Electronic Supplementary Information

Enhancing the Enthalpic Contribution of Hydrogen Bonds in FimH by Solvent Shielding

Jonathan Cramer,[a] Xiaohua Jiang,[a] Wojciech Schönemann,[a] Marleen Silbermann,[a] Pascal Zihlmann,[a] Stefan Siegrist,[a] Brigitte Fiege,[a] Roman Peter Jakob,[b] Said Rabbani,[a] Timm Maier[b] and Beat Ernst[a]*

[a] University of Basel, Department of Pharmaceutical Sciences, Institute of Molecular Pharmacy, Klingelbergstrasse 50, 4056, Basel, Switzerland

[b] University of Basel, Department Biozentrum, Focal Area Structural Biology, Klingelbergstr. 70, 4056 Basel, Switzerland

* Corresponding author. Tel.: +41 61 207 15 51; Fax: +41 61 207 15 52; E-mail: beat.ernst@unibas.ch (B. Ernst)

Table of Contents

1. Protein Expression and Purification

$FimH_{LD}$ from *E.coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a 6His-tag ($FimH_{LD}$-Th-6His, 173 residues) following a previously published protocol.[1] The clone containing the $FimH_{LD}$ construct was expressed in the protease-deficient *E.coli* HM125 strain at 30°C and 180 rpm in M9 minimal medium supplemented with 100 µg/mL ampicillin. The protein expression was induced by 1 mM IPTG at an OD_{600} of 0.8. The cells were further cultivated for 16 hrs, harvested by centrifugation for 20 min at 5'000 rpm and 4°C. The pellet was resuspended in lysis buffer containing 50 mM Tris pH 7.4, 150 mM NaCl, 5 mM EDTA, and 1 mg/mL polymyxin B sulfate. The supernatant containing the periplasmic extract was dialyzed against sodium phosphate buffer and purified on Ni-NTA columns. The protein was finally dialyzed against assay buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl and 1 mM $CaCl_2$. For long time storage the protein was frozen at -80°C. For production of uniformly ^{15}N-labeled $FimH_{LD}$-Th-6His for NMR experiments, *E.coli* HM125 was cultivated in M9 minimal medium containing 1 g/L $^{15}NH_4Cl$ (CortecNet, France) as the sole source of nitrogen. The labeled protein was purified as described above and dialyzed against 20 mM phosphate buffer pH 7. The exact molecular weight (18860.2 Da) was determined by mass spectrometry.

$FimH_{FL}$ in complex with a stabilizing donor strand of FimG - the DsG peptide - from the same *E. coli* strain was produced according to a previously published protocol.[2] The FimH and FimC were coexpressed in *E.coli* HM125 at 30°C in M9 minimal medium supplemented with 100 µg/mL ampicillin. The protein expression was induced by 1 mM IPTG at an OD_{600} of 1.5. The cells were further cultivated for 12–16 hrs, harvested by centrifugation for 20 min at 5'000 rpm and 4°C. The pellet was resuspended in lysis buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 5 mM EDTA and 1 mg/mL polymyxin B sulfate. The suspension was stirred at 4 °C for 1.5 h and cells and debris were pelleted. The supernatant containing the periplasmic extract was dialyzed against 20 mM Tris pH 8.0. All following purification steps were performed at 4 °C. The solution containing FimC–FimH was loaded onto a pre-equilibrated Uno Q column (Bio-Rad, California, USA). The fractions of the flow through containing the FimC–FimH complex were combined and dialyzed against 10 mM MOPS pH 7.0 buffer. The solution was loaded onto a pre-equilibrated Mono S column (GE Healthcare, Little Chalfront, UK). The complex was eluted with a linear gradient of NaCl (0–300 mM NaCl). Fractions containing the FimC–FimH complex were combined and dialyzed against

buffer containing 20 mM NaH_2PO_4 and 50 mM NaCl. The purified FimC–FimH complex was concentrated to 40 μM and incubated with a 3-fold molar excess of synthetic DsG peptide, corresponding to the N-terminal donor strand of FimG with an additional C-terminal arginine residue to improve solubility. The formed FimHFL–DsG complex was dialyzed against 20 mM acetic acid pH 4.5, loaded unto a Mono S column, and eluted using a linear NaCl gradient (0–400 mM). Finally, the purified protein was dialyzed against buffer containing 20 mM HEPES pH 7.4 and 150 mM NaCl. Concentration was determined by UV-Vis spectroscopy at 280 nm (extinction coefficient 35090 M^{-1} cm^{-1}).

2. Isothermal Titration Calorimetry

Isothermal titration calorimetric experiments with $FimH_{LD}$ or $FimH_{FL}$ were performed on an VP-ITC (Malvern instruments, Worcestershire, UK) or ITC200 (MicroCal, Northampton, USA) instrument at 25 °C using standard instrument settings (reference power 10 μcal s^{-1} (VP) / 6 μcal s^{-1} (ITC200), stirring speed 307 rpm (VP) / 750 rpm (ITC200), feedback mode high, filter period 2 s). Protein solutions were dialyzed against ITC buffer (20 mM HEPES, 150 mM NaCl) prior to the experiments and all samples were prepared using the dialysate buffer to minimize dilution effects. Protein concentrations were determined spectrophotometrically with the specific absorbance at 280 nm employing an extinction coefficient of 18600 M^{-1} cm^{-1} ($FimH_{LD}$) or 35090 M^{-1} cm^{-1} ($FimH_{FL}$). 2% DMSO was added as a co-solvent in all titrations. In a typical experiment, tenfold protein concentration was chosen as syringe concentration and 25 injections of 6–10 μL (VP) or 1.5 μL (ITC200) were performed. Baseline correction, peak integration, and non-linear regression analysis of experimental data was performed using either the AFFINImeter suite (v2.1802.5, S4SD - AFFINImeter, Santiago de Compostela, Spain) or the NITPIC (version 1.2.2.)[3] and sedphat (version 12.1b)[4] software packages. Typically, experiments were performed in duplicate or triplicate and the 68% confidence intervals from global fitting of multiple experiments were calculated as an estimate of experimental error. SEDPHAT was used for simulation of experimental data and calculation of error surface projections.

Due to high c-value conditions, the determination of the thermodynamics of **1** binding to $FimH_{LD}$ required a competitive titration setup. Binding enthalpy was determined from a direct titration of **1** into a solution containing 9 μM $FimH_{LD}$. Accurate determination of K_a was possible by pre-incubation of 9 μM $FimH_{LD}$ with an excess (600 μM) of weak binder **5**.

Data for compounds **1** and **3** binding to $FimH_{FL}$ were collected in a direct titration of **1** and **3** into 10 µM $FimH_{FL}$. A reliable extraction of thermodynamic data for the weak interaction of $FimH_{FL}$ with **5**–**8** required a competitive titration setup. For this, a solution containing 10 µM $FimH_{FL}$ was incubated with an excess of **5**–**8** and titrated with strong ligand **1**. The required compound concentration to sufficiently shift the observed binding isotherm of **1** depends on the affinity of the weak binder. It was found that insufficient saturation of $FimH_{FL}$ resulted in large errors and a high correlation between fitting parameters. Even for experiments with sufficiently high saturation, a flattening of the one-dimensional error surface projection of the fitting parameters was observed. Simulated data (Figure S19) revealed that this phenomenon was partly caused by experimental noise and could be alleviated by further increasing the concentration of the competitor. Reliable results were obtained when the protein was incubated with 5 mM of **5** (ca. 85% initial saturation) and **6** (ca. 80% initial saturation) and 11-12 mM of **7** (ca. 55% initial saturation) or **8** (ca. 70% initial saturation). One- and two-dimensional error surface projections for experiments performed with this setup are depicted in Figures S9–S18.

Table S1. Thermodynamic data from isothermal titration calorimetry experiments (LD stands for $FimH_{LD}$ and FL for $FimH_{FL}$).

Interaction	K_D [µM]	$\Delta G°$ [kJ mol^{-1}]	$\Delta H°$ [kJ mol^{-1}]	$-T\Delta S°$ [kJ mol^{-1}]	n
LD-**1**	3.2E-04 (2.9E-04 – 3.6E-04)	–54.2 (–54.4 to –53.9)	–69.1 (–69.1 to –69.0)	14.9 (14.6 to 15.2)	Direct: 1.06 Comp.: 1.08
LD-**8**	7408 (4831 – 13351)	–12.2 (–13.2 to –10.7)	5.1 (3.6 to 9.7)	–17.3 (–22.9 to –14.3)	1.01
FL-**1**	0.047 (0.034 – 0.062)	–41.9 (–42.6 to –41.2)	–80.6 (–82.8 to –78.6)	38.8 (36.0 to 41.6)	1.12
FL-**3**	0.764 (0.736 – 0.794)	–34.9 (–35.0 to –34.8)	–63.7 (–64.2 to –63.2)	28.8 (28.2 to 29.4)	1.08
FL-**5**	886 (718 – 1106)	–17.4 (–17.9 to –16.9)	–24.9 (–27.6 to –22.3)	7.5 (4.4 to 10.8)	0.95
FL-**6**	1431 (1071–1955)	–16.2 (–17.0 to –15.5)	–16.7 (–20.7 to –13.0)	0.5 (–4.0 to 5.3)	1.12
FL-**7**	8835 (6053–14054)	–11.7 (–12.7 to –10.6)	–2.6 (–8.6 to 2.3)	–9.2 (–14.9 to –1.9)	1.04
FL-**8**	4131 (2589 – 4131)	–13.6 (–14.8 to –12.2)	3.2 (–1.6 to 7.7)	–16.8 (–22.5 to –10.6)	1.10

Plots of titration experiments and fitted isotherms

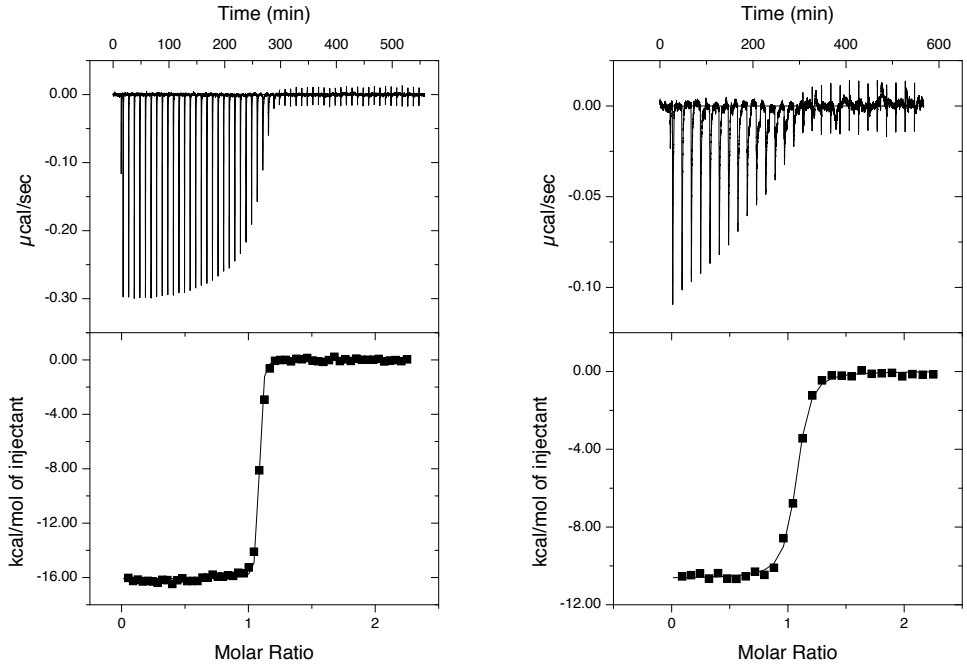

Figure S1. Direct (left) and competitive (right) titration of **1**. 100 µM **1** were titrated into 9 µM $FimH_{LD}$ (600µM **5** for the competitive experiment).

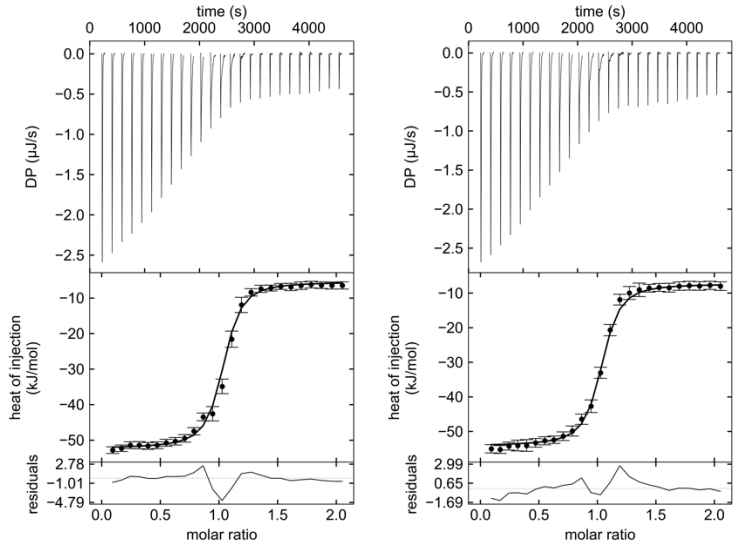

Figure S2. Competitive titration of **8**. 300 µM **4** were titrated into 30 µM $FimH_{LD}$ preincubated with 10 mM **8**.

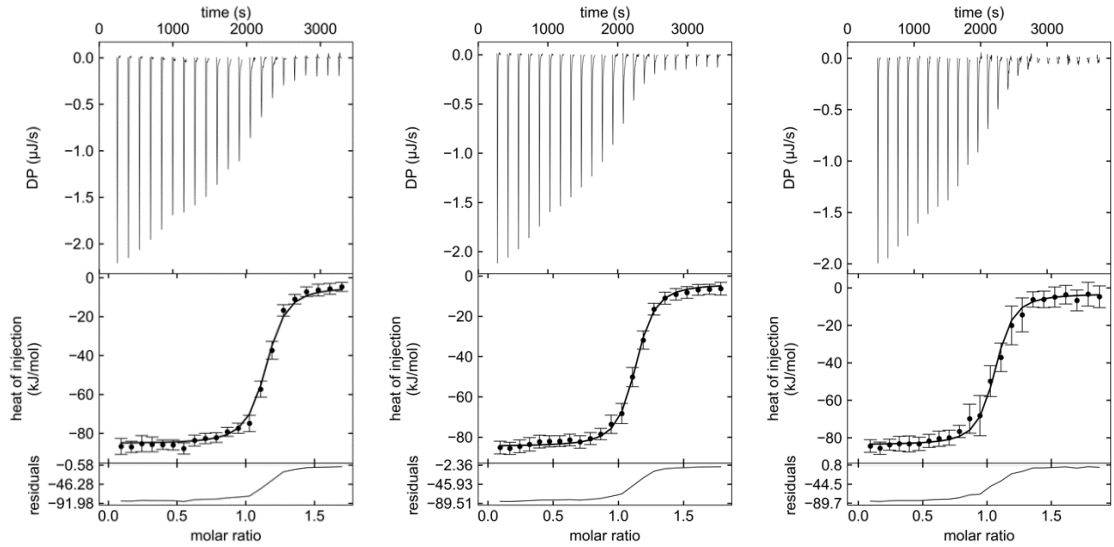

Figure S3. Direct titration of **1**. 100 µM **1** were titrated into 10 µM $FimH_{FL}$.

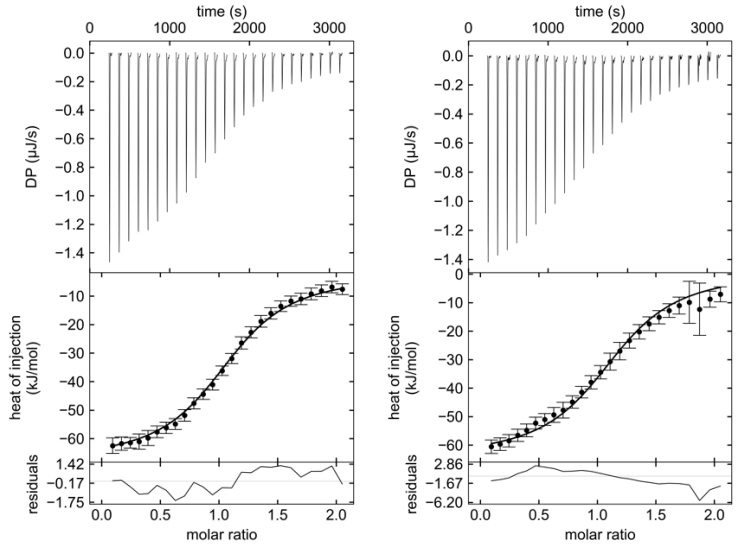

Figure S4. Direct titration of **3**. 100 µM **3** were titrated into 10 µM $FimH_{FL}$.

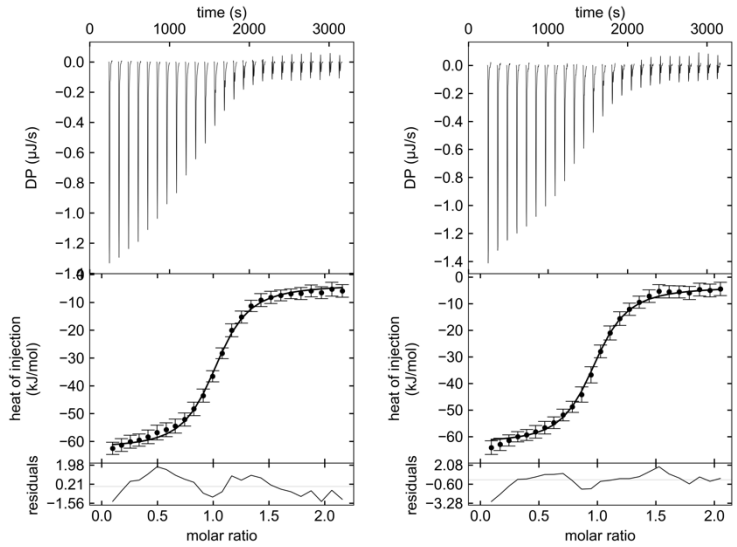

Figure S5. Competitive titration of **5**. 100 µM **1** were titrated into 10 µM $FimH_{FL}$ preincubated with 5 mM **5**.

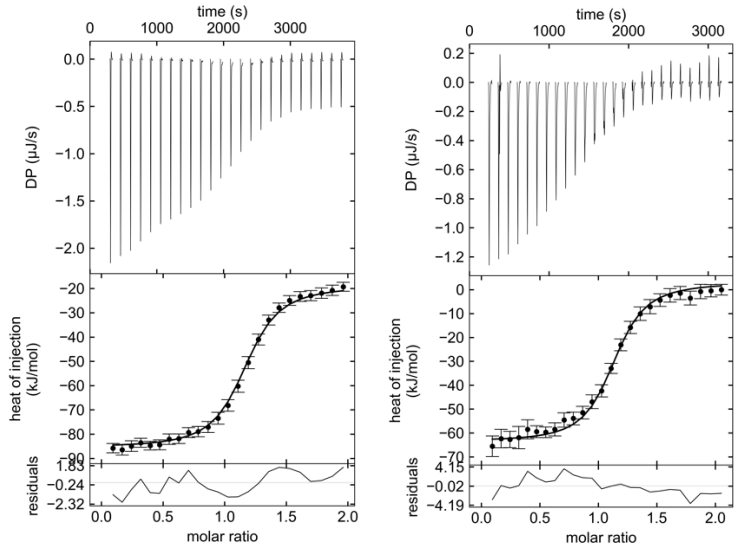

Figure S6. Competitive titration of **6**. 100 µM **1** were titrated into 10 µM $FimH_{FL}$ preincubated with 5 mM **6**.

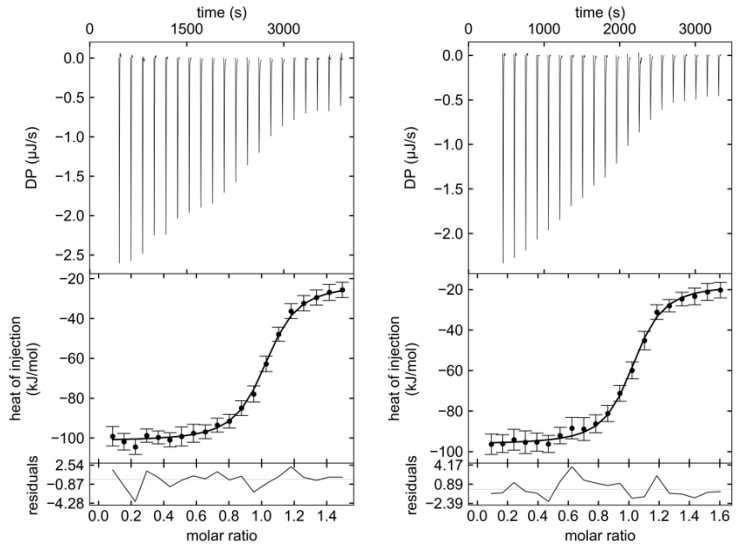

Figure S7. Competitive titration of **7**. 100 µM **1** were titrated into 10 µM $FimH_{FL}$ preincubated with 12 mM **7**.

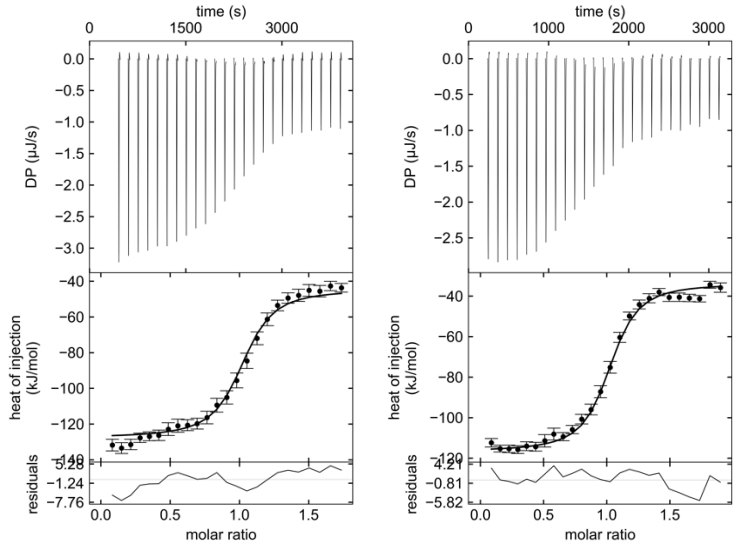

Figure S8. Competitive titration of **8**. 100 µM **1** were titrated into 10 µM $FimH_{FL}$ preincubated with 11 mM **8**.

Error surface projections

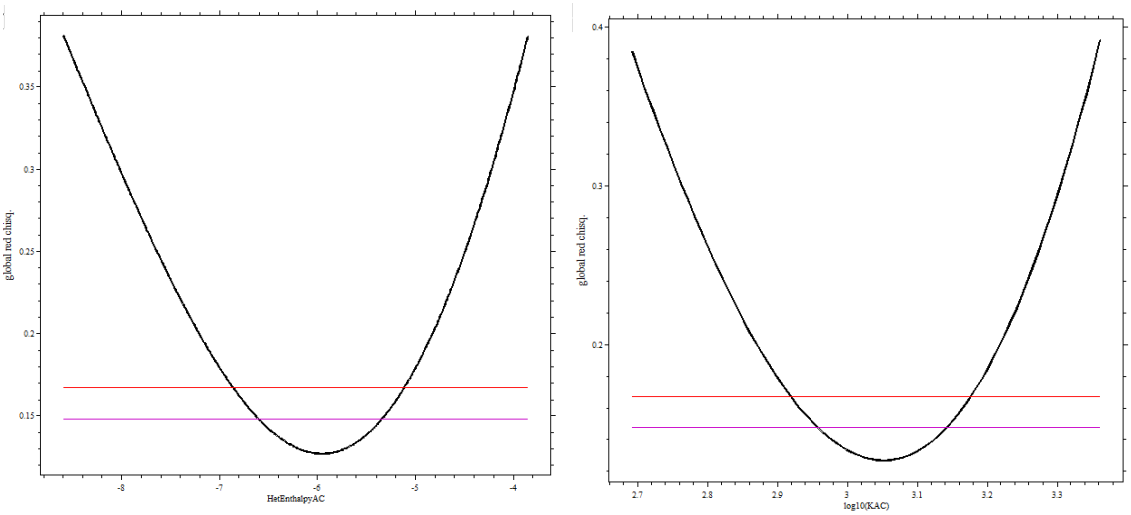

Figure S9. One-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **5**.

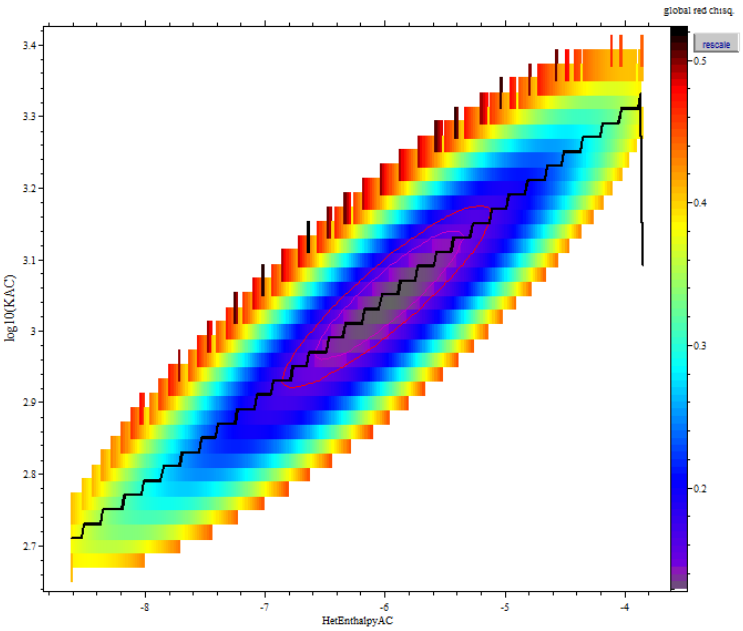

Figure S10. Two-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **5**.

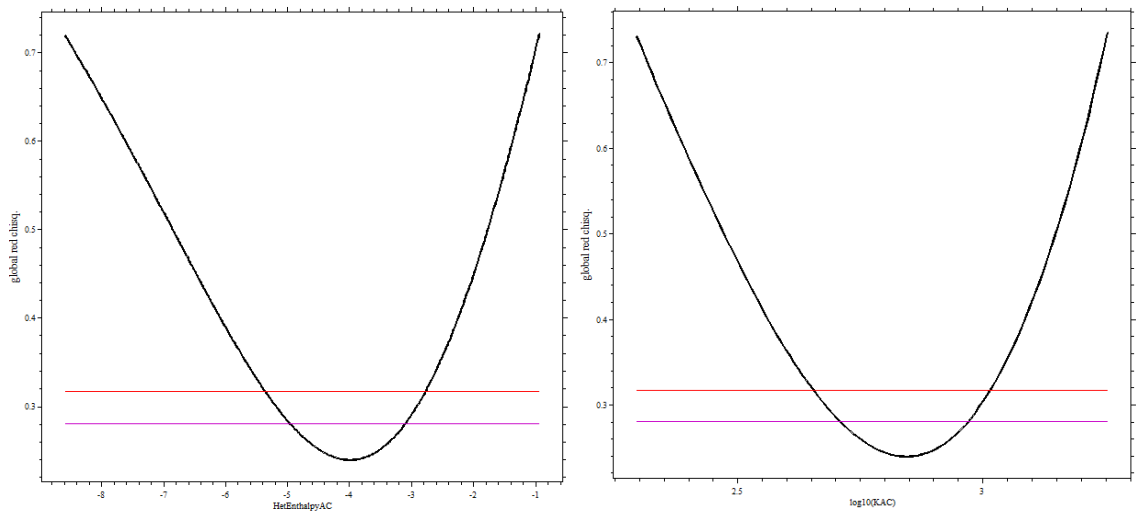

Figure S11. One-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **6**.

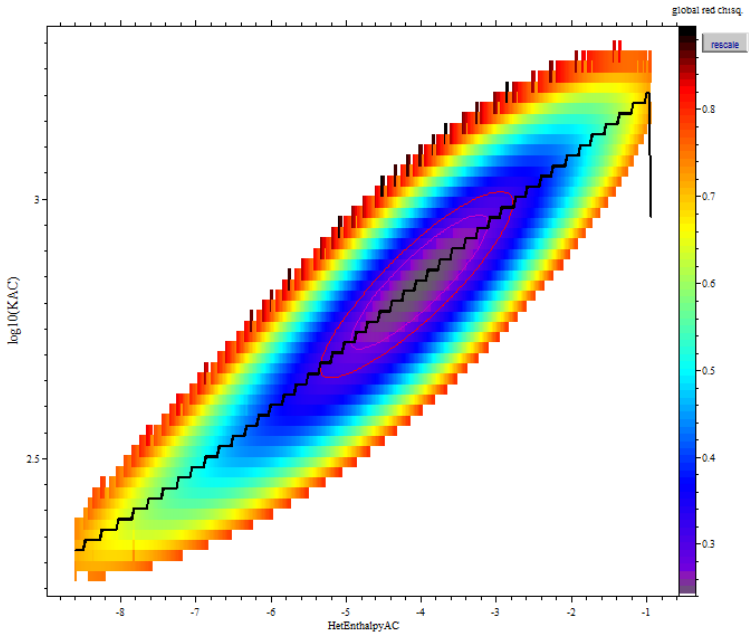

Figure S12. Two-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **6**.

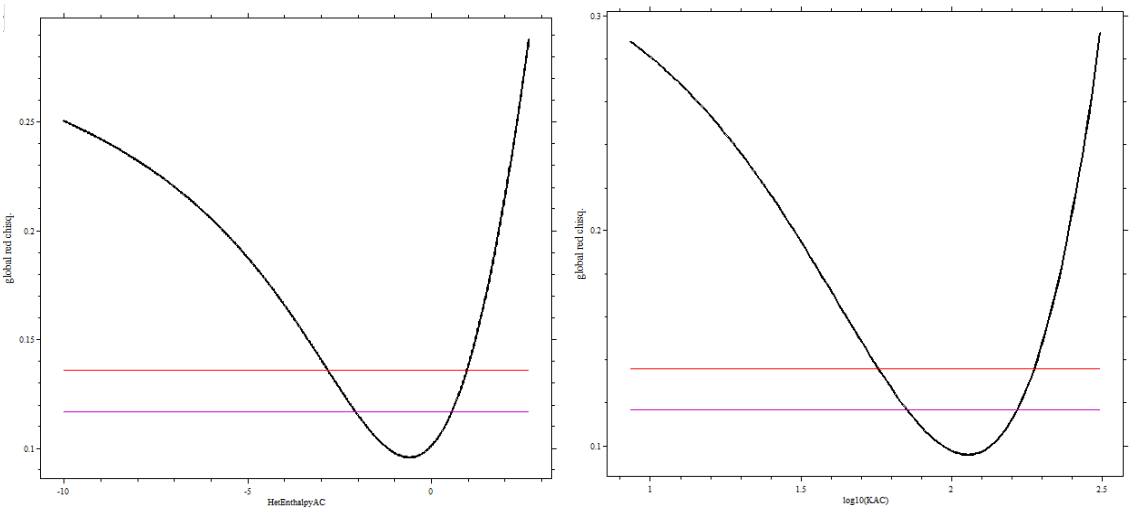

Figure S13. One-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **7**.

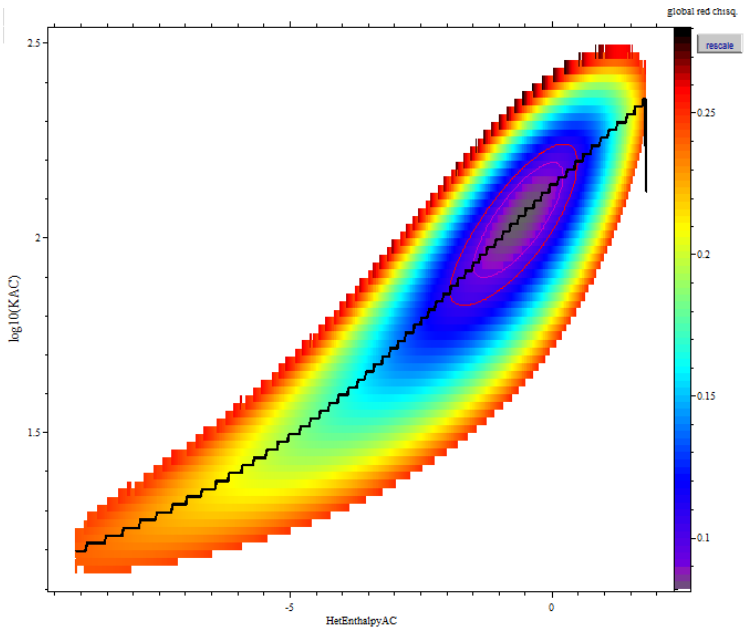

Figure S14. Two-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **7**.

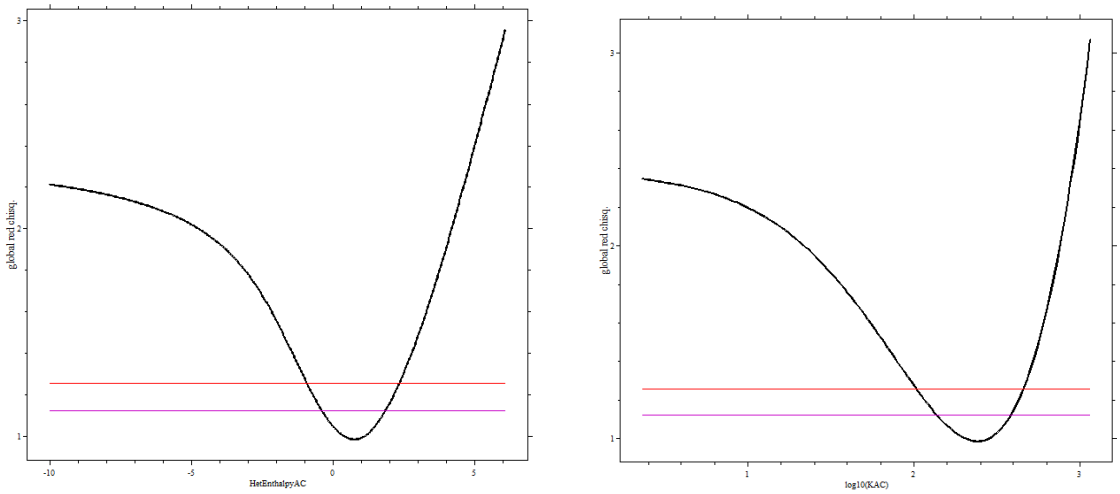

Figure S15. One-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **8**.

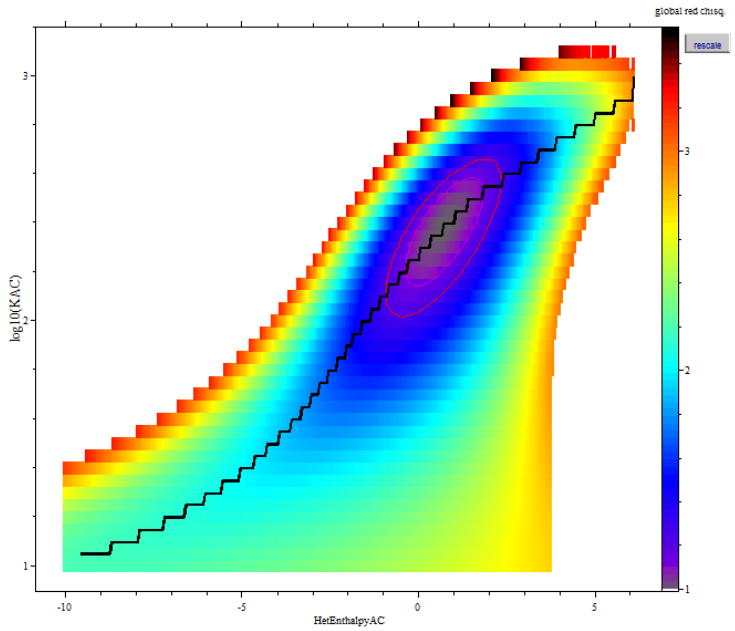

Figure S16. Two-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration with **8**.

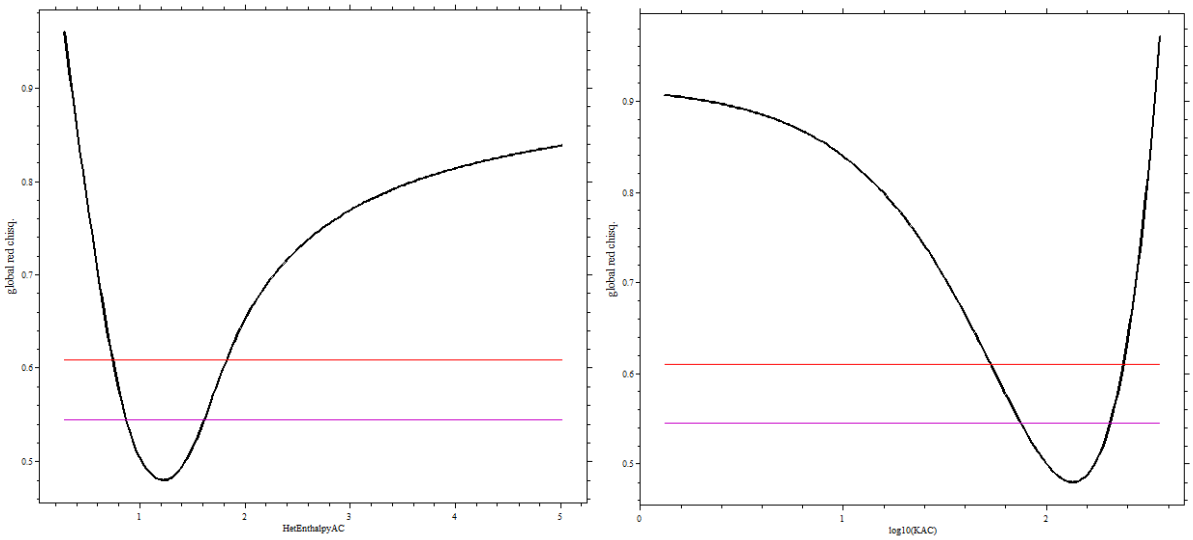

Figure S17. One-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration of $FimH_{LD}$ with **8**.

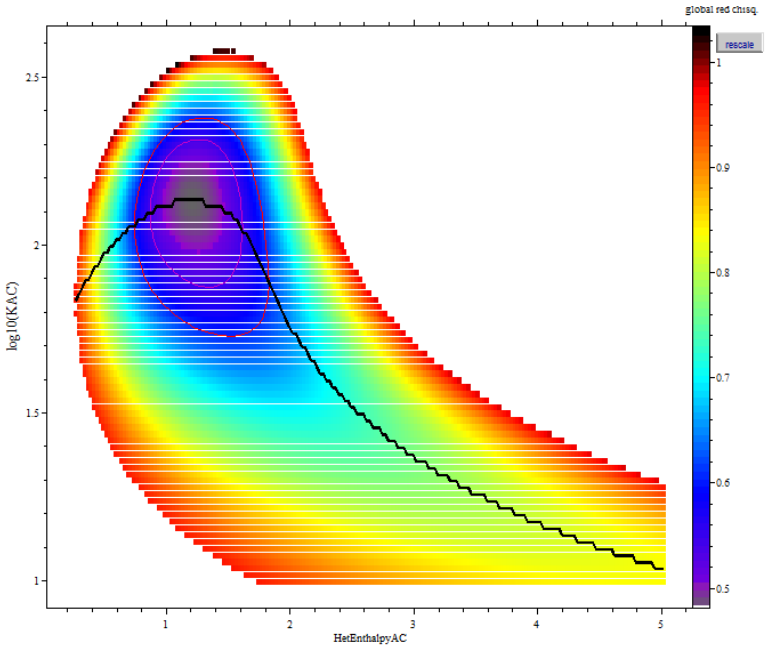

Figure S18. Two-dimensional error surface projection for fitting parameters $\Delta H°$ and $\log(K_a)$ in global fitting of competitive titration of $FimH_{LD}$ with **8**.

Evaluation of error surface contours from simulated data

To estimate the effect of experimental noise on the accuracy of fitting parameters a series of experiments were simulated using the SEDPHAT software package. A competitive model (B into AC) was chosen with parameters roughly corresponding to best fit parameters for the weakest binder **7** ($\log K_{aAB}$ = 7.43, $\Delta H°_{AB}$ = –19.21 kcal mol^{-1}, $\log K_{aAC}$ = 2.05, $\Delta H°_{AC}$ = –0.61 kcal mol^{-1}). In the simulated experiments, ca. 100 µM of B was titrated into a solution containing ca. 10 µM A and 10 mM C. The simulated experimental noise level was successively set to 100, 300, and 500. For the titration in Figure S19D, the concentration of C was set to 46 mM applying a noise level of 500.

The simulations show that flattening of the error surface projections is partly a result of experimental error. This correlates with a broadening of the confidence intervals and progressive correlation of fitting parameters $\Delta H°$ and K_a. The effect can be alleviated by increasing the competitor concentration (Figure S19D).

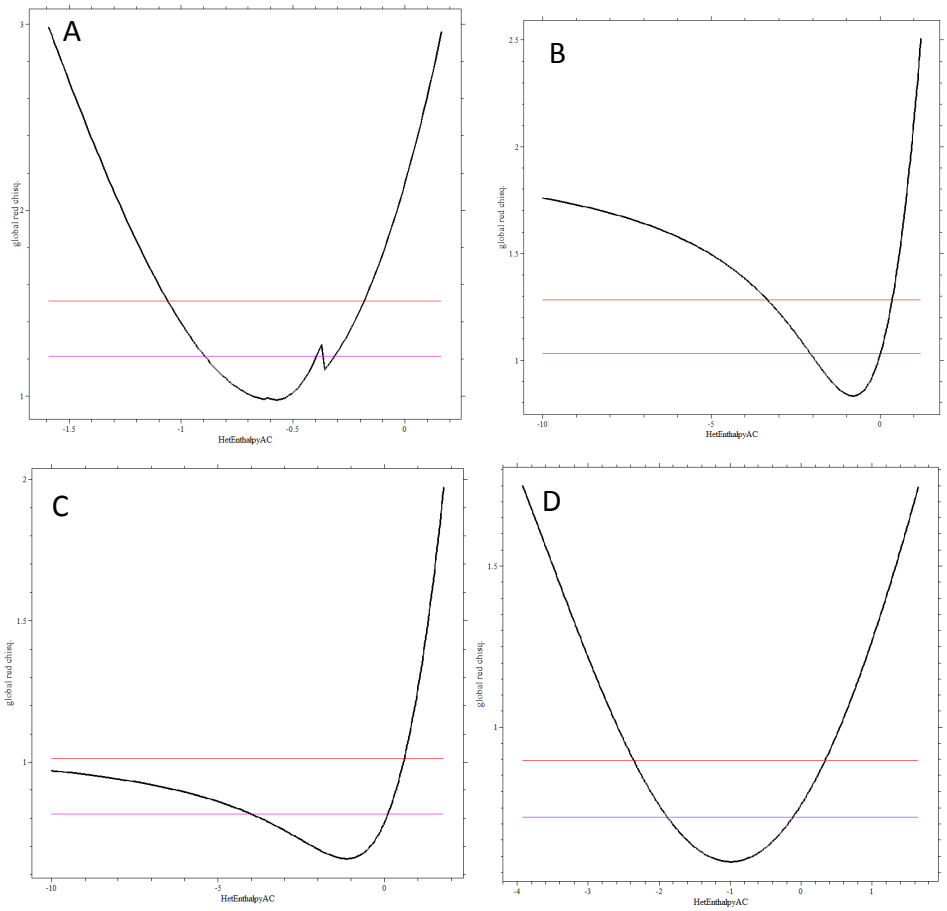

Figure S19. One-dimensional error surface projection for fitting parameter $\Delta H°$ in simulated competitive titrations at different noise levels. A) Noise level 100, [C] = 10 mM. B) Noise level 300, [C] = 10 mM. C) Noise level 500, [C] = 10 mM. D) Noise level 500, [C] = 46 mM.

3. X-ray Crystallography

For crystallization, $FimH_{LD}$ (residues 1-158)[1] was used at a final concentration of 12 mg/mL (ca. 0.8 mM) with a threefold molar excess of ligand (2.5 mM) in 20 mM HEPES buffer pH 7.4. Crystals were grown in sitting-drop vapor diffusion at 4°C, 12°C and 20°C in 0.2 M $(NH_4)_2SO_4$, 0.1 M HEPES pH 7 and 25-30% PEG3350. Plate like crystals appeared after 2 weeks, were cryopreserved by addition of 20% glycerol (v/v) and flash-cooled with liquid nitrogen. Data was collected at the SLS beamlines X06DA and X06SA of the Swiss Light Source (Paul Scherrer Institute, Switzerland) and indexed, integrated and scaled with XDS.[5,6] Structures were solved by molecular replacement with PHASER[7] using the $FimH_{LD}$-*n*-heptyl α-D-mannopyranoside complex (PDB code 4XO8) as search model. The structures were built using the COOT software[8] and periodically refined with the PHENIX and Buster-TNT software.[9,10] Geometric restraints for the ligands were generated with PRODRG[11] and Molprobity[12] was used for validation. The atomic coordinates have been deposited in the RCSB Protein Data Bank and are available under the accession code 5L4T, 5L4V, and 5L4X, respectively.

Table S2. Statistics on diffraction data and refinement of $FimH_{LD}$ and its ligand complexes.

	$FimH_{LD}$ 5	**$FimH_{LD}$ 6**	**$FimH_{LD}$ 7**
PDB Identifier	5L4T	5L4V	5L4X
Wavelength (Å)	1.00001	1.00004	1.00003
Resolution range (Å)	55.3 - 1.90 (2.01 - 1.90)*	41.8 - 3.0 (3.17 - 3.0) *	32.2 - 1.90 (1.98 - 1.90) *
Space group	P 21 2 21	P 1 21 1	P 21 21 21
Unit cell	67.76 68.57 96.11	44.83 95.34 70.80	61.08 61.38 95.63
α, β, γ (°)	90 90 90	90 105.0 90	90 90 90
Total reflections	195676 (30395)	51948 (7481)	214352 (12961)
Unique reflections	35539 (5439)	11564 (1748)	35551 (2143)
Multiplicity	5.5 (5.5)	4.4 (4.2)	6.0 (6.0)
Completeness (%)	98.5 (98.7)	98.6 (93.9)	99.9 (99.9)
Mean I/sigma(I)	7.0 (1.5)	4.5 (1.3)	8.8 (2.3)
Wilson B-factor	25.3	31.0	16.6
R-meas	0.175 (1.445)	0.301 (1.13)	0.062 (0.72)
CC1/2	0.995 (0.707)	0.860 (0.520)	0.997 (0.925)
R-work	0.205 (0.32)	0.248 (0.337)	0.178 (0.295)
R-free	0.221 (0.361)	0.276 (0.361)	0.205 (0.306)
RMS(bonds)	0.004	0.011	0.006
RMS(angles)	0.96	1.6	1.09
Ramachandran favored (%)	97.2	98.1	97.4
Ramachandran outliers (%)	0	0	0
Clashscore	1.2	1.5	2.7

*The values in parentheses correspond to the highest resolution shell

4. NMR Spectroscopy

$^{1}H,^{15}N$-HSQC NMR experiments were measured at 298 K on a Bruker Avance III 600 MHz NMR spectrometer equipped with a 5 mm TXI room temperature probe head. Samples contained 120 µM of ^{15}N-labeled $FimH_{LD}$ in 20 mM phosphate buffer pH 7.4 in water with 7% D_2O. Ligands were dissolved in H_2O at 10 to 20 mM concentrations and added stepwise up to 2- to 5-fold molar excess. NMR spectra were acquired and processed with Topspin 3.2 (Bruker BioSpin, Switzerland) and analyzed with CcpNmr Analysis (version 2.2).[13] The backbone assignment of $FimH_{LD}$ was available from previous studies.[14] Combined chemical shift differences, $\Delta\delta_{AV}$, between free and ligand-bound protein signals were calculated as in equation 1.[15]

$$\Delta\delta_{AV} = \sqrt{(\Delta\delta^{1}H)^{2} + (0.2\Delta\delta^{15}N)^{2}} \qquad \text{(eq. 1)}$$

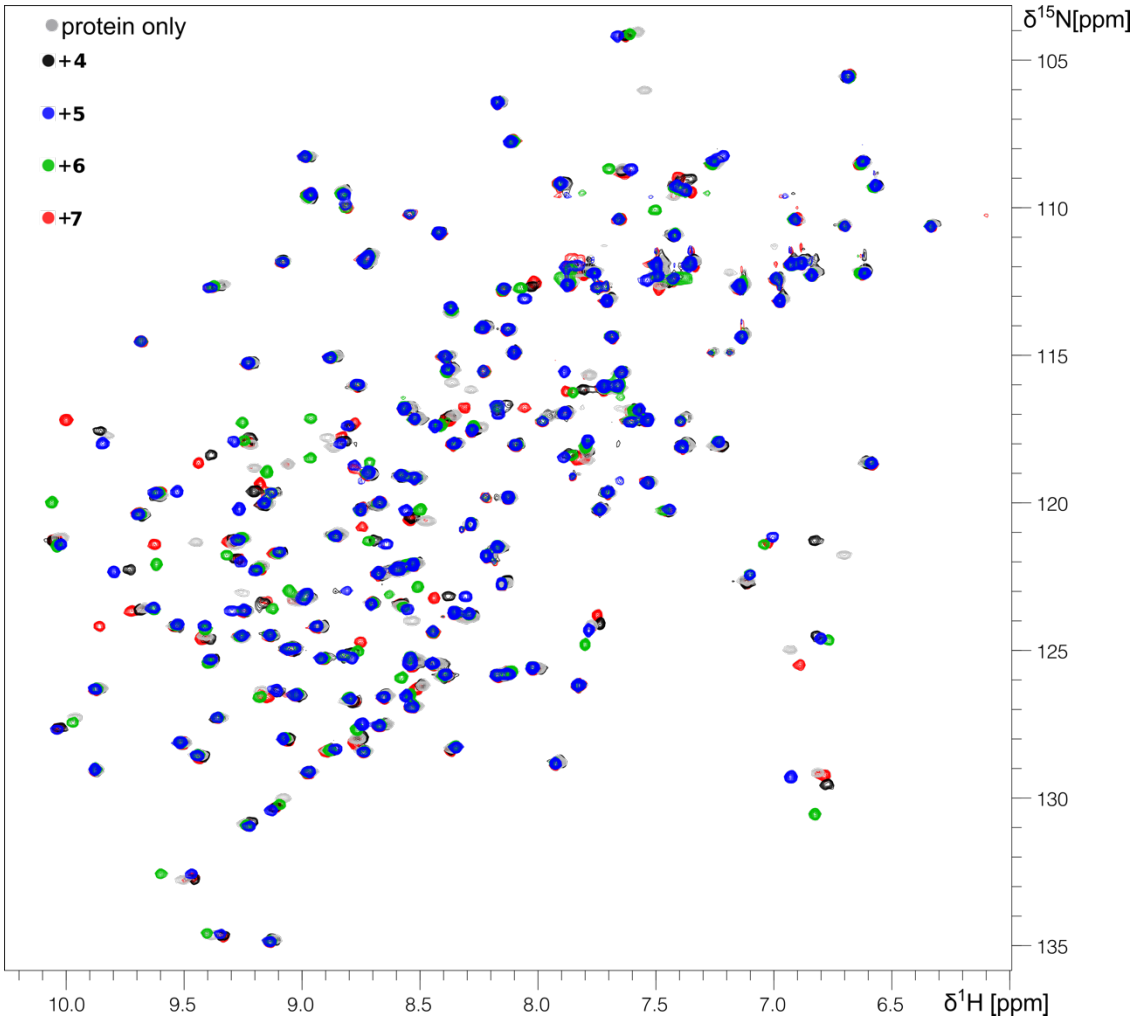

Figure S20. Overlap of $^{1}H,^{15}N$-HSQC spectra of $FimH_{LD}$ in absence of ligand (grey) and in presence of **4** (black), **5** (blue), **6** (green), and **7** (red).

5. References

[1] S. Rabbani, X. Jiang, O. Schwardt, B. Ernst, *Anal. Biochem.* **2010**, *407*, 188–195.

[2] M. M. Sauer, R. P. Jakob, J. Eras, S. Baday, D. Eriş, G. Navarra, S. Bernèche, B. Ernst, T. Maier, R. Glockshuber, *Nat. Commun.* **2016**, *7*, 10738.

[3] T. H. Scheuermann, C. A. Brautigam, *Methods* **2015**, *76*, 87–98.

[4] G. Piszczek, *Methods* **2015**, *76*, 137–148.

[5] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 133–144.

[6] W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 125–132.

[7] A. J. McCoy, *Acta Crystallogr D Biol Crystallogr* **2007**, *63*, 32–41.

[8] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

[9] P. D. Adams, P. V Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.

[10] E. Blanc, P. Roversi, C. Vonrhein, C. Flensburg, S. M. Lea, G. Bricogne, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2210–2221.

[11] D. M. van Aalten, R. Bywater, J. B. Findlay, M. Hendlich, R. W. Hooft, G. Vriend, *J. Comput. Aided. Mol. Des.* **1996**, *10*, 255–62.

[12] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 12–21.

[13] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins* **2005**, *59*, 687–696.

[14] B. Fiege, S. Rabbani, R. C. Preston, R. P. Jakob, P. Zihlmann, O. Schwardt, X. Jiang, T. Maier, B. Ernst, *ChemBioChem* **2015**, *16*, 1235–1246.

[15] M. Pellecchia, P. Sebbel, U. Hermanns, K. Wüthrich, R. Glockshuber, *Nat. Struct. Biol.* **1999**, *6*, 336–339.