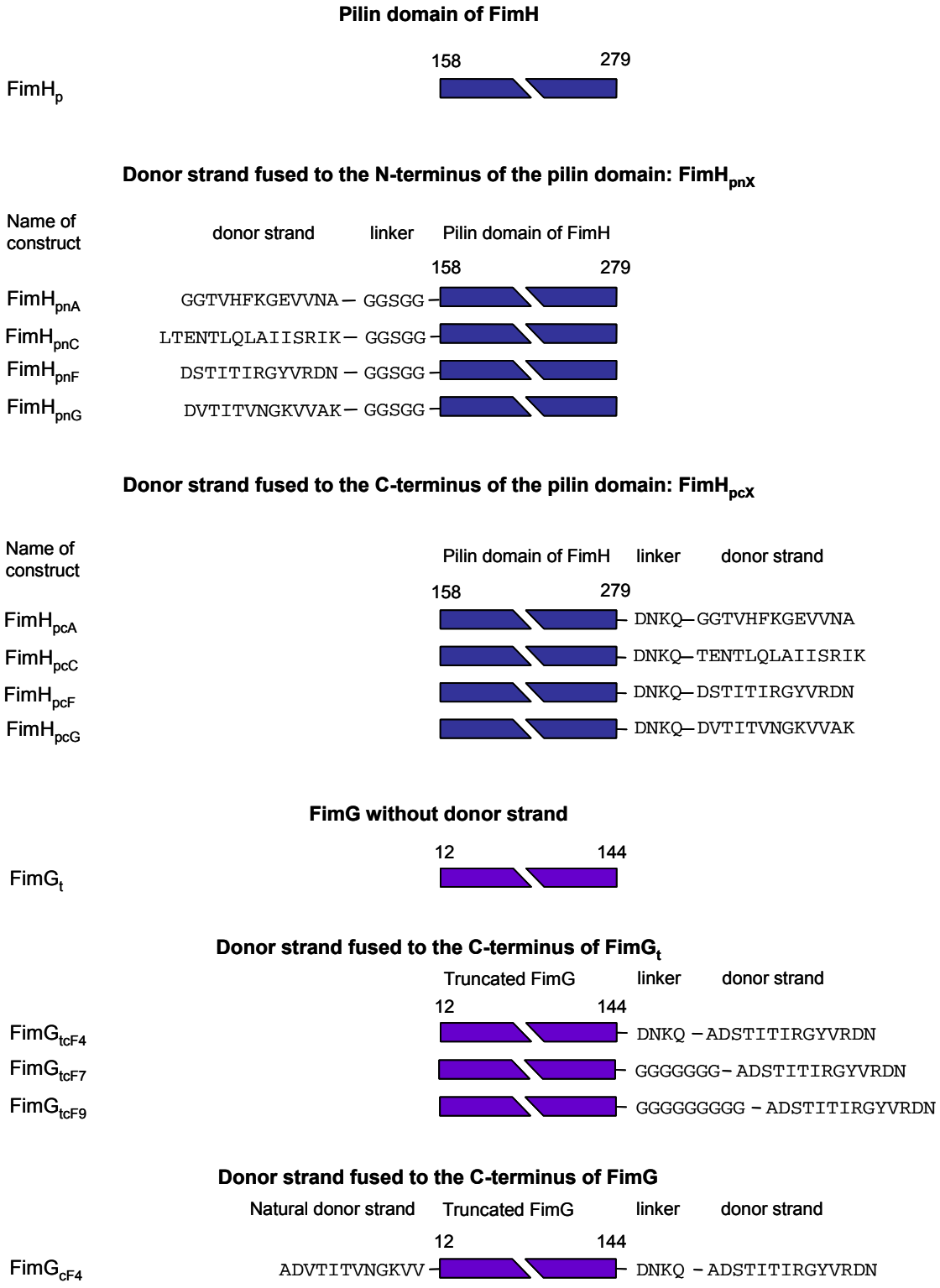


Figure 8: Overview of the different self-complemented constructs of $FimH_p$, $FimG_t$ and FimG.

4.3. Expression tests of the self-complemented $FimH_p$ variants

Expression tests were performed in the protease deficient *E. coli* strain HM125 (Meerman and Georgiou, 1994). Bacteria harboring the different expression plasmids were grown in 75 ml 2YT medium at 30 °C until an OD_{600} of 1 was reached. Production of the periplasmic constructs in presence and absence of a final concentration of 1 mM IPTG was tested. Samples were taken 0, 3, 6 and 24 hours after reaching OD_{600} of 1 (or induction). The bacteria were harvested and the periplasmic extract was isolated by selectively disrupting the outer membrane with polymyxin B sulfate (1 mg per ml of 50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA buffer). The amount of buffer that was added to the harvested cells was calculated according to the optical density of the cell cultures at the time of sample retrieval (6 µl per OD_{600} and ml of cell culture). The cell debris was then removed by centrifugation.

The periplasmic extracts were applied to 18% (w/v) SDS-PAGE gels (Laemmli, 1970) and the proteins subsequently blotted on to a PVDF membrane. A polyclonal anti-FimH rabbit antiserum was used for immunospecific detection of the $FimH_p$ variants. Staining was performed with sheep anti-rabbit antibodies conjugated with alkaline phosphatase as described (Blake et al., 1984).

4.4. Expression and protein purification

All self-complemented variants were expressed in the periplasm of protease deficient *E. coli* strain HM125 (Meerman and Georgiou, 1994).

4.4.1 Self-complemented $FimH_p$ variants

Espression tests revealed that the variants with FimA donor strand are produced at very low amounts. Therefore, only the $FimH_p$ variants with N- and C-terminal elongations with FimC, FimF and FimG donor strands were purified.

Bacteria carrying plasmids encoding $FimH_p$ variants with N-terminal elongations were grown in 2YT medium at 30 °C to an OD_{600} of 1 and induced with 1 mM IPTG (final concentration). Bacteria were then grown for additional 6 hours before harvesting. For the C-terminal elongated $FimH_p$ variants the same bacterial growth conditions were chosen but IPTG was omitted. Cells were harvested by

centrifugation and the periplasmic extract was isolated by selectively disrupting the outer cell membrane with polymyxin-B sulphate (1 mg ml^{-1} of periplasm extraction buffer). The cell debris was removed by centrifugation.

For all $FimH_p$ variants the same purification protocol was used. Periplasmic extract dialysed against 20 mM Tris/HCl, pH 8.0 was applied to an anion exchange column (QA-52 cellulose) at room temperature. 1 M acetate/NaOH, pH 4.2 was added to a final concentration of 100 mM acetate to the flow-through harbouring the protein of interest before dialysing against 20 mM acetate/NaOH, pH 4.2. The protein was then applied to a cation exchange column (SE-52 cellulose) at room temperature and eluted with 600 mM NaCl in a linear gradient. The fractions with pure protein were pooled, dialysed against 10 mM Mops/NaOH, pH 7.0 and stored at – 80 °C. Protein yields per litre of cell culture were 1.4 mg of $FimH_{pcG}$, 0.36 mg of $FimH_{pnG}$, 0.10 mg of $FimH_{pcC}$, 0.65 mg of $FimH_{pcF}$ and 0.36 mg of $FimH_{pnF}$. $FimH_{pnC}$ aggregated during purification.

4.4.2 Self-complemented $FimG_t$ and FimG variants

For production of $FimG_{tcF4}$ and $FimG_{cF4}$ bacteria transformed with pFimGtcF4 or pFimGcF4 were grown for 13 h at 30 °C in 2YT-Amp medium without induction. The cells were harvested by centrifugation, the pellets were resuspended in 50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA with 1 mg ml^{-1} polymyxin B sulphate (10 ml per l cell culture) and shaken for 2 h at 4 °C. After removal of the cell debris by centrifugation, the supernatant was dialysed against 20 mM Bis-Tris/HCl, pH 6.5, 5 mM EDTA and applied to a Q-sepharose fast flow column at 4 °C. The protein was eluted with a linear gradient from 0 to 500 mM NaCl. Fractions containing the protein of interest were pooled, ammonium sulphate was added to a final concentration of 1.2 M and the sample was applied to a Phenyl Sepharose column equilibrated in 20 mM Bis-Tris/HCl, pH 6.5, 0.1 mM EDTA, 1.2 M ammonium sulphate at 4 °C. Elution was performed with a linear gradient from 1.2 to 0.6 M ammonium sulphate. The fractions containing the protein of interest were pooled, concentrated and applied to a Superdex 75 HiLoad 26/60 gel-filtration column equilibrated in 20 mM NaPi, pH 7.4, 115 mM NaCl at 4 °C. Fractions containing the pure protein were pooled and dialysed against 10 mM Mops/NaOH, pH 7.0. The

yield for $FimG_{tcF4}$ was 1 mg per litre of bacterial culture, the yield for $FimG_{cF4}$ was 1.1 mg per litre of bacterial culture.

For both $FimG_{tcF7}$ and $FimG_{tcF9}$ HM125 cells carrying the corresponding plasmids were grown in 2YT-Amp medium at 30 °C. Bacteria for $FimG_{tcF7}$ production were grown to an OD_{600} of 0.8 and then induced with 1 mM IPTG (final concentration) for three hours. Bacteria for $FimG_{tcF9}$ production were grown for 13 hours without induction. After harvesting the periplasmic extracts were isolated as described for the $FimH_p$ constructs. The purification protocol for both $FimG_{tcF7}$ and $FimG_{tcF9}$ was identical. The periplasmic extracts were dialysed against 20 mM Bis-Tris/HCl, pH6.5, 5 mM EDTA and applied to an anion exchange column (QA-52 cellulose) at room temperature. 1 M acetate/NaOH, pH 4.5 was added to the flow-through with the protein of interest to a final concentration of 100 mM acetate. The protein solution was dialysed against 20 mM acetate/NaOH, pH 4.5 and applied to a cation exchange column (SE-52 cellulose) at room temperature. The flow-through with the almost pure protein was concentrated and applied to a Superdex 75 HiLoad 26/60 gel-filtration column equilibrated in 20 mM NaPi, pH 7.4, 115 mM NaCl at room temperature. Fractions containing $FimG_{tcF7}$ or $FimG_{tcF9}$ were pooled dialysed against 10 mM Mops/NaOH, pH 7.0, concentrated and shock-frozen for storage at – 80 °C. Protein yields per litre of bacterial culture were 1.3 mg of $FimG_{tcF7}$ and 1.6 mg of $FimG_{tcF9}$.

Type 1 pili were produced in *E. coli* strain AAEC189 carrying plasmid pSH2 (Orndorff and Falkow, 1984). Pili were purified according to Eshdat (Eshdat et al., 1981).

Production of the following proteins and protein complexes was performed as described: $FimC_{His}$-$FimG_t$ (Nishiyama et al., 2003), $FimC_{His}$-$FimH_p$ (Vetsch et al., 2002), $FimC_{YY}$ and $FimC_{His}$-$FimGt^*$ (Vetsch et al., 2004).

4.5. Protein and peptide concentrations

Protein concentrations were determined by the molar extinction coefficients at 280 nm of 7800 M^{-1} cm^{-1} for $FimH_{pcC}$, 9080 M^{-1} cm^{-1} for $FimH_{pcF}$, 9120 M^{-1} cm^{-1} for $FimH_{pnF}$, 7400 M^{-1} cm^{-1} for $FimH_{pcG}$, 7940 M^{-1} cm^{-1} for $FimH_{pnG}$, 7800 M^{-1} cm^{-1} for $FimH_p$, 13500 M^{-1} cm^{-1} for $FimG_{cF4}$ and $FimG_{tcF4}$, $FimG_{tcF7}$ and $FimG_{tcF9}$, 12200 M^{-1}

cm^{-1} for $FimG_t$ and $FimG_t^*$, 14300 M^{-1} cm^{-1} for $FimC_{YY}$ and 2560 M^{-1} cm^{-1} for FimA. Peptide concentrations were determined by the molar extinction coefficients at 205 nm of 54900 M^{-1} cm^{-1} for DsFimA, 47600 M^{-1} cm^{-1} for DsFimC, 40600 M^{-1} cm^{-1} for DsFimG and 49700 M^{-1} cm^{-1} for DsFimF and all the DsFimF variants.

4.6. Analytical size exclusion chromatography

Samples of $FimG_t$-DsFimF were incubated in GdmCl (concentration between 5 and 7.5 M) in 10 mM Mops/NaOH, pH 7.0 for 20 hours at 25 °C. Samples of Pili were incubated in GdmCl (concentration between 5.5 and 8 M) in 10 mM Mops/NaOH, pH 7.0 for 20 hours at 60 °C. The samples were then applied to a Superdex 75 HR 10/30 column equilibrated in buffers identical to the one that the complex was incubated in. Peak areas of $FimG_t$-DsFimF complex, $FimG_t$ and DsFimF or of unfolded FimA (in the whole-pilus experiment) in the elution profile were determined with PeakFit (SPSS Inc.) and corrected for the effective concentrations.

4.7. Production of $FimH_p$-$FimG_{cF4}$ and $FimH_p$-DsFimF complexes

For production of $FimH_p$-$FimG_{cF4}$ and $FimH_p$-DsFimF complex, $FimC_{His}$-$FimH_p$ was dissociated for four hours in 2.5 M GdmCl, 10 mM Mops/NaOH, pH 7.0 on ice. $FimC_{His}$ was removed by application to a Ni-NTA column equilibrated in the same buffer. Pure $FimH_p$ was in the flow through and was refolded in presence of either 2-fold excess $FimG_{cF4}$ or 10-fold excess DsFimF and concentrated. Both complexes were purified on a Superdex 75 HiLoad 26/60 size-exclusion column equilibrated in 20 mM NaPi, pH 7.4, 115 mM NaCl. Pure complexes were dialysed against 10 mM Mops/NaOH, pH 7.0 and concentrated.

4.8. Production of $FimG_t$-DsFimF and $FimC_{YY}$-$FimG_t$ complexes

Purified $FimG_t$ in 2.75 M GdmCl was obtained as described (Vetsch et al., 2004). For production of the $FimG_t$-DsFimF complex and the complexes with the different DsFimF variants, $FimG_t$ was refolded in presence of a 10-fold molar excess of DsFimF in 0.25 M GdmCl, 10 mM Mops/NaOH, pH 7.0, dialysed against 20 mM Tris/HCl, pH 8.0 and purified by anion exchange chromatography (1 ml Resource Q)

with a linear gradient between 0 and 500 mM NaCl. Pure complex was dialysed against 10 mM Mops/NaOH, pH 7.0.
$FimG_t$-$FimC_{YY}$ complex was obtained by refolding $FimG_t$ in presence of a 5-fold molar excess of $FimC_{YY}$ in 0.25 M GdmCl, 10 mM Mops/NaOH, pH 7.0.

4.9. GdmCl-dependent folding transitions

For unfolding transitions, protein solutions were rapidly diluted with buffer containing different GdmCl concentrations. For refolding transitions, protein was first incubated in 7 M GdmCl, 10 mM Mops/NaOH, pH 7.0 during 20 hours at 25 °C and then rapidly diluted into buffer containing different GdmCl concentrations. Folding transitions were measured at 25 °C by following the change in the far-UV CD signal at 210 nm (self-complemented $FimH_p$ variants, FimC-$FimH_p$ and $FimH_p$-$FimG_{cF4}$) or 230 nm ($FimG_{tcF4}$, $FimG_{tcF9}$ and $FimG_t$-DsFimF) on a JASCO J-715 spectropolarimeter with a peltier element for temperature control.
Transitions of $FimH_{pcF}$ and $FimH_{pcG}$ were measured in a 0.2 mm quartz cuvette at 30 μM protein concentration, all other transitions in a 1 mm quartz cuvette at 10 μM protein concentration. The GdmCl-induced refolding transition of $FimG_t$ was measured as described (Vetsch et al., 2004). Equilibrium unfolding and refolding transitions of $FimG_{tcF9}$ at 60 and 65 °C were measured by the change of the fluorescence signal at 330 nm (excitation at 280 nm) on a QM-7/2003 spectrofluorimeter (PTI) with a peltier element.
The unfolding transition of entire pili was performed by rapid dilution of the pilus samples into buffer containing different amounts of GdmCl and incubation at 60 °C during 20 h. The transition was followed by far-UV CD signal change at 230 nm.
All folding transitions were fitted according to the two-state model (Santoro and Bolen, 1988).

4.10. Thermal unfolding transitions

Thermal unfolding transitions of FimC-$FimH_p$ and $FimG_{cF4}$-$FimH_p$ in 10 mM Mops/NaOH, pH 7.0 were monitored by the change of the far-UV CD signal at 210 nm with a heating rate of 1 °C min^{-1} on a Jasco J-715 spectropolarimeter.

The thermal unfolding transition of $FimG_{tcF9}$ in 10 mM Mops/NaOH, pH 7.0 was monitored by the change of the far-UV CD signal at 230 nm with a heating rate of 0.2 °C min^{-1} on a Jasco J-715 spectropolarimeter. Five measurements were averaged and fitted according to the two-state model (Pace et al., 1998).

4.11. Unfolding kinetics

Unfolding kinetics of $FimG_{tcF4}$, $FimG_{tcF7}$, $FimG_{tcF9}$, $FimG_t$-DsFimF (with all different DsFimF peptide variants), $FimG_t$-$FimC_{YY}$, and intact pili were measured at 25 °C in 25 mM glycine/HCl, pH 2.0 with different GdmCl concentrations. $FimG_t$-$FimC_{YY}$ unfolding kinetics were measured on an SX-18MV stopped-flow spectrofluorimeter (Applied Photophysics) by monitoring the change of fluorescence emission at 330 nm with selective excitation of the tryptophan residue of $FimG_t$ at 295 nm. Unfolding kinetics of $FimG_t$-DsFimF (all DsFimF peptide variants), $FimG_{tcF4}$, $FimG_{tcF7}$ and $FimG_{tcF9}$, were measured on a QM-7/2003 spectrofluorimeter (PTI). The change of emitted fluorescence upon unfolding was monitored at 330 nm (excitation at 280 nm). Unfolding kinetics of type 1 pili were measured by the change of the far-UV CD signal at 230 nm on a Jasco J-715 spectropolarimeter.
All kinetic traces were fitted with a single exponential function.

4.12. Interrupted refolding experiments

All measurements were performed at 25 °C. Unfolded $FimG_t$ or unfolded $FimG_t$* (in 2.75 M GdmCl) was diluted 11-fold with 20 mM Tris/HCl, pH 8.0 containing $FimC_{YY}$, DsFimC, DsFimF or DsFimG. Final concentrations were 1 μM for $FimG_t$ and $FimG_t$*, 5 μM for $FimC_{YY}$ and 40 μM for DsFimC, DsFimF or DsFimG. For interrupted refolding of $FimG_t$ in the presence of $FimC_{YY}$, an SX-18MV stopped flow instrument (Applied Photophysics) was used. After various times of refolding (t_i), the GdmCl concentration was rapidly increased from 0.25 to 2.7 M, and the unfolding kinetics were monitored at 330 nm (excitation at 295 nm). The amplitudes corresponding to unfolding of native FimC-$FimG_t$ were normalized against the unfolding amplitudes of $FimG_t$ that had been completely refolded in the presence of $FimC_{YY}$. This yielded the fraction of native $FimG_t$ at t_i. Refolding of $FimG_t$* in the presence of $FimC_{YY}$ was measured accordingly. Refolding of 1 μM $FimG_t$ in the presence of 40 μM of DsFimC

was measured in the same manner as well, except that unfolding of the native $FimG_t$-DsFimC complex was induced by 4.1 M GdmCl, 50 mM glycine/HCl, pH 1.8, and that additional data points at 1800 and 3600 s were obtained by a combination of manual mixing (initiation of refolding) and stopped-flow mixing (detection of unfolding). In order to measure refolding of $FimG_t$* in the presence of 40 μM DsFimC, refolding was induced by manual mixing. A QM-7/2003 spectrofluorimeter (PTI) was used to record the unfolding reaction, which was induced by transferring the protein to 3.5 M GdmCl, 50 mM glycine/HCl, pH 1.8. The same procedure was performed to measure refolding of $FimG_t$ in presence of 40 μM of either DsFimF or DsFimG, with the difference that unfolding was induced with 2.75 M GdmCl, 100 mM histidine/HCl, pH 1.0 ($FimG_t$-DsFimG) or 6.6 M GdmCl, 100 mM histidine/HCl, pH 1.0 ($FimG_t$-DsFimF).

4.13. Quantitative unfolding for determination of protein-peptide affinities

For determination of the apparent dissociation constant (K_{Dapp}) of $FimG_t$-DsFimC, $FimG_t$-DsFimF and $FimG_t$-DsFimG, unfolded $FimG_t$ (in 2.75 M GdmCl, 20 mM Tris/HCl, pH 8.0) was refolded completely in presence of varying amounts of peptide and unfolded by transferring the protein to 3.25 M GdmCl, 100 mM histidine/HCl, pH 1.0 ($FimG_t$-DsFimC), 6.6 M GdmCl, 100 mM histidine/HCl pH 1.0 ($FimG_t$-DsFimF) or 2.75 M GdmCl, 100 mM histidine/HCl, pH 1.0. The unfolding rates were recorded on a QM-7/2003 spectrofluorimeter (PTI) and fitted with a single exponential function.

4.14. Formation of subunit-peptide complexes under native conditions

To test the displacement of FimC from the FimC-FimH complex by the different peptides, 10 μM FimC-FimH was incubated with a 1.1-fold excess of peptide at 25 °C in 10 mM Mops/NaOH, pH 7.0. The progress of displacement was determined with size exclusion runs on a Superdex 75 HR 10/30 column equilibrated in 20 mM NaPi, pH 7.4, 115 mM NaCl.

4.15. Crystallisation and structure determination

$FimG_t$-DsFimF was prepared as described and stored as a 8 mg/ml solution in 80 mM NaCl and 20 mM Tris/HCl, pH 8.0. Plate-like crystals were grown by vapour diffusion at 20° C, by mixing 2 μl of protein solution with an equal volume of a 4:1 mix of 27.5 % PEG 1500 and 0.1 M $CoCl_2$, and equilibration against 500 μl 27.5% PEG 1500 solution. Before direct flash-cooling in liquid nitrogen gas stream, crystals were cryoprotected in a solution of 10% ethylene glycol and 27.5 % PEG 1500.

X-ray data were collected at the X06SA beamline at the Swiss Light Source (Villigen PSI, Switzerland) (Table 3). Data from a triple wavelength multiple anomalous dispersion (MAD) experiment around the Co^{2+} K-edge (λ=1.6 Å, E= 7.7 keV) was processed using DENZO and SCALEPACK (Otwinowski and Minor, 1997) in $P2_1$ to 1.85 Å. The structure was solved using AutoSHARP (www.globalphasing.com), which autobuilt a model comprising 142 out of 147 residues of the asymmetric unit content. The final model was rebuilt using O (Jones et al., 1991) and was refined isotropically to 1.34 Å using CNS (Brunger et al., 1998). Map quality improved upon final anisotropic all-atom refinement using SHELX-97 (Shledrick and Schneider, 1997) and R_{free} dropped by 3% to 18.5 %. The final R_{work} value was 14.1 % (Table 3).

5. Results

5.1. Kinetic and thermodynamic stability of type 1 pilus subunits

Type 1 pili and other pilus structures are extremely stable assemblies of proteins. Harsh methods, e.g. acidic pH in combination with high temperature, are necessary for dissociation of these structures into single subunits (Eshdat et al., 1981; Krogfelt and Klemm, 1988). The subunits in the pilus interact very tightly with each other via a mechanism called donor strand complementation. The donor strand, which is provided by the adjacent subunit inserts antiparallel relative to the F-strand of the recipient subunit. This interaction is likely to contribute to the enormous stability of pili (Sauer et al., 2002; Zavialov et al., 2005). Donor strand complementation is also found in the transient periplasmic chaperone-subunit complexes. The donor strand of the chaperone, however, is oriented parallel relative to the F-strand of the recipient subunit (Choudhury et al., 1999; Sauer et al., 1999; Zavialov et al., 2003). To investigate the influence of donor strands on the stability of type 1 pilus subunits we designed self-complemented subunits, where the donor strand is attached via a linker to the subunit. These constructs were compared with intermolecular complemented complexes between subunits and peptides derived from donor strands.

5.1.1. Construction of donor strand complemented model proteins

The X-Ray structure of FimH in complex with FimC (Choudhury et al., 1999) reveals that FimC donates a β-strand to the pilin domain of FimH (termed $FimH_p$). N- and C-terminus of $FimH_p$ are in close proximity (Figure 9). Self-complemented variants of $FimH_p$ thus represent a good model system to determine the effect of different donor strands and their orientation on subunit stability. Attaching a donor strand to either

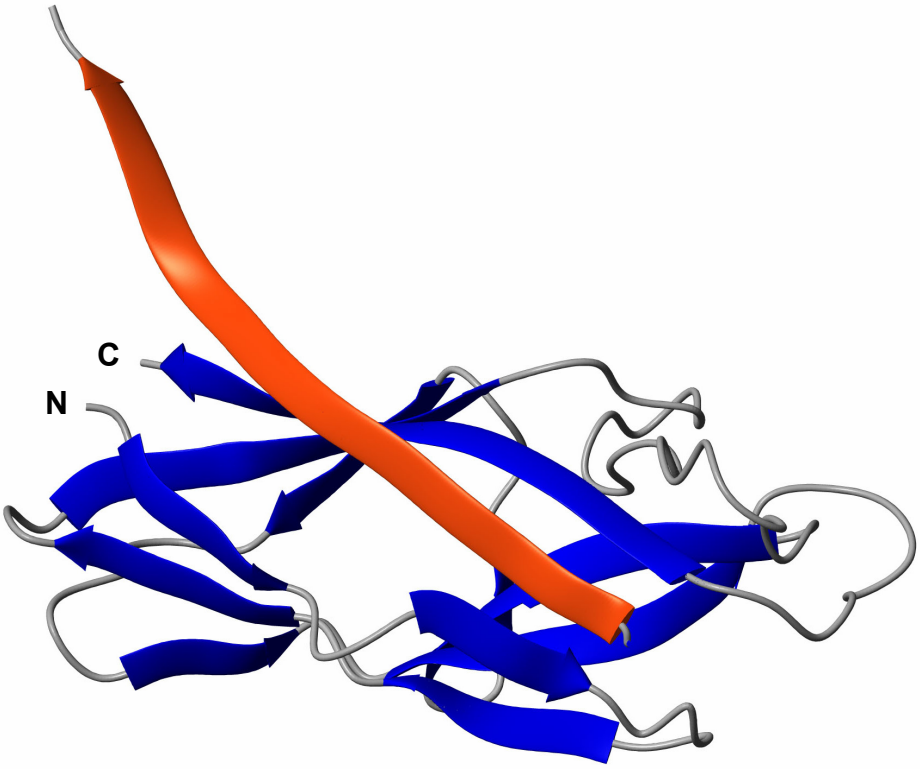


Figure 9: Pilin domain of FimH ($FimH_p$) with the donor strand of FimC, as observed in the X-ray structure of the FimC-FimH complex (Choudhury et al., 1999). For simplicity, the N-terminal lectin domain of FimH is omitted. N- and C-terminus of $FimH_p$ are labeled. $FimH_p$ and the donor strand of FimC are colored blue and orange, respectively.

terminus via a linker of sufficient length results in self-complemented variants of the pilin domain of FimH with the donor strand in opposite orientations. Attachment of a donor strand to the N-terminus results in a parallel orientation relative to the C-terminal β-strand of the subunit, similar to the situation found in chaperone subunit complexes (Choudhury et al., 1999; Sauer et al., 1999; Zavialov et al., 2003). Conversely, attachment of a donor strand to the C-terminus results in an anti-parallel

orientation as found between pilus subunits (Sauer et al., 2002; Zavialov et al., 2003; Zavialov et al., 2005).

Constructs with the donor strands from the subunits FimA, FimF and FimG and from the chaperone FimC attached to both termini were made (Figure 8 in Materials and Methods). For the constructs with N-terminal donor strand elongations a linker sequence of five amino acids was chosen (GGSGG). These constructs were named $FimH_{pnX}$, where $FimH_p$ stands for pilin domain of FimH, n for the site of attachment (N-terminus) and X describes the provenience of the donor strand (A for FimA, C for FimC, F for FimF or G for FimG). For the C-terminal elongated constructs a linker sequence of four amino acids (DNKQ) was used, a sequence often found in β-turns. These constructs were named $FimH_{pcX}$ in accordance to the N-terminal constructs, where c denotes the C-terminus as attachment site for the donor strand. Expression tests revealed that all variants, except for $FimH_{pcA}$, accumulate in the periplasm to levels exceeding the one of $FimH_p$ (Figure 10). Variants with C-terminal attached donor strand were generally expressed at higher levels than constructs with the same donor strand attached to the N-terminus (with exception of $FimH_{pcA}$). The

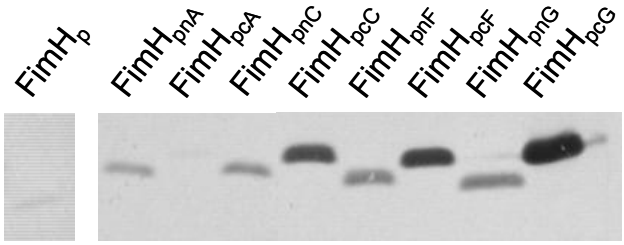


Figure 10: Expression tests of self-complemented $FimH_p$ variants. The western blot was stained with anti FimH polyclonal antibody. Periplasmic extracts retrieved six hours after the bacterial cell cultures have reached an OD_{600} of 1 are shown. Samples of C-terminal self-complemented $FimH_p$ variants are without IPTG induction, samples of N-terminal self-complemented $FimH_p$ variants are with IPTG induction.

constructs with the donor strand of FimG showed the highest expression levels, followed by the FimF, FimC and FimA donor strand constructs. This agrees with the assumption that FimG is the neighbor of FimH and complements its fold. Interestingly, the variants with C-terminal attached donor strands accumulate more if expression is not induced with IPTG.

The different expression levels might reflect differences in free energy of folding (ΔG_{N-U}). To test this hypothesis the constructs with the donor strands of FimC, FimF and FimG attached at either terminus were over-expressed and purified to homogeneity with exception of $FimH_{pnC}$. This variant with N-terminally attached FimC donor strand aggregated during purification and could therefore not be investigated further.

The pilin domain of $FimH_p$ has the disadvantage that it can only be investigated by circular dichroism. For this method high protein concentrations are needed and no fast kinetics can be performed. FimG has one tryptophane and can be investigated with fluorescence spectroscopy. We decided to additionally design variants of FimG in which the donor strand of FimF was attached with a linker to the C-terminus while the 12 N-terminal amino acids representing the natural donor strand of FimG were removed (termed $FimG_t$). To investigate the influence of the linker on folding of the self-complemented subunit, variants with different linkers were constructed. One variant had a linker consisting of the four amino acids DNKQ ($FimG_{tcF4}$), additional variants had linkers with seven and nine glycines ($FimG_{tcF7}$ and $FimG_{tcF9}$, respectively). The nomenclature was adapted from the $FimH_p$ variants with the numbers describing the length of the linker between subunit and donor strand. All $FimG_t$ variants were purified to homogeneity.

5.1.2. Effects of covalently bound donor strands on subunit stability

The different expression levels of the individual self-complemented variants indicate differences in stability. The stabilities of the different purified self-complemented $FimH_p$ variants were investigated *in vitro* by GdmCl-induced unfolding and refolding equilibrium transitions. All investigated variants are noticeable more stable than non-complemented $FimH_p$ (Figure 11). Interestingly, folding of some of the self-complemented $FimH_p$ variants was not in equilibrium after 2 days of incubation at 25 °C (Figure 11 A and B). High denaturant concentrations were necessary to unfold the two variants $FimH_{pcG}$ and $FimH_{pcF}$. Refolding in contrast occurred only at low GdmCl concentrations. The difference between unfolding and refolding is a phenomenon called hysteresis. The differences between unfolding and refolding branches are 3.65 M for $FimH_{pcG}$ and 2.63 for $FimH_{pcF}$ after 2 days of incubation at 25 °C. For $FimH_{pcG}$, the samples were re-measured after 5 and 12 days revealing

that the unfolding and refolding transitions move slowly towards each other with time (Figure 11 A). Interestingly $FimH_{pcG}$ exactly mimics the interaction between FimH

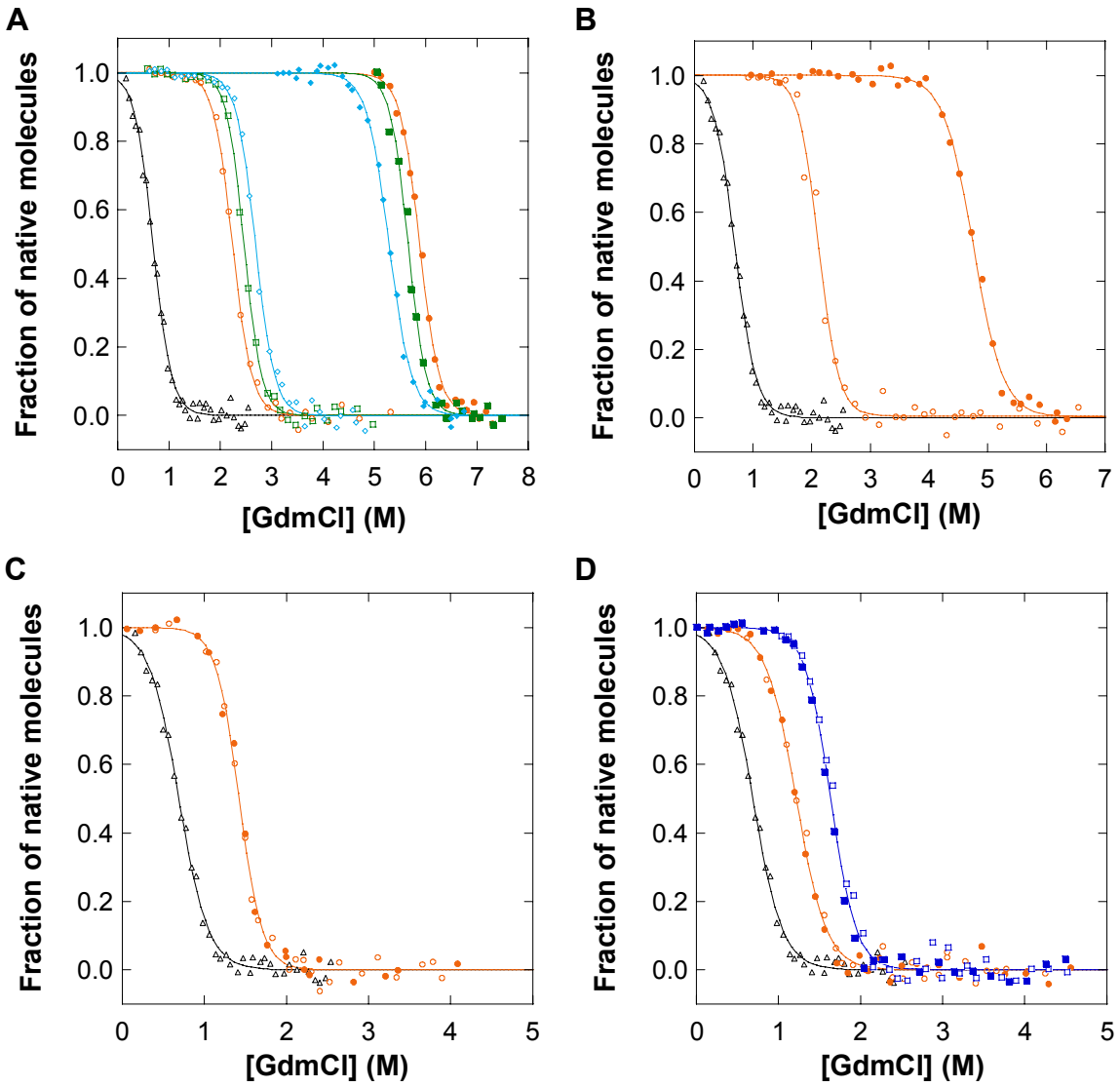


Figure 11: GdmCl induced equilibrium transitions of different donor strand complemented $FimH_p$ constructs. All transitions were monitored by the far-UV CD signal at 210 nm after incubation at 25 °C and pH 7.0. **(A)** Transitions of $FimH_{pcG}$ after 2 days of unfolding (●) and refolding (○), after 5 days of unfolding (■) and refolding (□) and after 12 days of unfolding (♦) and refolding (◊). **(B)** Transitions of $FimH_{pcF}$ after 2 days of incubation (unfolding ● and refolding ○). **(C)** Unfolding (●) and refolding (○) transitions of $FimH_{pcC}$ after 2 days. **(D)** Transitions of $FimH_{pnG}$ and $FimH_{pnF}$. Unfolding (filled symbols) and refolding transitions (open symbols) of the self complemented constructs of $FimH_p$ with the donor strands of FimG ($FimH_{pnG}$, squares) and FimF ($FimH_{pnF}$, circles) attached to the N-terminus after 2 days. In all panels, the refolding transition of non-complemented $FimH_p$ is shown for comparison (Δ). The GdmCl induced equilibrium transitions were evaluated according to the two-state model and normalised.

and FimG in the pilus (Figures 5 and 7), where the donor strand of FimG complements FimH in an anti-parallel orientation. The other variants, where the donor strand of FimC is attached to the C-terminus of $FimH_p$ (Figure 11 C) and the two constructs with the donor strands of FimF and FimG attached to the N-terminus of $FimH_p$ (Figure 11 D) are in equilibrium after 48 hours at 25 °C. They are all significantly more stable than $FimH_p$ (Table 1). For $FimH_{pcC}$ no hysteresis is observed although the donor strand has the same orientation as the donor strand in $FimH_{pcG}$ and $FimH_{pcF}$, which both exhibit a pronounced hystersis. Thus the sequence of the donor strand is an important determinant in subunit stability.

The hysteresis in the folding transitions of $FimH_{pcF}$ and $FimH_{pcG}$ could in principle be due to oligomeric complexes (Lai et al., 1997; Miller et al., 1998). These variants of $FimH_p$ were therefore analysed by size exclusion chromatography (Figure 12). They eluted at retention volumes expected for the monomeric species (at around 14 ml). Consequently, the hysteresis found for the two variants $FimH_{pcF}$ and $FimH_{pcG}$ reflects unusually high activation energies for the folding and unfolding reactions.

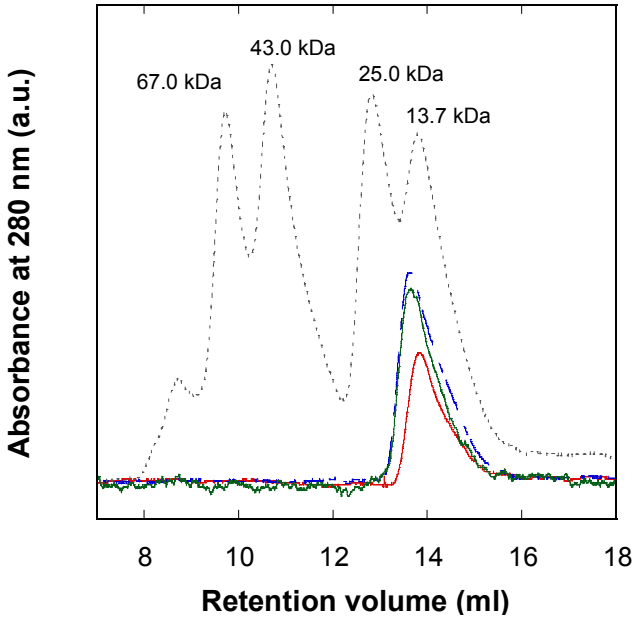


Figure 12: Size exclusion chromatography runs for determination of the oligomeric state of self-complemented $FimH_p$ variants. The profiles for $FimH_{pnG}$ (red), $FimH_{pcG}$ (blue) and $FimH_{pcC}$ (green) are shown. To estimate the approximate molecular weight of the $FimH_p$ variants, a standard profile of a mixture of proteins with the indicated molecular mass is also shown (dotted line).

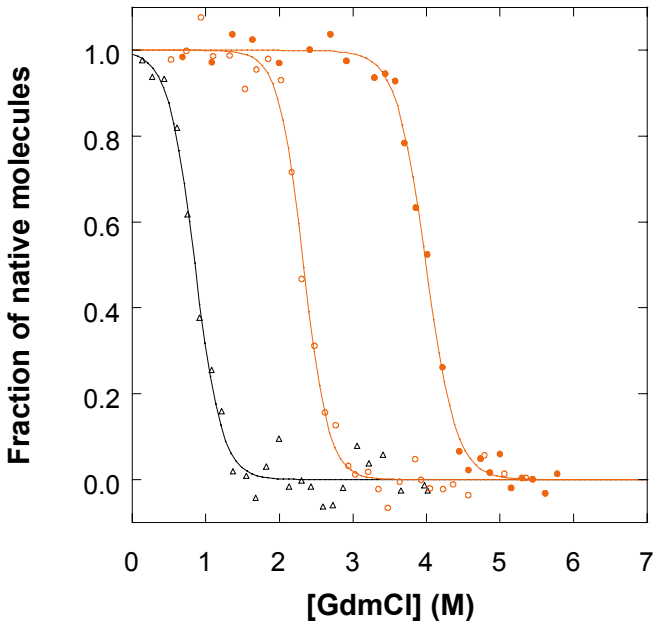


Figure 13: GdmCl dependent refolding and unfolding transitions of the self-complemented $FimG_{tcF4}$. Unfolding (●) and refolding (○) transitions of $FimG_{tcF4}$ after 20 hours of incubation at 25 °C, pH7.0 compared with the refolding transition of the non-complemented $FimG_t$ (Δ). Data were analysed according the two-state model and normalised although the transitions are not in equilibrium.

A similar behaviour was observed for $FimG_{tcF4}$ (Figure 13), a self-complemented variant of FimG. $FimG_{tcF4}$ has the donor strand of FimF attached to the C-terminus via a DNKQ linker. Compared to $FimH_{pcF}$ and $FimH_{pcG}$ (Figure 11 A and B) the hystersis of $FimG_{tcF4}$ is notably smaller by more than 1.1 M.

All GdmCl-induced equilibrium transitions were evaluated according to the two-state model. The thermodynamic data for the variants which have reached equilibrium are summarised in Table 1.

	ΔG (kJ mol^{-1})	Cooperativity of folding (kJ mol^{-1} M^{-1})	Midpoint of transition (M)
$FimH_p$	-9.5 ± 0.9 [a]	13.8 ± 1.2 [a]	0.69 ± 0.09 [a]
$FimH_{pnF}$	-16.4 ± 0.8	13.5 ± 0.6	1.2 ± 0.1
$FimH_{pnG}$	-26.5 ± 2.1	16.2 ± 1.2	1.6 ± 0.2
$FimH_{pcC}$	-26.8 ± 2.2	18.7 ± 1.3	1.4 ± 0.2
$FimG_t$	-10.9 ± 1.0 [a]	13.1 ± 0.8 [a]	0.84 ± 0.11 [a]
$FimH_{pcF}$	n. d.	n. d.	n. d.
$FimH_{pcG}$	n. d.	n. d.	n. d.
$FimG_{tcF4}$	n. d.	n. d.	n. d.

Table 1: Thermodynamic parameters of the different self-complemented variants and of non-complemented $FimH_p$ and $FimG_t$. For the variants with hysteresis the parameters could not be determined.

[a] Values for $FimH_p$ and $FimG_t$ are derived from refolding transitions only, as unfolding transitions were hampered by precipitation of the proteins at low denaturant concentrations.

5.1.3. Influence of FimC and FimG$_{cF4}$ on the stability of FimH$_p$

I have shown that sequence and orientation of complementing β-strands influence subunit stability profoundly. Next, we attempted to compare the stability of FimH$_p$ bound to the chaperone FimC (as in the periplasm) with the stability of FimH$_p$ bound to FimG (as in the final pilus structure). We prepared two defined complexes, the complex of FimC with FimH$_p$ and the complex between self-complemented FimG$_{cF4}$ and FimH$_p$. FimG$_{cF4}$ is equivalent to FimG$_{tcF4}$ but has its N-terminal wild-type donor strand. It thus has the ability to bind FimH$_p$ but, in contrast to FimG, can not form homo-oligomeric complexes because of the artificial donor strand at its C-terminus, which occupies the acceptor grove of FimG. The FimC-FimH$_p$ complex was produced *in vivo* by co-expression and purified. For the FimG$_{cF4}$-FimH$_p$ complex, FimH$_p$ was dissociated from FimC under denaturing conditions and then refolded in presence of FimG$_{cF4}$. The complex was finally purified by size exclusion chromatography.

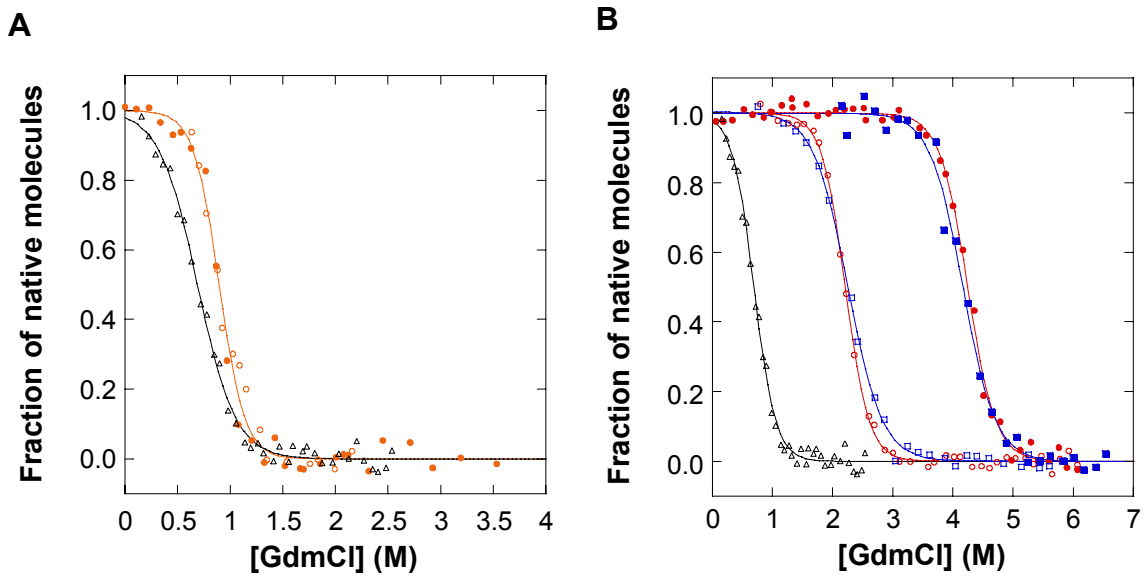


Figure 14: Determination of the effect of chaperone and subunit on the stability of FimH$_p$. **(A)** GdmCl induced unfolding (●) and refolding (○) transitions of FimH$_p$ in presence of FimC after 20 hours of incubation at 25 °C and pH 7.0. (Δ) refolding transition of non-complemented FimH$_p$. **(B)** GdmCl induced unfolding (filled symbols) and refolding (open symbols) transitions and of FimH$_p$ in presence of FimG$_{cF4}$ (circles) and of FimG$_{cF4}$ alone (squares) after 20 hours of incubation at 25 °C and pH 7.0. (Δ) refolding transition of non-complemented FimH$_p$. Data were normalised for the pre- and post-transition baselines.

The thermodynamic stabilities of $FimH_p$ in both complexes were determined with GdmCl-dependent unfolding and refolding equilibrium transitions. Unfolding and refolding transitions of the FimC-$FimH_p$ complex were in equilibrium after 20 hours of incubation at 25 °C and pH 7.0 (Figure 14 A). This is notable as the unfolding reaction is a uni-molecular process, whereas the refolding transition is a bi-molecular reaction. $FimH_p$ in the FimC-$FimH_p$ complex has a transition midpoint of 0.90 ± 0.09 M compared to the transition midpoint of 0.69 ± 0.09 M for $FimH_p$. GdmCl-induced unfolding and refolding transitions of $FimH_p$ in the $FimH_p$-$FimG_{cF4}$ complex exhibit a hysteresis and are very similar to the transitions of $FimG_{cF4}$ alone (Figure 14 B). $FimH_p$ is thus much more stable in complex with $FimG_{cF4}$ than in complex with FimC.

Both complexes were also analysed by thermal unfolding transitions (Figure 15). In contrast to the GdmCl-induced transitions (Figure 14), where the complexes show one transition, two transitions are visible in temperature induced unfolding transitions (red curves in Figure 15). For the FimC-$FimH_p$ complex the midpoint of the first transition is at higher temperature (4 °C) than for $FimH_p$ alone, whereas the second

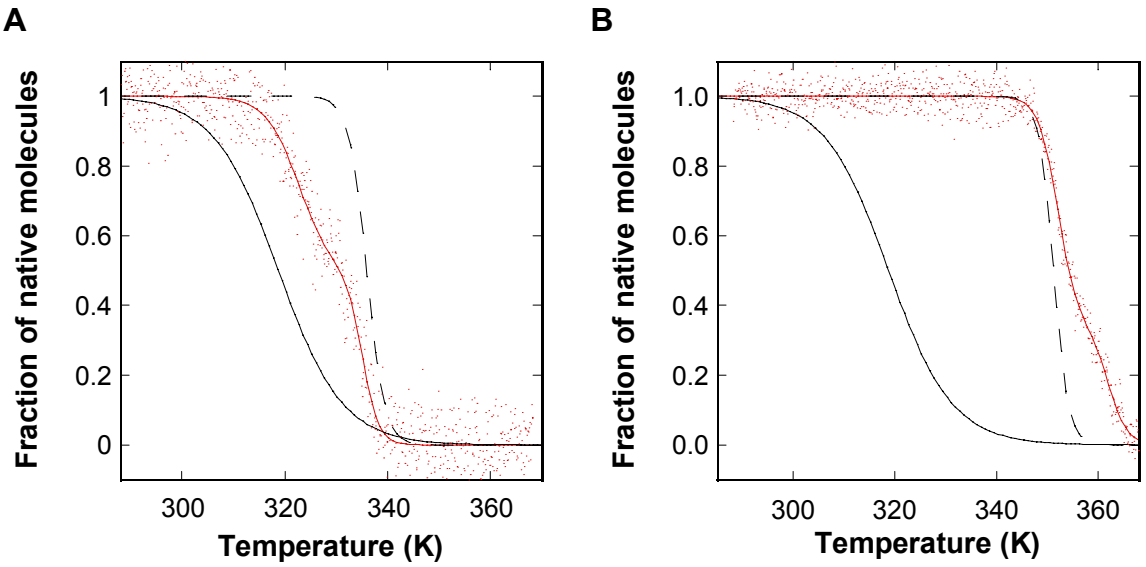


Figure 15: Thermal unfolding transitions of $FimH_p$ in complex with FimC and $FimG_{cF4}$ at pH 7.0. **(A)** unfolding transition of FimC-$FimH_p$ (red) compared to the unfolding transition of FimC (dashed black line) and $FimH_p$ (solid black line). **(B)** Unfolding transition of $FimH_p$-$FimG_{cF4}$ complex (red) compared with the transition of $FimG_{cF4}$ (dashed black line) and $FimH_p$ (solid black line). Transitions of monomeric proteins were analysed according the two-state model and normalized.

transition is very similar to the transition midpoint of monomeric FimC (0.5 °C deviation) (Figure 15 A). This suggests that $FimH_p$ is stabilised by the chaperone but unfolds at lower temperature than FimC. The midpoint of the first transition of the $FimH_p$-$FimG_{cF4}$ complex is in the range of the transition midpoint of $FimG_{cF4}$ alone (0.5 °C deviation), whereas the second transition midpoint is at higher temperatures (10 °C higher than $FimG_{cF4}$) (Figure 15 B). These data indicate that $FimG_{cF4}$ unfolds while its donor strand is still bound by $FimH_p$. Unfolding of the inter-molecular complemented $FimH_p$ proceeds only after further temperature elevation. Together, these results show that $FimH_p$ is much more stable in the complex with $FimG_{cF4}$ than in the complex with FimC. The difference in stability is thus very likely the driving force for pilus formation.

5.1.4. Formation of complexes between subunit and donor strand peptides

The results with self-complemented subunits indicate high transition barriers for folding and unfolding. To rule out that these high barriers are an artefact of the self-complemented variants, we wanted to extend our investigations to complexes between subunits and donor strand peptides. Peptides with sequences derived from the three subunits FimA, FimF and FimG and from the chaperone FimC (termed DsFimA, DsFimF, DsFimG and DsFimC) were chemically synthesized. To obtain subunit-peptide complexes we first tried to displace the chaperone from chaperone-subunit complex by addition of peptide. FimC-$FimH_p$ complex was incubated for 48 hours at 25 °C with a 1.1-fold excess of one of the four peptides and the amount of each species (FimC-$FimH_p$ complex, FimC and $FimH_p$-peptide complex) was analysed by size exclusion chromatography. Only the FimF donor strand peptide could displaced the chaperone FimC (Figure 16). After 48 hours a large fraction of $FimH_p$ is in complex with DsFimF (peak at 14 ml), indicating that the complex with DsFimF is more stable than the FimC-$FimH_p$ complex. The reason why DsFimG, the donor strand peptide with the sequence from the most-likely natural donor for FimH, did not bind could not be determined.

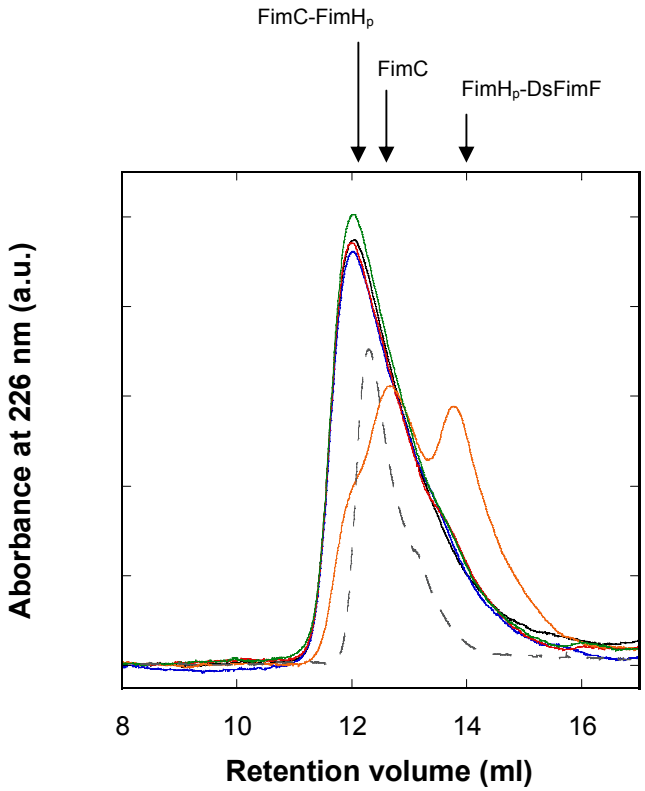


Figure 16: Size exclusion runs of FimC-FimH$_p$ after 48 hours of incubation at 25 °C and pH 7.0 in presence of different peptides. The profiles represent FimC-FimH$_p$ (black), FimC-FimH$_p$ with DsFimA (orange), FimC-FimH$_p$ with DsFimC (red), FimC-FimH$_p$ with DsFimG (green) and FimC-FimH$_p$ with DsFimF (blue). The profile of FimC is also shown (grey dashed). The FimC-FimH$_p$ concentration was 10 μM, the peptides were used at 1.1 fold excess.

For the production of FimH$_p$-DsFimF complex at a preparative scale another strategy was chosen as the displacement reaction is rather slow. FimC-FimH$_p$ was dissociated by addition of GdmCl to a final concentration of 2.5 M. FimC (bearing a C-terminal His$_6$-tag) was removed via a Ni-affinity column and FimH$_p$ was then refolded in presence of an excess of DsFimF. This complex was purified by size exclusion chromatography. The same procedure was also used to obtain the FimG$_t$-DsFimF complex, with the difference that anion exchange chromatography was used as the final purification step of the FimG$_t$-peptide complex.

5.1.5. Stability determination of subunit-peptide complexes

To determine the thermodynamic stability of the $FimH_p$-DsFimF and the $FimG_t$-DsFimF complex, GdmCl-dependent unfolding and refolding transitions were recorded. For the unfolding transitions the previously described complexes were used. For the refolding transitions the unfolded subunits were mixed with a 1.1 fold excess of DsFimF in the refolding buffer. For both complexes a very large hysteresis

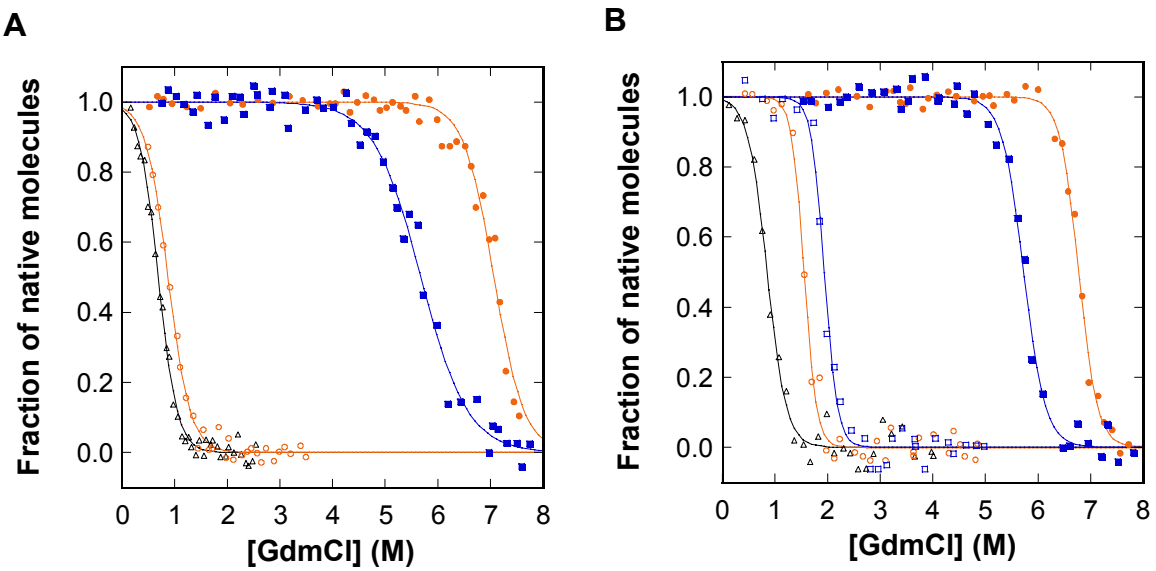


Figure 17: GdmCl dependent unfolding and refolding equilibrium transitions of subunit-peptide complexes at 25 °C and pH 7.0. **(A)** Unfolding (●) and refolding (○) transitions of $FimH_p$-DsFimF complex after 20 hours of incubation. (■) unfolding transition of $FimH_p$-DsFimF after 230 days of incubation. (Δ) refolding transition of $FimH_p$. **(B)** Unfolding (filled symbols) and refolding (open symbols) transitions of $FimG_t$-DsFimF complex after 20 hours (circles) and 50 days (squares) of incubation. (Δ) refolding transition of $FimG_t$. Data were fitted according the two state model although not in equilibrium, and normalised.

was observed after 20 hours of incubation at 25 °C (Figure 17). The distance between the folding and unfolding branches was 6.3 M for $FimH_p$-DsFimF and 5.9 M for $FimG_t$-DsFimF and thus much larger than for the variants where the donor strands are covalently linked to the subunits (Figures 11 A and B and Figure 13). For both transitions the measurements were repeated after long periods of incubation (≥ 50 days), but in neither case the transitions reached equilibrium. For $FimH_p$-DsFimF the refolding transition after 230 days could not be determined because unfolded $FimH_p$ precipitates during prolonged incubation in GdmCl-containing

buffers. Together, these results support our hypothesis of very high activation barriers for unfolding and refolding transitions of subunits incorporated into the pilus. The high kinetic stability against unfolding of subunits in the pilus might be of important biological relevance.

5.1.6 Influence of the linker length on stability of self-complemented subunits

In the self-complemented constructs the linker might cause conformational strain and their kinetic stability might therefore be lower than in intermolecular complemented subunits. To test this hypothesis we constructed self-complemented variants of

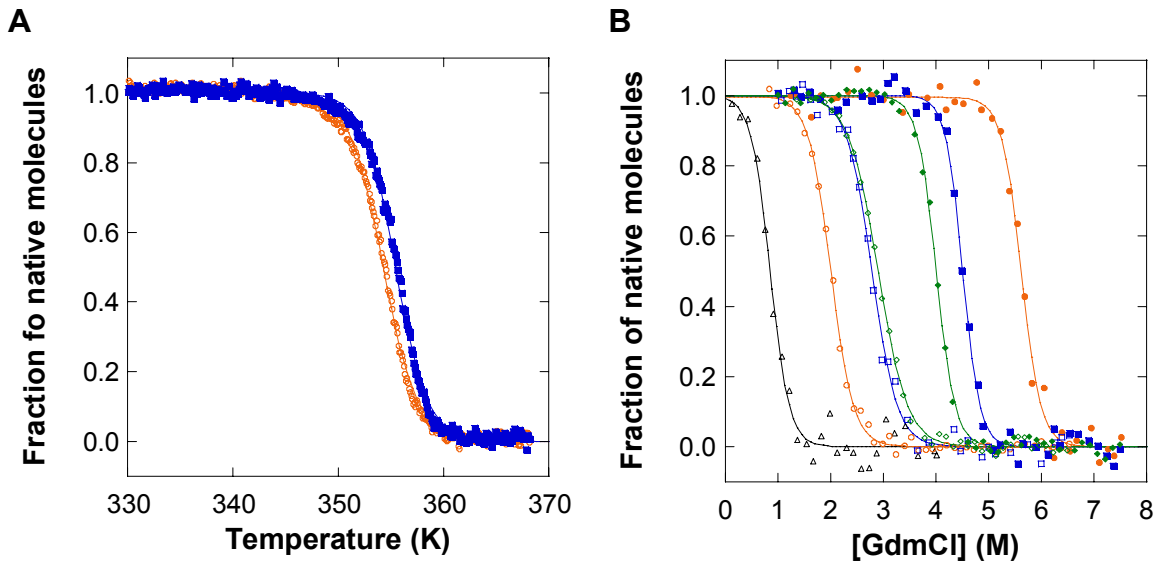


Figure 18: Transitions of self complemented $FimG_t$ with long linkers. **(A)** Thermal unfolding transitions of $FimG_{tcF7}$ (○) and $FimG_{tcF9}$ (■) at pH 7.0. **(B)** GdmCl induced unfolding (filled symbols) and refolding (open symbols) transitions of $FimG_{tcF9}$ after 20 hours (circles), 50 days (squares) and 240 days (diamonds) of incubation at 25 °C and pH 7.0.

$FimG_t$ with longer linkers of 7 and 9 glycines (termed $FimG_{tcF7}$ and $FimG_{tcF9}$, respectively). Thermal unfolding transitions of the two variants reveal that $FimG_{tcF9}$ is slightly more stable than $FimG_{tcF7}$ with a 2 °C higher melting temperature (Figure 18 A). GdmCl-induced unfolding and refolding equilibrium transitions of the more stable variant $FimG_{tcF9}$ reveal a larger hysteresis than for $FimG_{tcF4}$ after 20 hours of incubation at 25 °C (3.61 M) (Figure 18 B), but compared to the $FimG_t$-DsFimF complex the hysteresis is still smaller. The measurements after 50 and 240 days

reveal a slow movement of the transitions, with the unfolding transition moving faster than the refolding transition.

5.1.7. Determination of the change in free energy of folding for $FimG_{tcF9}$

To get an idea of the importance of the thermodynamic stability of subunits for the stability of pili, we wanted to determine the free energy of folding (ΔG_{N-U}) of a donor strand complemented subunit. The GdmCl-induced unfolding and refolding transitions of $FimH_{pcG}$, $FimH_{pcF}$ and $FimG_{tcF9}$ at 25 °C probably need years to reach equilibrium. There are other methods to determine ΔG_{N-U}, i.e. differential scanning calorimetry, thermal unfolding transitions with variation of the melting temperature by GdmCl, Urea or pH changes or the combination of GdmCl-induced equilibrium transitions at elevated temperature with thermal unfolding transitions. These experiments are reasonable under the prerequisite that the reactions are reversible and that the reaction is in equilibrium. We chose the self-complemented $FimG_{tcF9}$ for determination of the free energy of folding. For differential scanning calorimetry

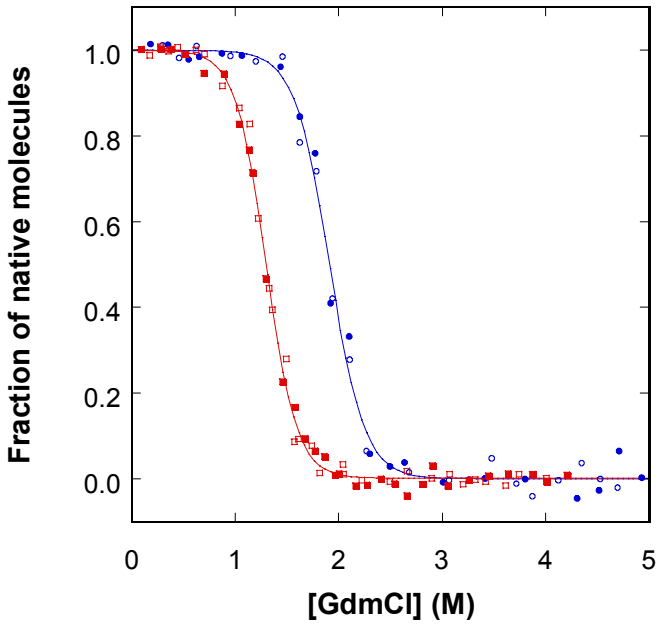


Figure 19: GdmCl induced equilibrium unfolding and refolding transitions of $FimG_{tcF9}$ at pH 7.0. Unfolding (filled symbols) and refolding (open symbols) are in equilibrium after 40 hours of incubation at 60 °C (circles) or 65 °C (squares). The GdmCl induced equilibrium transitions were followed by the emitted fluorescence at 330 nm. Data were evaluated according to the two-state model and normalised.

measurements the thermal unfolding of $FimG_{tcF9}$ is too slow and hence not in equilibrium unless very slow heating rates were used, which would result in prolonged exposure of the protein to high temperatures. Incubation of proteins at high temperatures for longer periods causes chemical modifications, which lead to irreversible thermal unfolding transitions of $FimG_{tcF9}$ (data not shown). Thermal unfolding of $FimG_{tcF9}$ in presence of Urea or GdmCl or at low pH was found to be irreversible.

ΔG_{N-U} of $FimG_{tcF9}$ was therefore determined by combining GdmCl-induced transitions at elevated temperature and thermal unfolding transitions.

GdmCl-induced unfolding and refolding transitions of $FimG_{tcF9}$ at temperatures above 55 °C reached equilibrium in two days and were reversible. Folding transitions at 60 and 65 °C were measured after incubation for 40 hours (Figure 19). Under these conditions the protein can be incubated for a certain period without the risk of modifications. Evaluation of the data according to the two-state model yielded a $\Delta G_{N-U,\ 60\ °C}$ of - 30.5 ± 1.2 kJ mol^{-1} and $\Delta G_{N-U,\ 65\ °C}$ of - 24.8 ± 0.8 kJ mol^{-1}.

The thermal unfolding transition of $FimG_{tcF9}$ was measured in absence of any denaturant with a slow heating rate (15 °C h^{-1}) in the range of the transition to allow

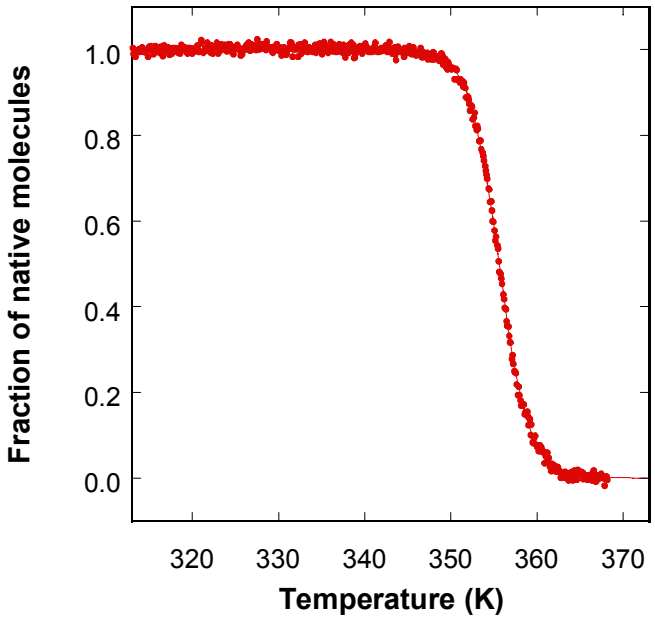


Figure 20: Thermal unfolding transitions of $FimG_{tcF9}$ at pH 7.0. Unfolding was followed by the CD signal change at 230 nm. Five individual measurements were averaged and fitted according the two state model.

equilibration during the reaction. To minimise the incubation time the pre-transition baseline was measured at a rate of 60 °C h^{-1}. The experiments were repeated 5 times. The averaged thermal unfolding transition yielded an enthalpy of folding of 611 ± 10 kJ mol^{-1} and a melting temperature (transition midpoint) of 355.6 ± 0.2 K (Figure 20). With the data from these experiments and the Gibbs-Helmholtz equation (5) the heat capacity change for unfolding (ΔC_p) was determined.

$$\Delta G_T = \Delta H_m \cdot \left(1 - \frac{T}{T_m}\right) - \Delta C_p \left[T_m - T + T \ln\left(\frac{T}{T_m}\right)\right] \qquad (5)$$

Consequently the free energy of folding at every temperature can be calculated. The calculated ΔC_p for unfolding is 11.2 ± 2.0 kJ mol^{-1} K^{-1} and $\Delta G_{N\text{-}U,\ 25\ °C}$ is - 43.6 ± 3.5 kJ mol^{-1}. The experimental ΔC_p value is in good accordance with the predicted value of 10.1 kJ mol^{-1} K^{-1} (Myers et al., 1995). According to the calculated dependence of $\Delta G_{N\text{-}U}$ on temperature (illustrated in Figure 21) the thermodynamic

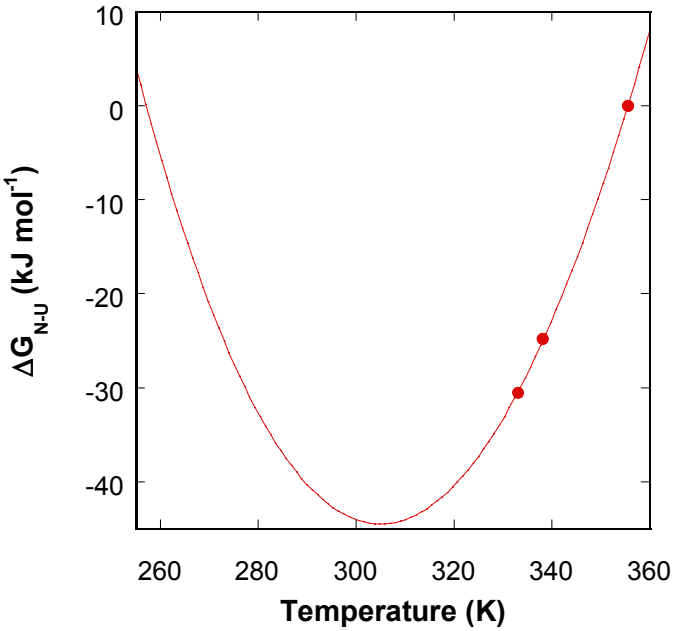


Figure 21: Calculated dependence of $\Delta G_{N\text{-}U}$ of $FimG_{tcF9}$ on temperature. The experimental values are also depicted (●)

stability of $FimG_{tcF9}$ is maximal at 32.0 ± 1.2 °C (- 44.6 ± 3.6 kJ mol^{-1}). The determined Gibbs free energy of folding for $FimG_{tcF9}$ thus lies in the range of – 20 to – 60 kJ mol^{-1} usually found for globular proteins (Fersht, 2000).

5.1.8. Determination of the half-times of unfolding of complemented FimG$_t$ and of FimA in the pilus structure

The thermodynamic stability of pilus subunits is not the only factor responsible for their stability. The hystereses found in the equilibrium unfolding and refolding transitions indicate high activation energies for folding and unfolding. To reveal the contributions of kinetic stability to the high stability of type 1 pili, it was therefore important to determine the kinetic stability of the different subunits with GdmCl-dependent unfolding kinetics, either of self-complemented subunits or subunits complemented by a donor strand peptide. The equilibrium transitions of the subunit-peptide complexes have revealed that unfolding at pH 7.0 is very slow even at high GdmCl concentrations (see e.g. Figure 11 A). Therefore the range for kinetic unfolding measurements at pH 7.0 would be very limited. In order to accelerate the unfolding reactions we decided to destabilise the subunits by performing unfolding kinetic experiments at pH 2.0. The unfolding traces were described with single exponential functions. The natural logarithm of the thus obtained unfolding rate constants was plotted against the GdmCl concentration (Figure 22). The rate

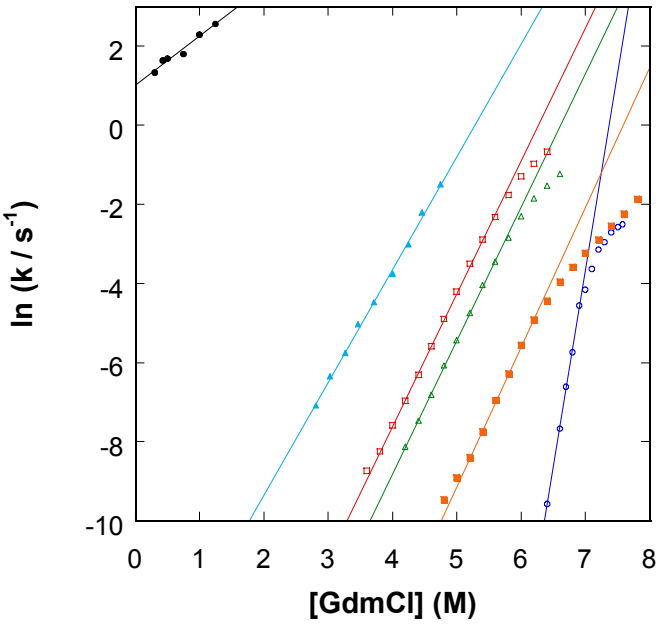


Figure 22: Evaluation of the GdmCl dependence of unfolding rates of donor strand complemented FimG$_t$ at 25 °C and pH 2.0. (●) unfolding kinetics of FimG$_t$-FimC, (▲) FimG$_{tcF4}$, (□) FimG$_{tcF7}$, (Δ) FimG$_{tcF9}$, (■) FimG$_t$-DsFimF and (○) pilus rod. The data were extrapolated to 0 M GdmCl (straight lines).

constants at 0 M denaturant concentration were obtained by linear extrapolation of the measured unfolding rate constants and the half-time of unfolding at 0 M GdmCl, pH 2.0 for every variant of donor strand complemented $FimG_t$ was determined (Table 2). The non-linear behaviour observed at high denaturant concentrations often occurs because of structural changes in the transition state (Otzen et al., 1999). The $FimG_t$-FimC complex has the shortest half-time of unfolding with 0.25 ± 0.03 s. The unfolding half-time for $FimG_t$-DsFimF on the other end of the scale is in the order of 10^{12} times slower (8900 years). The variants of self-complemented $FimG_t$ differing in linker-length showed marked differences in the half-times of unfolding, from days for the variant with the 4 amino acid linker to more than 100 years for the variant with the 9 amino acid linker. Nevertheless, $FimG_{tcF9}$ has still an about 80 times shorter half-time of unfolding than the $FimG_t$-DsFimF complex. The free activation energies of unfolding ($\Delta G^0_{‡\text{-}N}$) were calculated with equation (3) by applying a pre-exponential factor of 5 10^7 s^{-1} (Bieri et al., 1999). The values for $FimG_{tcF9}$ and the $FimG_t$-DsFimF complex are both above 100 kJ mol^{-1} (Table 2).

We wondered whether not only complemented FimG but also the main subunit FimA, incorporated into the pilus, would exhibit extraordinary kinetic stability and measured unfolding kinetics at pH 2.0 of purified assembled pili. These unfolding kinetics revealed that unfolding of FimA at 0 M GdmCl occurs with a half time of 1.5 10^{24} years. The reliability of the determined value is low because of the long extrapolation to 0 M GdmCl, but clearly unfolding is extremely slow.

In order to investigate the provenience of this very high kinetic stability, we designed six peptides where potential important amino acid side chains were exchanged (see Materials and Methods). For production of the $FimG_t$-peptide variant complexes the same protocol was applied as described for $FimG_t$-DsFimF. All of the six peptide variants were able to efficiently bind to $FimG_t$, with exception of $DsFimF_{I7R}$ where two major peaks are seen in the elution profile during purification of the subunit-peptide complex (Figure 23). We found that the $FimG_t$-$DsFimF_{I7R}$ complex obtained after

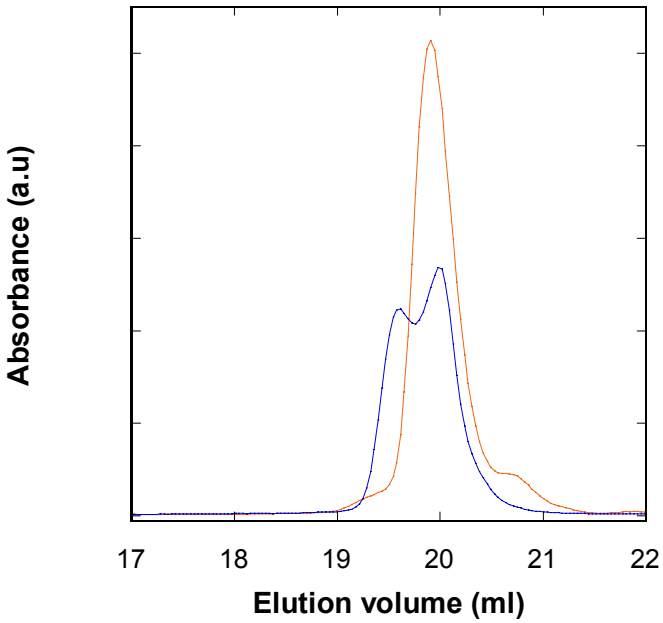


Figure 23: Purification of $FimG_t$-DsFimF complex. Complexes with the wild type DsFimF (orange) and with the I7R variant of DsFimF (blue) are shown.

purification was not stable enough for storage and thus this complex was not further investigated. GdmCl-dependent unfolding kinetics at pH 2.0 of all other complexes were measured (Figure 24). All of the peptide variants form kinetically less stable complexes with $FimG_t$ compared to the donor strand peptide with the wild type sequence. In one variant, termed $DsFimF_{triple}$, three amino acids (Ile5, Ile7 and Val11) were exchanged. The complex between $FimG_t$ and $DsFimF_{triple}$, the least stable one, had a half-time of unfolding of only about one minute (Table 2). This complex disintegrated after storage for two days. All other complexes have half-times of unfolding between two to thirty years, have thus a distinct lower activation barrier for unfolding than $FimG_t$ in complex with the peptide with wild type sequence. All the peptide variants probably affect the integrity of the hydrophobic core. We

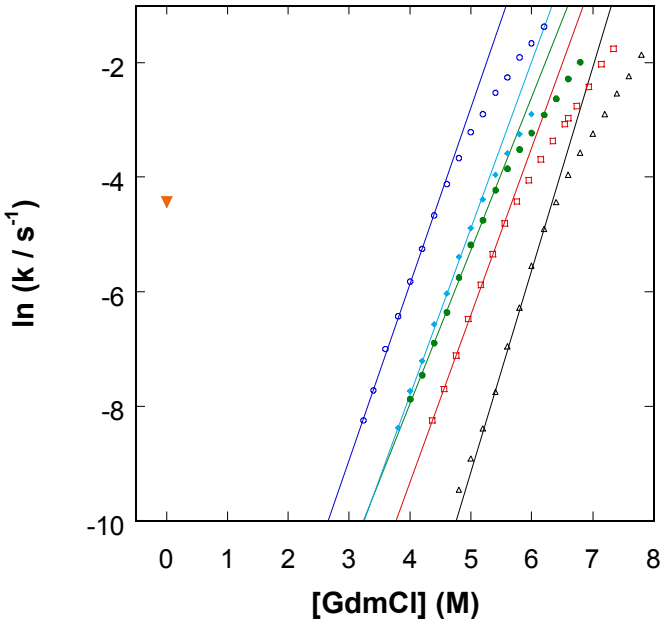


Figure 24: Evaluation of the GdmCl dependence of unfolding rates $FimG_t$ complemented with different donor strand peptide variants at 25 °C and pH 2.0: (△) $FimG_t$-DsFimF, (□) $FimG_t$-$DsFimF_{I7A}$, (●) $FimG_t$-$DsFimF_{V11A}$, (♦) $FimG_t$-$DsFimF_{G9A}$, (○) $FimG_t$-$DsFimF_{I5A}$ and (▼) $FimG_t$-$DsFimF_{triple}$. The data were linearly extrapolated to 0 M GdmCl (solid lines). Note that only one measurement of the $FimG_t$-$DsFimF_{triple}$ complex (at 0 M GdmCl) was made so far.

therefore postulate that a compact hydrophobic core is crucial for the high kinetic stability of the donor strand complemented subunits.

Construct	**k (s^{-1})**	**$\Delta G_{‡-N}$ (kJ mol^{-1})**	**$t_{½}$**
$FimG_t$-$FimC_{YY}$	2.76 ± 0.05	41.4 ± 2.0	0.25 ± 0.03 s
$FimG_{tcF4}$	$(2.92 \pm 0.13)\cdot10^{-7}$	81.2 ± 1.9	27.5 ± 1.4 d
$FimG_{tcF7}$	$(7.21 \pm 0.33)\cdot10^{-10}$	96.1 ± 2.0	30.5 ± 1.6 y
$FimG_{tcF9}$	$(2.09 \pm 0.12)\cdot10^{-10}$	99.2 ± 1.8	105.2 ± 3.9 y
$FimG_t$-DsFimF	$(2.41 \pm 0.14)\cdot10^{-12}$	110 ± 2.2	8894 ± 320 y
$FimG_t$-$DsFimF_{I5A}$	$(1.30 \pm 0.10)\cdot10^{-8}$	89.0 ± 1.9	1.7 ± 0.1 y
$FimG_t$-$DsFimF_{I7A}$	$(7.51 \pm 0.29)\cdot10^{-10}$	96.0 ± 2.1	29.2 ± 1.6 y
$FimG_t$-$DsFimF_{V11A}$	$(8.31 \pm 0.34)\cdot10^{-9}$	90.1 ± 2.0	2.6 ± 0.1 y
$FimG_t$-$DsFimF_{G9A}$	$(3.72 \pm 0.22)\cdot10^{-9}$	92.1 ± 1.8	5.9 ± 0.2 y
$FimG_t$-$DsFimF_{triple}$	$(1.18 \pm 0.10)\cdot10^{-2}$	55.0 ± 1.6	58.8 ± 2.7 s
Pilus rod	$(1.46 \pm 0.29)\cdot10^{-32}$	226 ± 3	$(1.5 \pm 0.1)\cdot10^{24}$ y

Table 2: Results from unfolding kinetic measurements at pH 2.0 extrapolated to 0 M GdmCl.

5.1.9. Unfolding is coupled to dissociation of pilus subunits

The high stability of the donor strand complemented subunits does not necessarily prevent pilus subunits within the pilus from dissociating, which would lead to breaking of the whole pilus. If this happened, the remaining pilus stump would be useless for the bacterial cell because the adhesin located at the tip of each newly formed pilus would be missing. Therefore, we wanted to test if spontaneous dissociation of pilus subunits in assembled pili occurs without unfolding of the subunits. To induce disassembly of type 1 pili, they were incubated in buffer containing different GdmCl concentrations at 60 °C for 20 hours. Then, the amount of monomeric FimA was analysed by size exclusion chromatography (Figure 25 A). Below 6.5 M GdmCl almost no monomeric FimA can be detected. At higher GdmCl concentrations a peak at a retention volume of 8 ml appears, characteristic for unfolded monomeric FimA. The area of the peak at 8 ml obtained under the different conditions was determined and compared to the unfolding transition of pili after

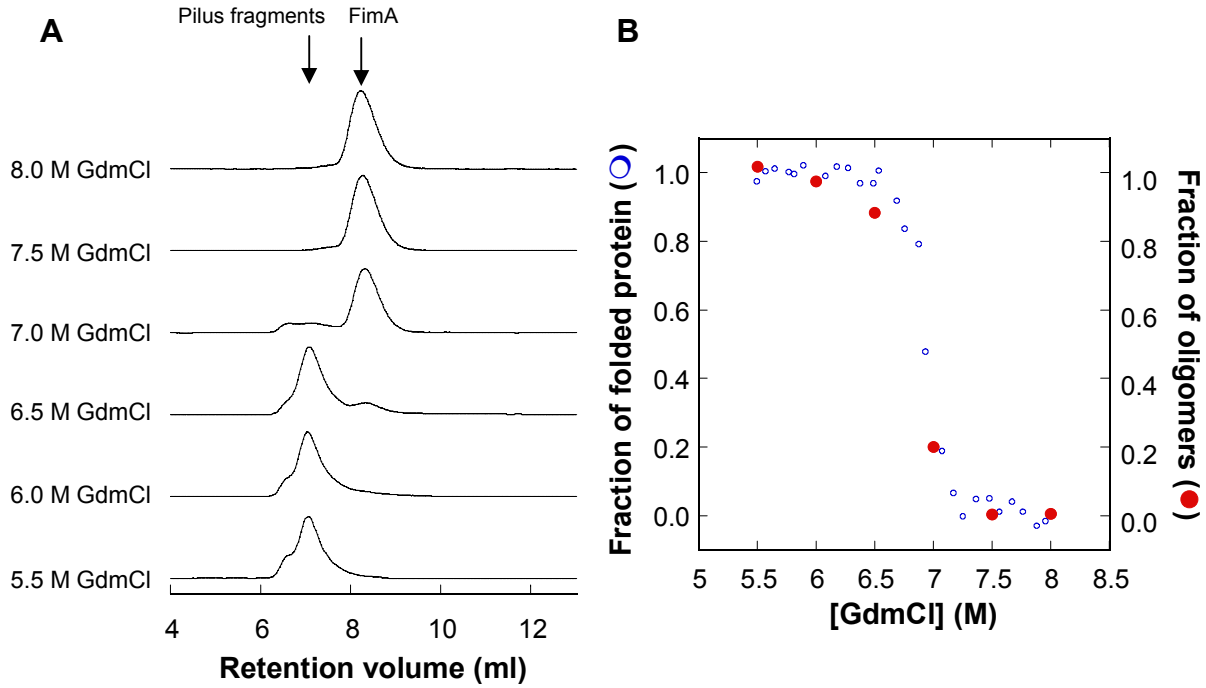


Figure 25: Dissociation and unfolding of oligomeric FimA at 60 °C, pH 7.0. **(A)** Size exclusion retention profiles after 20 h of incubation at 60 °C at the corresponding GdmCl concentrations. **(B)** Amount of monomeric FimA quantified by the peak areas shown in the left panel (●). The fraction of native protein (○) was determined by CD signal change at 230 nm after incubation of pili for 20 h at 60 °C.

20 hours of incubation at 60 °C (Figure 25 B). The unfolding transition was measured by following the far-UV CD signal change at 230 nm. The results from size exclusion and from the spectroscopical analysis superimpose almost perfectly, showing that in case of entire pili, unfolding is coupled to dissociation of the single subunits.

To see if the $FimG_t$-DsFimF complex, our highly defined model system for the interaction between two subunits, dissociates only upon unfolding as well, the same experiment was repeated with this complex. $FimG_t$-DsFimF was incubated for 20 hours at 25 °C and different GdmCl concentrations and then applied to a size exclusion column for analysis (Figure 26 A). At concentrations below 5.5 M GdmCl almost all $FimG_t$ has bound DsFimF and elutes at about 11 ml. For samples

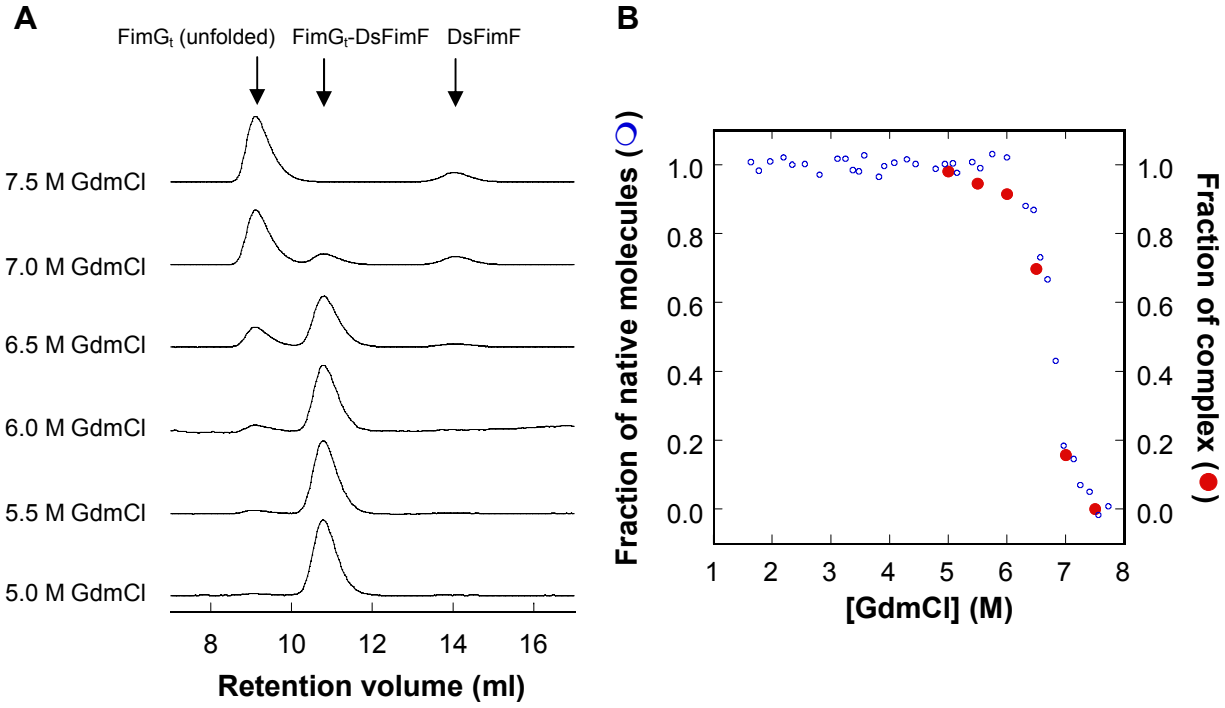


Figure 26: GdmCl-induced dissociation and unfolding of $FimG_t$-DsFimF complex at 25 °C and pH 7.0. **(A)** size exclusion runs of complex incubated at the indicated GdmCl concentrations. **(B)** Comparison of the size exclusion data (●) with the unfolding transition determined with far-UV CD signal change (○).

incubated at higher GdmCl concentrations an additional peak appears at a smaller retention volume. This apparently higher molecular weight is typical for unfolded proteins which are less compact than folded proteins. In addition a third peak at about 15 ml appears, representing the free peptide. The peak areas were again analysed to determine the relative amount of complex (Fig. 26 B). The unfolding transition of $FimG_t$-DsFimF complex followed by far UV CD signal change

superimposes with the size exclusion data, as seen before for the dissociation of entire pili.

5.1.10. Structure determination of the $FimG_t$-DsFimF complex

Kinetic stability is very important for the high stability of type 1 pilus subunits. In order to get insight into the structural basis of the high unfolding energy barriers we solved the crystal structure of the $FimG_t$-DsFimF complex (Table 3). The structure reveals the canonical immunoglobulin-like fold with two β-sheets, with the donor strand inserted in an anti-parallel orientation to the F-strand of $FimG_t$ (Figure 27 A and B). 12 of the 15 donor strand peptide residues interact in a specific way with $FimG_t$. Residues 3 – 12 are involved in β-sheet interactions sharing the maximal possible number of main chain H-bonds (Figure 27 C). Out of the three residues which do not specifically interact with $FimG_t$, Ala1 and Ala14 are solvent-exposed, whereas Arg15 forms an H-bond with Met36 of $FimG_t$. The latter is likely due to crystal packing.

Packing of the hydrophobic core has probably an effect on the stability of pilus subunits. The shape complementarity index (Sc) (Lawrence and Colman, 1993) is a measure for packing quality. For the complementarity of $FimG_t$-DsFimF and packing of the two β-sheets very high values are determined with 0.75 and 0.78, respectively. Five conserved residues of DsFimF (P1 - P5, Figure 27 D and 28) were detected to insert in the hydrophobic core of $FimG_t$. All these residues are hydrophobic except for Ser3 (P1), which forms two side-chain H-bonds with the $FimG_t$-residues Thr20 and Tyr141.

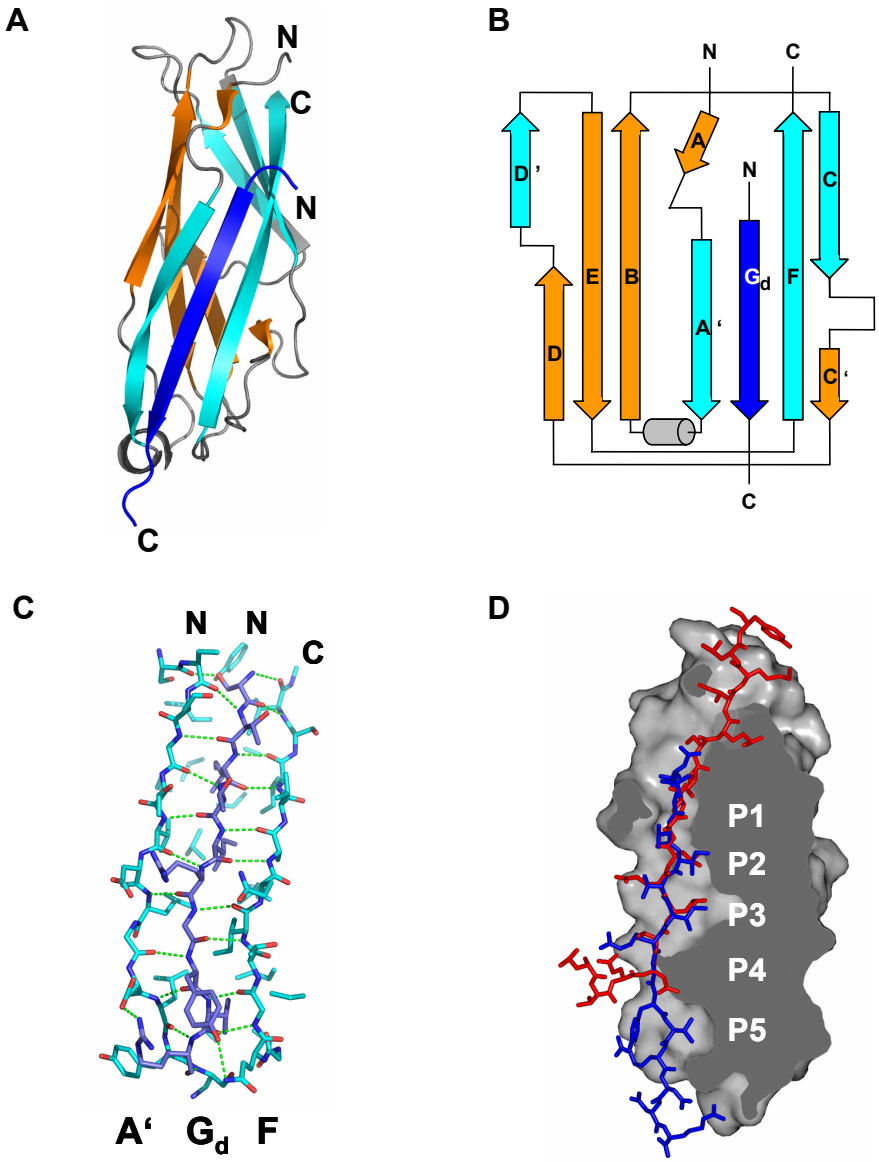


Figure 27: X-ray structure of the $FimG_t$-DsFimF complex. **(A)** Ribbon diagram of $FimG_t$-DsFimF showing the two β-sheets (orange and cyan) constituting the Ig-like subunit fold. The FimF donor strand peptide in blue is part of the cyan β-sheet. **(B)** Topology diagram of $FimG_t$-DsFimF. The FimF donor strand in blue (G_d) runs anti-parallel to the F strand of $FimG_t$ and complements the incomplete Ig-like fold in a canonical manner. **(C)** Hydrogen bond network between A' and F strands of $FimG_t$ (cyan) and DsFimF (blue). Peptide residues 3-12 form 20 main chain H-bonds, which is the highest possible number for ten residues in a β-sheet. **(D)** Slice-view through $FimG_t$, showing positions P1-P5 where residues from complementing strands interact with the hydrophobic groove of the subunit. The conserved glycine 9 by-passes Ile133 at P4 allowing Valine 11 to occupy P5. The G1 strand (red) of the chaperone FimC (Choudhury et al., 1999) is superimposed onto the FimF donor strand (blue). Figures prepared with PyMOL (www.pymol.org).

The kinetic stability of subunit-bound subunit is much higher than of chaperone-bound subunit (see Figure 14) (Zavialov et al., 2005). Compared to the FimC-FimH$_p$ complex where only the positions P1 to P4 are occupied, the conserved Gly 9 allows the FimF donor strand peptide to extend over the bulky Ile133 to position P5. Val11 at P5 is deeply buried in the hydrophobic groove.

The FimG$_t$-DsFimF complex has a total amount of 135 H-bonds. This high extent is

Figure 28: Alignment of the different donor strands of FimA, FimF, FimG and FimC. The positions involved in interaction with the hydrophobic core are indicated (P1 – P5).

probably only present after the subunit donor strand is bound and not in the chaperone bound form. Evaluation of the structures of chaperone-subunit and subunit-subunit complexes of the P-pilus system reveals an increase from 72 H-bonds in the PapD-PapE$_{Ntd}$ complex to 106 in the PapE$_{Ntd}$-PapK$_{Nte}$ complex (Sauer et al., 2002). In addition the secondary structure content increases by 10 %.

Table 3. Data Collection and Refinement

Data Collection

Data Set	Radiation (Å)	Resolution (Å)	Total/Unique Reflections	I/σ (I)	Completeness (%)	R_{sym}
Native	0.90002	30 – 1.34	62192/25442	13.4 (3.2)	96.3 (84.5)	0.064 (0.234)
Peak	1.60556	20 – 1.85	57204/17996	14.8 (4.8)	92.0 (71.7)	0.053 (0.135)
Inflection	1.60624	20 – 1.85	58354/17575	16.0 (3.7)	89.1 (54.6)	0.054 (0.165)
Remote	1.50238	20 – 1.85	56187/18341	12.9 (6.3)	93.1 (80.2)	0.059 (0.159)

Refinement

Data Set	Native
Resolution (Å)	22.0 – 1.34
Unique reflections	25434
Test Set	985
Working Set	24449
Atoms	1106
Heterogen Atoms	1
Water Atoms	190
R_{free}	18.5
R	14.1
RMSD Bonds (Å)	0.010
RMSD Angles (deg.)	0.025
Average B-factor ($Å^2$)	19.1
Ramachandran plot	92.2 % core 7.8 % allowed

Values for high-resolution shells in parentheses.

Values of R_{merge} are taken as R_{sym}.

5.2. FimC is a folding catalyst

5.2.1. FimC binds unfolded substrate

The type 1 pilus chaperone FimC is essential for pilus assembly (Jones et al., 1993; Kuehn et al., 1991; Lindberg et al., 1989). It forms stoichiometric complexes with the pilus subunits, prevents aggregation and mediates binding of the subunits to the outer membrane usher FimD (Nishiyama et al., 2005; Nishiyama et al., 2003). The subunits can be refolded *in vitro* from GdmCl-containing buffers (Vetsch et al., 2004; Vetsch et al., 2002), but the folding reaction is slow and the subunits are prone to aggregation. Aggregation is efficiently suppressed when the chaperone is present during refolding. When the N-terminally truncated subunit $FimG_t$ is refolded in presence of not only the chaperone FimC but also the self-complemented subunit $FimF_{cF4}$ (full-length FimF with the FimF donor strand attached to the C-terminus via a DNKQ peptide linker), $FimG_t$ first binds to FimC and is later quantitatively transferred to $FimF_{cF4}$. Although the subunit-subunit complex is more stable (~ 17 kJ mol^{-1}) the chaperone-subunit complex is formed first, indicative of an important role of the chaperone in the folding reaction (Vetsch et al., 2004).

When folding of $FimG_t$ is followed by fluorescence two phases are observed, with rate constants ($1/\tau$) of 0.14 ± 0.01 s^{-1} and 0.031 ± 0.001 s^{-1}. For investigation of the effect of FimC on the folding kinetics of $FimG_t$, a variant of FimC ($FimC_{YY}$) was used where the two tryptophans at positions 36 and 84 were substituted by tyrosines. Refolding of $FimG_t$ in presence of $FimC_{YY}$ can be described with the sum of four exponentials ((Vetsch et al., 2004), Figure S3 B). Two of the four folding phases are significantly faster than the two observed in folding of $FimG_t$ alone. The fast fluorescence increase for the refolding of $FimG_t$ observed only with $FimC_{YY}$ indicates that $FimC_{YY}$ interacts with unfolded $FimG_t$ or with an early folding intermediate. In contrast to the chaperone, donor strand peptides (DsFimC, DsFimF and DsFimG) had only a subtle influence on the kinetic trace of $FimG_t$ folding.

5.2.2. FimC accelerates the formation of native $FimG_t$

FimC binds to unfolded $FimG_t$ or to an early intermediate and can thus influence the folding pathway of $FimG_t$. It is possible that FimC accelerates the formation of an intermediate and that the subsequent transition to native $FimG_t$ is slow and

spectroscopical silent. Alternatively, FimC may accelerate the formation of native $FimG_t$. To distinguish between these possibilities, interrupted refolding experiments were performed, where $FimG_t$ was refolded for a defined time-period and unfolded again. This procedure allows for discrimination between native protein and folding intermediates, because the native protein has a characteristic unfolding rate at a given denaturant concentration which is slower than the unfolding rates of any intermediate (Schmid, 1983). The rate of formation of native protein can therefore be measured.

$FimG_t$ was refolded for various times in presence of 5-fold excess of $FimC_{YY}$ or 40-fold excess of DsFimC and then transferred to unfolding conditions. Figure 29 shows the evaluation of the interrupted refolding experiments of $FimG_t$ in presence of $FimC_{YY}$ and of the donor strand peptide DsFimC. The unfolding amplitudes after the different time points were determined and compared to the unfolding amplitude of completely folded $FimG_t$ in presence of the particular binding partner. Formation of native $FimG_t$ in presence of DsFimC can be fitted with an exponential function yielding a rate constant of 0.0041 ± 0.0001 s^{-1} (Figure 29), which is one order of

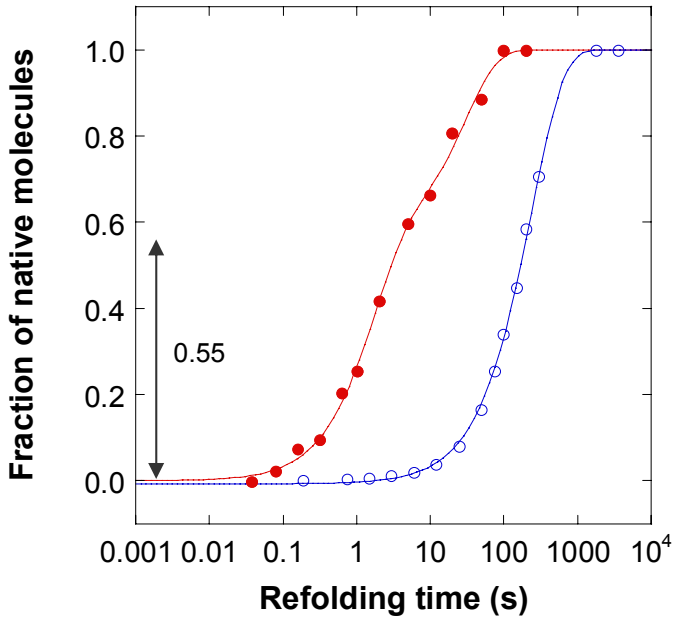


Figure 29: Interrupted folding of $FimG_t$. The effect of FimC on the folding rate of $FimG_t$ is shown in red, folding of $FimG_t$ in presence of DsFimC is shown in blue. The fraction of native molecules was determined by the unfolding amplitude of $FimG_t$. The data were fitted with a single exponential function ($FimG_t$-DsFimC) or with the sum of two exponential functions ($FimG_t$-$FimC_{YY}$). The fraction of fast folding $FimG_t$ molecules in presence of $FimC_{YY}$ is indicated.

magnitude slower than the slowest rate observed in the refolding fluorescence trace ((Vetsch et al., 2004), Figure S 3 D). This indicates that a spectroscopically silent, rate-limiting step leads to the native structure of $FimG_t$-DsFimC complex.

In presence of $FimC_{YY}$ the formation of native $FimG_t$ shows completely different kinetics. The major part of the molecules (about 55 ± 5 %) reaches the native state with a rate constant of 0.59 ± 0.08 s^{-1}, which is about 100 times faster than $FimG_t$ in presence of DsFimC (Figure 29). The second fraction (45 ± 5 %) folds with a rate constant of 0.033 ± 0.007 s^{-1}, which is 20 fold slower than the first fraction but still about 8 times faster than in presence of DsFimC. The rate of this slower folding fraction of proteins is similar to the rates for prolyl cis-trans isomerisation (Balbach and Schmid, 2000). To test whether prolyl isomerisation causes the slow phase in the formation of native $FimG_t$, cyclophilin, a periplasmic peptidyl prolyl isomerase from *E. coli* was used. Interrupted refolding experiments of 1 μM $FimG_t$ with 5 μM $FimC_{YY}$ and 1 μM cyclophilin in the refolding buffer were performed and the rate of formation of native $FimG_t$ was analysed (Figure 30). Cyclophilin has no effect on the folding of $FimG_t$ as the fractions of fast and slow folding are very similar (A_1 = 62 ± 5 % and A_2 = 38 ± 5 %). The reason for the absence of any effect of cyclophilin could

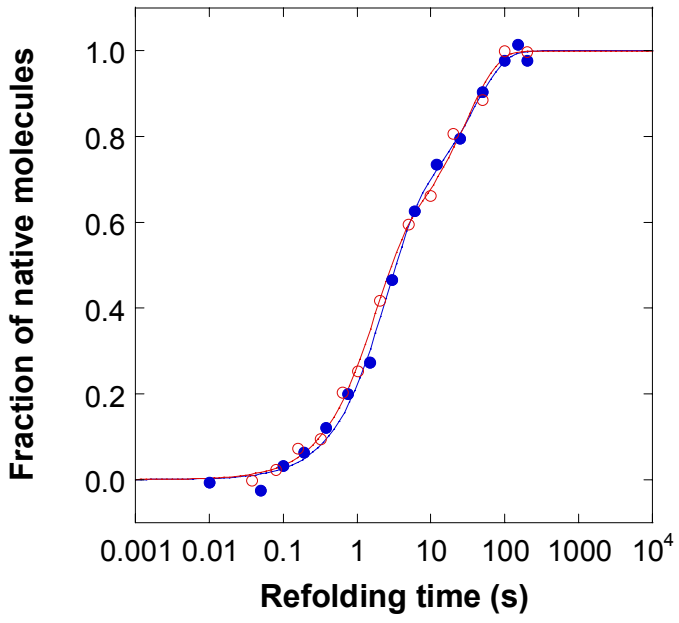


Figure 30: Interrupted refolding experiments of $FimG_t$ in presence of $FimC_{YY}$ with or without cyclophilin (○ and ●, respectively). The data were fitted with the sum of two exponential functions.

be the inaccessibility of the proline residues of $FimG_t$ in an early folding intermediate. To further investigate if the fraction of slow folding proteins is due to prolyl *cis-trans* isomerisation of non-native peptide bond conformations, a $FimG_t$ variant was designed where all three prolines (positions 15, 55 and 116 of mature FimG) were substituted by glycines. Expression tests revealed that the proline at position 55 is essential for the stability of $FimG_t$. However, the variant where only the two prolines at position 15 and 116 were substituted (termed $FimG_t^*$) could be expressed and purified. The refolding rate constants of $FimG_t^*$ in presence of 5 μM $FimC_{YY}$ or 40 μM DsFimC were again determined with interrupted refolding experiments (Figure 31). In both cases, with $FimC_{YY}$ and DsFimC, folding of $FimG_t^*$ is about three times faster than folding of $FimG_t$. The formation of native $FimG_t^*$ in presence of $FimC_{YY}$ has again to be fitted with the sum of two single exponentials. The fraction of slow folding

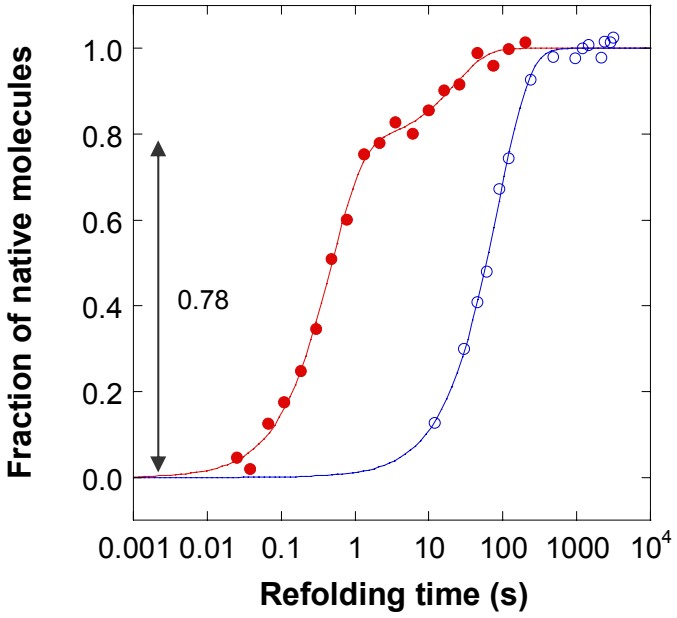
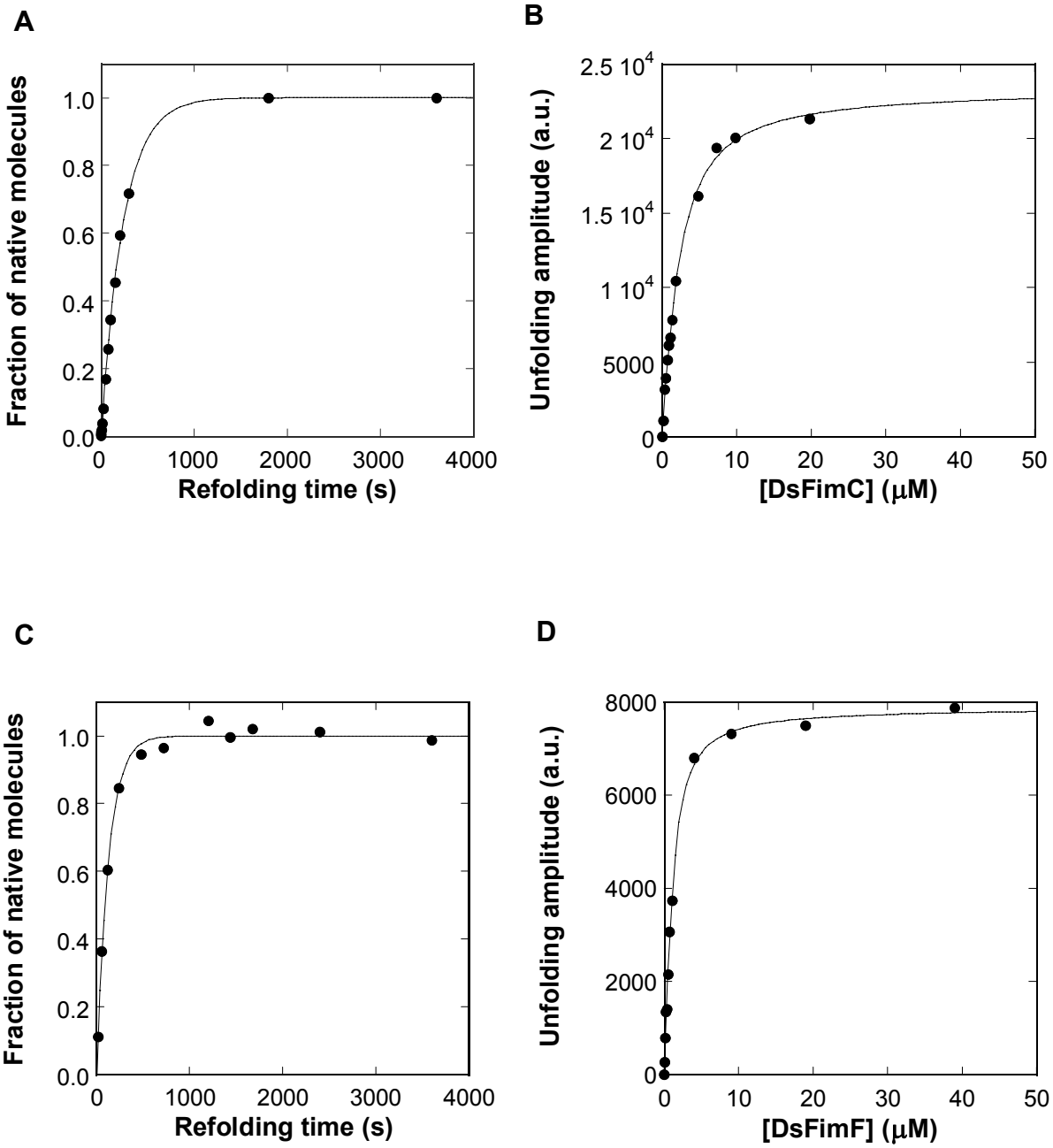


Figure 31: Interrupted refolding experiments for determination of the folding rate of $FimG_t^*$ ($FimG_t$ variant with P15G/P116G) in presence of $FimC_{YY}$ (●) or DsFimC (○). Data were fitted with a single exponential function ($FimG_t$-DsFimC) or with the sum of two exponential functions ($FimG_t$-$FimC_{YY}$). The fraction of fast folding $FimG_t^*$ molecules in presence of $FimC_{YY}$ is indicated.

molecules has however decreased from about 45 % for $FimG_t$ to 22 % for $FimG_t^*$. This shows that the slow phase in the folding of $FimG_t$ is caused by the prolyl *cis-trans* isomerisation. In agreement with this interpretation it was later found in the X-ray structure of the $FimG_t$-DsFimF complex that all three prolines of native $FimG_t$ are in trans.



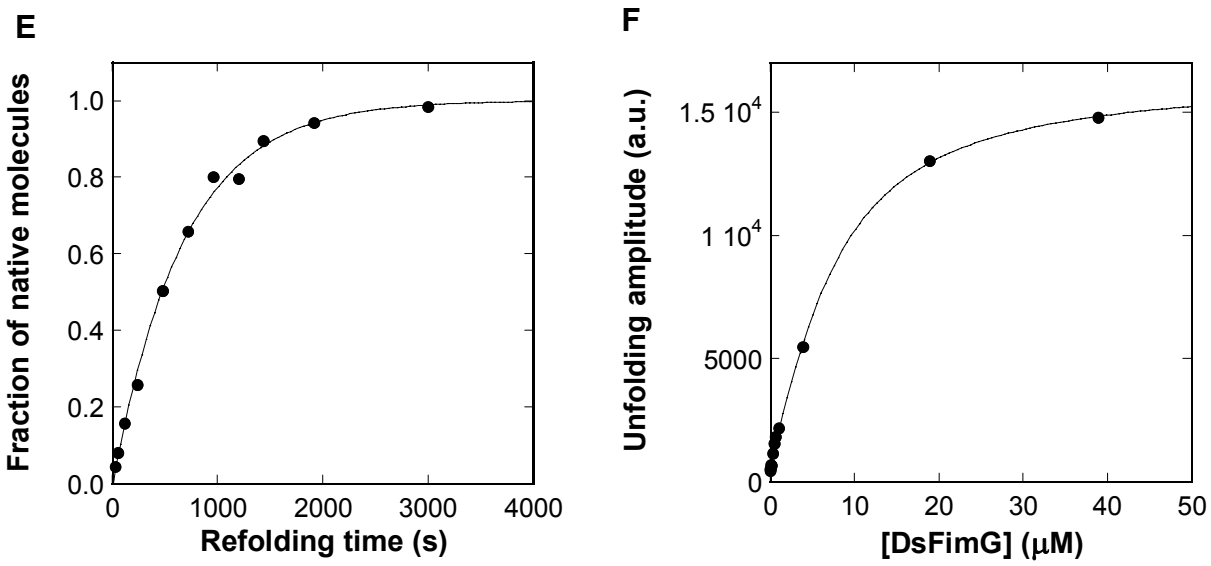


Figure 32: Interrupted folding experiments of $FimG_t$ in presence of different donor strand peptides and determination of the apparent affinities for the peptides. The interrupted refolding experiments were performed with 1 μM $FimG_t$ in presence of 40 μM DsFimC **(A)**, DsFimF **(C)** or DsFimG **(E)**. The affinity for $FimG_t$ for the different peptides was determined at 1 μM $FimG_t$ by refolding $FimG_t$ in presence of different amounts of peptide. After complete refolding the protein was transferred to unfolding conditions and the amount of complex was determined by the unfolding amplitude. Determination of the amount of bound DsFimC **(B)**, DsFimF **(D)** and DsFimG **(F)**. Data for the determination of the folding rate constant were fitted with a single exponential function, the data for determination of the affinity were fitted according the law of mass action.

We also determined the effect of other donor strand peptides derived from the subunits FimF and FimG on the folding rate constant of $FimG_t$. Interrupted refolding experiments of $FimG_t$ in presence of 40 μM DsFimC, DsFimF or DsFimG were performed. Folding in presence of DsFimF is fastest with a rate constant of 0.0078 ± 0.0006 s^{-1}, followed by folding with DsFimC ($1/\tau$ = 0.0041 ± 0.0001 s^{-1}) and the folding in presence of DsFimG as the slowest reaction ($1/\tau$ = 0.0015 ± 0.0001 s^{-1}) (Fig. 32 A, C and E). Folding of $FimG_t$ in presence of either peptide is at least 75 times slower than folding in presence of $FimC_{YY}$. The affinity of peptide binding to $FimG_t$ was tested with a method similar to interrupted refolding. Unfolded $FimG_t$ was folded in presence of different peptide concentrations. As difference in this double jump method $FimG_t$ was given the opportunity to fold completely before transferring to unfolding conditions. For the chosen conditions unfolding of native $FimG_t$-peptide

complex occurs with a characteristic rate constant and the unfolding amplitudes reflect the amount of bound peptide at a given peptide concentration. The thus obtained apparent dissociation constants (K_{Dapp}) were determined. DsFimF has the lowest K_{Dapp} with 0.64 ± 0.06 μM, followed by DsFimC with 1.9 ± 0.2 μM and DsFimG with 8.5 ± 0.8 μM (Fig. 32 B, D and F). The apparent dissociation constants correlate positively with the folding rate constants from the interrupted folding experiments.

6. Discussion

6.1. Kinetic stability is the reason for pilus stability

Type 1 pili are extraordinary stable organelles on the surface of uropathogenic *E. coli* strains. They are composed of up to 3000 subunit proteins, which interact with each other in a mechanism termed donor strand complementation. In this interaction the N-terminal extension of each subunit has an anti-parallel orientation relative to the C-terminal strand of the recipient subunit (Sauer et al., 2002; Zavialov et al., 2003). During assembly, transient complexes between the chaperone FimC and the different subunits are formed in the periplasm. Interactions in this complex involve donor strand complementation with the donor strand of FimC in a parallel orientation to the C-terminal strand of the subunit (Choudhury et al., 1999; Sauer et al., 1999; Zavialov et al., 2003). Because of a hydrophobic cleft that is solvent-exposed in absence of the donor strand, the subunits are only marginally stable (~ - 10 kJ mol^{-1}) and tend to aggregate (Vetsch et al., 2002). We found that upon binding to the chaperone FimC, the subunit $FimH_p$ is stabilised.

The orientation of the donor strand in the chaperone-subunit complex is opposite to the orientation in subunit-subunit complexes. The self-complemented $FimH_p$ variants that were made for investigation of the stability of type 1 pili account for both orientations. Attachment of the donor strand sequence to the N-terminus results in a parallel orientation to the C-terminal strand of the subunit, the same orientation as found in chaperone-subunit complexes. Conversely, attachment of the donor strand to the C-terminus gives an anti-parallel orientation in respect to the C-terminal strand of the subunit, as found in subunit-subunit complexes. The thermodynamic stability of the self-complemented variants was determined with GdmCl-induced unfolding and refolding transitions. All investigated variants have a higher thermodynamic stability than non-complemented $FimH_p$. Transitions of the variants $FimH_{pnG}$, $FimH_{pnF}$ and $FimH_{pcC}$ have reached equilibrium after 2 days at 25 °C. In contrast, the variants $FimH_{pcG}$ and $FimH_{pcF}$ with the same donor strand orientation as present in pili show very slow equilibration. Taking together, the results of the folding transitions of FimC-$FimH_p$ complex and of the self-complemented variants indicate that pilus assembly is very likely driven by the increase in thermodynamic stability of the subunits upon

incorporation. This was also postulated for assembly of the capsular F1 antigen from *Yersinia pestis* (Zavialov et al., 2005).

$FimH_{pcG}$ and $FimH_{pcF}$ show a hysteresis and thus have high activation energies for folding and unfolding. The same effect was also observed for all self-complemented $FimG_t$ variants with the donor strand of FimF covalently attached to the C-terminus, irrespective of the linker length.

To test whether the thermodynamic stability of the subunit variants showing a hysteresis is exceptional compared to the thermodynamic stability typically found for globular proteins, the Gibbs free energy of folding for $FimG_{tcF9}$ was determined by combining data from GdmCl equilibrium unfolding and refolding transitions at 60 and 65 °C and thermal unfolding transitions. The Gibbs free energy of folding lies in the range of – 40 to – 45 kJ mol^{-1} which is in coherence with the typical range of thermodynamic stability of globular proteins of – 20 to – 60 kJ mol^{-1} (Fersht, 2000). Thus the high stability of the pilus subunits is most likely not due to thermodynamic stability.

The hysteresis is not only an artefact of the intramolecular constructs. Subunits that have donor strand peptides non-covalently bound have even larger hystereses and the unfolding and refolding transitions show an even slower equilibration compared to self-complemented subunits, denoting even higher kinetic stability.

In order to determine the kinetic stability, we performed unfolding kinetics at pH 2.0 and calculated the activation free energy for unfolding at 0 M GdmCl ($\Delta G_{‡\text{-}N}$). With $\Delta G_{‡\text{-}N}$ of 100 kJ mol^{-1} and 110 kJ mol^{-1}, the self-complemented variant $FimG_{tcF9}$ and the $FimG_t$-DsFimF complex have an extraordinary high kinetic stability. Similar activation energies are only found for proteins of hyperthermophile organisms (Mukaiyama et al., 2004). Considering the fact that the activation energies were measured at pH 2.0, it can be assumed that at the pH optimum for stability, the activation energies are even higher.

The energetic differences between the self-complemented $FimG_{tcF9}$ and inter-molecular donor strand complemented $FimG_t$-DsFimF most probably derive from a steric strain caused by the linker between the subunit core and the donor strand. The donor strand might therefore not occupy the hydrophobic cleft in the optimal configuration. With self-complemented $FimG_t$ subunits bearing linkers of different length it was shown that the kinetic stability is influenced by the linker length. The

variant with the shortest linker has the lowest activation energy and the one with the longest linker has the highest activation energy for unfolding.

The extraordinary high kinetic stability of the donor strand complemented subunits mimicking the interactions in the pilus implicates that the unfolding probability of the subunits is very low. The half-times of unfolding are in the range of hundreds to thousands of years.

Yet, the most extraordinary properties were found for assembled type 1 pili. The unfolding rate constant of the high molecular assemblies is so slow that it takes astronomic timescales to unfold a pilus structure.

The sequence of the donor strand is important for kinetic stability. This can be seen from the self-complemented $FimH_p$ variants with the donor strands attached to the C-terminus. The variant with the donor strand of FimG, most likely the natural donor for FimH, has the largest hysteresis, whereas the variant with the donor strand of the chaperone FimC shows no hysteresis. More systematic investigations were performed with different variants of the FimF donor strand peptide in complex with $FimG_t$. The peptides were designed on the basis of the interactions found in the crystal structure of the $FimG_t$-DsFimF complex. The complexes of $FimG_t$ with the DsFimF peptide variants do not have the same high kinetic stability as the complex with the wild-type donor strand peptide. The free activation energies for unfolding range from 55 to 92 kJ mol^{-1} compared to 110 kJ mol^{-1} for the peptide with the wild type sequence. The unfolding half-times are between 300 and 5 10^9 times faster than for the $FimG_t$-DsFimF complex. According to the structure of the $FimG_t$-DsFimF complex, the DsFimF peptide variants with non-wild type sequence all, except $DsFimF_{G9A}$, introduce cavities in the hydrophobic core that is formed between $FimG_t$ and the donor strand peptide.

In the structure of the $FimG_t$-DsFimF complex high packing values (or Sc-values) between the subunit and the peptide (0.75) as well as between the two β-sheets (0.78) were found. The subunits reach the full extent of interactions only after the incorporation into the pilus. This is first reflected in the shape correlation between the two β-sheets compared in the subunit-peptide and chaperone-subunit complex. (Zavialov et al., 2003) report for the Capsular F1 Antigen an increase in shape correlation from 0.58 for the β-sheets of the subunit in the chaperone-subunit complex to 0.71 in the subunit-subunit complex. At the same time the number of

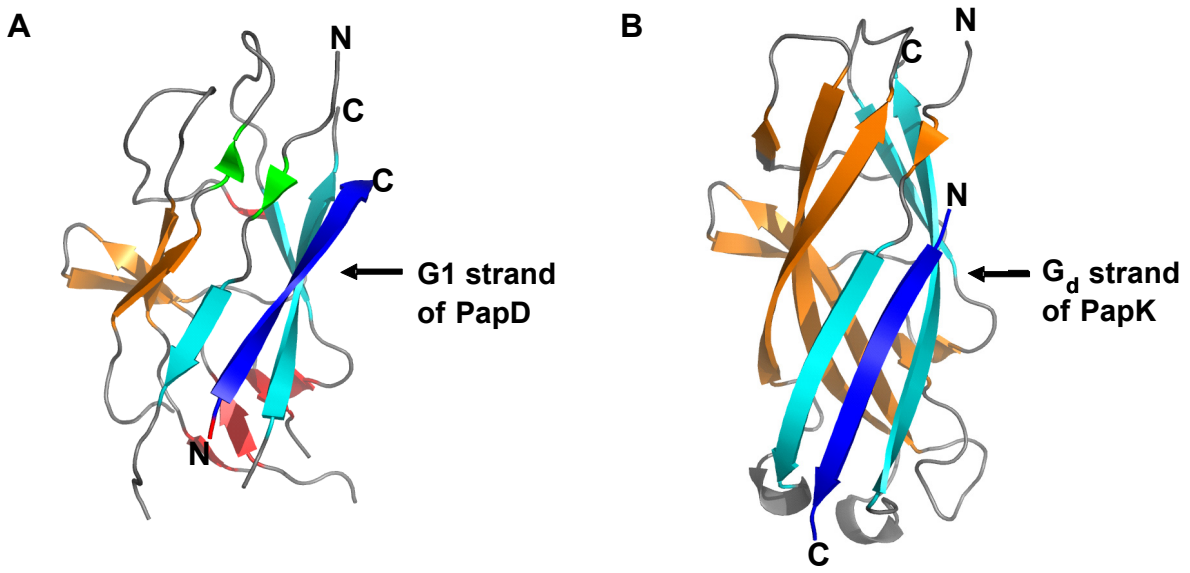


Figure 33: Secondary structure extent of $PapE_{Ntd}$ different complexes. The structures are adapted from (Sauer et al., 2002) for analysis. **(A)** $PapE_{Ntd}$ (orange and cyan) with the donor strand of the chaperone PapD (blue). **(B)** $PapE_{Ntd}$ (orange and cyan) with the donor strand of the subunit PapK (blue).

H-bonds increases drastically and the extent of secondary structure is augmented (Figure 33). All these factors lead to increased stability of the subunit, but the most important is probably the extraordinary tight packing of the hydrophobic core, as the results with the DsFimF peptide variants suggest.

For the function of type 1 pili as adhesive organelles, it is important not to dissociate. The localisation of the adhesin at the tip of the pilus renders the pilus unfunctional if the pilus breaks because of dissociation of the subunits. We have shown for the $FimG_t$-DsFimF complex and for entire type 1 pili that dissociation of subunits depends on unfolding. The donor strand is locked in the hydrophobic groove without the possibility to dissociate unless the whole acceptor subunit unfolds.

A well characterised example of a kinetically stabilised protein is the extracellular α-lytic protease from *Lysobacter enzymogenes*. Similar to type 1 pilus subunits α-lytic protease needs assistance for folding, which is in contrast to the type 1 pilus system provided by a covalently bound pro-region (Ikemura et al., 1987; Silen and Agard, 1989). After folding, the pro-peptide is auto-catalytically cleaved and the mature α-lytic protease is trapped in the folded, active state by a high activation

barrier (Cunningham et al., 1999). This prevents unfolding and thus degradation of the protease.

Comparatively, type 1 pili are stabilised by an extraordinary high activation barrier against unfolding and dissociation, which is important to retain binding competence of the adhesive organelle.

Kinetic stabilisation is probably a widespread mechanism for protection against unfolding for extra-cellular proteins or for proteins that are exposed to extreme and harmful conditions.

6.2. FimC catalyses folding of type 1 pilus subunits

Molecular chaperones have the function to increase the folding efficiency and minimise the amount of unproductive folding and aggregation. The type 1 pilus chaperone FimC prevents aggregation in the periplasm. The *in vitro* interrupted refolding experiments with $FimG_t$ in presence of chaperone show that FimC in addition accelerates the folding reaction by more than a factor of 100, compared to refolding of $FimG_t$ in presence of the DsFimC donor strand peptide. In addition, the donor strand peptides derived from the subunits FimF and FimG were found to have only a marginal effect on $FimG_t$ folding. Acceleration of folding of the subunits thus requires additional interactions of FimC other than the contacts with the donor strand.

FimC thus represents a new type of folding catalyst that recognises non-native subunits, assists and accelerates their folding and retains the subunits in an activated, assembly competent conformation.

7. References

Anfinsen, C.B. (1973) Principles that govern the folding of protein chains. *Science*, **181**, 223-230.

Ansari, A., Jones, C.M., Henry, E.R., Hofrichter, J. and Eaton, W.A. (1992) The role of solvent viscosity in the dynamics of protein conformational changes. *Science*, **256**, 1796-1798.

Arkowitz, R.A., Joly, J.C. and Wickner, W. (1993) Translocation can drive the unfolding of a preprotein domain. *Embo J*, **12**, 243-253.

Bader, M., Muse, W., Ballou, D.P., Gassner, C. and Bardwell, J.C. (1999) Oxidative protein folding is driven by the electron transport system. *Cell*, **98**, 217-227.

Baker, D., Shiau, A.K. and Agard, D.A. (1993) The role of pro regions in protein folding. *Curr Opin Cell Biol*, **5**, 966-970.

Baker, D., Silen, J.L. and Agard, D.A. (1992) Protease pro region required for folding is a potent inhibitor of the mature enzyme. *Proteins*, **12**, 339-344.

Balbach, J. and Schmid, F.X. (2000) *Protein Folding: Frontiers in Molecular Biology*, **(ed. Pain, R.)**.

Bardwell, J.C., Lee, J.O., Jander, G., Martin, N., Belin, D. and Beckwith, J. (1993) A pathway for disulfide bond formation in vivo. *Proc Natl Acad Sci U S A*, **90**, 1038-1042.

Behrens, S., Maier, R., de Cock, H., Schmid, F.X. and Gross, C.A. (2001) The SurA periplasmic PPIase lacking its parvulin domains functions in vivo and has chaperone activity. *Embo J*, **20**, 285-294.

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.

Blake, M.S., Johnston, K.H., Russell-Jones, G.J. and Gotschlich, E.C. (1984) A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on Western blots. *Anal Biochem*, **136**, 175-179.

Brady, G.P. and Sharp, K.A. (1997) Entropy in protein folding and in protein-protein interactions. *Curr Opin Struct Biol*, **7**, 215-221.

Brunger, A.T., Adams, P.D., Clore, G.M., 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. *Acta Crystallogr D Biol Crystallogr*, **54 (Pt 5)**, 905-921.

Chan, H.S. and Dill, K.A. (1998) Protein folding in the landscape perspective: chevron plots and non-Arrhenius kinetics. *Proteins*, **30**, 2-33.

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.

Cheng, H.N. and Bovey, F.A. (1977) Cis-trans equilibrium and kinetic studies of acetyl-L-proline and glycyl-L-proline. *Biopolymers*, **16**, 1465-1472.

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.

Cunningham, E.L., Jaswal, S.S., Sohl, J.L. and Agard, D.A. (1999) Kinetic stability as a mechanism for protease longevity. *Proc Natl Acad Sci U S A*, **96**, 11008-11014.

Dailey, F.E. and Berg, H.C. (1993) Mutants in disulfide bond formation that disrupt flagellar assembly in Escherichia coli. *Proc Natl Acad Sci U S A*, **90**, 1043-1047.

Dill, K.A. (1990) Dominant forces in protein folding. *Biochemistry*, **29**, 7133-7155.

Dill, K.A. and Chan, H.S. (1997) From Levinthal to pathways to funnels. *Nat Struct Biol*, **4**, 10-19.

Dobson, C.M. (2003) Protein folding and misfolding. *Nature*, **426**, 884-890.

Dobson, C.M. and Karplus, M. (1999) The fundamentals of protein folding: bringing together theory and experiment. *Curr Opin Struct Biol*, **9**, 92-101.

Dodd, D.C. and Eisenstein, B.I. (1984) Kinetic analysis of the synthesis and assembly of type 1 fimbriae of Escherichia coli. *J Bacteriol*, **160**, 227-232.

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.

Ellis, R.J. (2000) Chaperone substrates inside the cell. *Trends Biochem Sci*, **25**, 210-212.

Ellis, R.J. (2001) Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr Opin Struct Biol*, **11**, 114-119.

Ellis, R.J. and Hartl, F.U. (1999) Principles of protein folding in the cellular environment. *Curr Opin Struct Biol*, **9**, 102-110.

Eshdat, Y., Silverblatt, F.J. and Sharon, N. (1981) Dissociation and reassembly of Escherichia coli type 1 pili. *J Bacteriol*, **148**, 308-314.

Fersht, A. (2000) *Structure and mechanism in protein science a guide to enzyme catalysis and protein folding*. Freeman, New York.

Fersht, A.R., Matouschek, A. and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J Mol Biol*, **224**, 771-782.

Grathwohl, C. and Wuthrich, K. (1976a) Nmr studies of the molecular conformations in the linear oligopeptides H-(L-Ala)n-L-Pro-OH. *Biopolymers*, **15**, 2043-2057.

Grathwohl, C. and Wuthrich, K. (1976b) The X-Pro peptide bond as an nmr probe for conformational studies of flexible linear peptides. *Biopolymers*, **15**, 2025-2041.

Gray, A.M. and Mason, A.J. (1990) Requirement for activin A and transforming growth factor--beta 1 pro-regions in homodimer assembly. *Science*, **247**, 1328-1330.

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.

Hänggi, P., Talkner, P. and Borkovec, M. (1990) Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.*, **62**, 251.

Herzberg, O. and Moult, J. (1991) Analysis of the steric strain in the polypeptide backbone of protein molecules. *Proteins*, **11**, 223-229.

Hultgren, S.J., Normark, S. and Abraham, S.N. (1991) Chaperone-assisted assembly and molecular architecture of adhesive pili. *Annu Rev Microbiol*, **45**, 383-415.

Hung, D.L., Knight, S.D., Woods, R.M., Pinkner, J.S. and Hultgren, S.J. (1996) Molecular basis of two subfamilies of immunoglobulin-like chaperones. *Embo J*, **15**, 3792-3805.

Ikemura, H., Takagi, H. and Inouye, M. (1987) Requirement of pro-sequence for the production of active subtilisin E in Escherichia coli. *J Biol Chem*, **262**, 7859-7864.

Jacob-Dubuisson, F., Pinkner, J., Xu, Z., Striker, R., Padmanhaban, A. and Hultgren, S.J. (1994a) PapD chaperone function in pilus biogenesis depends on oxidant and chaperone-like activities of DsbA. *Proc Natl Acad Sci U S A*, **91**, 11552-11556.

Jacob-Dubuisson, F., Striker, R. and Hultgren, S.J. (1994b) Chaperone-assisted self-assembly of pili independent of cellular energy. *J Biol Chem*, **269**, 12447-12455.

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.

Jaenicke, R. (1991) Protein stability and molecular adaptation to extreme conditions. *Eur J Biochem*, **202**, 715-728.

Jones, C.H., Danese, P.N., Pinkner, J.S., Silhavy, T.J. and Hultgren, S.J. (1997) The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems. *Embo J*, **16**, 6394-6406.

Jones, C.H., Pinkner, J.S., Nicholes, A.V., Slonim, L.N., Abraham, S.N. and Hultgren, S.J. (1993) FimC is a periplasmic PapD-like chaperone that directs assembly of type 1 pili in bacteria. *Proc Natl Acad Sci U S A*, **90**, 8397-8401.

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, **47 (Pt 2)**, 110-119.

Kobayashi, T. and Ito, K. (1999) Respiratory chain strongly oxidizes the CXXC motif of DsbB in the Escherichia coli disulfide bond formation pathway. *Embo J*, **18**, 1192-1198.

Kramers, H.A. (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **4**, 284.

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.

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.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lai, Z., McCulloch, J., Lashuel, H.A. and Kelly, J.W. (1997) Guanidine hydrochloride-induced denaturation and refolding of transthyretin exhibits a marked hysteresis: equilibria with high kinetic barriers. *Biochemistry*, **36**, 10230-10239.

Lawrence, M.C. and Colman, P.M. (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol*, **234**, 946-950.

Lee, B. (1991) Solvent reorganization contribution to the transfer thermodynamics of small nonpolar molecules. *Biopolymers*, **31**, 993-1008.

Levinthal, C. (1968) Are there pathways for protein folding? *J Chim Phys*, **65**, 44-45.

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.

MacArthur, M.W. and Thornton, J.M. (1991) Influence of proline residues on protein conformation. *J Mol Biol*, **218**, 397-412.

MacIntyre, S., Zyrianova, I.M., Chernovskaya, T.V., Leonard, M., Rudenko, E.G., Zav'Yalov, V.P. and Chapman, D.A. (2001) An extended hydrophobic interactive surface of Yersinia pestis Caf1M chaperone is essential for subunit binding and F1 capsule assembly. *Mol Microbiol*, **39**, 12-25.

Martinez, J.J., Mulvey, M.A., Schilling, J.D., Pinkner, J.S. and Hultgren, S.J. (2000) Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *Embo J*, **19**, 2803-2812.

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.

Miller, S., Schuler, B. and Seckler, R. (1998) Phage P22 tailspike protein: removal of head-binding domain unmasks effects of folding mutations on native-state thermal stability. *Protein Sci*, **7**, 2223-2232.

Missiakas, D., Georgopoulos, C. and Raina, S. (1993) Identification and characterization of the Escherichia coli gene dsbB, whose product is involved in the formation of disulfide bonds in vivo. *Proc Natl Acad Sci U S A*, **90**, 7084-7088.

Missiakas, D. and Raina, S. (1997) Protein folding in the bacterial periplasm. *J Bacteriol*, **179**, 2465-2471.

Mukaiyama, A., Takano, K., Haruki, M., Morikawa, M. and Kanaya, S. (2004) Kinetically robust monomeric protein from a hyperthermophile. *Biochemistry*, **43**, 13859-13866.

Mulvey, M.A., Lopez-Boado, Y.S., Wilson, C.L., Roth, R., Parks, W.C., Heuser, J. and Hultgren, S.J. (1998) Induction and evasion of host defenses by type 1-piliated uropathogenic Escherichia coli. *Science*, **282**, 1494-1497.

Myers, J.K., Pace, C.N. and Scholtz, J.M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci*, **4**, 2138-2148.

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.

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.

Orndorff, P.E. and Falkow, S. (1984) Organization and expression of genes responsible for type 1 piliation in Escherichia coli. *J Bacteriol*, **159**, 736-744.

Otwinowski, Z. and Minor, W. (1997) *Processing of X-ray diffraction data collected in oscillation mode*.

Otzen, D.E., Kristensen, O., Proctor, M. and Oliveberg, M. (1999) Structural changes in the transition state of protein folding: alternative interpretations of curved chevron plots. *Biochemistry*, **38**, 6499-6511.

Pace, C.N., Hebert, E.J., Shaw, K.L., Schell, D., Both, V., Krajcikova, D., Sevcik, J., Wilson, K.S., Dauter, Z., Hartley, R.W. and Grimsley, G.R. (1998) Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2 and Sa3. *J Mol Biol*, **279**, 271-286.

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.

Peters, R.J., Shiau, A.K., Sohl, J.L., Anderson, D.E., Tang, G., Silen, J.L. and Agard, D.A. (1998) Pro region C-terminus:protease active site interactions are critical in catalyzing the folding of alpha-lytic protease. *Biochemistry*, **37**, 12058-12067.

Randall, L.L. and Hardy, S.J. (1986) Correlation of competence for export with lack of tertiary structure of the mature species: a study in vivo of maltose-binding protein in E. coli. *Cell*, **46**, 921-928.

Rietsch, A., Bessette, P., Georgiou, G. and Beckwith, J. (1997) Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. *J Bacteriol*, **179**, 6602-6608.

Rouviere, P.E. and Gross, C.A. (1996) SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes Dev*, **10**, 3170-3182.

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.

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.

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.

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.

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.

Schmid, F.X. (1983) Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry*, **22**, 4690-4696.

Schmid, F.X. (1993) Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions. *Annu Rev Biophys Biomol Struct*, **22**, 123-142.

Shinde, U. and Inouye, M. (2000) Intramolecular chaperones: polypeptide extensions that modulate protein folding. *Semin Cell Dev Biol*, **11**, 35-44.

Shledrick, G.M. and Schneider, T.R. (1997) *SHELXL: High-resolution refinement*.

Silen, J.L. and Agard, D.A. (1989) The alpha-lytic protease pro-region does not require a physical linkage to activate the protease domain in vivo. *Nature*, **341**, 462-464.

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.

Spiess, C., Beil, A. and Ehrmann, M. (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell*, **97**, 339-347.

Stewart, D.E., Sarkar, A. and Wampler, J.E. (1990) Occurrence and role of cis peptide bonds in protein structures. *J Mol Biol*, **214**, 253-260.

Symons, M.C. (1975) Water structure and hydration. *Philos Trans R Soc Lond B Biol Sci*, **272**, 13-28.

Vetsch, M. and Glockshuber, R. (2005) *Formation of Adhesive Pili by the Chaperone-Usher Pathway*. WILEY-VCH, Weinheim.

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.

Vetsch, M., Sebbel, P. and Glockshuber, R. (2002) Chaperone-independent folding of type 1 pilus domains. *J Mol Biol*, **322**, 827-840.

Wolynes, P.G., Onuchic, J.N. and Thirumalai, D. (1995) Navigating the folding routes. *Science*, **267**, 1619-1620.

Wunderlich, M. and Glockshuber, R. (1993) Redox properties of protein disulfide isomerase (DsbA) from Escherichia coli. *Protein Sci*, **2**, 717-726.

Zapun, A., Missiakas, D., Raina, S. and Creighton, T.E. (1995) Structural and functional characterization of DsbC, a protein involved in disulfide bond formation in Escherichia coli. *Biochemistry*, **34**, 5075-5089.

Zav'yalov, V.P., Chernovskaya, T.V., Chapman, D.A., Karlyshev, A.V., MacIntyre, S., Zavialov, A.V., Vasiliev, A.M., Denesyuk, A.I., Zav'yalova, G.A., Dudich, I.V., Korpela, T. and Abramov, V.M. (1997) Influence of the conserved disulphide bond, exposed to the putative binding pocket, on the structure and function of the immunoglobulin-like molecular chaperone Caf1M of Yersinia pestis. *Biochem J*, **324 (Pt 2)**, 571-578.

Zavialov, A.V., Berglund, J., Pudney, A.F., Fooks, L.J., Ibrahim, T.M., MacIntyre, S. and Knight, S.D. (2003) Structure and biogenesis of the capsular F1 antigen from Yersinia pestis: preserved folding energy drives fiber formation. *Cell*, **113**, 587-596.

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.

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.

Zwanzig, R., Szabo, A. and Bagchi, B. (1992) Levinthal's paradox. *Proc Natl Acad Sci U S A*, **89**, 20-22.

Curriculum vitae

Name: **Chasper Puorger**

Date of birth: 01.07.1976

Place of birth: Samedan, Switzerland

Nationality: Swiss

Education:

1983 – 1989 Primarschule, Silvaplana

1989 – 1996 Gymnasium, Zuoz (Matura Typus C)

1996 – 2001 Study of biology, ETH Zürich (Dipl. Natw. ETH)

2001 – 2006 Ph. D. studies at the Institute for Molecular Biology and Biophysics, ETH Zürich

Ich danke Rudi Glockshuber, dass er mir ermöglicht hat auf diesem faszinierenden Projekt zu arbeiten und dass er mir Freiraum für die Verwirklichung eigener Ideen gegeben hat.
Eilika Weber-Ban möchte ich für die Hilfe in technischen aber auch wissenschaftlichen Belangen danken. Ausserdem hat sie sich kurzfristig als Korreferentin zur Verfügung gestellt.
Michael Vetsch danke ich für anregende (nicht nur Wissenschafts-bezogene) Diskussionen und das Einbringen von wertvollen Ideen. Ausserdem bin ich für sein kritisches Begutachten meiner Arbeit sehr dankbar.
Allen anderen Mitstreitern auf dem Pilus Projekt und den anderen „Mitbewohnern“ im E2 möchte ich für die gute Zusammenarbeit und die positive Stimmung im Labor danken.
Mein Dank geht auch an alle anderen ehemaligen und derzeitigen Mitarbeiter der Glockshuber-Gruppe für das angenehme Arbeitsumfeld und die vielfältige Hilfeleistung. Im Besonderen möchte ich Andrea Fritz dafür danken, dass sie mich bei der Proteinproduktion und Reinigung unterstützt hat und auch sonst immer ein offenes Ohr für mich hatte.
Bedanken möchte ich mich auch bei Oliv Eidam für die sehr gute Zusammenarbeit bei der Kristallisation und Strukturaufklärung und für die anregenden Diskussionen.
Ein grosser Dank geht auch an Florian Dworkowski, Verena Finder, Patrick Frei und Heike Summer, die mich immer aufgemuntert haben und die auch ausserhalb des Labors immer wieder für Unternehmungen zu haben waren.
Mein ganz besonderer Dank geht an meine Eltern und meine Schwester. Dank ihrer Unterstützung war es mir vergönnt diesen Weg einzuschlagen und mein Ziel zu erreichen.

Ün cordiel grazcha fich a tuots.