

Supplementary Material

Structure, Folding and Stability of FimA, the Main Structural Subunit of Type 1 Pili from Uropathogenic *Escherichia coli* Strains

Chasper Puorger¹, Michael Vetsch^{1,2}, Gerhard Wider¹, and Rudi Glockshuber¹

¹ ETH Zürich, Institute of Molecular Biology and Biophysics, 8093 Zurich, Switzerland

²Present address: M. Vetsch, Novartis Pharma AG, Biotechnology Development, 4002 Basel, Switzerland

FimAwt

1	10	20	30	40	50	60
AATTVNGGTVHFKGEVVNAACA VDAGSVDQTVQLGQVRTASLAQEGATSSAVGFNIQLND						
	70	80	90	100	110	120
CDTNVASKAAVAFLGTAIDAGHTNVLALQSSAAGSATNVGVQILDRTGAAL TLDGATFSS						
	130	140	150			
ETTLNNGTNTIPFQARYFATGAATPGAANADATFKVQYQ						

FimAa

1	10	20	30	40	50	60
AATTVNGGTVHFKGEVVNAACA VDAGSVDQTVQLGQVRTASLAQEGATSSAVGFNIQLND						
	70	80	90	100	110	120
CDTNVASKAAVAFLGTAIDAGHTNVLALQSSAAGSATNVGVQILDRTGAAL TLDGATFSS						
	130	140	150	160	170	180
ETTLNNGTNTIPFQARYFATGAATPGAANADATFKVQYQGGGGGAATTVNGGTVHFKGE						

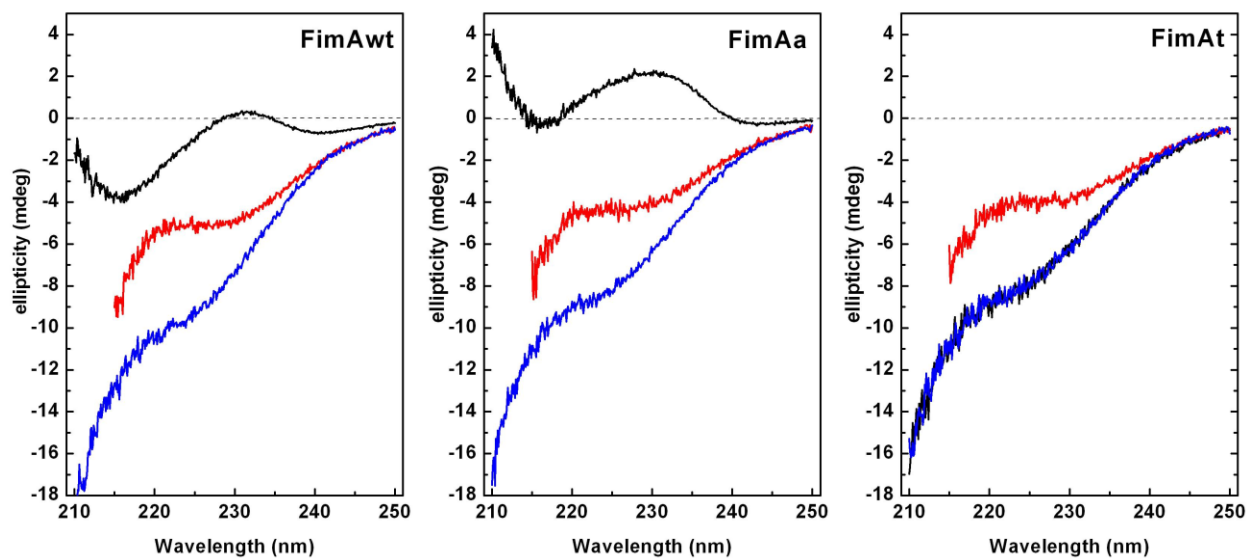
VVNA

FimAt

1	10	20	30	40	50	60
	<i>MHHHHH</i>	GEVVNAACA VDAGSVDQTVQLGQVRTASLAQEGATSSAVGFNIQLND				
	70	80	90	100	110	120
CDTNVASKAAVAFLGTAIDAGHTNVLALQSSAAGSATNVGVQILDRTGAAL TLDGATFSS						
	130	140	150			
ETTLNNGTNTIPFQARYFATGAATPGAANADATFKVQYQ						

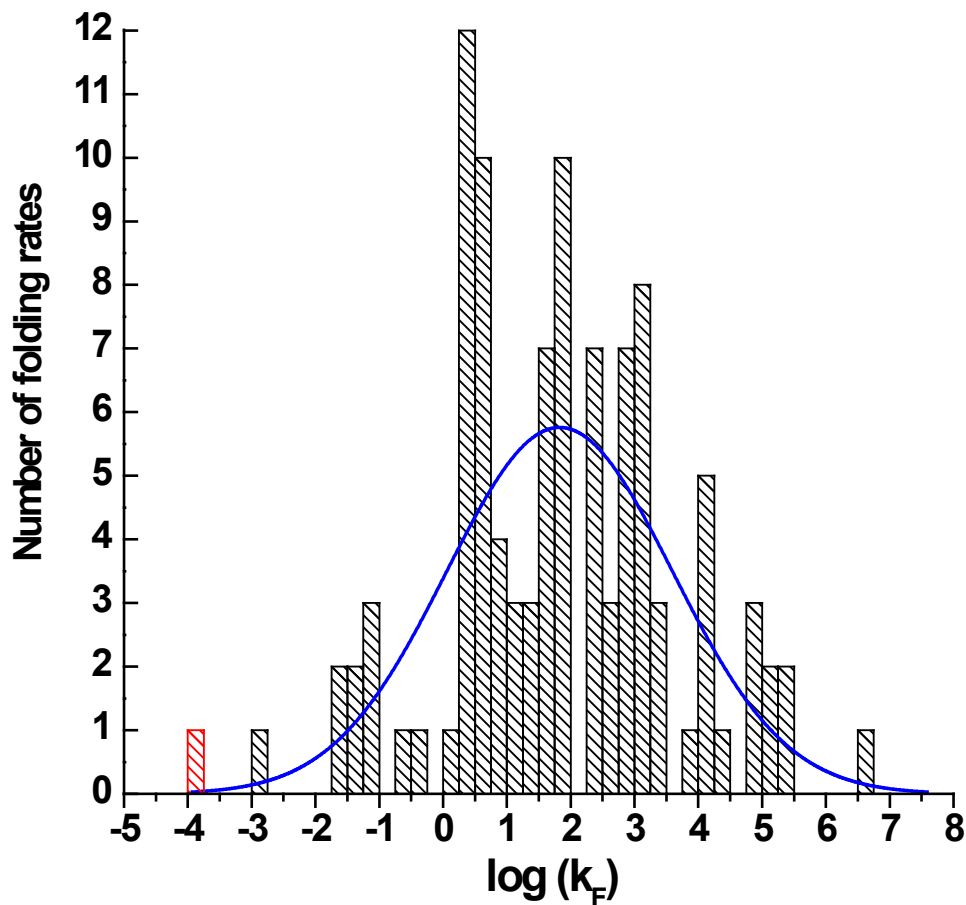
Supplementary figure 1:

Amino acid sequences of FimAwt, FimAa and FimAt. The N-terminal (His)₆ tag in FimAt and the (Gly)₆-linker between the natural FimA C-terminus and the engineered, C-terminal donor strand in FimAa are indicated in italics, the C-terminal donor strand in FimAa is underlined. The amino acid numbering is according to the sequence of mature FimAwt.



Supplementary figure 2:

Far-UV CD spectra of FimAwt (a), FimAa (b) and FimAt (c) measured right after dilution (1:20) of unfolded protein (in 6 M GdmCl) with 10 mM sodium phosphate, pH 7.0, 200 mM NaCl buffer with (red) or without (blue) 6 M GdmCl. The measurement of one spectrum took less than three minutes. The black lines represent the corresponding spectra after several days of incubation in refolding buffer.



Supplementary figure 3:

Frequency of reported folding rates of model proteins. 104 rate constants are covered in this plot (55 folding rates of two-state folders, 39 rates of formation of the native state and 10 rates of intermediate formation for non-two-state folders).³⁸⁻⁴⁴ The folding rate of FimA is indicated by the red bar. The solid blue line represents a fit according to a Gaussian function (fit does not include FimA).

Table S1:

Contact order (CO), absolute contact order (ACO), and the number of sequence distant native pairs (Q_d) describing the topological complexity of the FimAa structure[#].

	Value
CO (averaged over all structures)	13.5 %
CO (contacts present in all structures)	18.2 %
ACO (averaged over all structures)	22.6
ACO (contacts present in all structures)	30.4
Q_d (averaged over all structures)	201
Q_d (contacts present in all structures)	145

[#] Since NMR structures are represented by a number of slightly different conformers representing the structure it is not straightforward to calculate the contact parameters. We used two approaches: calculation of the parameter for all conformers individually and take its average, as well as considering only those contacts that are present in all structures.