Electronic Supplementary Information

The Price of Flexibility – A Case Study on Septanoses as Pyranose Mimetics

Christoph P. Sager[a], Brigitte Fiege[a], Pascal Zihlmann[a] Raghu Vannam[b], Said Rabbani[a], Roman P. Jacob[c], Roland C. Preston[a], Adam Zalewski[a], Timm Maier[c], Mark W. Peczuh[b*], Beat Ernst[a*]

[a] *University of Basel, Institute of Molecular Pharmacy, Pharmacenter of the University of Basel, Klingelbergstrasse 50, 4056, Basel, Switzerland*
Email: beat.ernst@unibas.ch

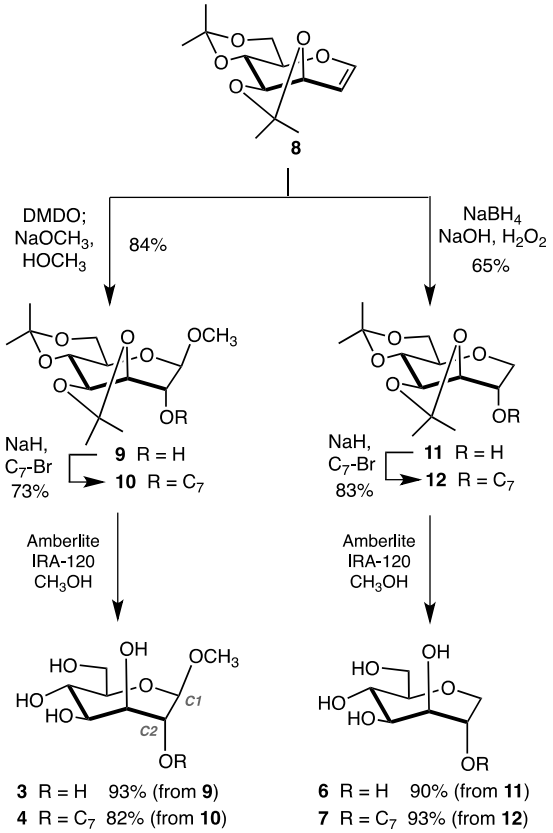
[b] *Department of Chemistry, University of Connecticut, 55 N. Eagleville Road U3060, Storrs, CT, 06279 USA*
Email: mark.peczuh@uconn.edu

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

Total pages: 31

Experimental Procedures

Synthesis and Characterization of New Compounds. Details on the synthesis of compounds **4**, **6** and **7**, including characterization data, are provided below. For the synthesis of mannoside **2**, septanoside **3** and oxepane **5** see references Duff *et al.* (2011)[1] and Markad *et al.* (2008)[2].



Scheme S1. Synthesis of new FimH ligands.

Note:

• Purity of all new compounds was assessed to be >95% by TLC and ^{13}C NMR.

Methyl 2-*O*-*n*-heptyl-3,4;5,7-diacetonide-D-*glycero*-D-galactoseptanoside (10). To a solution of **9**[2] (0.024 g, 0.078 mmol) in dry DMF (1 mL), were added tetrabutylammoniumiodide (TBAI) (0.0028 g, 0.0078 mmol) and NaH (0.093 g, 0.39 mmol) at 0 °C. After stirring this solution at 0 °C for 30 min, heptyl bromide (0.069 g, 0.39 mmol) was added. After 12 h, the mixture was quenched with several drops of CH_3OH at 0 °C and then concentrated under reduced pressure. The residue was redissolved in EtOAc (20 mL), washed with water (2×10 mL) and brine (1x10 mL). The organic layer was dried over Na_2SO_4, concentrated and purified by column chromatography by using 5:1 Hex: EtOAc as

eluent to give **10** (0.023 g, 73%) as colorless oil. R_f 0.25 (8:2 Hex: EtOAc). $[\alpha]_D$ +35.0 (c = 0.004 in $CHCl_3$); ^{1}HNMR (400 MHz, $CDCl_3$) δ (ppm), 4.28-4.17 (m, 3H), 3.94-3.85 (m, 2H), 3.79-3.70 (m, 2H), 3.67 (m, 1H), 3.46 (s, 3H), 3.40-3.32 (m, 2H), 1.61-1.56 (m, 3H), 1.51 (s, 6H), 1.45 (s, 3H), 1.40 (s, 3H), 1.29-1.32 (m, 8H), 0.89 (t, *J* = 6.6 Hz, 3H); ^{13}CNMR (100 MHz, $CDCl_3$) δ (ppm) 109.0, 105.8, 99.2, 82.3, 78,8, 78.8, 78.6, 73.2, 73.0, 69.1, 62.5, 56.6, 32.1, 30.2, 29.3, 28.5, 28.1, 26.2, 25.2, 22.9, 19.7, 14.3; DART HRMS m/z calcd for $C_{21}H_{39}O_7$ $[M+H]^+$ 403.2696, found 403.2700.

Methyl 2-*O*-*n*-heptyl-D-*glycero*-D-galactoseptanoside (4). To a solution of **10** (0.020 g, 0.049 mmol) in EtOH/H_2O (3 mL, 1:3) was added Amberlite H^+ IR-120 (0.023 g) and refluxed it for 4 h at 110 °C. The mixture was allowed to cool to rt and then filtered through a short pad of celite with washing (3x3 mL MeOH). The filtrate was concentrated under reduced pressure and purified by column chromatography by using 10:1 DCM: MeOH as eluent to give pure **4** (0.013 g, 82%) as colorless oil. R_f0.5 (DCM : MeOH 8:2). $[\alpha]_D$ +53.1 (c = 0.002 in CH_3OH); ^{1}HNMR (400 MHz, CD_3OD) δ (ppm) 4.27 (d, *J* = 7.1 Hz, 1H), 3.94-3.88 (m, 2H), 3.73-3.67(m, 3H), 3.63-3.59 (m, 3H), 3.48(s, 3H), 3.38-3.32 (m, 2H), 1.59-1.54 (m, 2H), 1.39-1.30 (m, 8H), 0.92 (t, *J* = 7.1 Hz, 3H); ^{13}CNMR (100Hz, CD_3OD) δ (ppm) 109.9, 83.8, 83.2, 74.4, 71.2, 70.2, 63.1, 55.2, 31.9, 29.9, 29.1, 26.0, 22.5, 13.3; DART HRMS m/z calcd for $C_{15}H_{30}O_7$ $[M+H]^+$ 323.2070, found 323.2082.

3,4;5,7-diacetonide-1,6-anhydro-D-*glycero*-D-galactitol (11). To a solution of **8**[3] (0.0300 g, 0.11 mmol) in THF (0.5 mL) was added excess BH_3THF (0.4 mL, 1M) at 0 °C. After stirring 2h at 0 °C, excess of BH_3THF was quenched with H_2O (0.2 mL) at 0 °C. To this solution 3M NaOH (0.2 mL) and 30% H_2O_2 (0.5 mL) were added at 0 °C. After 1h the solution was diluted with Et_2O (30 mL) and washed with water (2×10 mL), brine (10 mL), dried over Na_2SO_4 and concentrated under reduced pressure. The resulting residue was purified by column chromatography by using 3:10 (Hex: EtOAc) as eluent to give **11** (0.0200 g, 65%) as a colorless oil. R_f0.25 (3:10 Hex: EtOAc). $[\alpha]_D$+20.03 (c = 0.004 in $CHCl_3$); ^{1}HNMR (400 MHz, $CDCl_3$) δ (ppm) 4.38 (dd, *J* = 7.7, 7.7 Hz, 1H), 4.12-4.08 (dd, *J* = 9.6, 7.5 Hz, 1H), 3.98-3.88 (m, 3H), 3.75 (ddd, *J* = 12.3, 9.7, 2.6 Hz, 1H), 3.64 (dd, *J*=12.1, 6.2 Hz, 1H), 3.31-3.22 (m, 2H), 2.46 (broad s, OH), 1.54 (s, 3H), 1.50 (s, 3H), 1.43 (s, 3H), 1.42 (s, 3H); ^{13}CNMR (100 MHz, $CDCl_3$) δ (ppm) 110.1, 99.5, 81.4, 79.6, 74.3, 73.5, 71.9, 71.2, 62.8, 28.0, 26.9, 25.1, 20.2; DART HRMS m/z calcd for $C_{13}H_{22}O_6$ $[M+H]^+$ 275.1496, found 275.1496.

2-*O*-*n*-heptyl-3,4;5,7-diacetonide-1,6-anhydro-D-*glycero*-D-galactitol (12). To **11** (0.0180 g, 0.065 mmol) in a flask were added DMF (1 mL) and TBAI (0.0023 g, 0.0065 mmol). After cooling the mixture to 0 °C, NaH (0.0062 g, 0.26 mmol) was added. After stirring at 0 °C for 0.5 h, heptylbromide (0.046 g, 0.26 mmol) was added. The mixture was allowed to warm to rt and stirred for an additional 12 h. The reaction was then quenched with CH_3OH at 0 °C and the resulting mixture was concentrated under reduced pressure. The residue was redissolved in EtOAc (20 mL), washed with water (2×10 mL) and brine (1x10 mL). The organic layer was dried over Na_2SO_4, concentrated and purified by column chromatography by using 5:1 Hex: EtOAc as eluent to give **12** (0.0200 g, 83%) as a colorless oil. R_f 0.2 (8:2 Hex: EtOAc). $[\alpha]_D$ +8.5600 (c 0.04 in $CHCl_3$); ^{1}HNMR (400 MHz, $CDCl_3$) δ (ppm) 4.26 (m, 2H), 3.94-3.87 (m, 2H), 3.76 (ddd, *J* = 8.4, 8.4, 4.1 Hz, 1H), 3.57-3.53(m, 3H), 3.46(ddd, *J* = 12.0, 9.6, 4.1 Hz, 1H), 3.27-3.21(m, 2H), 1.58(m, 2H), 1.52(s, 3H), 1.48(s, 3H), 1.43(s,3H), 1.41(s, 3H), 1.30(m, 8H), 0.89(t, *J* = 6.7 Hz, 3H).^{13}CNMR (100 MHz, $CDCl_3$) δ (ppm) 109.5, 99.3, 80.8, 79.6, 79.1, 74.3, 73.8, 71.3, 70.2, 62.7, 32.0, 30.1, 29.3, 27.9, 27.4, 26.2, 24.8, 22.8, 20.5, 14.3. DART HRMS m/z calcd for $C_{20}H_{37}O_6$ $[M+H]^+$ 373.2591, found 373.2596.

1,6-anhydro-D-*glycero*-D-galactitol (6). To a solution of **11** (0.018 g, 0.065 mmol) in EtOH/H_2O (4 mL, 1:3) was added Amberlite H^+ IR-120 resin (0.030 g) and heated at 110 °C for 4 h, cooled to rt and filtered through Celite. The filtrate was concentrated under reduced pressure to give compound **6** (0.011 g, 90%); R_f 0.2 (3:2 DCM: MeOH); ^{1}HNMR (400 MHz, CD_3OD) δ (ppm) 4.12-4.08 (m, 1H), 3.94-3.93 (dd, J= 6.2,6.2 Hz, 1H), 3.90-3.84 (m,2H), 3.75-3.70 (dd J=11.7, 2.8 Hz, 1H), 3.67-3.57 (m, 2H), 3.38-3.33(m, 2H); ^{13}CNMR (100 MHz, $CDCl_3$) δ (ppm) 85.4, 75.2, 74.7, 71.1, 70.5, 70.2, 63.5; DART HRMS m/z calcd for $C_{14}H_{28}O_6$ $[M+H]^+$ 195.0869, found 195.0873.

2-*O*-*n*-heptyl-1,6-anhydro-D-*glycero*-D-galactitol (7). To a solution of **12** (0.018 g, 0.048 mmol) in EtOH/H_2O (3 mL 1:3) was added Amberlite H^+ IR-120 resin (0.023 g) and the mixture was heated at 110 °C for 4h and then allowed to cool to rt. The mixture was filtered through celite with additional washings of MeOH (3 x 3 mL). The combined solvents were removed under reduced pressure and the residue was purified by column chromatography using 10:1 DCM: CH_3OH as eluent to give pure **7** (0.013 g, 93%) as a colorless oil. R_f 0.6 (DCM: CH_3OH 8:2). $[\alpha]_D$ +53.06 (c = 0.002 in CH_3OH). ^{1}HNMR (400 MHz, CD_3OD) δ (ppm) 4.17-4.12 (dd, *J* = 12.8, 5.1 Hz, 1H), 4.03-4.02 (dd, *J* = 5.7, 1.6 Hz, 1H), 3.88-3.85

(dd, J = 11.6, 2.8 Hz, 1H), 3.78-3.75 (dd, J = 11.6, 2.8 Hz, 1H), 3.65-3.51 (m, 4H), 3.45 (dd, J = 12.8, 6.5 Hz, 1H), 3.37 (ddd, J = 10.2, 7.4, 2.7 Hz, 1H), 1.61-1.54 (qnt, J = 6.9 Hz, 2H), 1.40-1.30 (m, 8H), 0.92 (t, J = 6.9 Hz, 3H); ^{13}CNMR (100 MHz, CD_3OD) δ (ppm) 84.2, 79.3, 74.7, 73.3, 70.5, 70.3, 69.3, 63.3, 31.8, 30.0, 29.1, 26.0, 22.5, 13.2. DART HRMS m/z calcd for $C_{14}H_{28}O_6$ $[M+H]^+$ 293.1965, found 293.1936.

Protein preparation. $FimH_{LD}$ from *E. coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a His_6-tag ($FimH_{LD}$-Th-His_6, 173 residues) following a previously published protocol.[3] Briefly, the clone containing the $FimH_{LD}$ construct was expressed in the protease-deficient *E. coli* HM 125 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 h, harvested by centrifugation for 20 min at 2000 *g* 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-His_6 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.

Cell-free competitive binding assay. To determine the affinity of FimH antagonists, a cell-free binding assay described previously was applied.[3] Microtiter plates (F96 MaxiSorp, Nunc) were coated with 100 µL/well of a 10 µg/mL solution of $FimH_{LD}$-Th-His_6 in assay buffer overnight at 4 °C. The coating solution was discarded, and the wells were blocked with 150 µL/well of 3% BSA in assay buffer for 2 h at 4 °C. After three washing steps with assay buffer (150 µL/well), a 4-fold serial dilution of the test compound (50 µL/well) in assay buffer containing 5% DMSO and streptavidin-peroxidase coupled biotinylated polyacrylamide (PAA) glycopolymers [Manα1–3(Manα1–6)Manβ1–4Glc*N*Acβ1–4Glc*N*Acβ-PAA-biotin, TM-PAA] (50 µL/well of a 0.5 µg/mL solution) was added. The plates were incubated for 3 h at 25 °C and 350 rpm and then carefully washed four times with 150 µL/well assay buffer. After the addition of 100 µL/well of the horseradish peroxidase substrate 2,2'-azino-di(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), the colorimetric reaction was allowed to develop for 4 min, then stopped by the addition of 2% aqueous oxalic acid before the optical density (OD) was measured at 415 nm on a microplate-reader (Spectramax 190, Molecular Devices, CA). The IC_{50} values of the compounds were calculated with Prism 5 (GraphPad Software, Inc., La Jolla, CA). The IC_{50} value defines the

molar concentration of the test compound that reduces the maximal specific binding of TM-PAA polymer to $FimH_{LD}$ by 50%.

Isothermal titration calorimetry. ITC experiments were performed at 15 °C, 25 °C and 37 °C using a VP-ITC (Malvern Instruments, Worcestershire, UK) with an injection volume between 6 and 10 µl, a reference power of 10 □cal/sec, a stirring speed of 307 rpm, high feedback, a spacing time of 500 sec to 600 sec and a filter period of 2 sec. Preceding the measurements, $FimH_{LD}$-Th-His_6 was dialyzed against 10 mM HEPES pH 7.4 containing 150 mM NaCl. The ligands were dissolved in the same buffer. Protein concentration was determined by NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, MA, USA) using an extinction coefficient of 24'180 M^{-1} cm^{-1}.[4] The concentrations used in every ITC experiment are given in Table S1. All compounds (**2**, **4** and **7**) were titrated directly and two independent experiments evaluated the consistency of the measurements. Baseline adjustment and peak integration were carried out using Origin 7.0 as described by the manufacturer (OriginLab, Northampton, MA, USA). The parameters N (stoichiometry), K_A (association constant) and $\Delta H°_{obs}$ (change in enthalpy) are measured by ITC. The three-parameter nonlinear least-square fitting and the calculation of 95% confidence intervals were determined by performing a global fit analysis of multiple ITC experiments by SEDPHAT software.[5] $\Delta G°_{obs}$ (free energy of binding) and $\Delta S°_{obs}$ (change in entropy) were calculated as $\Delta G°_{obs} = \Delta H°_{obs} - T\Delta S°_{obs} = -RT\ln K_A$, with T being the absolute temperature (in K) and R being the universal gas constant (8.314 J/mol K).

NMR experiments. All NMR experiments were performed on a Bruker Avance III 600 MHz NMR spectrometer equipped with a 5 mm TXI RT probe head, or on a Bruker Avance III 500 MHz NMR spectrometer equipped with a BBO RT probe head at a temperature of 298 K. A backbone resonance assignment of $FimH_{LD}$ was available from previous publications.[6-7] Compounds **2** and **7** were dissolved in D_2O at a concentration of 100 mM and 10 mM, respectively. ^1H,^{15}N-HSQC experiments were performed with samples containing ca. 120 µM uniformly ^{15}N-labeled $FimH_{LD}$-Th-His_6 in absence of antagonists ("apo") or in presence of 180 µM compound **2** or 250 µM compound **7** in 20 mM phosphate buffer pH 7 with 7% D_2O. 0.1 mM TSP-*d4* (3-(trimethylsilyl)-2,2',3,3'-tetradeuteropropionic acid, Armar Chemicals, Switzerland) was added as internal reference. Spectra were acquired and processed with Topspin 2.1 (Bruker BioSpin, Switzerland) and analyzed with CcpNmr Analysis (version 2.2).[8] As both compounds bound in slow exchange to $FimH_{LD}$, backbone

amide signals of the bound state were assigned from chemical shift proximity. Combined chemical shift changes of backbone amide signals were calculated as

$$\Delta\delta_{AV}=\sqrt{\left(\Delta\delta^{1}H^{N}\right)^{2}+\left(0.2\Delta\delta^{15}N\right)^{2}}. \quad \text{[ref. 9]}$$

For coupling constant analysis of compound **7**, an NMR sample of **7** at a concentration of ca. 5 mM was prepared in D_2O. A 1D ^{1}H NMR spectrum was measured on a Bruker Avance III 900 MHz NMR spectrometer equipped with a TCI cryogenic probe at 298 K. Coupling constant analysis was performed with Topspin 3.2 (Bruker BioSpin, Switzerland).

MD simulations. The crystal structure of **2** and **7** were used as input structures for 4.8 ns molecule dynamics simulations using Desmond[10-13] and the OPLS 2005 force field. Default parameters were applied unless stated otherwise. TIP3P was selected as water model and a physiological salt concentration (0.15 M) was added. An energy barrier of 5 kcal/mol restricted backbone movement. The Simulation Interaction Diagram, implemented in Schrödinger Release 2015-4, was used to analyze the obtained trajectories. Metadynamics simulations were carried out for the ligands in explicit TIP3P water and default parameters (Gaussian height: 0.03 kcal/mol; interval 0.09 ps) for 4.8 ns using Desmond.[10] The monitored variables for compound **2** were defined as the angle between O1-C1-C4 and the dihedral torsion of O1-C1-O5-C5; for compound **7** the corresponding angle O2-C2-C5 and dihedral torsion O2-C2-C1-O6 were defined.

Ab initio calculations. To compare the van-der-Waals interactions of the tyrosine-gate and different ligands bearing an *n*-heptyl aglycone the protein-ligand complexes were subjected to energy calculations using the density functional theory (DFT) with the empirical corrected B3LYP-MM functional, which has been specifically parameterized to describe non-covalent interactions well, and the cc-pVDZ++ basis-set in the gas-phase as implemented in Jaguar 9.0.[14-16]

FimH co-crystallization and structure refinement. For crystallization, $FimH_{LD}$ (residues 1-158) at a final concentration of 10 mg/mL (ca. 0.7 mM) with a threefold molar excess of ligand **7** (2.5 mM) in 20 mM HEPES buffer pH 7.4 was used. Crystals were grown in sitting-drop vapor diffusion at 4°C in 1.5 M $(NH_4)_2SO_4$, 0.1 M BisTrisPropane pH 7 and 2% PEG3350. Crystals appeared after 3 months and were cryopreserved by addition of 20 % glycerol (v/v) and flash-cooled with liquid nitrogen. Data was collected at the PX beamline of

the Swiss Light Source (Paul Scherrer Institute, Switzerland) and indexed, integrated and scaled with XDS[17] The structure was solved by molecular replacement with PHASER[18] using the $FimH_{LD}$-*n*-butyl α-D-mannopyranoside complex (PDB code 1UWF) as search model. The structure was built using the COOT software[19] and periodically refined with the PHENIX software.[20] Geometric restraints for the ligands were generated with PRODRG.[21] Molprobity[22] was used for validation and proton addition for distance calculation. The structure was deposited in the Protein Data Bank with PDB code 5CGB.

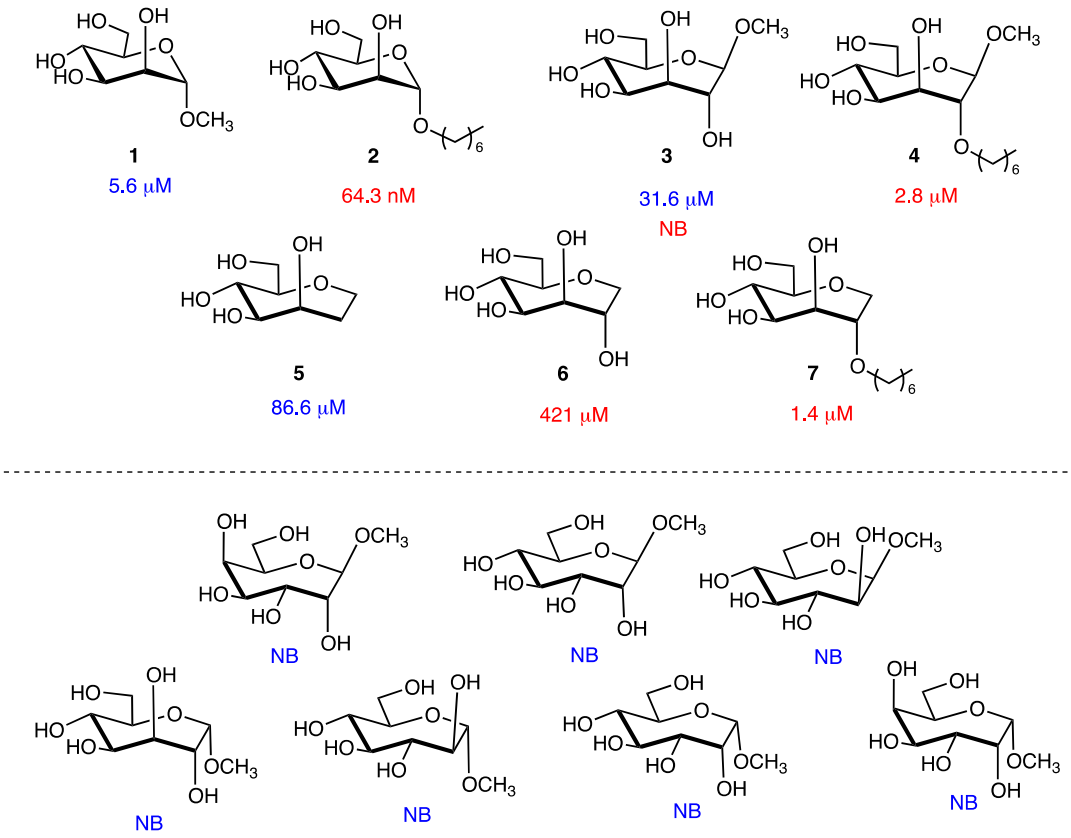


Figure S1. Septanoside and septanose derivatives evaluated for binding to $FimH_{LD}$ in a cell-free competitive binding assay.[1,2]

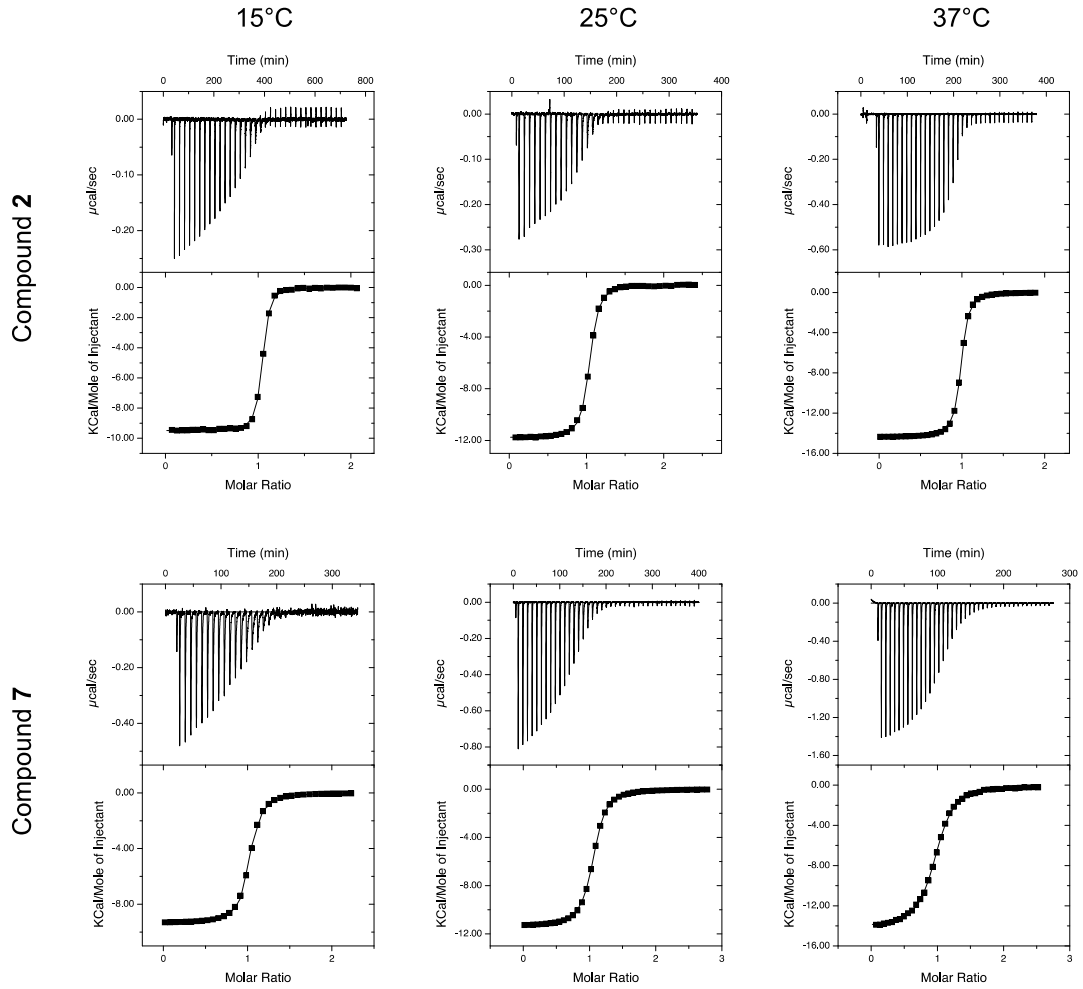


Figure S2. Calorimetric titrations of compounds **2** (top) and **7** (bottom) binding to $FimH_{LD}$ at 15°C (left), 25°C (center) and 37°C (right).

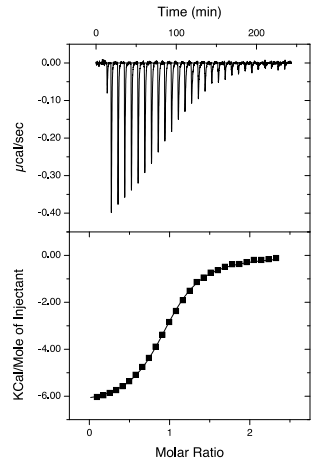


Figure S3. Calorimetric titration of compound **4** binding to $FimH_{LD}$ at 25°C.

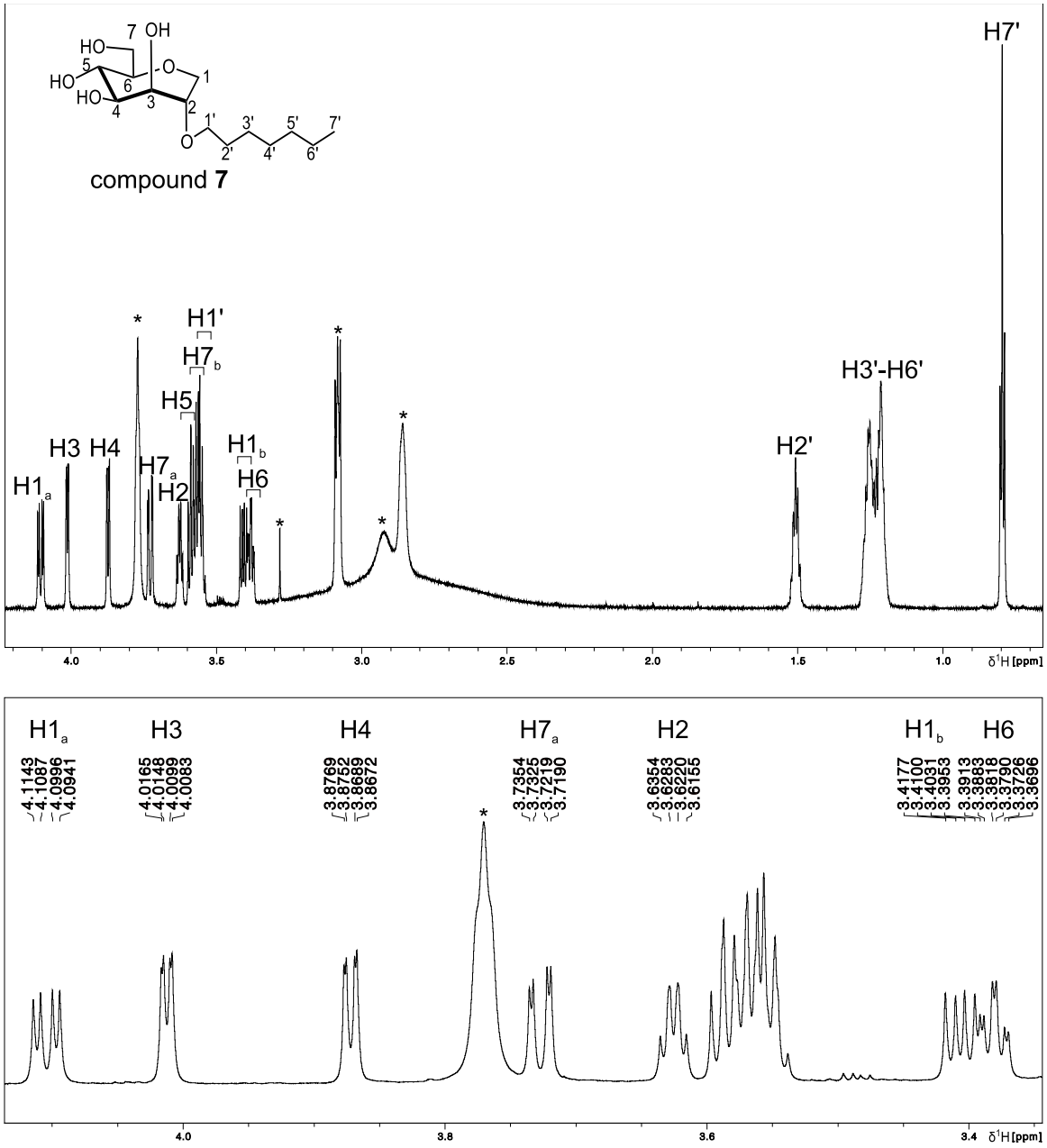


Figure S4. 1D ^{1}H NMR spectrum of compound **7** (ca. 5 mM in D_2O) measured at 900 MHz and a temperature of 298 K. The extracted coupling constants are shown in Table S2. Contaminations in the sample are indicated by asterisks.

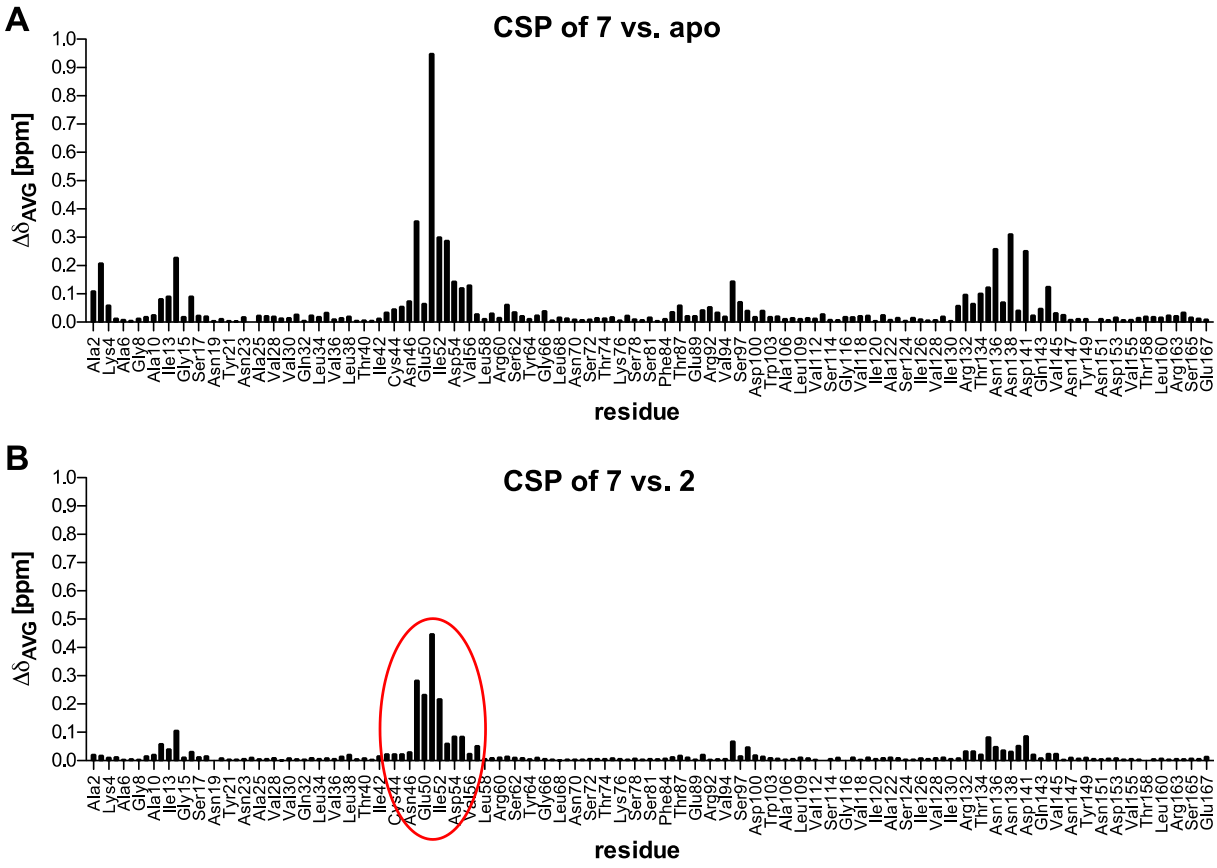


Figure S5. Chemical shift perturbation (CSP) of $FimH_{LD}$ upon binding to compound **7**. A: Absolute combined CSP effects ($\Delta\delta_{AVG}$ [ppm]) of $FimH_{LD}$ backbone amide resonances in presence of **7** relative to the apo protein. B: Absolute combined CSP effects ($\Delta\delta_{AVG}$ [ppm]) of $FimH_{LD}$ in complex with **7** relative to the complex with *n*-heptyl α-D-mannopyranoside **2**; deviations of $\Delta\delta_{AVG}$ values in the binding loop containing Tyr48 are marked by a red oval and are due to slightly different side chain torsion angles of Tyr48 resulting in differential aromatic ring current effects of the nearby loop residues.

Table S1. Temperature dependent thermodynamic quantities for ligands **2**, **4** and **7** to $FimH_{LD}$ obtained from isothermal titration calorimetry (ITC). The lines in bold letters represent the global fits of two measurements. All experiments were carried out in 10 mM HEPES buffer adjusted to pH 7.4 containing 150 mM NaCl.

Ligand	Temp	Titration	Ligand [µM]	Protein [µM]	K_D [nM]	$\Delta G°_{obs}$ [kJ/mol]	$\Delta H°_{obs}$ [kJ/mol]	$-T\Delta S°_{obs}$ [kJ/mol]	*N*	c-value
2	15°C	direct	150	15.0	19.4 (17.1 – 21.9)	-42.5	-40.8 (-40.5 - -41.1)	-1.8	0.98	774
2	15°C	direct	150	14.5	21.6 (17.7 – 25.9)	-42.3	-41.0 (-40.5 - -41.5)	-1.3	1.01	671
2	**15°C**				**20.6 (17.5 – 23.6)**	**-42.4**	**-40.9 (-40.6 - -41.3)**	**-1.5**	**1.00 ± 0.02**	**722 ± 73**
2	25°C	direct	100	10.0	29.1 (25.7 – 32.9)	-43.0	-49.8 (-49.4 - -50.3)	6.8	0.98	343
2	25°C	direct	100	8.6	29.0(26.7 – 31.4)	-43.0	-50.9 (-50.5 - -51.2)	7.8	1.01	296
2	**25°C**				**28.9 (25.8 – 32.3)**	**-43.0**	**-50.3** (-50.2 - -50.7)	**7.3**	**1.00 ± 0.02**	**320 ± 34**
2	37°C	direct	150	16.4	34.4 (30.4 – 38.8)	-44.3	-61.7 (-61.1 - -62.3)	17.4	0.97	478
2	37°C	direct	150	16.4	38.6 (33.4 – 44.4)	-44.0	-60.1 (-59.5 - -60.8)	16.1	0.96	423
2	**37°C**				**36.2 (30.8 – 42.6)**	**-44.2**	**-60.9 (-60.2 - -61.6)**	**16.7**	**0.96 ± 0.01**	**451 ± 38**
4	25°C	direct	430	32.6	2146.8 (2026.7 – 2274.0)	-32.4	-27.6 (-27.2 - -28.0)	-4.8	0.98	15
4	25°C	direct	400	35.2	2287.7 (2135.0 – 2451.3)	-32.2	-28.3 (-27.9 - -28.7)	-3.9	0.97	15
4	**25°C**				**2203.9 (2075.9 – 2339.9)**	**-32.3**	**-27.9 (-27.5 - -28.3)**	**-4.4**	**0.98 ± 0.01**	**15 ± 0**
7	15°C	direct	450	36.0	186.1 (162.1 – 212.7)	-38.4	-38.9 (-38.4 - -39.4)	0.5	1.00	193
7	15°C	direct	450	35.0	154.4 (137.8 – 174.3)	-38.9	-40.0 (-39.6 - -40.4)	1.1	1.03	225
7	**15°C**				**168.4 (150.5 – 188.1)**	**-38.7**	**-40.0 (-39.6 - -40.4)**	**1.3**	**1.02 ± 0.02**	**209 ± 22**
7	25°C	direct	390	29.1	263.7 (249.5 – 278.0)	-37.6	-48.9 (-48.6 - -49.2)	11.4	1.04	111
7	25°C	direct	245	30.2	275.5 (254.2 – 297.9)	-37.4	-50.1 (-49.5 - -50.6)	12.6	0.97	110
7	**25°C**				**264.2 (245.4 – 284.4)**	**-37.5**	**-49.4 (-48.9 - -49.8)**	**11.8**	**1.00 ± 0.04**	**110 ± 1**
7	37°C	direct	450	28.2	575.3 (541.9 – 609.4)	-37.1	-59.3 (-58.8 - -59.7)	22.2	1.03	49
7	37°C	direct	450	37.0	750.9 (688.0 – 817.7)	-36.4	-60.4 (-59.7 - -61.2)	24.1	0.96	49
7	**37°C**				**660.8 (606.9 – 719.6)**	**-36.7**	**-59.9 (-59.2 - -60.5)**	**23.2**	**1.00 ± 0.05**	**49 ± 0**

Table S2. Coupling constant analysis of compound **7**: ^{2}J and ^{3}J coupling constants of compound **7** extracted from a 1D ^{1}H NMR spectrum at 900 MHz.

NMR coupling constants of compound **7** (5 mM solved in D_2O) were determined from a 900 MHz NMR spectrum (Table S2, Figure S4). The measured coupling constants were compared to the expected coupling constants for the two main conformations of **7** with the *n*-heptyl aglycone either in the axial or equatorial conformation. The informative dihedral angles (i.e. with drastically different values in the two main conformations) of the lowest energy conformations from metadynamics simulations are H2-C2-C3-H3 and H3-C3-C4-H4 (ca. 45° and 63° in the axial conformation; ca. 160° and 31° in the equatorial conformation). The expected vicinal coupling constants (medium $^{3}J_{2,3}$ and small $^{3}J_{3,4}$ for axial conformation; very large $^{3}J_{2,3}$ and large $^{3}J_{3,4}$ for equatorial conformation)[23] indicate that only the axial conformation is in qualitative agreement with the measured coupling constants.

Coupling constant	Value
$^{2}J_{1a,1b}$	13.2
$^{3}J_{1a,2}$	5.0
$^{3}J_{1b,2}$	7.0
$^{3}J_{2,3}$	5.9
$^{3}J_{3,4}$	1.4
$^{3}J_{4,5}$	7.3
$^{3}J_{5,6}$	8.4
$^{3}J_{6,7a}$	2.6
$^{3}J_{6,7b}$	8.4
$^{2}J_{7a,7b}$	12.2

Table S3. MD simulations: H-bond occupancy of analyzed MD simulations, subdivided by ligand-protein interactions.

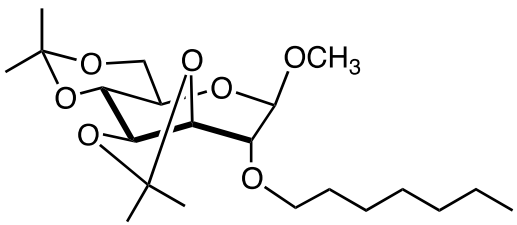
Two independent 4.8 ns molecular dynamics simulations were run for the two crystal structures of compound **2** and **7** (PDB codes: 4BUQ and 5CGB, respectively). Trajectories were analyzed for hydrogen bonds between ligand and protein and summarized in Table S3.

				H-Bond Occupancy	
	Donor	⋯	**Acceptor**	**5CBG**	**4BUQ**
2-OH	Phe1	⋯	OH	0.98	0.99
	OH	⋯	Water	0.73	0.67
3-OH	Gln133	⋯	OH	0.99	0.99
	OH	⋯	Asp140	0.99	1.00
4-OH	Asn135	⋯	OH	0.97	0.98
	OH	⋯	Asp54	1.00	1.00
6-OH	Phe1	⋯	OH	1.00	0.99
	Asp47	⋯	OH	0.99	0.99
	OH	⋯	Asp54	0.99	1.00

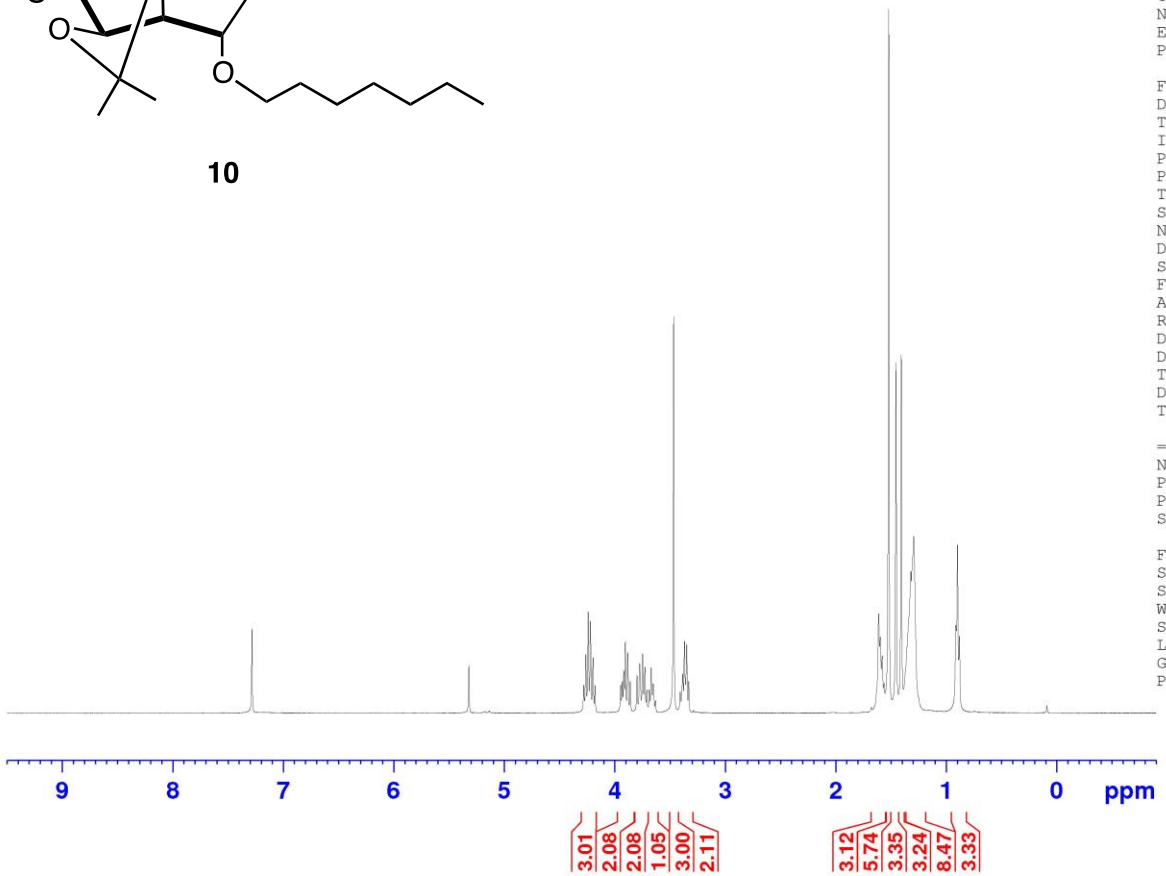
Table S4. Statistics on diffraction data and structure refinement of the $FimH_{LD}$–**7** complex.

Data set	**$FimH_{LD}$/7**[a]
PDB ID	5CGB
wavelength	1.00000
space group	$P2_12_12_1$
unit cell	63.58 68.55 95.78 90 90 90
resolution (Å)	47.89-1.60 (1.69-1.60)
total reflections	456,089
unique reflections	55,686
multiplicity	8.1 (8.0)
completeness (%)	99.5 (98.8)
mean I/sigma(I)	19.7 (1.6)
Wilson B-factor	17.2
R-meas	0.076 (0.76)
$CC_{1/2}$	0.99(0.93)
R-work	0.160 (0.272)
R-free	0.179 (0.282)
number of atoms	5315
macromolecules	4739
water	489
Protein residues	316
RMS(bonds)	0.006
RMS(angles)	1.2
Ramachandran favored (%)	97.5
Ramachandran outliers (%)	0.0
clash score	1.3
Average B-factor	
macromolecules	22.1
ligands	28.9
solvent	38.7

[a]Values in parentheses are for the highest resolution shell.

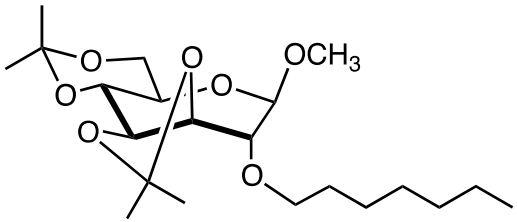


10



BRUKER

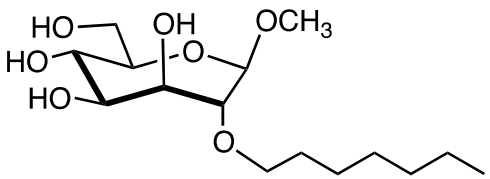
```
Current Data Parameters
NAME                  FimH_10
EXPNO                       1
PROCNO                      1

F2 - Acquisition Parameters
Date_              20120511
Time                  15.19
INSTRUM               spect
PROBHD    5 mm PABBO BB/
PULPROG                zg30
TD                    65536
SOLVENT               CDCl3
NS                       16
DS                        2
SWH                8012.820 Hz
FIDRES             0.122266 Hz
AQ                 4.0894465 sec
RG                       101
DW                    62.400 usec
DE                     6.50 usec
TE                    298.0 K
D1              1.00000000 sec
TD0                         1

======== CHANNEL f1 ========
NUC1                       1H
P1                    15.00 usec
PLW1           12.00000000 W
SFO1           400.1464711 MHz

F2 - Processing parameters
SI                    65536
SF             400.1440000 MHz
WDW                       EM
SSB       0
LB                     0.30 Hz
GB        0
PC                     1.00
```



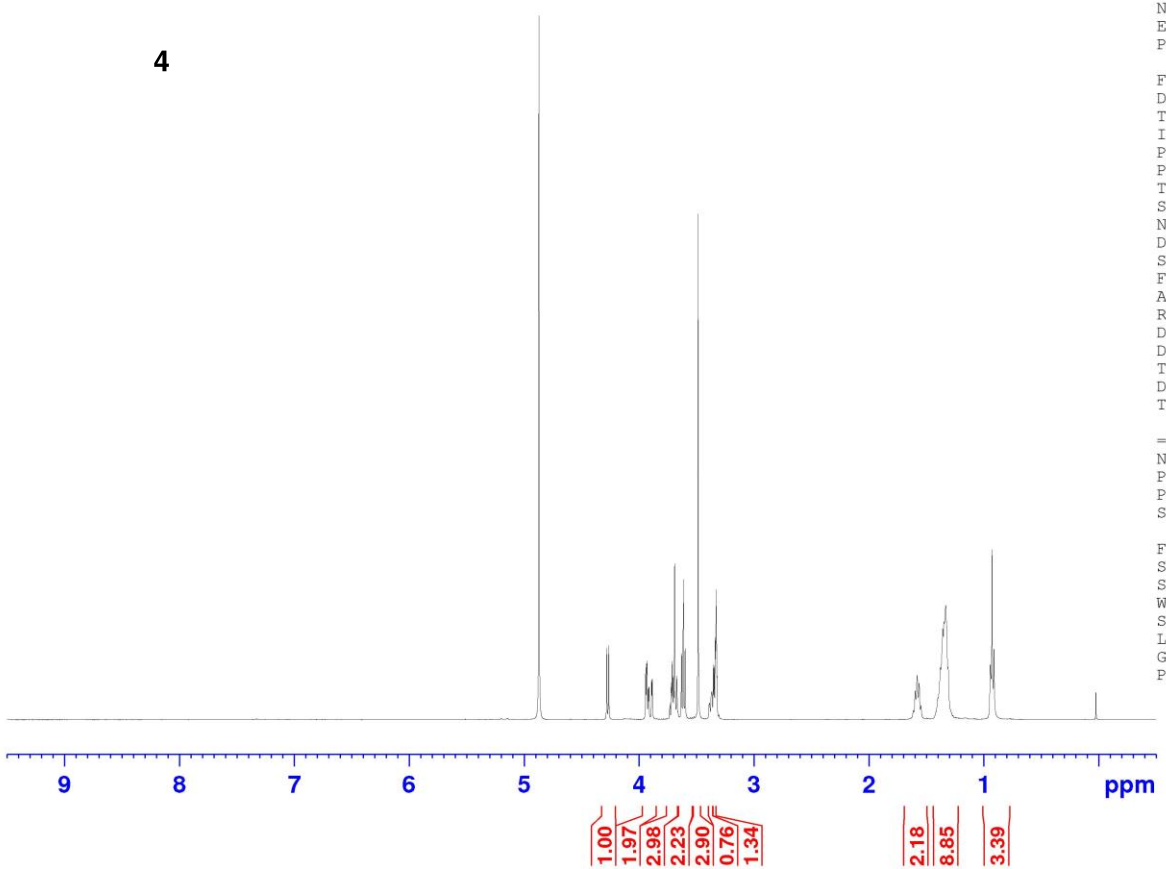
10

BRUKER

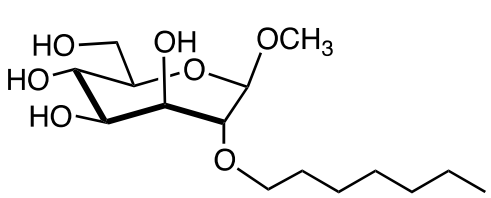
```
Current Data Parameters
NAME                  FimH_10
EXPNO                       2
PROCNO                      1

F2 - Acquisition Parameters
Date_              20120511
Time                  15.26
INSTRUM               spect
PROBHD    5 mm PABBO BB/
PULPROG             zgpg30
TD                   65536
SOLVENT              CDCl3
NS                     221
DS                       4
SWH               24038.461 Hz
FIDRES             0.366798 Hz
AQ                 1.3631488 sec
RG                     203
DW                  20.800 usec
DE                   6.50 usec
TE                  298.0 K
D1                2.00000000 sec
d11               0.03000000 sec
DELTA             1.89999998 sec
TD0                        1
SFO1            100.6263500 MHz
NUC1                    13C
P1                  10.00 usec
PLW1             50.00000000 W
SFO2            400.1456006 MHz
NUC2                    1H
CPDPRG[2             waltz16
PCPD2                90.00 usec
PLW2             12.00000000 W
PLW12             0.33333001 W
PLW13             0.27000001 W

F2 - Processing parameters
SI                     32768
SF              100.6162656 MHz
WDW                       EM
SSB       0
LB                   1.00 Hz
GB        0
PC                   1.40
```



4



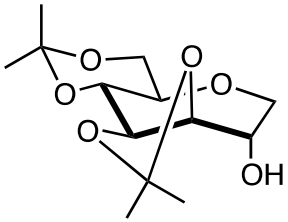

```
Current Data Parameters
NAME                  FimH_4
EXPNO                      1
PROCNO                     1

F2 - Acquisition Parameters
Date_              20130326
Time                 18.49
INSTRUM              spect
PROBHD    5 mm PABBO BB/
PULPROG              zg30
TD                  48076
SOLVENT               MeOD
NS                      16
DS                       2
SWH                8012.820 Hz
FIDRES             0.166670 Hz
AQ                 2.9999423 sec
RG                      64
DW                    62.400 usec
DE                     6.50 usec
TE                    298.0 K
D1                 2.00000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                      1H
P1                     15.20 usec
PLW1            12.00000000 W
SFO1            400.1464009 MHz

F2 - Processing parameters
SI                    131072
SF              400.1440000 MHz
WDW                       EM
SSB       0
LB                     0.50 Hz
GB        0
PC                     1.00
```



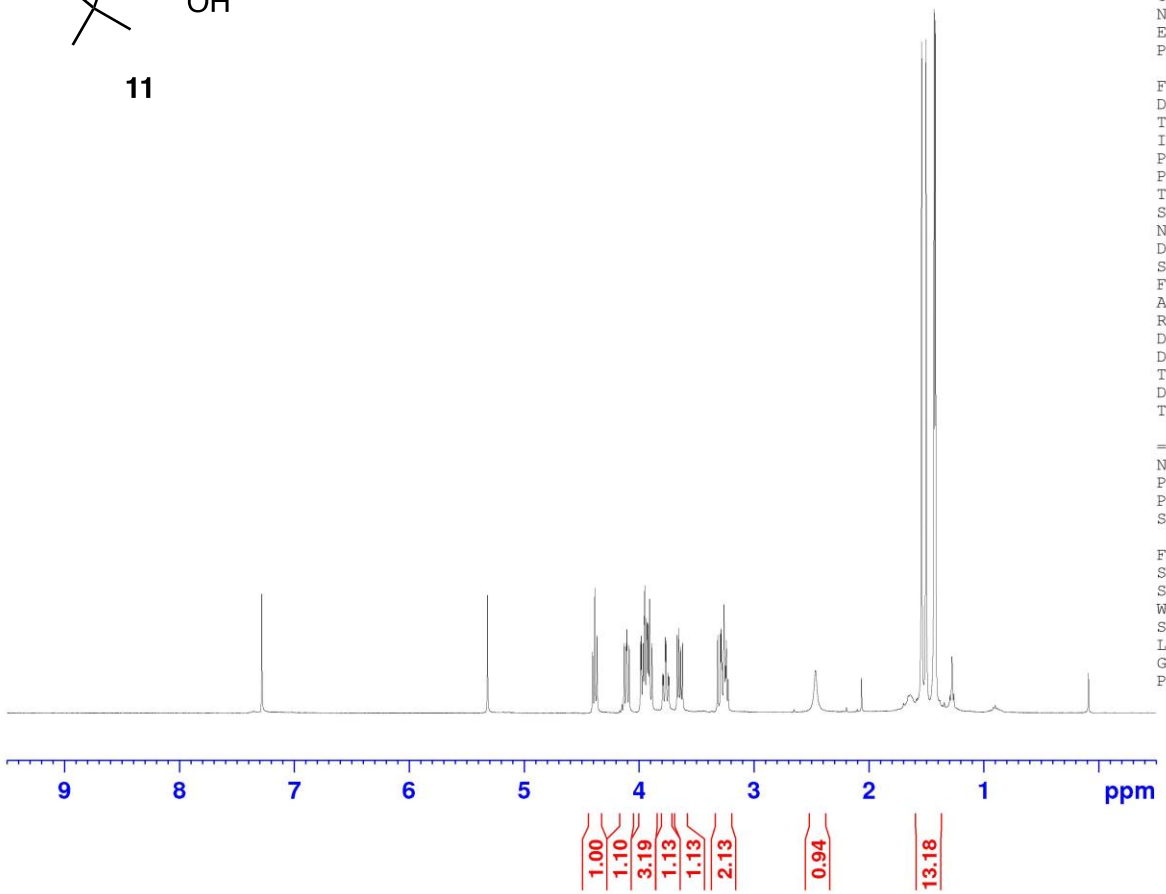
4

BRUKER

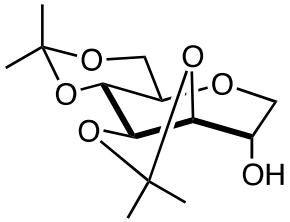
```
Current Data Parameters
NAME                 FimH_4
EXPNO                     2
PROCNO                    1

F2 - Acquisition Parameters
Date_              20130326
Time                 18.51
INSTRUM              spect
PROBHD    5 mm PABBO BB/
PULPROG              zgpg30
TD                    43698
SOLVENT                MeOD
NS                      109
DS                        4
SWH              24038.461 Hz
FIDRES            0.550104 Hz
AQ                0.9089184 sec
RG                      114
DW                   20.800 usec
DE                    6.50 usec
TE                   298.1 K
D1              1.50000000 sec
D11             0.03000000 sec
TD0                       1

======== CHANNEL f1 ========
NUC1                    13C
P1                   10.00 usec
PLW1           50.00299835 W
SFO1          100.6263500 MHz

======== CHANNEL f2 ========
CPDPRG[2            waltz16
NUC2                    1H
PCPD2                90.00 usec
PLW2           12.00000000 W
PLW12           0.33333001 W
PLW13           0.27000001 W
SFO2          400.1456006 MHz

F2 - Processing parameters
SI                   131072
SF            100.6162656 MHz
WDW                     EM
SSB          0
LB                    2.50 Hz
GB           0
PC                    1.40
```



11



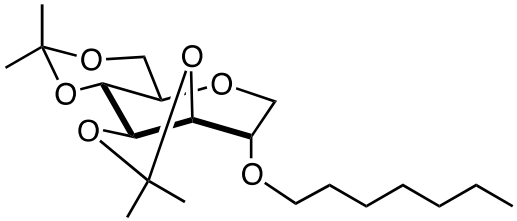

```
Current Data Parameters
NAME                  FimH_11
EXPNO                       1
PROCNO                      1

F2 - Acquisition Parameters
Date_                20120906
Time                    15.57
INSTRUM                 spect
PROBHD   5 mm PABBO BB/
PULPROG                 zg30
TD                     52882
SOLVENT                CDCl3
NS                        16
DS                         2
SWH                 8012.820 Hz
FIDRES              0.151523 Hz
AQ                3.2998369 sec
RG                      101
DW                   62.400 usec
DE                    6.50 usec
TE                   300.0 K
D1                1.00000000 sec
TD0                         1

======== CHANNEL f1 ========
NUC1                       1H
P1                      15.22 usec
PLW1             12.00000000 W
SFO1             400.1463943 MHz

F2 - Processing parameters
SI                     262144
SF               400.1440000 MHz
WDW                        EM
SSB       0
LB                      0.30 Hz
GB        0
PC                      1.00
```



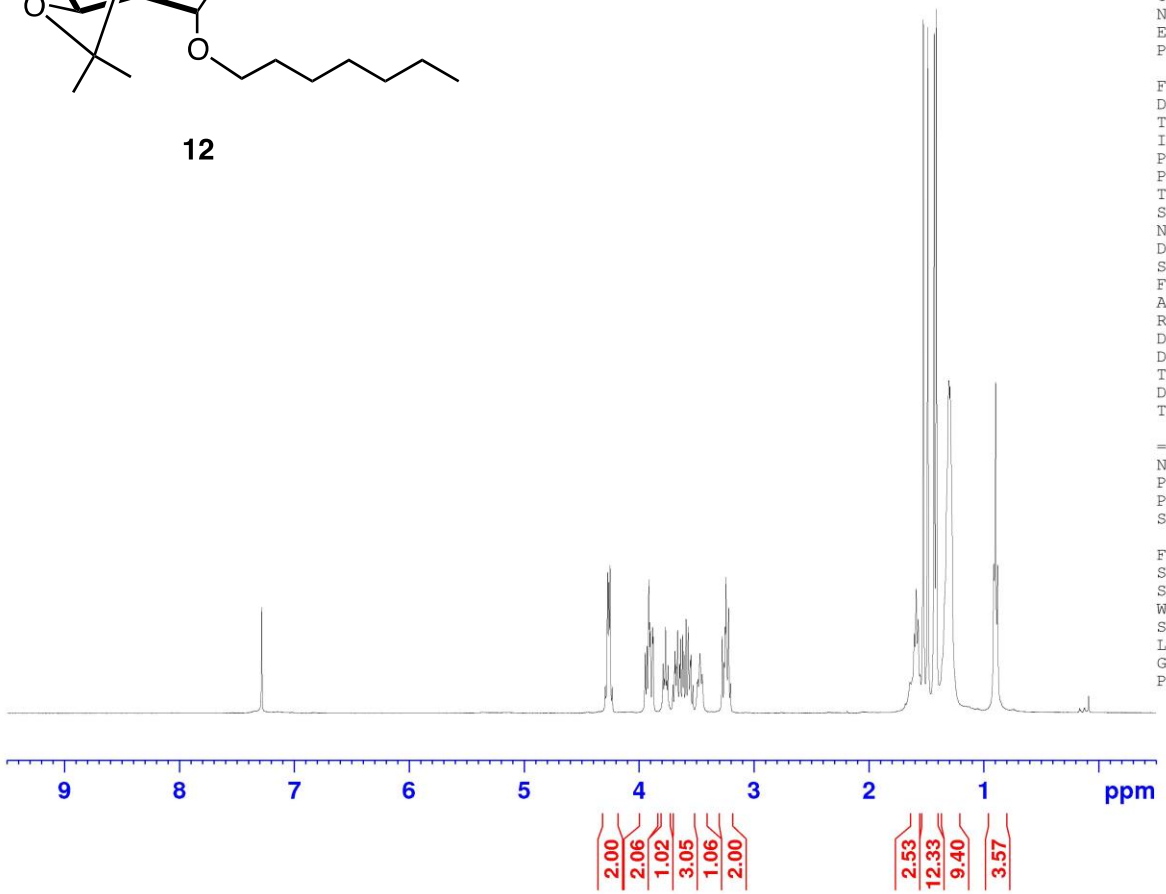
11

BRUKER

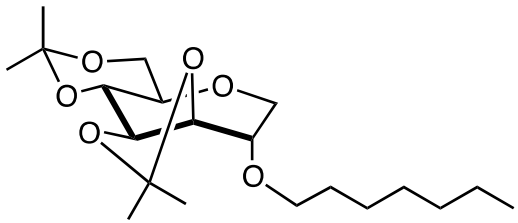
```
Current Data Parameters
NAME                  FimH_11
EXPNO                       2
PROCNO                      1

F2 - Acquisition Parameters
Date_              20120906
Time                 16.01
INSTRUM              spect
PROBHD    5 mm PABBO BB/
PULPROG            zgpg30
TD                  43698
SOLVENT             CDCl3
NS                    134
DS                      4
SWH              24038.461 Hz
FIDRES            0.550104 Hz
AQ                0.9089184 sec
RG                    203
DW                  20.800 usec
DE                   6.50 usec
TE                  300.0 K
D1               1.50000000 sec
d11              0.03000000 sec
DELTA            1.39999998 sec
TD0                     1
SFO1           100.6263500 MHz
NUC1                  13C
P1                  10.00 usec
PLW1           50.00000000 W
SFO2           400.1456006 MHz
NUC2                   1H
CPDPRG[2           waltz16
PCPD2                90.00 usec
PLW2           12.00000000 W
PLW12           0.33333001 W
PLW13           0.27000001 W

F2 - Processing parameters
SI                  131072
SF             100.6162656 MHz
WDW                    EM
SSB       0
LB                   2.50 Hz
GB        0
PC                   1.40
```



12



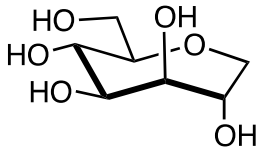

```
Current Data Parameters
NAME                  FimH_12
EXPNO                       1
PROCNO                      1

F2 - Acquisition Parameters
Date_                20120921
Time                    12.46
INSTRUM                 spect
PROBHD   5 mm PABBO BB/
PULPROG                  zg30
TD                      52882
SOLVENT                 CDCl3
NS                         16
DS                          2
SWH                  8012.820 Hz
FIDRES               0.151523 Hz
AQ                   3.2998369 sec
RG                         64
DW                     62.400 usec
DE                      6.50 usec
TE                     300.0 K
D1                  1.00000000 sec
TD0                         1

======== CHANNEL f1 ========
NUC1                       1H
P1                     15.22 usec
PLW1               12.00000000 W
SFO1               400.1463943 MHz

F2 - Processing parameters
SI                      262144
SF                 400.1440000 MHz
WDW                        EM
SSB          0
LB                      0.30 Hz
GB           0
PC                      1.00
```



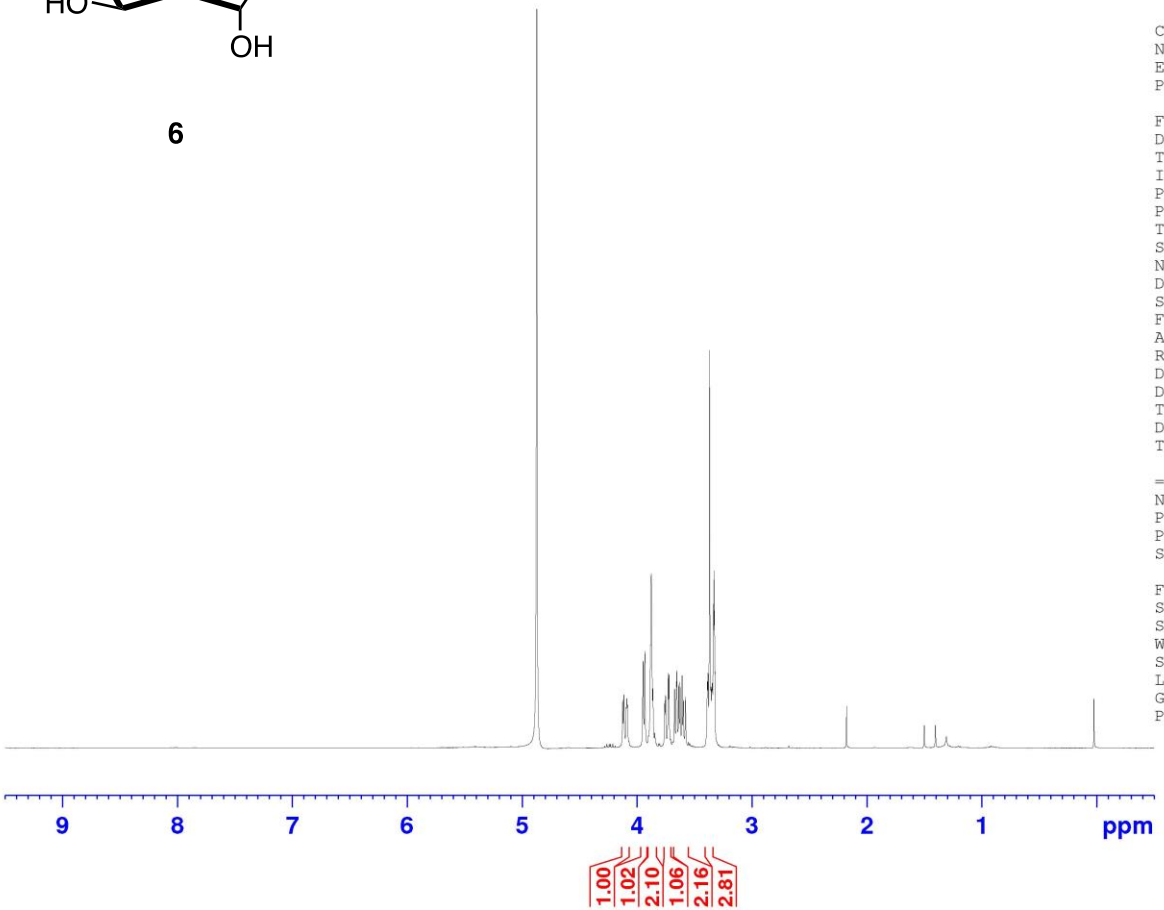
12

BRUKER

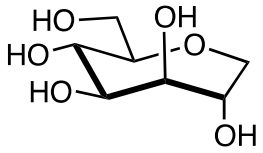
```
Current Data Parameters
NAME                  FimH_12
EXPNO                       2
PROCNO                      1

F2 - Acquisition Parameters
Date_              20120921
Time                  12.48
INSTRUM               spect
PROBHD    5 mm PABBO BB/
PULPROG             zgpg30
TD                   43698
SOLVENT              CDCl3
NS                      18
DS                       4
SWH              24038.461 Hz
FIDRES            0.550104 Hz
AQ                0.9089184 sec
RG                      203
DW                 20.800 usec
DE                  6.50 usec
TE                  300.1 K
D1              1.50000000 sec
d11             0.03000000 sec
DELTA           1.39999998 sec
TD0                     1
SFO1           100.6263500 MHz
NUC1                  13C
P1                 10.00 usec
PLW1           50.00000000 W
SFO2           400.1456006 MHz
NUC2                   1H
CPDPRG[2           waltz16
PCPD2               90.00 usec
PLW2           12.00000000 W
PLW12           0.33333001 W
PLW13           0.27000001 W

F2 - Processing parameters
SI                   131072
SF             100.6162656 MHz
WDW                    EM
SSB       0
LB                  2.50 Hz
GB        0
PC                  1.40
```



6



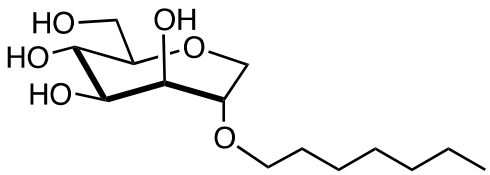

```
Current Data Parameters
NAME                  FimH_6
EXPNO                      1
PROCNO                     1

F2 - Acquisition Parameters
Date_              20130326
Time                 18.41
INSTRUM              spect
PROBHD    5 mm PABBO BB/
PULPROG              zg30
TD                  48076
SOLVENT              MeOD
NS                     16
DS                      2
SWH               8012.820 Hz
FIDRES               0.166670 Hz
AQ                2.9999423 sec
RG                     80.6
DW                   62.400 usec
DE                    6.50 usec
TE                   298.0 K
D1             2.00000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                      1H
P1                    15.20 usec
PLW1          12.00000000 W
SFO1          400.1464009 MHz

F2 - Processing parameters
SI                    131072
SF            400.1440000 MHz
WDW                       EM
SSB       0
LB                     0.50 Hz
GB        0
PC                     1.00
```



6

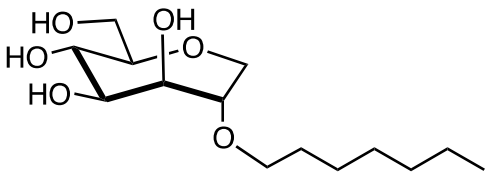
BRUKER

```
Current Data Parameters
NAME                  FimH_6
EXPNO                      2
PROCNO                     1

F2 - Acquisition Parameters
Date_              20130326
Time                  18.43
INSTRUM               spect
PROBHD    5 mm PABBO BB/
PULPROG              zgpg30
TD                    43698
SOLVENT                MeOD
NS                       58
DS                        4
SWH               24038.461 Hz
FIDRES             0.550104 Hz
AQ                 0.9089184 sec
RG                      114
DW                   20.800 usec
DE                    6.50 usec
TE                   298.0 K
D1                1.50000000 sec
D11               0.03000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                     13C
P1                     10.00 usec
PLW1             50.00299835 W
SFO1            100.6263500 MHz

======== CHANNEL f2 ========
CPDPRG[2              waltz16
NUC2                      1H
PCPD2                  90.00 usec
PLW2             12.00000000 W
PLW12             0.33333001 W
PLW13             0.27000001 W
SFO2            400.1456006 MHz

F2 - Processing parameters
SI                    131072
SF               100.6162656 MHz
WDW                       EM
SSB          0
LB                     2.50 Hz
GB           0
PC                     1.40
```



7


```
Current Data Parameters
NAME                  FimH_7
EXPNO                      1
PROCNO                     1

F2 - Acquisition Parameters
Date_              20130326
Time                 18.59
INSTRUM              spect
PROBHD    5 mm PABBO BB/
PULPROG              zg30
TD                  48076
SOLVENT              MeOD
NS                     16
DS                      2
SWH               8012.820 Hz
FIDRES            0.166670 Hz
AQ                2.9999423 sec
RG                     80.6
DW                  62.400 usec
DE                   6.50 usec
TE                  298.0 K
D1              2.00000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                      1H
P1                    15.20 usec
PLW1           12.00000000 W
SFO1           400.1464009 MHz

F2 - Processing parameters
SI                    131072
SF             400.1440000 MHz
WDW                       EM
SSB       0
LB                    0.50 Hz
GB        0
PC                    1.00
```



7

BRUKER

```
Current Data Parameters
NAME                  FimH_7
EXPNO                      2
PROCNO                     1

F2 - Acquisition Parameters
Date_              20130326
Time                 19.02
INSTRUM              spect
PROBHD    5 mm PABBO BB/
PULPROG            zgpg30
TD                  43698
SOLVENT               MeOD
NS                     130
DS                       4
SWH              24038.461 Hz
FIDRES            0.550104 Hz
AQ               0.9089184 sec
RG                     114
DW                20.800 usec
DE                 6.50 usec
TE                 298.0 K
D1              1.50000000 sec
D11             0.03000000 sec
TD0                      1

======== CHANNEL f1 ========
NUC1                   13C
P1                  10.00 usec
PLW1           50.00299835 W
SFO1          100.6263500 MHz

======== CHANNEL f2 ========
CPDPRG[2           waltz16
NUC2                   1H
PCPD2               90.00 usec
PLW2           12.00000000 W
PLW12           0.33333001 W
PLW13           0.27000001 W
SFO2          400.1456006 MHz

F2 - Processing parameters
SI                   131072
SF            100.6162656 MHz
WDW                    EM
SSB       0
LB                  2.50 Hz
GB        0
PC                  1.40
```

References

1. M. R. Duff, Jr.; W. S. Fyvie, S. D. Markad, A. E. Frankel, C. V. Kumar, J. A. Gascón, M. W. Peczuh, *Org. Biomol. Chem.* 2011, **9**, 154.
2. S. D. Markad, S. Xia, N. L. Snyder, B. Surana, M. D. Morton, C. M. Hadad, M. W. Peczuh, *J. Org. Chem.* 2008, **73**, 6341.
3. S. Rabbani, X. Jiang, O. Schwardt, B. Ernst, *Anal. Biochem.* 2010, **407**, 188.
4. H. Edelhoch, *Biochemistry* 1967, **6**, 1948.
5. J. C. Houtman, P. H. Brown, B. Bowden, H. Yamaguchi, E. Appella, L. E. Samelson, P. Schuck, *Protein Sci.* 2007, **16**, 30.
6. B. Fiege, S. Rabbani, R. C. Preston, R. P. Jakob, P. Zihlmann, O. Schwardt, X. Jiang, T. Maier, B. Ernst, B. *ChemBioChem* 2015, **16**, 1235.
7. S. Vanwetswinkel, A. N. Volkov, Y. G. J. Sterckx, A. Garcia-Pino, L. Buts, W. F. Vranken, J. Bouckaert, R. Roy, L. Wyns, N. A. J. van Nuland, *J. Med. Chem.* 2014, **57**, 1416.
8. 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.
9. M. Pellecchia, P. Sebbel, U. Hermanns, K. Wuthrich, R. Glockshuber, *Nat. Struct. Biol.* 1999, **6**, 336.
10. Schrödinger Suite 2015-4: Desmond Molecular Dynamics System, D. E. Shaw Research, Maestro-Desmond Interoperability Tools, Version 4.4 ed, Schrödinger, LLC, New York, NY, 2015.
11. Z. Guo, U. Mohanty, J. Noehre, T. K. Sawyer, W. Sherman, G. Krilov, *Chem. Biol. Drug Des.* 2010, **75**, 348.
12. D. Shivakumar, J. Williams, Y. Wu, W. Damm, J. Shelley, W. Sherman, *J. Chem. Theory Comput.* 2010, **6**, 1509.
13. J. K. Bowers, R. O. Dror, D. E. Shaw, *J. Chem. Phys.* 2006, **124**, 184109.
14. Schrödinger Suite 2015-4, Jaguar, Version 9.0, Schrödinger, LLC, New York, NY, 2015.
15. S. T. Schneebeli, A. D. Bochevarov, R. A. Friesner, *J. Chem. Theory Comput.* 2011, **7**, 658.

16. A. D. Bochevarov, E. Harder, T. F. Hughes, J. R. Greenwood, D. A. Braden, D. M. Philipp, D. Rinaldo, M. D. Halls, J. Zhang, R. A. Friesner, *Int. J. Quantum Chem.* 2013, **113**, 2110.

17. W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.* 2010, **66**, 125.

18. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* 2007, **40**, 658.

19. P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* 2004, **60**, 2126.

20. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. D Biol. Crystallogr.* 2010, **66**, 213.

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

22. V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. D Biol. Crystallogr.* 2010, **66**, 12.

23. Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870-2871.