

## Supporting Information

# **An $\alpha$ -helical peptidomimetic scaffold for dynamic combinatorial library formation**

Nathalie Busschaert,<sup>ab</sup> Sam Thompson<sup>\*bc</sup> and Andrew D. Hamilton<sup>\*ab</sup>

<sup>a</sup> Department of Chemistry, New York University, 100 Washington Square East, New York,  
NY 10003, USA. E-mail: andrew.hamilton@nyu.edu

<sup>b</sup> Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA,  
UK. E-mail: sam.thompson@chem.ox.ac.uk

<sup>c</sup> Chemistry, University of Southampton, Southampton, SO17 1BJ, UK.  
E-mail: sam.thompson@soton.ac.uk

## Table of contents

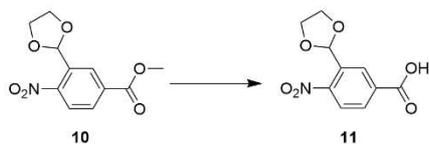
1. Synthesis and characterisation .....	S3
1.1. General Experimental.....	S3
1.2. Synthetic route to 3-(1,3-Dioxolan-2-yl)-4-nitrobenzoic acid ( <b>11</b> ) .....	S4
1.3. Synthetic route to methyl capped $\alpha$ -helix mimetics ( <b>1a-6a</b> ) .....	S5
1.4. Synthetic route to PEG capped $\alpha$ -helix mimetics ( <b>1b-9b</b> ).....	S10
2. 1D NMR spectra .....	S22
3. Conformational analysis .....	S54
3.1. Single crystal X-ray diffraction (solid state).....	S54
3.2. Computational Modelling .....	S61
3.3. Dilution studies (solution).....	S68
3.4. DMSO-d <sub>6</sub> and CDCl <sub>3</sub> comparison (solution).....	S70
3.5. Variable temperature NMR (solution) .....	S72
3.6. Deuterium exchange (solution) .....	S76
3.7. NOESY/ROESY NMR (solution).....	S81
4. Library synthesis .....	S108
5. References.....	S118

# 1. Synthesis and characterisation

## 1.1. General Experimental

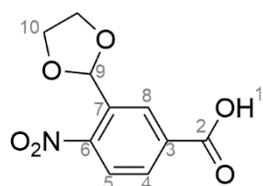
Reactions were carried out under a nitrogen or argon atmosphere in oven-dried glassware unless otherwise stated. Solvents and reagents were used directly as received from commercial suppliers. Anilines produced through the reduction of nitro groups were found to be unstable and were used immediately without purification. Flash column chromatography was carried out using Merck 60 silica gel. Column chromatography of imines was performed using Sigma-Aldrich aluminium oxide (activated, basic). Thin-layer chromatography was carried out using Merck Kieselgel 60 F254 (230-400 mesh) fluorescent treated silica, visualized under UV light (254 nm) or by staining with ninhydrin solution.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded using a Bruker 600 or 400 MHz spectrometer running TopSpin<sup>TM</sup> software and are quoted in ppm for measurement against residual solvent peaks as internal standards (i.e. DMSO- $d_6$  ( $\delta = 2.50$  ( $^1\text{H}$ ) and 39.51 ppm ( $^{13}\text{C}$ )) or  $\text{CDCl}_3$  ( $\delta = 7.27$  ( $^1\text{H}$ ) and 77.00 ppm ( $^{13}\text{C}$ )).  $^{13}\text{C}$  NMR spectra were always collected proton decoupled. ACD/Labs version 12.01 was used for processing and viewing NMR data. Chemical shifts ( $\delta$ ) are given in parts per million (ppm), and coupling constants ( $J$ ) are given in Hertz (Hz). The  $^1\text{H}$  NMR spectra are reported as follows:  $\delta$  / ppm (multiplicity, coupling constant  $J$  / Hz (where appropriate), number of protons, assignment). Multiplicity is abbreviated as follows: s = singlet, br = broad, d = doublet, t = triplet, m = multiplet. Peaks that could not be assigned in the  $^1\text{H}$  spectra due to the similarity of alkyl protons are indicated by 'PEG-H' or 'Ar-H' for aromatic protons. Compound names are those generated by ChemBioDraw<sup>TM</sup> Ultra version 13.0.2.3021 (CambridgeSoft) following IUPAC nomenclature. However, the NMR assignment numbering used is arbitrary and does not follow any particular convention. The  $^{13}\text{C}$  NMR spectra are reported in  $\delta$  / ppm. Where necessary or appropriate, two-dimensional (COSY, HSQC, HMBC, NOESY or ROESY) NMR experiments were used to assist the assignment of signals in the  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra. In some cases, complete assignment of spectra was not possible due to spectral coincidence; in these cases only a partial assignment is reported. IR spectra were recorded on a Bruker Tensor 27 FT-IR or Thermo Scientific Nicolet 6700 FT-IR spectrometer from a thin film deposited onto a diamond ATR module. Only selected maximum absorbances ( $\nu_{\text{max}}$ ) of the most intense peaks are reported ( $\text{cm}^{-1}$ ). Low-resolution mass spectra were recorded on a Waters LCT premier XE Micromass spectrometer (ESI). High-resolution mass spectra were recorded on a Bruker MicroTof mass spectrometer (ESI) by the internal service at the Department of Chemistry, University of Oxford. Compound **10** was synthesised from methyl 3-formyl-4-nitrobenzoate according to a previously reported method.<sup>1</sup>

## 1.2. Synthetic route to 3-(1,3-Dioxolan-2-yl)-4-nitrobenzoic acid (**11**)



Conditions: 1 M NaOH, THF, r.t., 5 h, 94% yield.

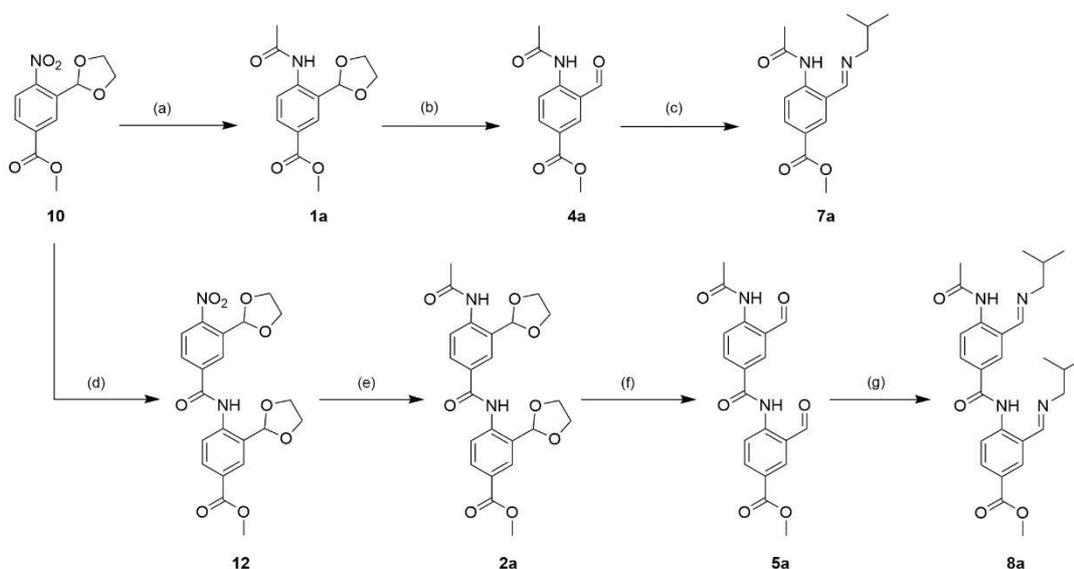
### 3-(1,3-Dioxolan-2-yl)-4-nitrobenzoic acid (**11**)



Carboxylic acid **10** (1.600 g, 6.32 mmol) was dissolved in THF (64 mL) and a 1 M NaOH solution was added (16 mL). The mixture was allowed to stir for 5 h. The organic solvent was removed *in vacuo* and 1 M HCl was added dropwise to the remaining aqueous solution until no more precipitate was

formed (~ pH 3). The obtained precipitate was filtered, washed with water and concentrated *in vacuo* to give *title compound 11* as a pale yellow solid (1.424 g, 5.95 mmol, 94% yield);  $\delta_{\text{H}}$  (400 MHz, DMSO- $d_6$ ) 13.72 (br. s., 1 H, *H1*), 8.23 (d, *J* 2.0, 1 H, *H8*), 8.14 (dd, *J* 8.3, 1.7, 1 H, *H4*), 8.05 (d, *J* 8.3, 1 H, *H5*), 6.31 (s, 1 H, *H9*), 3.88-4.05 (m, 4 H, *H10*);  $\delta_{\text{C}}$  (101 MHz, DMSO- $d_6$ ) 165.6, 150.8, 134.5, 132.6, 131.0, 128.2, 124.8, 98.5, 64.9; HRMS (ESI) for  $\text{C}_{10}\text{H}_8\text{NO}_6$   $[\text{M}-\text{H}]^-$   $m/z$  238.03571 (calc), 238.03557 (found); IR (neat) 2963 (br), 2896 (br), 2663 (br), 2541 (br), 1700, 1527, 1293, 1254, 1200, 1103, 1063, 969, 936, 914, 846, 746.

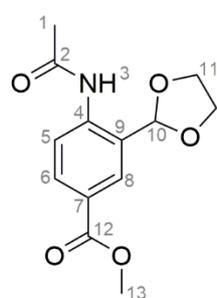
### 1.3. Synthetic route to methyl capped $\alpha$ -helix mimetics (**1a-6a**)



**Conditions:** (a) *i.* THF,  $K_2CO_3$  (0.2 eq), Pd/C (cat), r.t., 6 h,  $H_2$ -atm, *ii.* DCM, pyridine (2.5 eq),  $Ac_2O$  (1.3 eq), DMAP (0.05 eq), r.t., overnight,  $N_2$ -atm, 76% yield over two steps; (b) THF, 10% HCl, r.t., 2.5 h, 87% yield; (c) isobutylamine (1.5 eq), toluene, 3 Å MS, reflux, overnight,  $N_2$ -atm, 68% yield; (d) *i.* THF,  $K_2CO_3$  (0.2 eq), Pd/C (cat), r.t., overnight,  $H_2$ -atm, *ii.* **11** (1.3 eq), DCM, DIPEA (3.1 eq), 2-chloro-1-methylpyridinium iodide (1.55 eq), 45°C, overnight,  $N_2$ -atm, 68% yield over two steps; (e) *i.* 1:1 ethanol:THF,  $K_2CO_3$  (0.2 eq), Pd/C (cat), r.t., 24 h,  $H_2$ -atm, *ii.* DCM, pyridine (15 eq),  $Ac_2O$  (10 eq), DMAP (0.2 eq), 45°C, overnight,  $N_2$ -atm, 76% yield over two steps; (f) THF, 10% HCl, r.t., 2.5 h, 91% yield; (g) isobutylamine (3.6 eq), *p*-TsOH (cat), toluene, 3 Å MS, reflux, overnight,  $N_2$ -atm, 42% yield;

**Note:** Trimeric compounds **3a**, **6a** and **9a** could not be synthesized due to insolubility in common solvents.

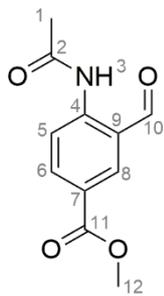
#### Methyl 4-acetamido-3-(1,3-dioxolan-2-yl)benzoate (**1a**)



According to the method of Ikeda *et al.*,<sup>1</sup> nitro aromatic **10** (1.0 g, 3.95 mmol, 1.0 eq) was dissolved in THF (50 mL) and  $K_2CO_3$  (109 mg, 0.79 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to stir for 6 h. Pyridine (330  $\mu$ L) was added and the solution was filtered over Celite™, washed with ethyl acetate and subsequently concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI)  $m/z$  224.1  $[M+H]^+$ ). To the obtained aniline were successively added dichloromethane (10 mL), pyridine (795  $\mu$ L, 9.83 mmol, 2.5 eq), acetic anhydride (485  $\mu$ L, 5.13 mmol, 1.3 eq) and DMAP (24 mg, 0.20 mmol, 0.05 eq) and the mixture was stirred for 16 h. Water was added and subsequently extracted with dichloromethane (x 1) and diethyl ether (x 2). The combined organic phases were dried with  $MgSO_4$  and concentrated *in vacuo*. The residue was purified by flash column chromatography (2:3 hexane:EtOAc,  $R_f \approx 0.3$ ) to give the title compound **1a** as a white crystalline solid (0.795 g, 3.00 mmol, 76% yield);  $\delta_H$  (400 MHz,  $CDCl_3$ ) 8.76 (br. s., 1 H,  $H_3$ ), 8.38 (d,  $J$  8.1, 1 H,  $H_5$ ), 8.12 (d,  $J$  2.0, 1 H,  $H_8$ ), 8.04 (dd,  $J$  8.6, 2.0, 1 H,  $H_6$ ), 5.91 (s, 1 H,  $H_{10}$ ), 4.03-4.26 (m, 4 H,  $H_{11}$ ), 3.91 (s, 3 H,  $H_{13}$ ), 2.21 (s, 3 H,  $H_1$ );  $\delta_H$  (400 MHz,  $DMSO-d_6$ ) 9.43 (s, 1 H,  $H_3$ ), 8.06 (d,  $J$  1.9, 1 H,  $H_8$ ), 7.87-7.99 (m, 2 H,  $H_5+H_6$ ), 6.00 (s, 1 H,  $H_{10}$ ), 4.05-4.16 (m, 2 H,

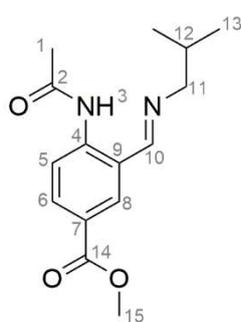
*H11*), 3.94-4.03 (m, 2 H, *H11'*), 3.84 (s, 3 H, *H13*), 2.11 (s, 3 H, *H1*);  $\delta_{\text{C}}$  (101 MHz,  $\text{CDCl}_3$ ) 168.3, 166.4, 140.6, 131.4, 128.7, 125.2, 125.0, 121.2, 102.4, 65.0, 52.0, 25.1;  $\delta_{\text{C}}$  (101 MHz,  $\text{DMSO-}d_6$ ) 168.7, 165.6, 140.8, 130.3, 128.8, 128.0, 125.1, 123.7, 99.3, 64.8, 52.1, 23.9; HRMS (ESI) for  $\text{C}_{13}\text{H}_{15}\text{NO}_5\text{Na}$  [ $\text{M}+\text{Na}$ ] $^+$   $m/z$  288.08424 (calc), 288.08419 (found); IR (neat) 3376 (br), 2956 (br), 2900 (br), 1418, 1693, 1511, 1283, 1201, 1077, 996, 957, 907, 767. X-ray diffraction confirmed the structure of compound **1a** (see Section 3.1).

#### Methyl 4-acetamido-3-formylbenzoate (**4a**)



Acetal **1a** (500 mg, 1.88 mmol) was dissolved in THF (18 mL) and a 10% HCl solution was added (2 mL). The mixture was allowed to stir for 2.5 h. A saturated  $\text{NaHCO}_3$  solution was added and extracted 3 times with diethyl ether. The combined organic phases were dried with  $\text{MgSO}_4$  and concentrated *in vacuo* to approximately 50 mL. The remaining diethyl ether was allowed to evaporate at room temperature leading to the formation of a crystalline precipitate. *Title compound 4a* was obtained as colourless needles after filtration and washing with cold diethyl ether (360 mg, 1.63 mmol, 87% yield);  $\delta_{\text{H}}$  (400 MHz,  $\text{CDCl}_3$ ) 11.32 (br. s., 1 H, *H3*), 9.99 (s, 1 H, *H10*), 8.83 (d,  $J$  8.8, 1 H, *H5*), 8.40 (d,  $J$  1.8, 1 H, *H8*), 8.25 (dd,  $J$  8.8, 2.0, 1 H, *H6*), 3.96 (s, 3 H, *H12*), 2.30 (s, 3 H, *H1*);  $\delta_{\text{H}}$  (400 MHz,  $\text{DMSO-}d_6$ ) 11.00 (s, 1 H, *H3*), 10.05 (s, 1 H, *H10*), 8.45 (d,  $J$  2.0, 1 H, *H8*), 8.35 (d,  $J$  8.7, 1 H, *H5*), 8.20 (dd,  $J$  8.8, 2.1, 1 H, *H6*), 3.88 (s, 3 H, *H12*), 2.20 (s, 3 H, *H1*);  $\delta_{\text{C}}$  (101 MHz,  $\text{CDCl}_3$ ) 195.1, 169.8, 165.4, 144.3, 137.8, 137.1, 124.6, 120.9, 119.5, 52.4, 25.5;  $\delta_{\text{C}}$  (101 MHz,  $\text{DMSO-}d_6$ ) 194.3, 169.7, 165.0, 143.5, 135.7, 134.9, 124.3, 123.2, 120.4, 52.3, 24.6; HRMS (ESI) for  $\text{C}_{11}\text{H}_{10}\text{NO}_4$  [ $\text{M}-\text{H}$ ] $^-$   $m/z$  220.06153 (calc), 220.06127 (found); IR (neat) 3260 (br), 2960, 1708, 1665, 1589, 1504, 1457, 1276, 1195, 1180, 762. X-ray diffraction confirmed the structure of compound **4a** (see Section 3.1).

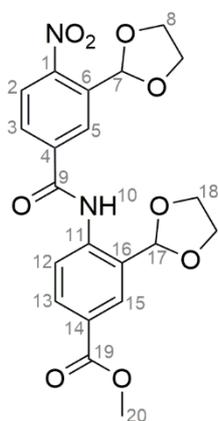
#### Methyl (*E*)-4-acetamido-3-((isobutylimino)methyl)benzoate (**7a**)



Aldehyde **4a** (120 mg, 0.54 mmol, 1 eq) and isobutylamine (81  $\mu\text{L}$ , 0.81 mmol, 1.5 eq) were dissolved in toluene (10 mL) and 3 Å activated molecular sieves were added. The mixture was refluxed for 16 h, after which the molecular sieves were filtered off and washed with ethyl acetate. The crude mixture was concentrated *in vacuo* and purified by flash column chromatography over alumina (basic, activated, 15% ethyl acetate in hexane,  $R_f \approx 0.3$ ) to give *the title compound 7a* as a white solid (102 mg, 0.37 mmol, 68% yield);  $\delta_{\text{H}}$  (400 MHz,  $\text{CDCl}_3$ ) 13.03 (br. s., 1 H, *H3*), 8.78 (d,  $J$  9.4, 1 H, *H5*), 8.38 (s, 1 H, *H10*), 8.06 (d,  $J$  1.8, 1 H, *H8* or *H6*), 8.05 (dd,  $J$  7.8, 2.0, 1 H, *H6* or *H8*), 3.93 (s, 3 H, *H15*), 3.50

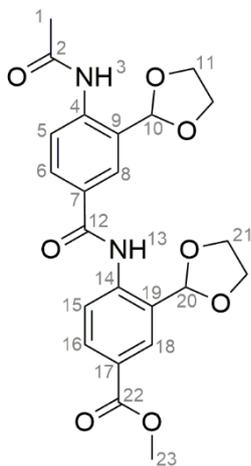
(dd,  $J$  6.3, 1.3, 2 H,  $H_{11}$ ), 2.22 (s, 3 H,  $H_1$ ), 1.95-2.05 (m, 1 H,  $H_{12}$ ), 1.04 (d,  $J$  6.8, 6 H,  $H_{13}$ );  $\delta_H$  (400 MHz, DMSO- $d_6$ ) 13.01 (s, 1 H,  $H_3$ ), 8.65 (d,  $J$  8.6, 1 H,  $H_5$ ), 8.62 (s, 3 H,  $H_{10}$ ), 8.17 (d,  $J$  1.5, 1 H,  $H_8$ ), 8.00 (dd,  $J$  8.7, 1.6, 1 H,  $H_6$ ), 3.85 (s, 3 H,  $H_{15}$ ), 3.49 (d,  $J$  6.3, 2 H,  $H_{11}$ ), 2.14 (s, 3 H,  $H_1$ ), 1.82-2.02 (m, 1 H,  $H_{12}$ ), 0.99 (d,  $J$  6.8, 6 H,  $H_{13}$ );  $\delta_C$  (101 MHz, CDCl<sub>3</sub>) 169.8, 166.2, 163.7, 143.8, 134.7, 132.6, 123.9, 120.1, 119.1, 69.2, 52.1, 29.8, 25.4, 20.6;  $\delta_C$  (101 MHz, DMSO- $d_6$ ) 169.2, 165.3, 164.6, 143.3, 134.5, 132.1, 123.3, 120.1, 118.4, 67.8, 52.0, 29.2, 25.0, 20.4; HRMS (ESI) for C<sub>15</sub>H<sub>21</sub>N<sub>2</sub>O<sub>3</sub> [M+H]<sup>+</sup>  $m/z$  277.15467 (calc), 277.15473 (found); IR (neat) 2958, 2926 (br), 2867, 2848, 1701, 1636, 1583, 1519, 1277, 1228, 1198, 1112, 1040, 804, 770. X-ray diffraction confirmed the structure of compound **7a** (see Section 3.1).

### Methyl 4-(3-(1,3-dioxolan-2-yl)-4-nitrobenzamido)-3-(1,3-dioxolan-2-yl)benzoate (**12**)



Nitro aromatic **10** (600 mg, 2.37 mmol, 1 eq) was dissolved in THF (50 mL) and K<sub>2</sub>CO<sub>3</sub> (66 mg, 0.48 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to stir for 16 h. Pyridine (200  $\mu$ L) was added and the solution was filtered over Celite™, washed with ethyl acetate and subsequently concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI)  $m/z$  224.1 [M+H]<sup>+</sup>). Compound **11** (734 mg, 3.07 mmol, 1.3 eq) was dissolved in dichloromethane (20 mL) and *N,N*-diisopropylethylamine (1.3 mL, 7.46 mmol, 3.1 eq) was added dropwise, followed by the addition of 2-chloro-1-methylpyridinium iodide (940 mg, 3.68 mmol, 1.55 eq). The mixture was stirred at 45 °C for 30 mins. Then, a dichloromethane solution of the aniline obtained in the first step was added and the mixture was stirred for 20 h at 45 °C. A saturated NH<sub>4</sub>Cl solution was added and subsequently extracted with dichloromethane (x 3). The combined organic phases were dried with MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (1.25% methanol in dichloromethane,  $R_f \approx 0.7$ ), followed by recrystallization from hot methanol (precipitation occurs during sonication and cooling). Concentration *in vacuo* gave the *title compound 12* as a pale yellow solid (712 mg, 1.60 mmol, 68% yield);  $\delta_H$  (400 MHz, CDCl<sub>3</sub>) 10.00 (s, 1 H,  $H_{10}$ ), 8.63 (d,  $J$  8.6, 1 H,  $H_{12}$ ), 8.30 (d,  $J$  2.0, 1 H,  $H_5$ ), 8.16 (d,  $J$  2.2, 1 H,  $H_{15}$ ), 8.14 (dd,  $J$  8.6, 2.0, 1 H,  $H_{13}$ ), 8.11 (dd,  $J$  8.3, 2.0, 1 H,  $H_3$ ), 8.04 (d,  $J$  8.3, 1 H,  $H_2$ ), 6.53 (s, 1 H,  $H_7$ ), 5.95 (s, 1 H,  $H_{17}$ ), 4.15-4.27 (m, 4 H,  $H_{18}$ ), 4.04-4.14 (m, 4 H,  $H_8$ ), 3.94 (s, 3 H,  $H_{20}$ );  $\delta_C$  (101 MHz, CDCl<sub>3</sub>) 162.2, 162.8, 150.5, 140.5, 138.5, 134.0, 131.8, 129.6, 129.0, 125.9, 125.7, 125.2, 124.9, 120.9, 103.4, 99.2, 65.4, 65.0, 52.2; HRMS (ESI) for C<sub>21</sub>H<sub>20</sub>N<sub>2</sub>O<sub>9</sub>Na [M+Na]<sup>+</sup>  $m/z$  467.10610 (calc), 467.10584 (found); IR (neat) 3400 (br), 2954 (br), 2880 (br), 1687, 1530, 1291, 1272, 1230, 1198, 1108, 944, 843, 770, 736.

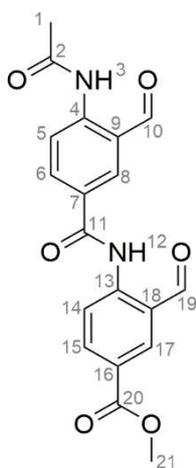
### Methyl 4-(4-acetamido-3-(1,3-dioxolan-2-yl)benzamido)-3-(1,3-dioxolan-2-yl)benzoate (**2a**)



Nitro aromatic **12** (600 mg, 1.35 mmol, 1 eq) was dissolved in 1:1 ethanol:THF (50 mL) and  $K_2CO_3$  (38 mg, 0.27 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to for 24 h. The solution was filtered over Celite™, washed with ethyl acetate and subsequently concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI)  $m/z$  415.2  $[M+H]^+$ ). To the obtained aniline were successively added dichloromethane (30 mL), pyridine (1.63 mL, 20.1 mmol, 15 eq), acetic anhydride (1.28 mL, 13.5 mmol, 10 eq) and DMAP (33 mg, 0.27 mmol, 0.2 eq) and the mixture was stirred for 16 h at 45 °C.

Water was added and subsequently extracted with dichloromethane (x 3). The combined organic phases were dried over  $MgSO_4$  and concentrated *in vacuo*. The residue was purified by flash column chromatography (1:4 hexane:EtOAc,  $R_f \approx 0.1$ , to 100% EtOAc,  $R_f \approx 0.4$ ) to give *the title compound 2a* as an off-white solid (464 mg, 1.02 mmol, 76% yield);  $\delta_H$  (400 MHz,  $CDCl_3$ ) 9.80 (s, 1 H, *H13*), 8.74 (br. s., 1 H, *H3*), 8.62 (d, *J* 8.6, 1 H, *H15*), 8.43 (d, *J* 8.6, 1 H, *H5*), 8.14 (d, *J* 2.0, 1 H, *H18*), 8.10 (dd, *J* 8.6, 2.3, 1 H, *H16*), 8.07 (d, *J* 2.3, 1 H, *H8*), 7.90 (dd, *J* 8.6, 2.0, 1 H, *H6*), 5.96 (s, 1 H, *H10*), 5.93 (s, 1 H, *H20*), 4.11-4.26 (m, 8 H, *H11+H21*), 3.92 (s, 3 H, *H23*), 2.24 (s, 3 H, *H1*);  $\delta_C$  (101 MHz,  $CDCl_3$ ) 168.3, 166.4, 164.2, 141.1, 139.8, 131.7, 129.4, 128.4, 126.1, 124.9, 124.7, 121.7, 103.3, 102.1, 65.1, 65.0, 52.0, 25.0 (overlapping C signals); HRMS (ESI) for  $C_{23}H_{24}N_2O_8Na$   $[M+Na]^+$   $m/z$  479.14249 (calc), 479.14208 (found); IR (neat) 3336, 3234 (br), 2996 (br), 2957 (br), 2896, 1719, 1680, 1653, 1593, 1519, 1397, 1315, 1273, 1102, 1062, 967, 910, 766.

### Methyl 4-(4-acetamido-3-formylbenzamido)-3-formylbenzoate (**5a**)

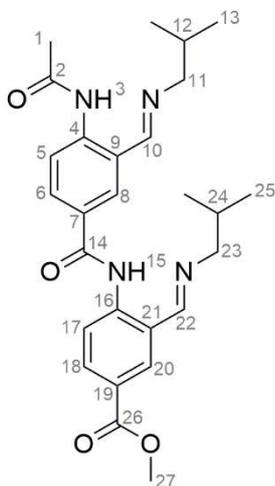


Compound **2a** (350 mg, 0.77 mmol) was dissolved in THF (31.5 mL) and a 10% HCl solution was added (3.5 mL). The mixture was allowed to stir for 2.5 h. A precipitate was formed, which was filtered and washed with diethyl ether. The solid was then dissolved in dichloromethane and extracted with a saturated  $NaHCO_3$  solution (x 1) and brine (x 1). The combined organic phases were dried over  $MgSO_4$  and concentrated *in vacuo* to give *the title compound 5a* as an off-white solid (256 mg, 0.70 mmol, 91% yield);  $\delta_H$  (400 MHz,  $CDCl_3$ ) 12.37 (s, 1 H, *H12*), 11.33 (s, 1 H, *H3*), 10.08 (s, 2 H, *H10+H19*), 9.02 (d, *J* 8.8, 1 H, *H14*), 8.96 (d, *J* 8.8, 1 H, *H5*), 8.48 (d, *J* 1.8, 1 H, *H17*), 8.46 (d, *J* 2.0, 1 H, *H8*), 8.34

(dd, *J* 8.8, 2.0, 1 H, *H15*), 8.31 (dd, *J* 8.8, 2.0, 1 H, *H6*), 3.99 (s, 3 H, *H21*), 2.32 (s, 3 H, *H1*);  $\delta_C$  (151 MHz,  $CDCl_3$ ) 195.6, 195.1, 169.8, 165.4, 164.2, 144.5, 144.1, 137.9, 137.3, 136.1, 134.6, 128.0, 125.1, 121.5, 121.3, 120.1, 119.7, 52.4, 25.5; HRMS (ESI) for  $C_{19}H_{15}N_2O_6$   $[M-H]^-$   $m/z$

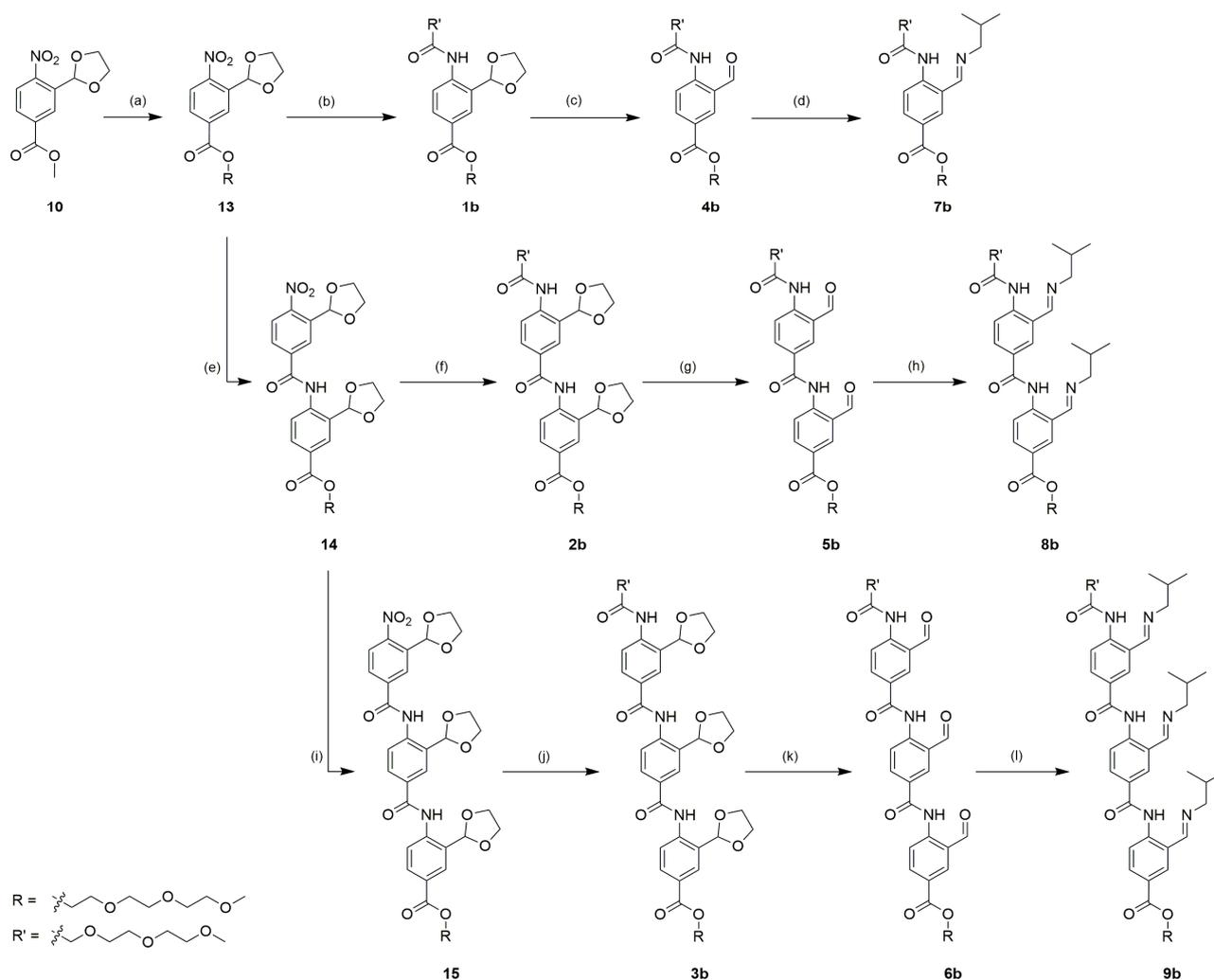
367.09356 (calc), 367.09406 (found); IR (neat) 3252 (br), 3130 (br), 2958, 1659, 1592, 1520, 1392, 1299, 1280, 1261, 1229, 1176, 1112, 764, 692.

**Methyl 4-(4-acetamido-3-((*E*)-(isobutylimino)methyl)benzamido)-3-((*E*)-(isobutylimino)methyl)benzoate (**8a**)**



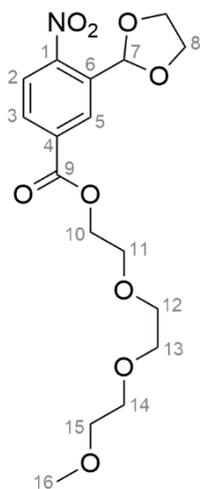
Aldehyde **5a** (50 mg, 0.136 mmol, 1.0 eq) and isobutylamine (50  $\mu$ L, 0.50 mmol, 3.6 eq) were dissolved in toluene (10 mL) and *p*-toluenesulfonic acid (catalytic) and 3 Å activated molecular sieves were added. The mixture was refluxed for 16 h, after which the molecular sieves were filtered off and washed with dichloromethane. The residue was concentrated *in vacuo* and purified by flash column chromatography over alumina (basic, activated, 20% ethyl acetate in hexane,  $R_f \approx 0.4$ ) to give the title compound **8a** as a yellow solid (27.4 mg, 0.057 mmol, 42% yield);  $\delta_H$  (400 MHz,  $CDCl_3$ ) 13.73 (s, 1 H, *H*15), 13.03 (br. s., 1 H, *H*3), 9.00 (d, *J* 9.4, 1 H, *H*17), 8.84 (d, *J* 8.8, 1 H, *H*5), 8.43 (s, 2 H, *H*10+*H*22), 8.09-8.17 (m, 3 H, *H*8+*H*18+*H*20), 8.03 (dd, *J* 8.8, 2.0, 1 H, *H*6), 3.95 (s, 3 H, *H*27), 3.57 (d, *J* 6.6, 2 H, *H*11 or *H*23), 3.52 (d, *J* 6.1, 2 H, *H*11 or *H*23), 2.24 (s, 3 H, *H*1), 1.95-2.13 (m, 2 H, *H*12+*H*24), 1.05 (d, *J* 6.8 Hz, 6 H, *H*13 or *H*25), 0.97 (d, *J* 6.8, 6 H, *H*13 or *H*25);  $\delta_C$  (101 MHz,  $CDCl_3$ ) 169.9, 166.2, 165.6, 164.4, 163.8, 144.2, 143.1, 134.9, 133.6, 132.8, 130.3, 128.5, 124.1, 120.8, 120.4, 119.5, 119.2, 69.7, 69.3, 52.1, 29.8, 29.7, 25.3, 20.7, 20.6; HRMS (ESI) for  $C_{27}H_{35}N_4O_4$  [ $M+H$ ] $^+$   $m/z$  479.26528 (calc), 479.26510 (found); IR (neat) 2956 (br), 2925 (br), 2867 (br), 1706, 1678, 1637, 1591, 1520, 1281, 1231, 1195, 1043, 772, 755. X-ray diffraction confirmed the structure of compound **8a** (see Section 3.1).

#### 1.4. Synthetic route to PEG capped $\alpha$ -helix mimetics (**1b-9b**)



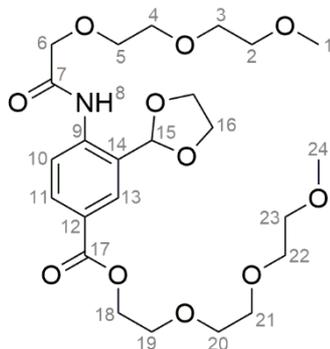
**Conditions:** (a) Methyltriglycol (4 eq), NaOMe (0.05 eq), 75°C, overnight, N<sub>2</sub>-atm, 73% yield; (b) *i.* EtOAc, K<sub>2</sub>CO<sub>3</sub> (0.2 eq), Pd/C (cat), r.t., overnight, H<sub>2</sub>-atm, *ii.* 2-[2-(2-methoxyethoxy)ethoxy]acetic acid (2.0 eq), DCM, DIPEA (2.6 eq), 2-chloro-1-methylpyridinium iodide (2.4 eq), 45°C, 20 h, N<sub>2</sub>-atm, 81% yield over two steps; (c) THF, 10% HCl, r.t., 2.5 h, 88% yield; (d) isobutylamine (2.4 eq), toluene, 3 Å MS, reflux, overnight, N<sub>2</sub>-atm, 43% yield; (e) *i.* EtOAc, K<sub>2</sub>CO<sub>3</sub> (0.2 eq), Pd/C (cat), r.t., overnight, H<sub>2</sub>-atm, *ii.* **11** (1.4 eq), DCM, DIPEA (2.5 eq), 2-chloro-1-methylpyridinium iodide (1.56 eq), 45°C, 48 h, N<sub>2</sub>-atm, 86% yield over two steps; (f) *i.* EtOAc, K<sub>2</sub>CO<sub>3</sub> (0.2 eq), Pd/C (cat), 45°C, overnight, H<sub>2</sub>-atm, *ii.* 2-[2-(2-methoxyethoxy)ethoxy]acetic acid (2.0 eq), DCM, DIPEA (2.6 eq), 2-chloro-1-methylpyridinium iodide (2.4 eq), 45°C, 48 h, N<sub>2</sub>-atm, 68% yield over two steps; (g) THF, 10% HCl, r.t., 2.5 h, 92% yield; (h) isobutylamine (3.0 eq), toluene, 3 Å MS, reflux, overnight, N<sub>2</sub>-atm, 58% yield; (i) *i.* EtOAc, K<sub>2</sub>CO<sub>3</sub> (0.2 eq), Pd/C (cat), 45°C, overnight, H<sub>2</sub>-atm, *ii.* **11** (1.4 eq), DCM, DIPEA (2.5 eq), 2-chloro-1-methylpyridinium iodide (1.56 eq), 45°C, 48 h, N<sub>2</sub>-atm, 74% yield over two steps; (j) *i.* EtOAc, K<sub>2</sub>CO<sub>3</sub> (0.2 eq), Pd/C (cat), 45°C, 5 h, H<sub>2</sub>-atm, *ii.* 2-[2-(2-methoxyethoxy)ethoxy]acetic acid (2.0 eq), DCM, DIPEA (2.5 eq), 2-chloro-1-methylpyridinium iodide (2.4 eq), 45°C, 48 h, N<sub>2</sub>-atm, 52% yield over two steps; (k) THF, 10% HCl, r.t., 2.5 h, 91% yield; (l) isobutylamine (4.5 eq), toluene, 3 Å MS, reflux, overnight, N<sub>2</sub>-atm, 42% yield.

### 2-(2-(2-Methoxyethoxy)ethoxy)ethyl 3-(1,3-dioxolan-2-yl)-4-nitrobenzoate (**13**)



Compound **10** (2.0 g, 7.90 mmol, 1 eq), methyltriglycol (5 mL, 31.24 mmol, 4.0 eq) and sodium methoxide (21 mg, 0.39 mmol, 0.05 eq) were stirred and heated to 75 °C with removal of methanol (Vigreux) for 16 h. Dichloromethane was added and the mixture was extracted with saturated NH<sub>4</sub>Cl (x 1) and brine (x 2). The organic phase was dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (1.25% methanol in dichloromethane, R<sub>f</sub> ≈ 0.45) to give *the title compound 13* as a pale yellow oil (2.231 g, 5.79 mmol, 73% yield); δ<sub>H</sub> (400 MHz, CDCl<sub>3</sub>) 8.44 (d, *J* 1.8, 1 H, *H5*), 8.17 (dd, *J* 8.3, 1.8, 1 H, *H3*), 7.91 (d, *J* 8.3, 1 H, *H2*), 6.45 (s, 1 H, *H7*), 4.49-4.59 (m, 2 H, *H10*), 3.98-4.12 (m, 4 H, *H8*), 3.81-3.92 (m, 2 H, *H11*), 3.60-3.78 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.50-3.57 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.37 (s, 3 H, *H16*); δ<sub>C</sub> (101 MHz, CDCl<sub>3</sub>) 164.5, 151.3, 133.8, 133.6, 130.9, 129.2, 124.4, 99.2, 71.9, 70.7, 70.6, 70.6, 69.0, 65.4, 64.9, 59.0; HRMS (ESI) for C<sub>17</sub>H<sub>24</sub>NO<sub>9</sub> [M+H]<sup>+</sup> *m/z* 386.14456 (calc), 386.14390 (found); IR (neat) 3228 (br), 2877 (br), 1722, 1684, 1669, 1584, 1521, 1394, 1278, 1263, 1178, 1099.

### 2-(2-(2-Methoxyethoxy)ethoxy)ethyl 3-(1,3-dioxolan-2-yl)-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzoate (**1b**)

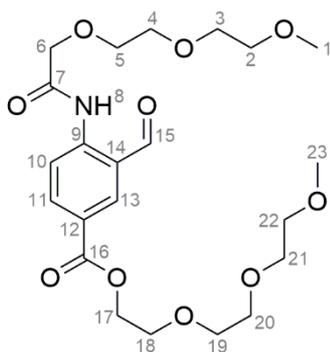


Nitro aromatic **13** (500 mg, 1.30 mmol, 1.0 eq) was dissolved in ethyl acetate (25 mL) and K<sub>2</sub>CO<sub>3</sub> (36 mg, 0.26 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to stir for 16 h. The solution was filtered over Celite™, washed with ethyl acetate and subsequently concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI) *m/z* 356.2 [M+H]<sup>+</sup>). 2-[2-(2-Methoxyethoxy)ethoxy]acetic acid (462

mg, 2.59 mmol, 2.0 eq) was dissolved in dichloromethane (20 mL) and *N,N*-diisopropylethylamine (587 μL, 3.37 mmol, 2.6 eq) was added dropwise, followed by the addition of 2-chloro-1-methylpyridinium iodide (795 mg, 3.11 mmol, 2.4 eq). The mixture was stirred at 45 °C for 30 min. Then, a dichloromethane solution of the aniline obtained in the first step was added and the mixture was stirred for 20 h at 45 °C. A saturated NH<sub>4</sub>Cl solution was added and subsequently extracted with dichloromethane (x 3). The combined organic phases were washed with brine dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (2% methanol in dichloromethane, R<sub>f</sub> ≈ 0.35) to give *the title compound 1b* as a colourless oil (541 mg, 1.05 mmol, 81% yield); δ<sub>H</sub> (400 MHz, CDCl<sub>3</sub>) 9.85 (s, 1 H, *H8*), 8.49 (d, *J* 8.6, 1 H, *H10*), 7.98-8.17 (m, 2 H, *H11+H13*), 5.80 (s, 1 H, *H15*), 4.45 (m, 2 H, *H18*), 4.19-4.26 (m, 2 H, *H16*), 4.14 (s,

2 H, *H6*), 4.02-4.09 (m, 2 H, *H16'*), 3.76-3.83 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.68-3.73 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.61-3.67 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.50-3.55 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.35 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.34 (s, 3 H, *PEG-CH<sub>3</sub>*);  $\delta_{\text{H}}$  (400 MHz, DMSO-*d*<sub>6</sub>) 9.82 (s, 1 H, *H8*), 8.41 (d, *J* 9.4, 1 H, *H10*), 7.92-8.11 (m, 2 H, *H11+H13*), 5.84 (s, 1 H, *H15*), 4.33-4.49 (m, 2 H, *H18*), 4.16-4.21 (m, 2 H, *H16*), 4.14 (s, 2 H, *H6*), 3.99-4.06 (m, 2 H, *H16'*), 3.69-3.78 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.47-3.65 (m, 10 H, *PEG-CH<sub>2</sub>*), 3.37-3.46 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.22 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.21 (s, 3 H, *PEG-CH<sub>3</sub>*);  $\delta_{\text{C}}$  (101 MHz, CDCl<sub>3</sub>) 168.0, 165.7, 140.4, 131.8, 129.9, 125.0, 124.7, 120.9, 103.5, 71.8, 71.1, 71.0, 70.6, 70.5, 70.5, 70.3, 69.1, 65.0, 64.0, 59.0, 58.9 (overlapping of PEG-C);  $\delta_{\text{C}}$  (101 MHz, DMSO-*d*<sub>6</sub>) 168.1, 164.9, 140.5, 131.2, 129.7, 125.1, 124.5, 120.6, 102.4, 71.2, 70.4, 70.2, 69.9, 69.7, 69.6, 68.3, 64.8, 64.0, 58.0, 58.0 (overlapping of PEG-C); HRMS (ESI) for C<sub>24</sub>H<sub>37</sub>NO<sub>11</sub>Na [M+Na]<sup>+</sup> *m/z* 538.22588 (calc), 538.22498 (found); IR (neat) 3577 (br), 3526 (br), 3340 (br), 2888 (br), 1714, 1700, 1590, 1521, 1280, 1241, 1196, 1098.

## 2-(2-(2-Methoxyethoxy)ethoxy)ethyl 3-formyl-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzoate (**4b**)

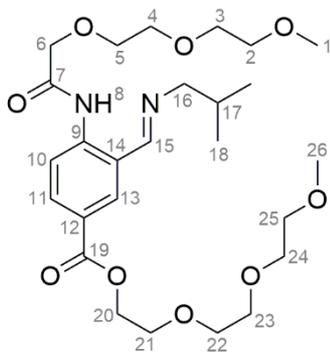


Compound **1b** (425 mg, 0.82 mmol) was dissolved in THF (18 mL) and a 10% HCl solution was added (2 mL). The mixture was allowed to stir for 2.5 h. A saturated NaHCO<sub>3</sub> solution was added and extracted 3 times with ethyl acetate. The combined organic phases were washed with brine and subsequently dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (3% methanol in ethyl acetate, *R<sub>f</sub>* ≈ 0.3) to give the title compound **4b** as a

colourless oil (342 mg, 0.72 mmol, 88% yield);  $\delta_{\text{H}}$  (400 MHz, CDCl<sub>3</sub>) 12.07 (s, 1 H, *H8*), 10.02 (s, 1 H, *H15*), 8.89 (d, *J* 8.8, 1 H, *H10*), 8.43 (d, *J* 2.0, 1 H, *H13*), 8.28 (dd, *J* 8.8, 2.0, 1 H, *H11*), 4.44-4.60 (m, 2 H, *H17*), 4.22 (s, 2 H, *H6*), 3.81-3.96 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.63-3.77 (m, 8 H, *PEG-CH<sub>2</sub>*), 3.51-3.62 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.38 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.37 (s, 3 H, *PEG-CH<sub>3</sub>*);  $\delta_{\text{H}}$  (400 MHz, DMSO-*d*<sub>6</sub>) 11.91 (s, 1 H, *H8*), 10.09 (s, 1 H, *H15*), 8.74 (d, *J* 8.6, 1 H, *H10*), 8.53 (d, *J* 2.0, 1 H, *H13*), 8.24 (dd, *J* 8.8, 2.0, 1 H, *H11*), 4.36-4.49 (m, 2 H, *H17*), 4.19 (s, 2 H, *H6*), 3.73-3.79 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.68-3.72 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.58-3.62 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.48-3.57 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.37-3.45 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.21 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.21 (s, 3 H, *PEG-CH<sub>3</sub>*);  $\delta_{\text{C}}$  (101 MHz, CDCl<sub>3</sub>) 194.3, 170.2, 164.9, 143.2, 137.7, 136.9, 124.9, 121.8, 119.7, 71.9, 71.6, 71.3, 70.7, 70.7, 70.6, 70.6, 69.1, 59.0 (overlapping of PEG-C);  $\delta_{\text{C}}$  (101 MHz, DMSO-*d*<sub>6</sub>) 195.9, 170.2, 164.4, 142.6, 137.3, 136.1, 124.3, 122.1, 119.1, 71.2, 71.2, 70.8, 70.5, 69.9, 69.7, 69.7, 69.6, 68.3, 64.2, 58.0 (overlapping of PEG-C); HRMS (ESI) for C<sub>22</sub>H<sub>34</sub>NO<sub>10</sub> [M+H]<sup>+</sup> *m/z* 472.21772 (calc),

472.21793 (found); IR (neat) 3608 (br), 3534 (br), 3250 (br), 2877 (br), 1706, 1677, 1581, 1516, 1468, 1272, 1181, 1099, 766.

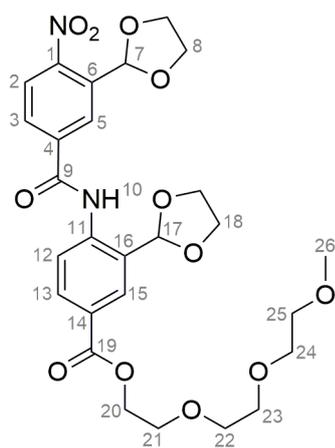
**2-(2-(2-Methoxyethoxy)ethoxy)ethyl (E)-3-((isobutylimino)methyl)-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzoate (7b)**



Aldehyde **4b** (100 mg, 0.21 mmol, 1.0 eq) and isobutylamine (50  $\mu$ L, 0.50 mmol, 2.4 eq) were dissolved in toluene (10 mL) and 3 Å activated molecular sieves were added. The mixture was refluxed for 16 h, after which the molecular sieves were filtered off and washed with ethyl acetate. The mixture was concentrated *in vacuo* and purified by flash column chromatography over alumina (basic, activated, 2:1 EtOAc:hexane,  $R_f \approx 0.3$ ) to give *the title compound 7b* as a colourless oil (48.3 mg, 0.09 mmol, 43% yield);  $\delta_H$  (400 MHz,  $CDCl_3$ ) 13.25 (s, 1 H, *H*8), 8.87 (d, *J* 8.6, 1 H, *H*10), 8.35 (s, 1 H, *H*15), 7.99-8.17 (m, 2 H, *H*13+*H*11), 4.42-4.57 (m, 2 H, *H*20), 4.22 (s, 2 H, *H*6), 3.77-3.89 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.60-3.77 (m, 10 H, *PEG-CH*<sub>2</sub>), 3.52-3.60 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.48 (d, *J* 6.6, 2 H, *H*16), 3.38 (s, 3 H, *PEG-CH*<sub>3</sub>), 3.37 (s, 3 H, *PEG-CH*<sub>3</sub>), 1.95-2.17 (m, 1 H, *H*17), 0.98 (d, *J* 6.8, 6 H, *H*18);  $\delta_H$  (400 MHz,  $DMSO-d_6$ ) 13.24 (s, 1 H, *H*8), 8.80 (d, *J* 8.8, 1 H, *H*10), 8.58 (s, 1 H, *H*15), 8.18 (d, *J* 2.0, 1 H, *H*13), 8.01 (dd, *J* 8.7, 2.2, 1 H, *H*11), 4.37-4.42 (m, 2 H, *H*20), 4.16 (s, 2 H, *H*6), 3.73-3.78 (m, 2 H, *H*21), 3.66-3.71 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.56-3.64 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.49-3.54 (m, 6 H, *PEG-CH*<sub>2</sub>), 3.46 (d, *J* 6.3, 2 H, *H*16), 3.37-3.43 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.22 (s, 3 H, *PEG-CH*<sub>3</sub>), 3.20 (s, 3 H, *PEG-CH*<sub>3</sub>), 1.93-2.05 (m, 1 H, *H*17), 0.93 (d, *J* 6.6, 6 H, *H*18);  $\delta_C$  (101 MHz,  $CDCl_3$ ) 170.5, 165.7, 163.2, 143.0, 134.7, 132.6, 124.3, 121.2, 119.7, 71.9, 71.0, 70.7, 70.7, 70.6, 69.7, 69.2, 64.1, 59.1, 59.0, 29.5, 20.6 (overlapping of *PEG-C*);  $\delta_C$  (101 MHz,  $DMSO-d_6$ ) 170.2, 164.8, 164.1, 142.5, 134.5, 132.0, 123.8, 121.1, 118.9, 71.2, 71.0, 70.4, 69.8, 69.7, 69.6, 69.6, 68.3, 68.3, 64.0, 58.0, 58.0, 28.9, 20.3 (overlapping of *PEG-C*); HRMS (ESI) for  $C_{26}H_{43}N_2O_9$  [ $M+H$ ]<sup>+</sup>  $m/z$  527.29631 (calc), 527.29584 (found); IR (neat) 2952 (br), 2923 (br), 2870 (br), 1716, 1695, 1641, 1582, 1521, 1277, 1193, 1101, 1040, 849. 768.

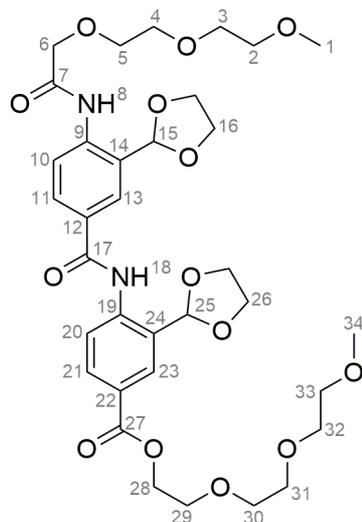
**2-(2-(2-Methoxyethoxy)ethoxy)ethyl 4-(3-(1,3-dioxolan-2-yl)-4-nitrobenzamido)-3-(1,3-dioxolan-2-yl)benzoate (14)**

Nitro aromatic **13** (1.0 g, 2.59 mmol, 1 eq) was dissolved in ethyl acetate (50 mL) and  $K_2CO_3$  (72 mg, 0.52 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to stir for 16 h. The solution was filtered over Celite™, washed with ethyl acetate and concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI)  $m/z$  356.2 [ $M+H$ ]<sup>+</sup>). Compound **11** (850 mg, 3.55 mmol, 1.4 eq) was dissolved in dichloromethane



(30 mL) and *N,N*-di-isopropylethylamine (1.15 mL, 6.60 mmol, 2.5 eq) was added dropwise, followed by the addition of 2-chloro-1-methylpyridinium iodide (1.032 g, 4.04 mmol, 1.56 eq). The mixture was stirred at 45 °C for 30 min. Then, a dichloromethane solution of the aniline obtained in the first step was added and the mixture was stirred for 48 h at 45 °C. A saturated NH<sub>4</sub>Cl solution was added and subsequently extracted with dichloromethane (x 3). The combined organic phases were dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (1.25% methanol in dichloromethane, R<sub>f</sub> ≈ 0.3) to give *the title compound 14* as a pale yellow solid (1.280 g, 2.22 mmol, 86% yield); δ<sub>H</sub> (400 MHz, CDCl<sub>3</sub>) 10.00 (s, 1 H, *H*10), 8.63 (d, *J* 8.5, 1 H, *H*12), 8.29 (d, *J* 1.9, 1 H, *H*5), 8.12-8.18 (m, 2 H, *H*15+*H*13), 8.10 (dd, *J* 8.4, 2.0, 1 H, *H*3), 8.03 (d, *J* 8.4, 1 H, *H*2), 6.52 (s, 1 H, *H*7), 5.94 (s, 1 H, *H*17), 4.46-4.52 (m, 2 H, *H*20), 4.14-4.26 (m, 4 H, *H*18), 4.03-4.13 (m, 4 H, *H*8), 3.83-3.89 (m, 2 H, *H*21), 3.71-3.75 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.64-3.71 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.53-3.58 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.38 (s, 3 H, *H*26); δ<sub>C</sub> (101 MHz, CDCl<sub>3</sub>) 165.7, 162.8, 150.6, 140.6, 138.5, 134.0, 132.0, 129.8, 129.0, 125.9, 125.7, 125.1, 124.7, 120.8, 103.5, 99.2, 71.9, 70.7, 70.7, 70.6, 69.2, 65.4, 65.0, 64.2, 59.0; HRMS (ESI) for C<sub>27</sub>H<sub>32</sub>N<sub>2</sub>O<sub>12</sub>Na [M+Na]<sup>+</sup> *m/z* 599.18475 (calc), 599.18421 (found); IR (neat) 2289 (br), 2890 (br), 1705, 1690, 1598, 1525, 1342, 1279, 1197, 1099, 1050, 948, 916, 846, 767. X-ray confirmed the structure of compound **14** (see Section 3.1).

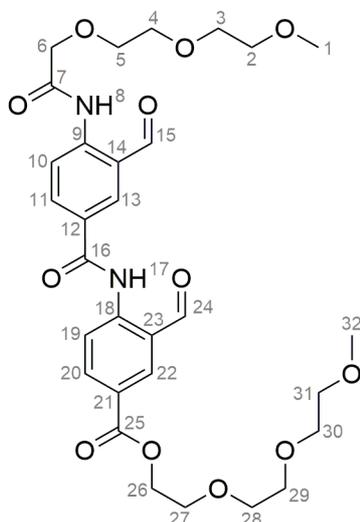
### 2-(2-(2-Methoxyethoxy)ethoxy)ethyl 4-(3-(1,3-dioxolan-2-yl)-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzamido)-3-(1,3-dioxolan-2-yl)benzoate (**2b**)



Nitro aromatic **14** (700 mg, 1.21 mmol, 1.0 eq) was dissolved in ethyl acetate (30 mL) and K<sub>2</sub>CO<sub>3</sub> (35 mg, 0.25 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to 16 h at 45 °C. The solution was filtered over Celite™, washed with ethyl acetate and subsequently concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI) *m/z* 547.3 [M+H]<sup>+</sup>). 2-[2-(2-Methoxyethoxy)ethoxy]acetic acid (433 mg, 2.43 mmol, 2.0 eq) was dissolved in dichloromethane (50 mL) and *N,N*-diisopropylethylamine (550 μL, 3.16 mmol, 2.6 eq) was added dropwise, followed by the addition of 2-chloro-1-methylpyridinium iodide (745 mg, 2.92 mmol, 2.4 eq). The mixture was stirred at 45 °C for 30 min. Then, a dichloromethane solution of the aniline obtained in the first step was added and

the mixture was stirred for another 48 h at 45 °C. A saturated NH<sub>4</sub>Cl solution was added and subsequently extracted with dichloromethane (x 3). The combined organic phases were washed with brine dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (2% methanol in dichloromethane, R<sub>f</sub> ≈ 0.2, then 4% methanol in ethyl acetate, R<sub>f</sub> ≈ 0.2) to give *the title compound 2b* as a white solid (588 mg, 0.83 mmol, 68% yield); δ<sub>H</sub> (400 MHz, CDCl<sub>3</sub>) 9.84 (s, 1 H, H8), 9.81 (s, 1 H, H18), 8.62 (d, J 8.6, 1 H, H20), 8.55 (d, J 8.6, 1 H, H10), 8.15 (d, J 2.0, 1 H, H23), 8.12 (dd, J 8.7, 2.2, 1 H, H21), 8.05 (d, J 2.3, 1 H, H13), 7.90 (dd, J 8.6, 2.3, 1 H, H11), 5.93 (s, 1 H, H25), 5.90 (s, 1 H, H15), 4.43-4.55 (m, 2 H, H28), 4.07-4.27 (m, 10 H, H6+H16+H26), 3.79-3.89 (m, 4 H, PEG-CH<sub>2</sub>), 3.71-3.77 (m, 4 H, PEG-CH<sub>2</sub>), 3.64-3.70 (m, 6 H, PEG-CH<sub>2</sub>), 3.52-3.58 (m, 4 H, PEG-CH<sub>2</sub>), 3.38 (s, 3 H, PEG-CH<sub>3</sub>), 3.37 (s, 3 H, PEG-CH<sub>3</sub>); δ<sub>H</sub> (400 MHz, DMSO-*d*<sub>6</sub>) 10.08 (s, 1 H, H18), 9.79 (s, 1 H, H8), 8.42 (d, J 8.3, 1 H, H10), 8.11 (d, J 2.02, 1 H, H23), 8.07 (d, J 8.5, 1 H, H20), 7.97-8.04 (m, 3 H, H21+H13+H11), 6.07 (s, 1 H, H25), 5.87 (s, 1 H, H15), 4.38-4.44 (m, 2 H, H28), 4.17-4.23 (m, 2 H, H16 or H26), 4.15 (s, 2 H, H6), 4.08-4.00 (m, 6 H, H16 or H26), 3.72-3.78 (m, 4 H, PEG-CH<sub>2</sub>), 3.62-3.66 (m, 2 H, PEG-CH<sub>2</sub>), 3.58-3.62 (m, 2 H, PEG-CH<sub>2</sub>), 3.49-3.57 (m, 6 H, PEG-CH<sub>2</sub>), 3.43-3.46 (m, 2 H, PEG-CH<sub>2</sub>), 3.39-3.42 (m, 2 H, PEG-CH<sub>2</sub>), 3.23 (s, 3 H, PEG-CH<sub>3</sub>), 3.22 (s, 3 H, PEG-CH<sub>3</sub>); δ<sub>C</sub> (101 MHz, CDCl<sub>3</sub>) 168.2, 165.8, 164.2, 141.2, 139.6, 131.9, 129.6, 129.6, 128.5, 127.3, 125.9, 124.9, 124.7, 121.5, 120.8, 103.4, 103.0, 71.9, 71.1, 70.7, 70.7, 70.6, 70.6, 70.4, 69.2, 65.1, 65.0, 64.1, 59.0, 59.0 (overlapping of PEG-C); δ<sub>C</sub> (101 MHz, DMSO-*d*<sub>6</sub>) 168.0, 165.1, 164.1, 140.9, 139.5, 130.6, 129.3, 129.2, 128.8, 128.4, 127.9, 125.8, 125.4, 123.9, 120.9, 102.5, 100.4, 71.2, 70.4, 70.2, 69.9, 69.7, 69.6, 68.3, 64.8, 64.8, 64.1, 58.1, 58.0 (overlapping of PEG-C); HRMS (ESI) for C<sub>34</sub>H<sub>46</sub>N<sub>2</sub>O<sub>14</sub>Na [M+Na]<sup>+</sup> m/z 729.28413 (calc), 729.28335 (found); IR (neat) 3358, 3338, 2892 (br), 1718, 1688, 1593, 1516, 1317, 1291, 1271, 1233, 1197, 1120, 1101, 1059, 767.

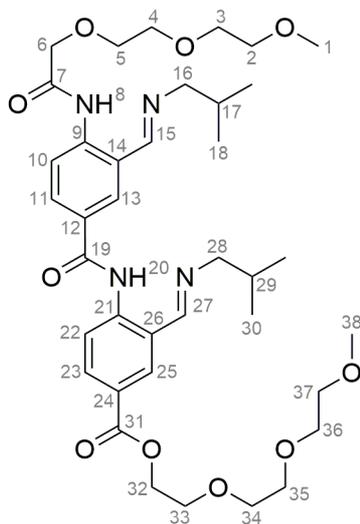
**2-(2-(2-Methoxyethoxy)ethoxy)ethyl 3-formyl-4-(3-formyl-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzamido)benzoate (5b)**



Compound **2b** (450 mg, 0.64 mmol) was dissolved in THF (45 mL) and a 10% HCl solution was added (5 mL). The mixture was allowed to stir for 2.5 h. A saturated NaHCO<sub>3</sub> solution was added and extracted 3 times with ethyl acetate. The combined organic phases were washed with brine dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (2% methanol in ethyl acetate, R<sub>f</sub> ≈ 0.4) to give *the title compound 4b* as a white solid (364 mg, 0.59 mmol, 92% yield); δ<sub>H</sub> (400 MHz,

CDCl<sub>3</sub>) 12.39 (s, 1 H, *H17*), 12.09 (s, 1 H, *H8*), 10.10 (s, 1 H, *H15*), 10.09 (s, 1 H, *H24*), 9.02 (d, *J* 8.8, 2 H, *H10+H19*), 8.50 (d, *J* 1.8, 1 H, *H22*), 8.48 (d, *J* 2.0, 1 H, *H13*), 8.36 (dd, *J* 8.8, 2.0, 1 H, *H20*), 8.31 (dd, *J* 8.8, 2.3, 1 H, *H11*), 4.47-4.60 (m, 2 H, *H26*), 4.25 (s, 2 H, *H6*), 3.83-3.96 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.63-3.79 (m, 8 H, *PEG-CH<sub>2</sub>*), 3.52-3.62 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.39 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.38 (s, 3 H, *PEG-CH<sub>3</sub>*); δ<sub>H</sub> (400 MHz, DMSO-*d*<sub>6</sub>) 11.95 (s, 1 H, *H17*), 11.86 (s, 1 H, *H8*), 10.15 (s, 1 H, *H24*), 10.11 (s, 1 H, *H15*), 8.79 (d, *J* 8.8, 1 H, *H10*), 8.59 (d, *J* 8.6, 1 H, *H19*), 8.57 (d, *J* 2.3, 1 H, *H13*), 8.54 (d, *J* 2.0, 1 H, *H22*), 8.28 (dd, *J* 8.6, 2.3, 1 H, *H20*), 8.27 (dd, *J* 8.8, 2.3, 1 H, *H11*), 4.40-4.49 (m, 2 H, *H26*), 4.21 (s, 2 H, *H6*), 3.74-3.82 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.68-3.74 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.59-3.64 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.48-3.59 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.42-3.47 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.38-3.42 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.22 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.22 (s, 3 H, *PEG-CH<sub>3</sub>*); δ<sub>C</sub> (101 MHz, CDCl<sub>3</sub>) 195.7, 194.2, 170.3, 164.8, 164.2, 144.5, 142.9, 138.0, 137.4, 135.9, 134.3, 128.4, 125.0, 122.2, 121.4, 120.3, 119.7, 71.9, 71.6, 71.3, 70.7, 70.7, 70.6, 70.6, 69.1, 64.5, 59.0 (overlapping of *PEG-C*); δ<sub>C</sub> (101 MHz, DMSO-*d*<sub>6</sub>) 195.6, 195.4, 170.1, 164.4, 164.0, 143.5, 142.0, 136.0, 135.7, 135.6, 134.0, 128.1, 124.8, 123.5, 122.3, 120.6, 119.4, 71.2, 71.2, 70.8, 70.5, 69.9, 69.7, 69.7, 69.6, 68.3, 64.3, 58.0 (overlapping of *PEG-C*); HRMS (ESI) for C<sub>30</sub>H<sub>38</sub>N<sub>2</sub>O<sub>12</sub>Na [M+Na]<sup>+</sup> *m/z* 641.23170 (calc), 641.23079 (found); IR (neat) 3226 (br), 2877 (br), 1726, 1665, 1582, 1517, 1392, 1275, 1260, 1175, 1101, 767, 703.

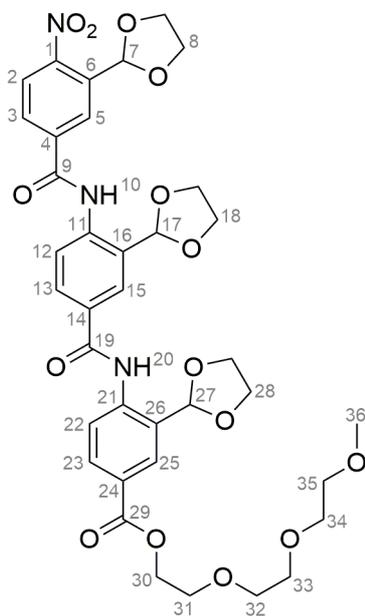
**2-(2-(2-Methoxyethoxy)ethoxy)ethyl 3-((*E*)-(isobutylimino)methyl)-4-(3-((*E*)-(isobutylimino)methyl)-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzamido)benzoate (**8b**)**



Aldehyde **5b** (60 mg, 0.097 mmol, 1.0 eq) and isobutylamine (29 μL, 0.29 mmol, 3.0 eq) were dissolved in toluene (10 mL) and 3 Å activated molecular sieves were added. The mixture was refluxed for 16 h, after which the molecular sieves were filtered off and washed with hot ethyl acetate. The mixture was concentrated *in vacuo* and purified by flash column chromatography over alumina (basic, activated, 3:1 EtOAc:hexane, *R<sub>f</sub>* ≈ 0.2) to give the title compound **8b** as a yellow oil that crystallized upon standing (41 mg, 0.056 mmol, 58% yield); δ<sub>H</sub> (400 MHz, CDCl<sub>3</sub>) 13.75 (s, 1 H, *H20*), 13.24 (s, 1 H, *H8*), 9.00 (d, *J* 9.3, 1 H, *H22*), 8.93 (d, *J* 8.6, 1 H, *H10*), 8.44 (s, 1 H, *H27*), 8.39 (s, 1 H, *H15*), 8.09-8.20 (m, 3 H, *H13+H23+H25*), 8.04 (dd, *J* 8.7, 1.9, 1 H, *H11*), 4.44-4.57 (m, 2 H, *H32*), 4.24 (s, 2 H, *H6*), 3.84-3.89 (m, 2 H, *H33*), 3.79-3.84 (m, 2 H, *PEG-CH<sub>2</sub>*) 3.72-3.77 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.63-3.72 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.53-3.61 (m, 6 H, *H28+PEG-CH<sub>2</sub>*), 3.50 (d, *J* 6.6, 2 H, *H16*), 3.38 (s, 6 H, *H1+H38*), 2.00-2.13 (m, 2 H, *H17+H29*), 0.99 (d, *J* 6.6, 6 H, *H18*), 0.97 (d, *J* 6.8, 6 H, *H30*); δ<sub>H</sub> (400 MHz, DMSO-*d*<sub>6</sub>) 13.86 (s, 1 H, *H20*), 13.17 (s, 1 H, *H8*),

8.87 (d, *J* 8.8, 1 H, *H*22), 8.84 (d, *J* 8.8, 1 H, *H*10), 8.69 (s, 1 H, *H*27), 8.55 (s, 1 H, *H*15), 8.22 (d, *J* 2.0, 1 H, *H*25), 8.11 (d, *J* 2.0, 1 H, *H*13), 8.07 (dd, *J* 8.8, 2.0, 1 H, *H*23), 7.99 (dd, *J* 8.8, 2.0, 1 H, *H*11), 4.36-4.47 (m, 2 H, *H*32), 4.17 (s, 2 H, *H*6), 3.73-3.81 (m, 2 H, *H*33), 3.67-3.73 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.56-3.66 (m, 6 H, *H*28+*PEG-CH*<sub>2</sub>), 3.46-3.56 (m, 8 H, *H*16+*PEG-CH*<sub>2</sub>), 3.38-3.45 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.22 (s, 3 H, *PEG-CH*<sub>3</sub>), 3.21 (s, 3 H, *PEG-CH*<sub>3</sub>), 1.90-2.07 (m, 2 H, *H*17+*H*29), 0.95 (d, *J* 6.6, 6 H, *H*18), 0.88 (d, *J* 6.8, 6 H, *H*30);  $\delta_{\text{C}}$  (101 MHz, CDCl<sub>3</sub>) 170.5, 165.7, 165.8, 164.4, 163.3, 144.2, 142.2, 135.0, 133.4, 133.0, 130.1, 129.0, 124.0, 121.5, 120.7, 119.8, 119.4, 71.9, 71.0, 70.7, 70.7, 70.6, 69.8, 69.6, 69.3, 64.2, 59.1, 59.0, 29.8, 29.6, 20.7, 20.6 (overlapping of *PEG-C*);  $\delta_{\text{C}}$  (101 MHz, DMSO-*d*<sub>6</sub>) 170.2, 165.3, 164.8, 164.7, 163.8, 143.8, 141.8, 134.8, 132.8, 132.3, 130.2, 128.4, 123.7, 121.1, 120.7, 119.0, 118.7, 71.2, 71.0, 70.4, 69.9, 69.7, 69.7, 69.6, 68.6, 68.3, 68.1, 64.1, 58.1, 58.0, 29.3, 29.0, 20.4, 20.4 (overlapping of *PEG-C*); HRMS (ESI) for C<sub>38</sub>H<sub>57</sub>N<sub>4</sub>O<sub>10</sub> [M+H]<sup>+</sup> *m/z* 729.40692 (calc), 729.40588 (found); IR (neat) 2957, 2925 (br), 2870 (br), 1710, 1684, 1635, 1596, 1577, 1514, 1279, 1191, 1102, 1038, 851, 771. X-ray confirmed the structure of compound **8b** (see Section 3.1).

**2-(2-(2-Methoxyethoxy)ethoxy)ethyl 4-(4-(3-(1,3-dioxolan-2-yl)-4-nitrobenzamido)-3-(1,3-dioxolan-2-yl)benzamido)-3-(1,3-dioxolan-2-yl)benzoate (15)**

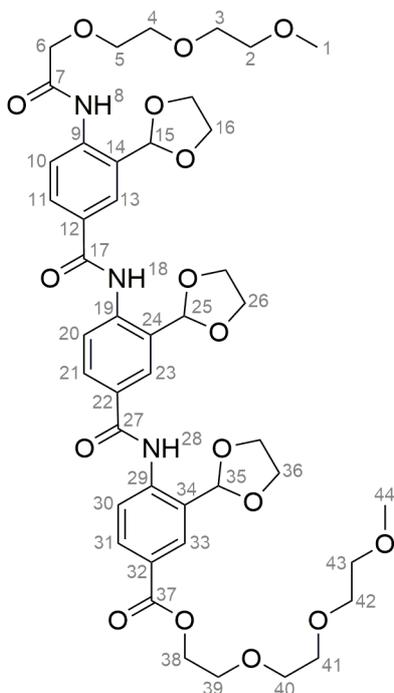


Nitro aromatic **14** (700 mg, 1.21 mmol, 1 eq) was dissolved in ethyl acetate (30 mL) and K<sub>2</sub>CO<sub>3</sub> (35 mg, 0.25 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to stir for 16 h at 45 °C. The solution was filtered over Celite™, washed with ethyl acetate and subsequently concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI) *m/z* 547.3 [M+H]<sup>+</sup>). Compound **11** (397 mg, 1.66 mmol, 1.4 eq) was dissolved in dichloromethane (50 mL) and *N,N*-diisopropylethylamine (529  $\mu$ L, 3.04 mmol, 2.5 eq) was added dropwise, followed by the addition of 2-chloro-1-methylpyridinium iodide (484 mg, 1.89 mmol, 1.56 eq). The mixture was stirred at 45 °C for 30 min. Then, a dichloromethane solution of the aniline

obtained in the first step was added and the mixture was stirred for another 48 h at 45 °C. A saturated NH<sub>4</sub>Cl solution was added and subsequently extracted with dichloromethane (x 3). The combined organic phases were dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (1.25% methanol in dichloromethane, *R*<sub>f</sub>  $\approx$  0.2) to give the title compound **15** as a pale yellow solid (690 mg, 0.90 mmol, 74% yield);  $\delta_{\text{H}}$  (400 MHz, CDCl<sub>3</sub>) 9.97 (s, 1 H, *H*10), 9.86 (s, 1 H, *H*20), 8.66 (d, *J* 8.6, 1 H, *H*12), 8.62 (d, *J* 8.6, 1 H, *H*22),

8.30 (d, *J* 2.0, 1 H, *H*5), 8.15 (d, *J* 2.0, 1 H, *H*25), 8.08-8.14 (m, 3 H, *H*23+*H*3+*H*15), 8.03 (d, *J* 8.6, 1 H, *H*2), 7.96 (dd, *J* 8.6, 2.2, 1 H, *H*13), 6.52 (s, 1 H, *H*7), 6.00 (s, 1 H, *H*17), 5.94 (s, 1 H, *H*27), 4.45-4.51 (m, 2 H, *H*30), 4.03-4.27 (m, 12 H, *H*8+*H*18+*H*28), 3.82-3.87 (m, 2 H, *H*31), 3.71-3.75 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.64-3.70 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.52-3.57 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.38 (s, 3 H, *H*36);  $\delta_c$  (101 MHz, CDCl<sub>3</sub>) 165.8, 164.0, 162.8, 150.6, 141.1, 139.7, 138.3, 134.0, 131.9, 130.1, 129.6, 128.9, 128.6, 127.1, 126.0, 125.8, 125.2, 125.0, 124.7, 121.4, 120.8, 103.4, 103.0, 99.2, 71.9, 70.7, 70.6, 70.6, 69.2, 65.4, 65.1, 65.0, 64.1, 59.0; HRMS (ESI) for C<sub>37</sub>H<sub>41</sub>N<sub>3</sub>O<sub>15</sub>Na [M+Na]<sup>+</sup> *m/z* 790.24299 (calc), 790.24243 (found); IR (neat) 3346 (br), 2950 (br), 2890 (br), 1716, 1682, 1594, 1514, 1272, 1231, 1195, 1095, 1060, 974, 940, 847, 765.

**2-(2-(2-Methoxyethoxy)ethoxy)ethyl 4-(4-(3-(1,3-dioxolan-2-yl)-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzamido)-3-(1,3-dioxolan-2-yl)benzamido)-3-(1,3-dioxolan-2-yl)benzoate**  
(**3b**)

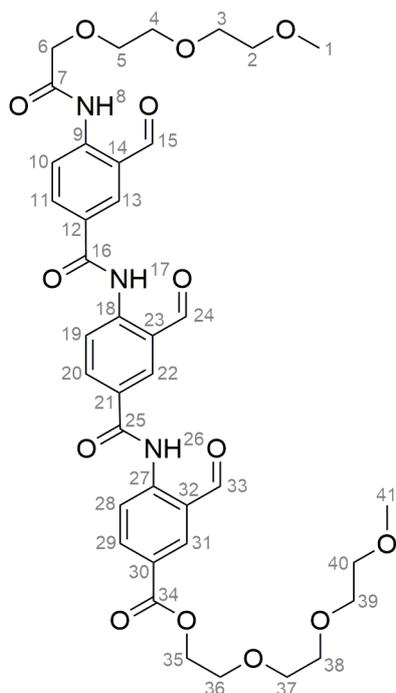


Nitro aromatic **15** (600 mg, 0.78 mmol, 1 eq) was dissolved in ethyl acetate (30 mL) and K<sub>2</sub>CO<sub>3</sub> (21 mg, 0.15 mmol, 0.2 eq) and a catalytic amount of Pd/C were added. The mixture was degassed, put under a hydrogen atmosphere and allowed to stir at 45 °C for 5 h. The solution was filtered over Celite™, washed with ethyl acetate and subsequently concentrated *in vacuo* to give the corresponding aniline (LRMS (ESI) *m/z* 738.5 [M+H]<sup>+</sup>). 2-[2-(2-Methoxyethoxy)ethoxy]acetic acid (276 mg, 1.55 mmol, 2.0 eq) was dissolved in dichloromethane (30 mL) and *N,N*-diisopropylethylamine (342  $\mu$ L, 1.96 mmol, 2.5 eq) was added dropwise, followed by the addition of 2-chloro-1-methylpyridinium iodide (480 mg, 1.88 mmol, 2.4 eq). The mixture was stirred at 45 °C for 30 min. Then, a dichloromethane

solution of the aniline obtained in the first step was added and the mixture was stirred for another 48 h at 45 °C. A saturated NH<sub>4</sub>Cl solution was added and subsequently extracted with dichloromethane (x 3). The combined organic phases were washed with brine dried over MgSO<sub>4</sub> and concentrated *in vacuo*. The residue was purified by flash column chromatography (2.5% methanol in dichloromethane, *R<sub>f</sub>*  $\approx$  0.3, then 6% methanol in ethyl acetate, *R<sub>f</sub>*  $\approx$  0.2) to give *the title compound 3b* as an off-white solid (372 mg, 0.41 mmol, 52% yield);  $\delta_H$  (400 MHz, CDCl<sub>3</sub>) 9.86 (s, 1 H, *H*8), 9.84 (s, 1 H, *H*28), 9.77 (s, 1 H, *H*18), 8.66 (d, *J* 8.3, 1 H, *H*20), 8.64 (d, *J* 8.6, 1 H, *H*30), 8.56 (d, *J* 8.6, 1 H, *H*10), 8.15 (d, *J* 2.0, 1 H, *H*33), 8.12 (dd, *J* 8.3, 2.3, 1 H, *H*31), 8.09 (d, *J* 2.0, 1 H, *H*23), 8.07 (d, *J* 2.3, 1 H, *H*13), 7.94 (dd, *J* 8.3, 2.3, 1 H, *H*21), 7.91 (dd, *J* 8.6, 2.3, 1 H, *H*11), 6.00 (s, 1

H, *H25*), 5.93 (s, 1 H, *H35*), 5.90 (s, 1 H, *H15*), 4.45-4.51 (m, 2 H, *H38*), 4.10-4.27 (m, 14 H, *H6+H16+H26+H36*), 3.80-3.88 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.72-3.77 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.65-3.71 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.53-3.57 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.38 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.37 (s, 3 H, *PEG-CH<sub>3</sub>*);  $\delta_{\text{H}}$  (400 MHz, DMSO-*d*<sub>6</sub>) 10.14 (s, 1 H, *H28*), 10.09 (s, 1 H, *H18*), 9.80 (s, 1 H, *H8*), 8.43 (d, *J* 8.3, 1 H, *Ar-H*), 8.08-8.15 (m, 3 H, *Ar-H*), 7.99-8.06 (m, 5 H, *Ar-H*), 6.10 (s, 1 H, *H25*), 6.08 (s, 1 H, *H35*), 5.88 (s, 1 H, *H15*), 4.39-4.44 (m, 2 H, *H38*), 4.02-4.23 (m, 14 H, *H6+H16+H26+H36*), 3.72-3.78 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.62-3.67 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.59-3.62 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.50-3.57 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.44-3.46 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.39-3.42 (m, 2 H, *PEG-CH<sub>2</sub>*), 3.24 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.22 (s, 3 H, *PEG-CH<sub>3</sub>*);  $\delta_{\text{C}}$  (101 MHz, CDCl<sub>3</sub>) 168.2, 165.9, 164.2, 164.2, 141.3, 140.3, 139.6, 131.9, 129.7, 129.5, 129.4, 128.6, 128.5, 127.4, 126.8, 125.9, 125.8, 124.9, 124.6, 121.5, 121.4, 120.8, 103.5, 103.1, 102.9, 71.9, 71.1, 70.7, 70.7, 70.6, 70.6, 70.4, 69.2, 65.2, 65.1, 65.0, 64.1, 59.0, 59.0 (overlapping of PEG-C);  $\delta_{\text{C}}$  (101 MHz, DMSO-*d*<sub>6</sub>) 168.0, 165.1, 164.2, 164.2, 140.9, 139.8, 139.4, 130.6, 130.1, 130.0, 129.3, 129.0, 128.9, 128.7, 128.5, 128.0, 126.3, 125.6, 125.3, 124.5, 123.6, 120.8, 102.5, 100.6, 100.3, 71.2, 70.4, 70.2, 69.9, 69.7, 69.6, 68.3, 64.8, 64.8, 64.8, 64.1, 58.1, 58.0 (overlapping of PEG-C); HRMS (ESI) for C<sub>44</sub>H<sub>55</sub>N<sub>3</sub>O<sub>17</sub>Na [M+Na]<sup>+</sup> *m/z* 920.34237 (calc), 920.34131 (found); IR (neat) 3350 (br), 2887 (br), 1416, 1679, 1593, 1510, 1397, 1315, 1271, 1233, 1197, 1096, 2062, 975, 914, 765.

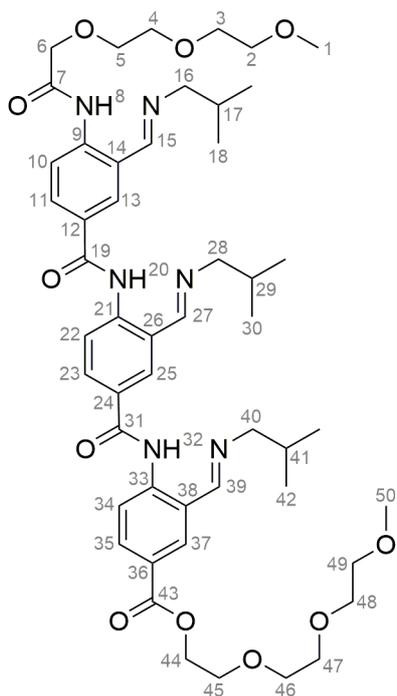
**2-(2-(2-Methoxyethoxy)ethoxy)ethyl 3-formyl-4-(3-formyl-4-(3-formyl-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzamido)benzamido)benzoate (6b)**



Compound **3b** (300 mg, 0.33 mmol) was dissolved in THF (45 mL) and a 10% HCl solution was added (5 mL). The mixture was allowed to stir for 2.5 h. A precipitate was formed, which was filtered off and washed with THF (x1), ethyl acetate (x1), methanol (x1) and diethyl ether (x2). Concentration *in vacuo* gave the title compound **6b** as a white solid (227 mg, 0.30 mmol, 91% yield);  $\delta_{\text{H}}$  (400 MHz, CDCl<sub>3</sub>) 12.41 (s, 1 H, *H26*), 12.38 (s, 1 H, *H17*), 12.09 (s, 1 H, *H8*), 10.17 (s, 1 H, *H24*), 10.11 (s, 1 H, *H15*), 10.09 (s, 1 H, *H33*), 9.13 (d, *J* 9.1, 1 H, *H19*), 9.01 (d, *J* 8.8, 2 H, *H10+H28*), 8.53 (d, *J* 2.0, 1 H, *H22*), 8.50 (d, *J* 1.8, 1 H, *H31*), 8.48 (d, *J* 2.0, 1 H, *H13*), 8.38 (dd, *J* 8.6, 2.0, 1 H, *H20*), 8.36 (dd, *J* 8.3, 1.8, 1 H, *H29*), 8.31 (dd, *J* 8.8, 2.0, 1 H, *H11*), 4.47-4.61 (m, 2 H, *H35*), 4.24 (s, 2 H, *H6*), 3.83-3.94 (m, 6 H, *PEG-CH<sub>2</sub>*), 3.65-3.78 (m, 8 H, *PEG-CH<sub>2</sub>*), 3.52-3.62 (m, 4 H, *PEG-CH<sub>2</sub>*), 3.39 (s, 3 H, *PEG-CH<sub>3</sub>*), 3.38 (s, 3 H, *PEG-CH<sub>3</sub>*);  $\delta_{\text{H}}$  (600 MHz, DMSO-*d*<sub>6</sub>) 12.01 (s, 1 H, *H26*), 11.89 (s, 1 H, *H8*), 11.87 (s, 1 H, *H17*),

10.19 (s, 1 H, *H*24), 10.17 (s, 1 H, *H*33), 10.14 (s, 1 H, *H*15), 8.82 (d, *J* 8.8, 1 H, *H*10), 8.62 (d, *J* 2.6, 1 H, *H*13), 8.62 (d, *J* 8.4, 1 H, *H*28), 8.61 (d, *J* 2.7, 1 H, *H*22), 8.59 (d, *J* 8.6, 1 H, *H*19), 8.57 (d, *J* 2.2, 1 H, *H*31), 8.36 (dd, *J* 8.8, 2.20, 1 H, *H*20), 8.33 (dd, *J* 8.8, 2.2, 1 H, *H*11), 8.31 (dd, *J* 8.6, 1.8, 1 H, *H*29), 4.43-4.47 (m, 2 H, *H*35), 4.22 (s, 2 H, *H*6), 3.76-3.79 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.70-3.73 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.60-3.62 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.56-3.58 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.53-3.56 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.51-3.53 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.43-3.45 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.40-3.42 (m, 2 H, *PEG-CH*<sub>2</sub>), 3.23 (s, 3 H, *PEG-CH*<sub>3</sub>), 3.22 (s, 3 H, *PEG-CH*<sub>3</sub>);  $\delta_c$  (101 MHz, CDCl<sub>3</sub>) 195.8, 195.7, 194.2, 170.3, 164.8, 164.1, 164.0, 144.4, 144.1, 142.9, 138.0, 137.5, 136.3, 136.0, 134.8, 134.3, 128.4, 128.3, 125.0, 122.2, 121.8, 121.4, 120.3, 120.2, 119.7, 71.9, 71.6, 71.3, 70.7, 70.7, 70.6, 69.1, 64.5, 59.0 (overlapping of *PEG-C*); HRMS (ESI) for C<sub>38</sub>H<sub>43</sub>N<sub>3</sub>O<sub>14</sub>Na [M+Na]<sup>+</sup> *m/z* 788.26372 (calc), 788.26373 (found); IR (neat) 3229 (br), 2874 (br), 1720, 1667, 1584, 1517, 1394, 1262, 1177, 1099, 841, 784, 763, 743, 694.

**2-(2-(2-Methoxyethoxy)ethoxy)ethyl 3-((*E*)-(isobutylimino)methyl)-4-(3-((*E*)-(isobutylimino)methyl)-4-(3-((*E*)-(isobutylimino)methyl)-4-(2-(2-(2-methoxyethoxy)ethoxy)acetamido)benzamido)benzoate (**9b**)**



Compound **6b** (60 mg, 0.078 mmol, 1.0 eq) and isobutylamine (35  $\mu$ L, 0.352 mmol, 4.5 eq) were dissolved in toluene (20 mL) and 3 Å activated molecular sieves were added. The mixture was refluxed for 16 h, after which the molecular sieves were filtered off and washed with hot ethyl acetate. The mixture was concentrated *in vacuo* and purified by flash column chromatography over alumina (basic, activated, 3:1 EtOAc:hexane, *R<sub>f</sub>*  $\approx$  0.2) to give the title compound **9b** as an off-white solid (30.6 mg, 0.033 mmol, 42% yield);  $\delta_H$  (400 MHz, CDCl<sub>3</sub>) 13.79 (s, 1 H, *H*32), 13.76 (s, 1 H, *H*20), 13.25 (s, 1 H, *H*8), 9.07 (d, *J* 8.8, 1 H, *H*22), 9.02 (d, *J* 9.3, 1 H, *H*34), 8.95 (d, *J* 8.8, 1 H, *H*10), 8.49 (s, 1 H, *H*27), 8.46 (s, 1 H, *H*39), 8.41 (s, 1 H, *H*15), 8.09-8.20 (m, 5 H, *H*13+*H*23+*H*25+*H*35+*H*37). 8.06 (dd, *J* 8.8, 1.8, 1 H, *H*11), 4.45-4.57 (m, 2 H, *H*44), 4.25 (s, 2 H, *H*6), 3.79-3.94 (m, 4 H, *PEG-CH*<sub>2</sub>), 3.63-3.79 (m, 10 H, *PEG-CH*<sub>2</sub>), 3.53-3.63 (m, 8 H, *H*28+*H*40+*PEG-CH*<sub>2</sub>), 3.51 (d, *J* 6.6, 2 H, *H*16), 3.39 (s, 6 H, *PEG-CH*<sub>3</sub>), 2.01-2.15 (m, 3 H, *H*17+*H*29+*H*41), 0.96-1.05 (m, 18 H, *H*18+*H*30+ *H*42);  $\delta_H$  (400 MHz, DMSO-*d*<sub>6</sub>) 13.84 (s, 1 H, *H*32), 13.75 (s, 1 H, *H*20), 13.15 (s, 1 H, *H*8), 8.90 (d, *J* 8.8, 1 H, *H*22), 8.87 (d, *J* 8.6, 1 H, *H*34), 8.83 (d, *J* 8.8, 1 H, *H*10), 8.69 (s, 1 H, *H*39), 8.64 (s, 1 H, *H*27), 8.54 (s, 1 H, *H*15), 8.20 (d, *J* 2.0, 1 H, *H*37), 8.14 (d, *J* 1.8, 1 H, *H*25), 8.09 (d, *J* 2.0, 1 H, *H*13), 8.05 (dd, *J* 8.8, 2.0, 1

H, *H35*), 8.03 (dd, *J* 8.8, 2.0, 1 H, *H23*), 7.98 (dd, *J* 8.7, 1.9, 1 H, *H11*), 4.37-4.43 (m, 2 H, *H44*), 4.14-4.21 (m, 2 H, *H6*), 3.74-3.79 (m, 2 H, *PEG-CH2*), 3.68-3.73 (m, 2 H, *PEG-CH2*), 3.58-3.66 (m, 8 H, *H28+H40+PEG-CH2*) 3.46-3.57 (m, 8 H, *H16+PEG-CH2*), 3.38-3.46 (m, 4 H, *PEG-CH2*), 3.24 (s, 3 H, *PEG-CH3*), 3.22 (s, 3 H, *PEG-CH3*), 1.91-2.07 (m, 3 H, *H17+H29+H41*), 0.88 - 0.97 (m, 18 H, *H18+H30+H42*);  $\delta_C$  (101 MHz,  $CDCl_3$ ) 170.5, 165.7, 165.6, 165.5, 164.4, 163.2, 144.2, 143.4, 142.2, 135.0, 133.8, 133.4, 133.0, 130.5, 130.2, 129.0, 128.7, 124.0, 121.5, 121.1, 120.7, 119.8, 119.5, 119.4, 71.9, 71.9, 71.0, 70.7, 70.7, 70.6, 70.6, 69.9, 69.8, 69.7, 69.3, 64.2, 59.1, 59.1, 29.9, 29.9, 29.7, 20.7, 20.7, 20.6 (overlapping of carbonyl-C and PEG-C);  $\delta_C$  (101 MHz,  $DMSO-d_6$ ) 170.2, 165.3, 165.0, 164.8, 164.6, 164.5, 163.8, 143.6, 142.8, 141.8, 134.8, 133.1, 132.8, 132.3, 130.5, 130.2, 128.3, 128.1, 123.7, 121.1, 120.7, 120.6, 119.0, 118.8, 118.7, 71.3, 71.0, 70.4, 69.9, 69.7, 69.7, 69.6, 68.6, 68.3, 68.1, 64.0, 58.1, 58.0, 29.4, 29.4, 29.0, 20.4, 20.4, 20.3 (overlapping of PEG-C); HRMS (ESI) for  $C_{50}H_{70}N_6O_{11}Na$   $[M+Na]^+$   $m/z$  953.49948 (calc), 953.49829 (found); IR (neat) 2956, 2923 (br), 2868 (br), 1716, 1672, 1635, 1595, 1511, 1277, 1193, 1101, 1039, 860, 761.

## 2. 1D NMR spectra

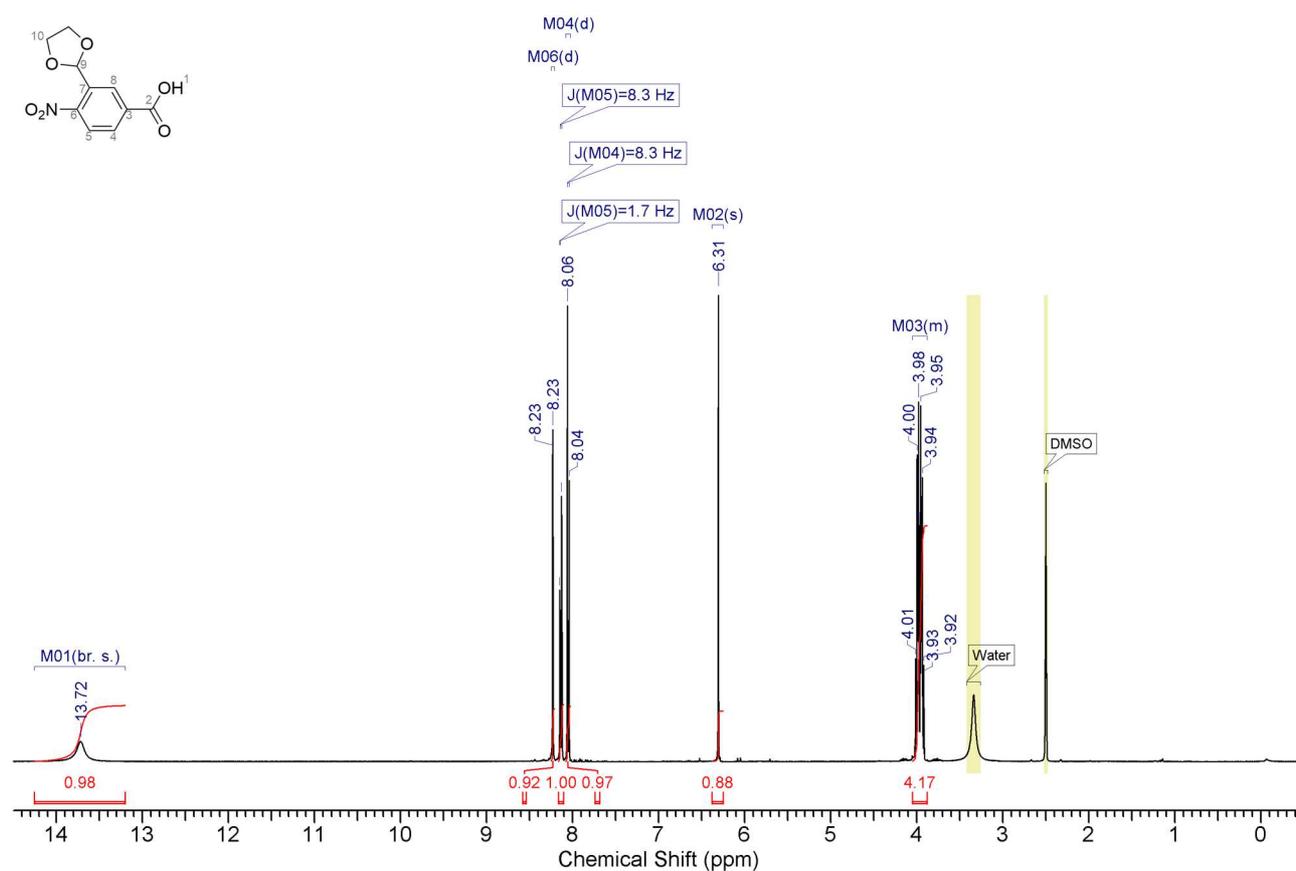


Figure S1. <sup>1</sup>H NMR spectrum (400 MHz) of compound **11** in DMSO-*d*<sub>6</sub> at 298 K.

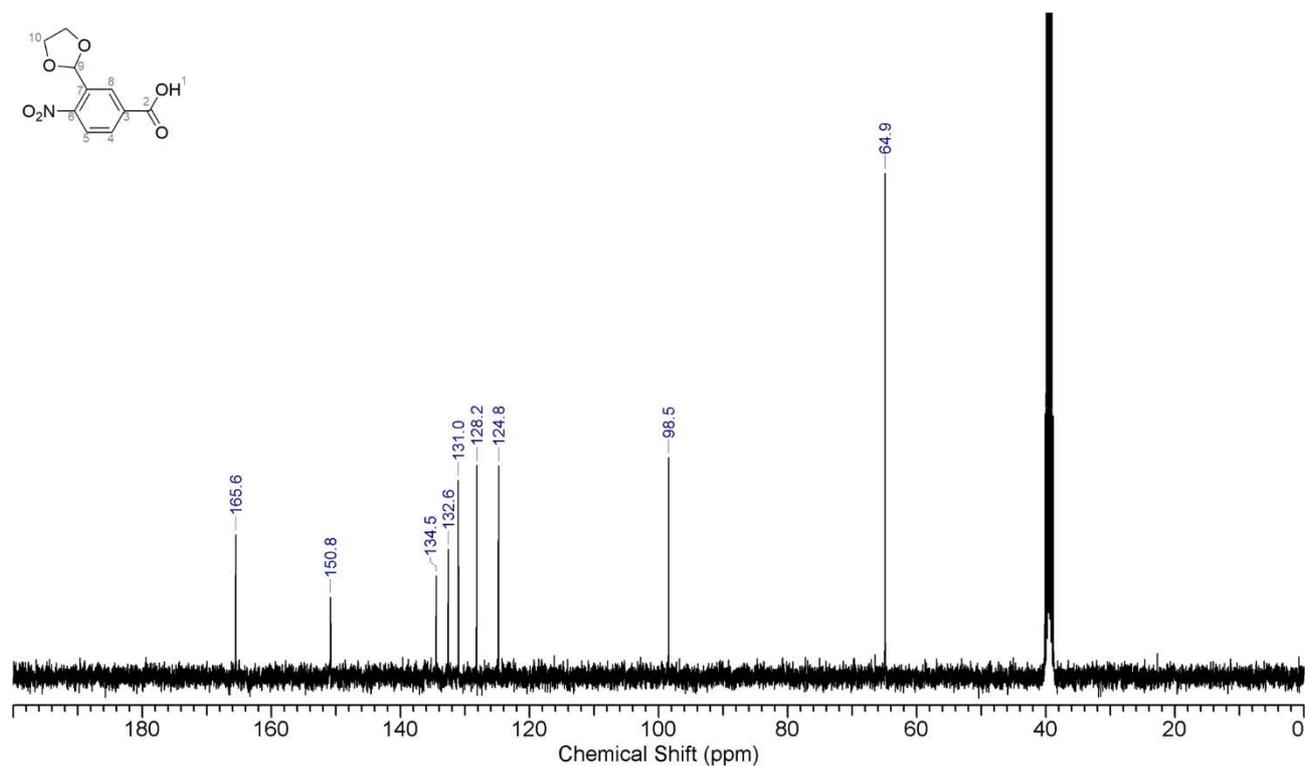
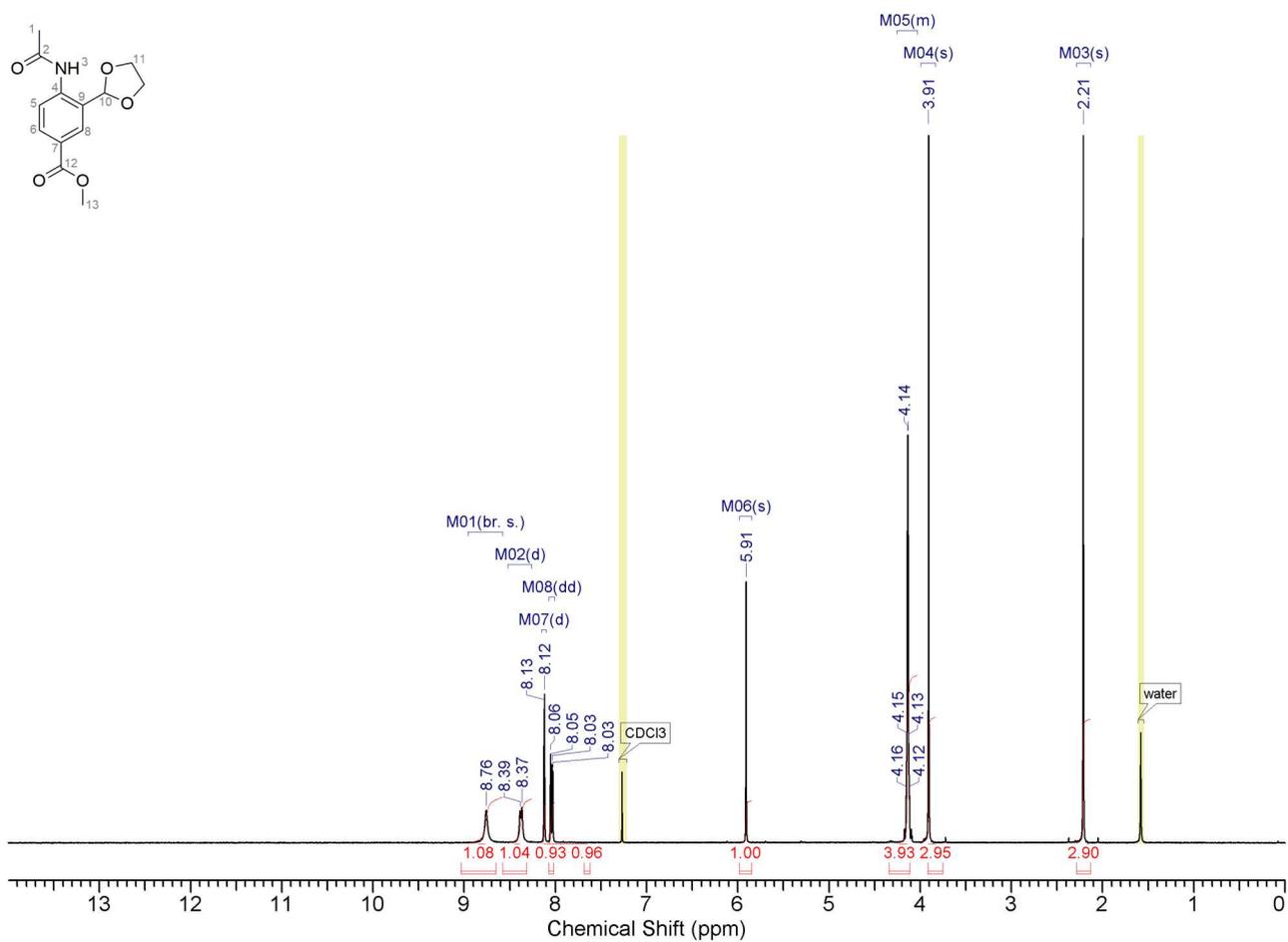
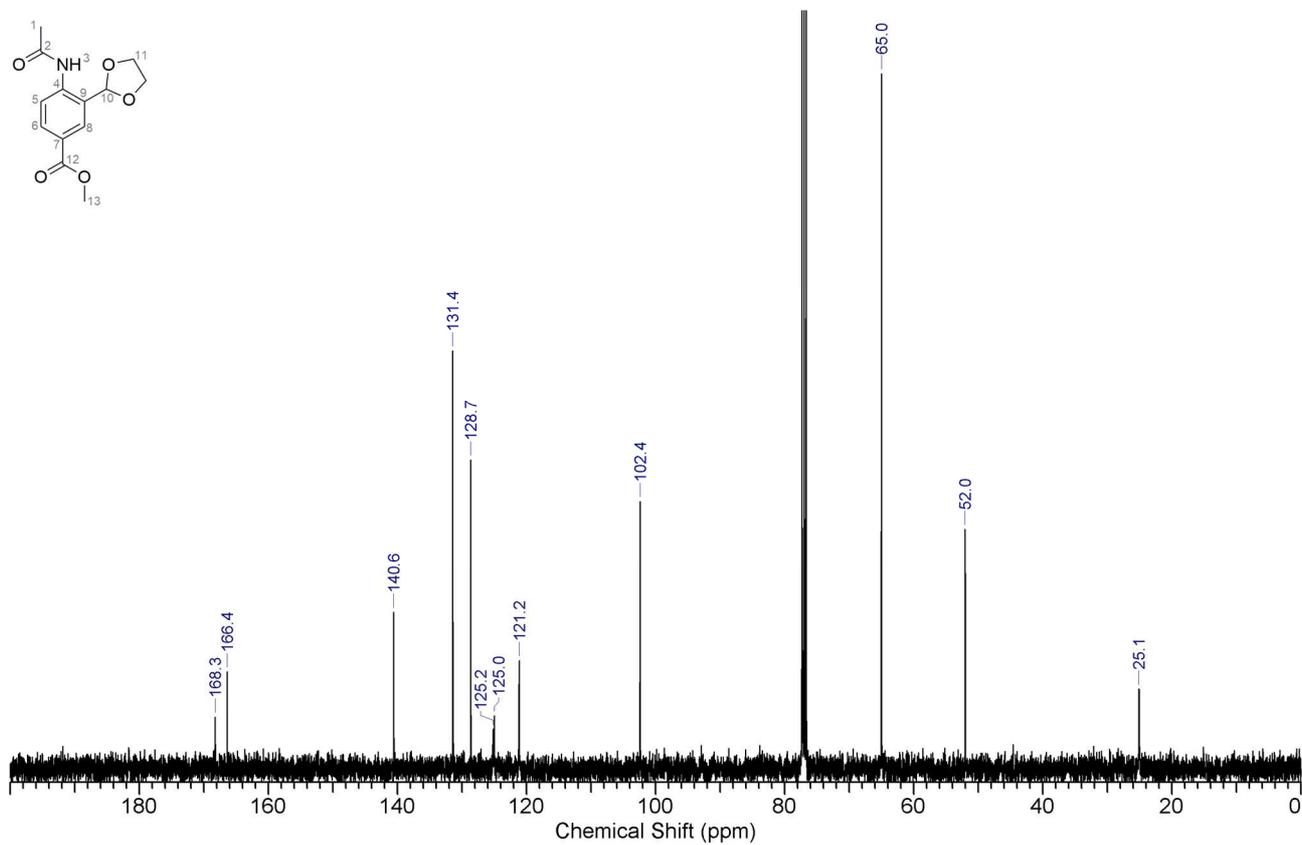


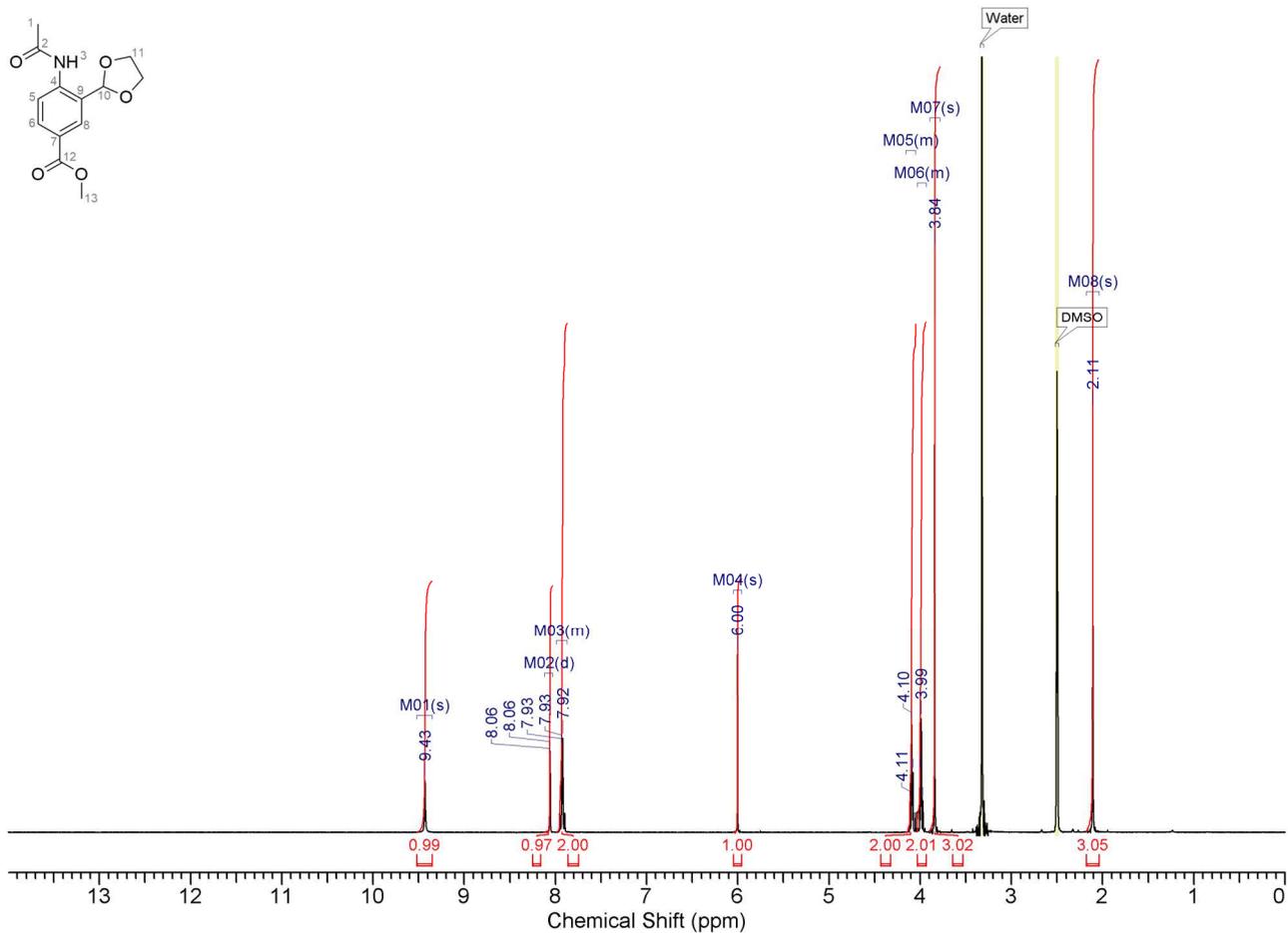
Figure S2. <sup>13</sup>C NMR spectrum (101 MHz) of compound **11** in DMSO-*d*<sub>6</sub> at 298 K.



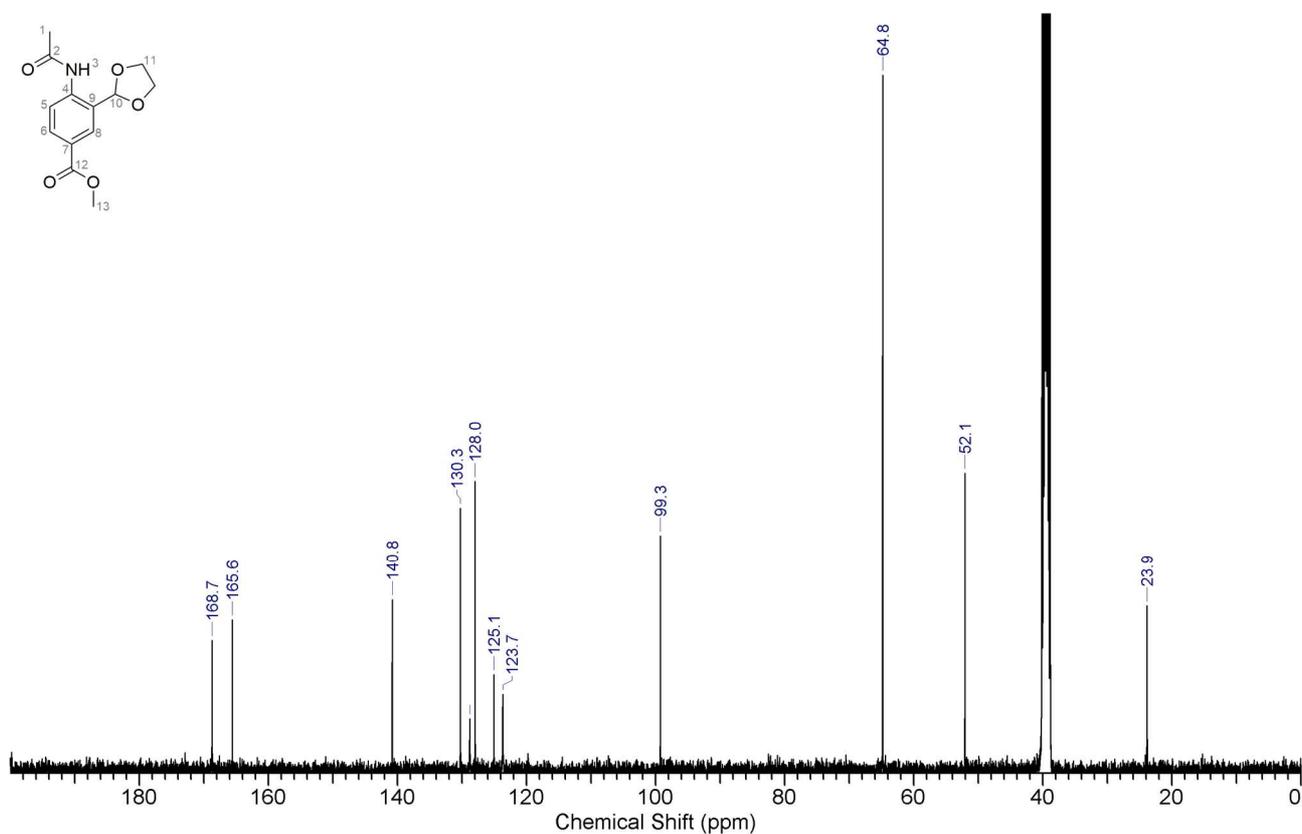
**Figure S3.** <sup>1</sup>H NMR spectrum (400 MHz) of compound **1a** in CDCl<sub>3</sub> at 298 K.



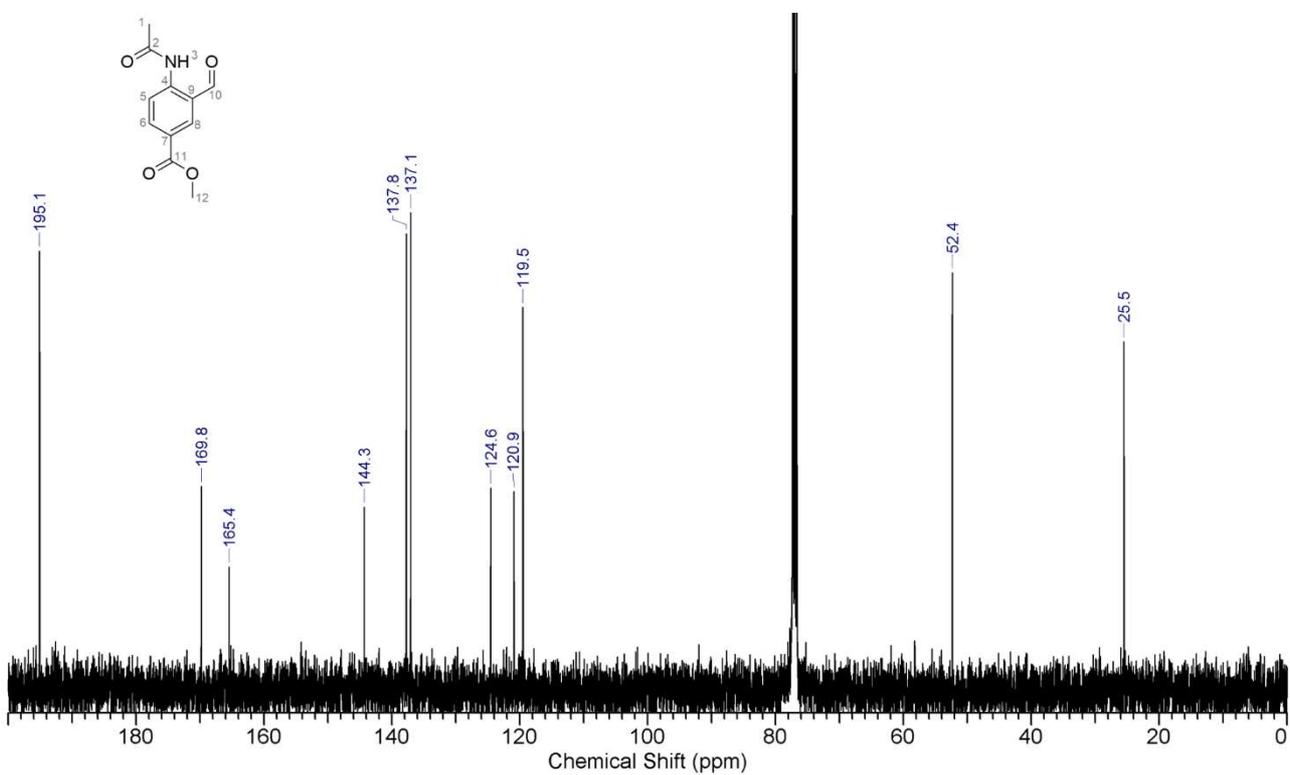
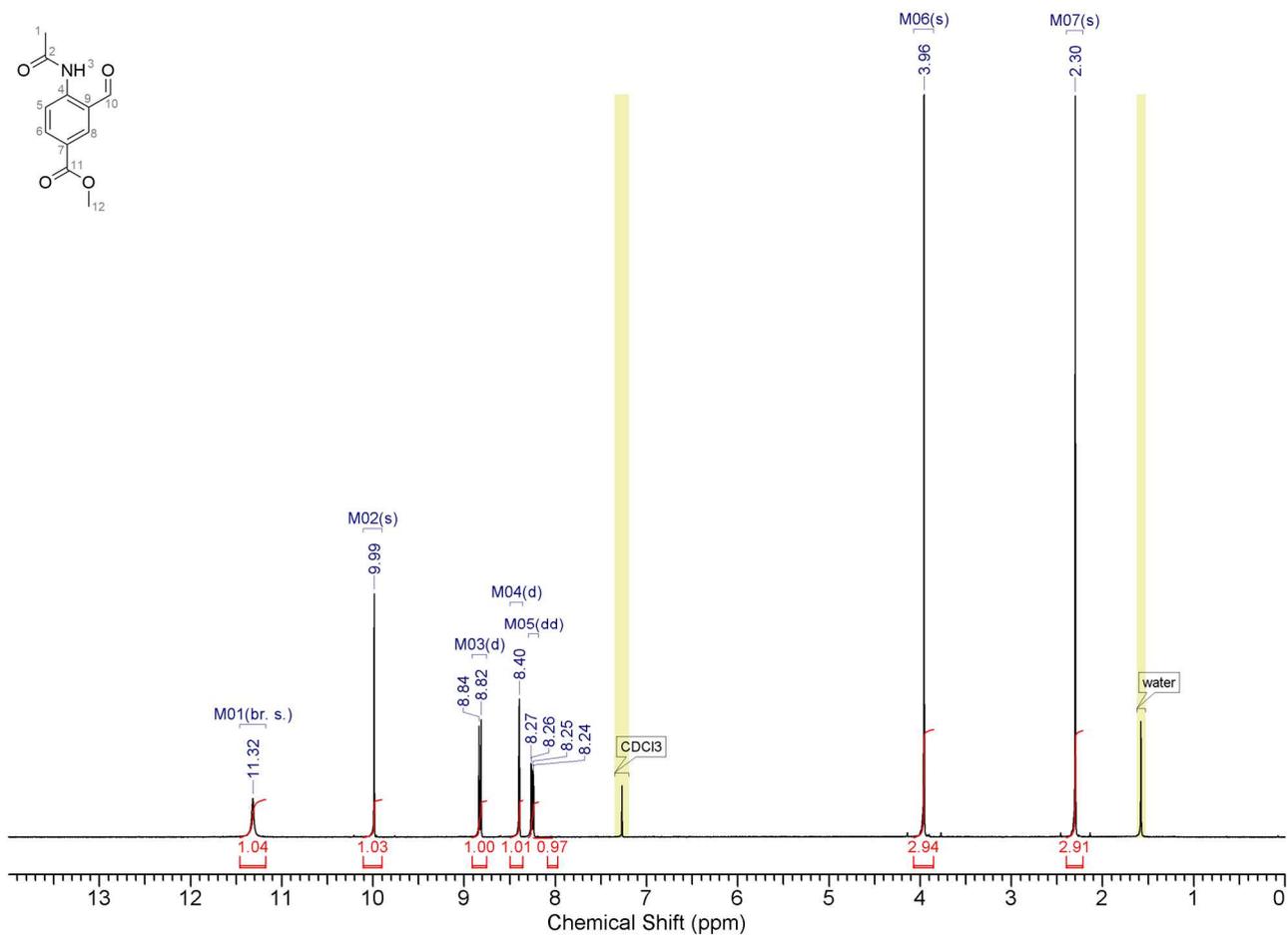
**Figure S4.** <sup>13</sup>C NMR spectrum (101 MHz) of compound **1a** in CDCl<sub>3</sub> at 298 K.

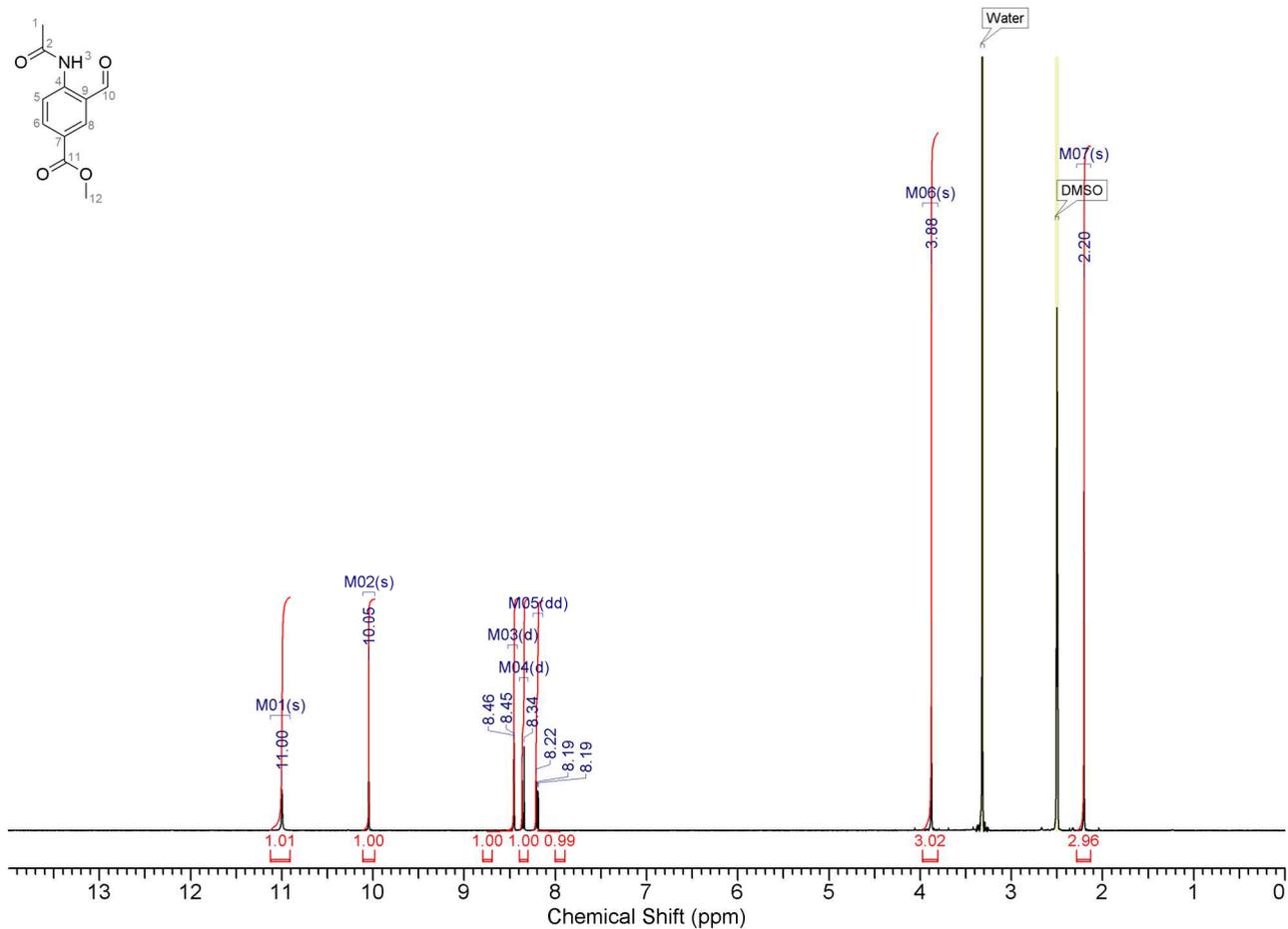


**Figure S5.** <sup>1</sup>H NMR spectrum (400 MHz) of compound **1a** in DMSO-*d*<sub>6</sub> at 298 K.

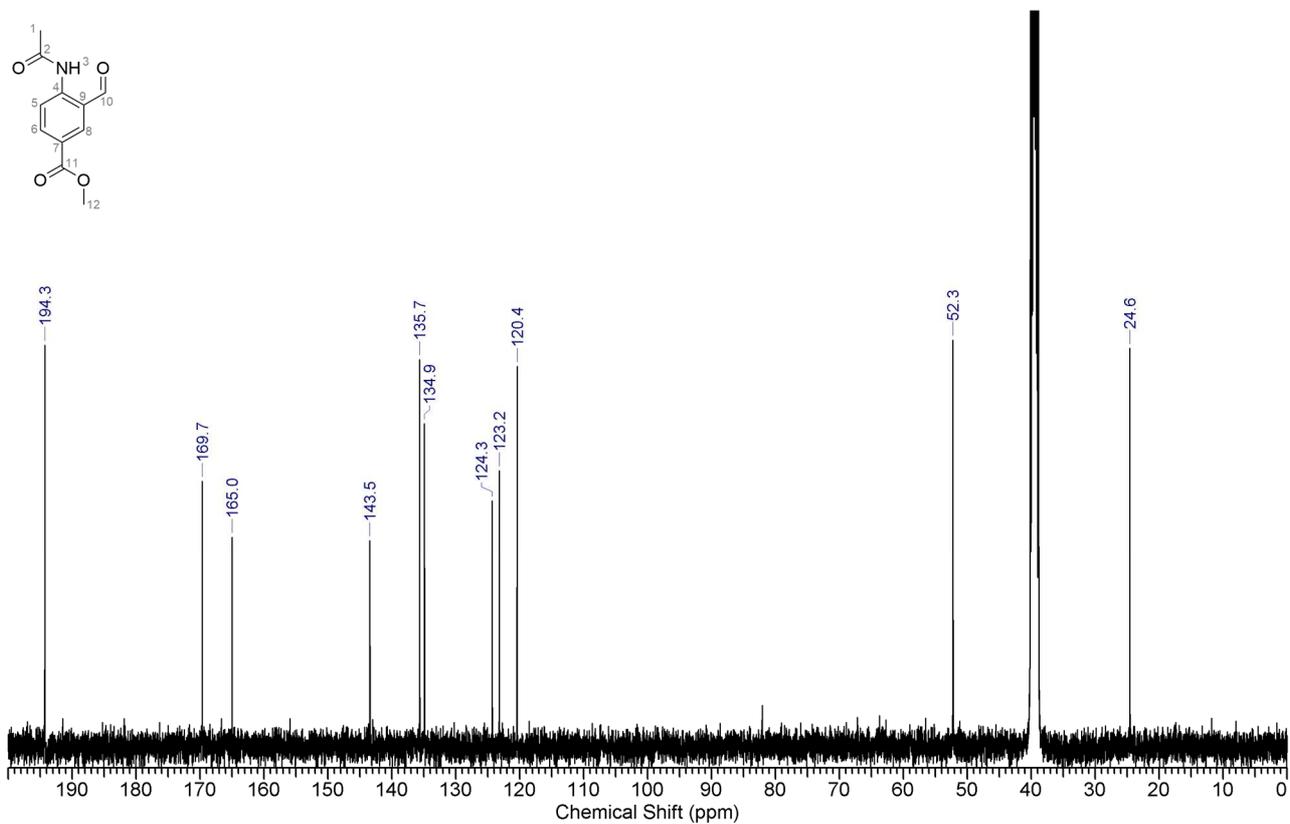


**Figure S6.** <sup>13</sup>C NMR spectrum (101 MHz) of compound **1a** in DMSO-*d*<sub>6</sub> at 298 K.

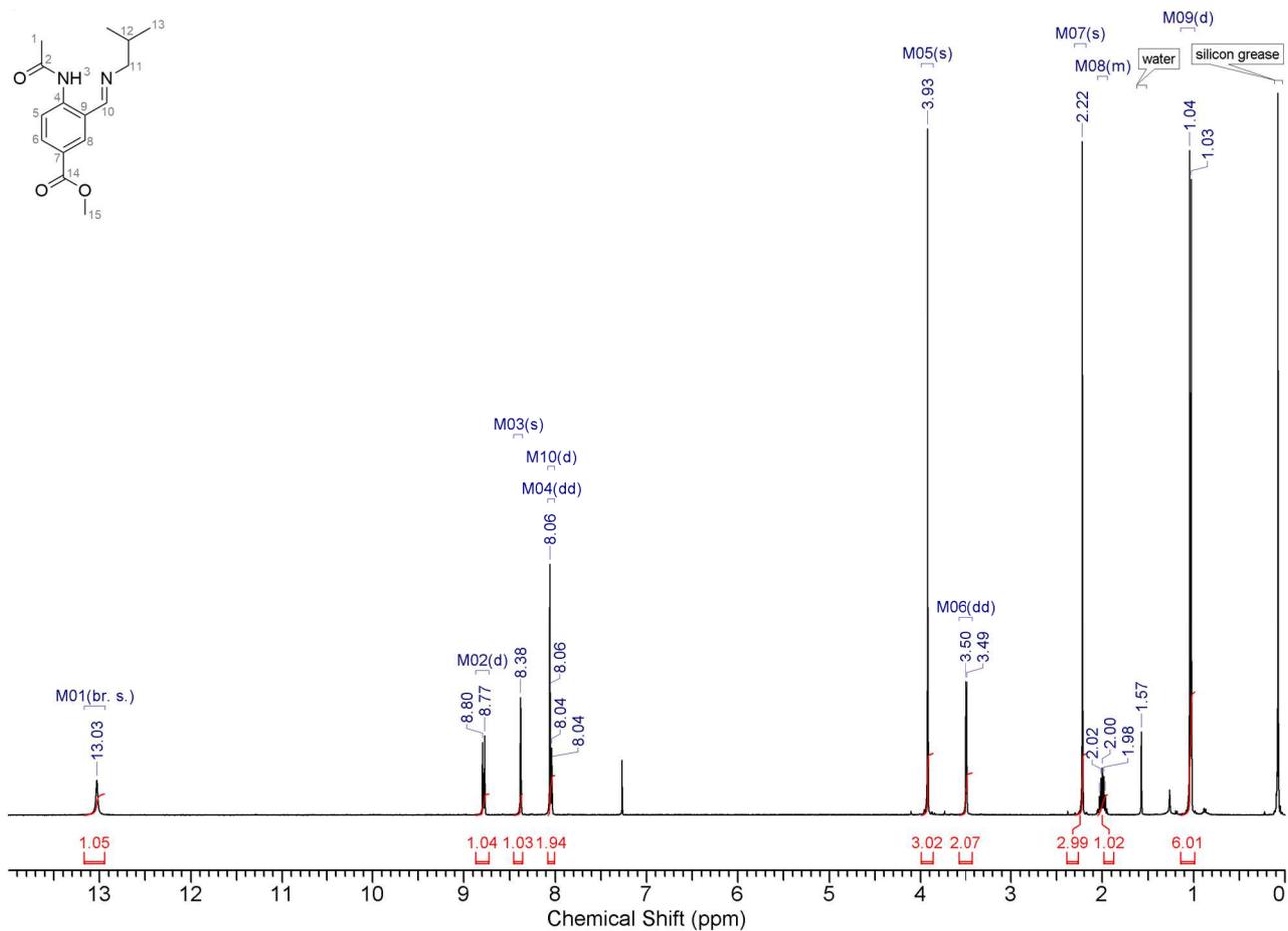




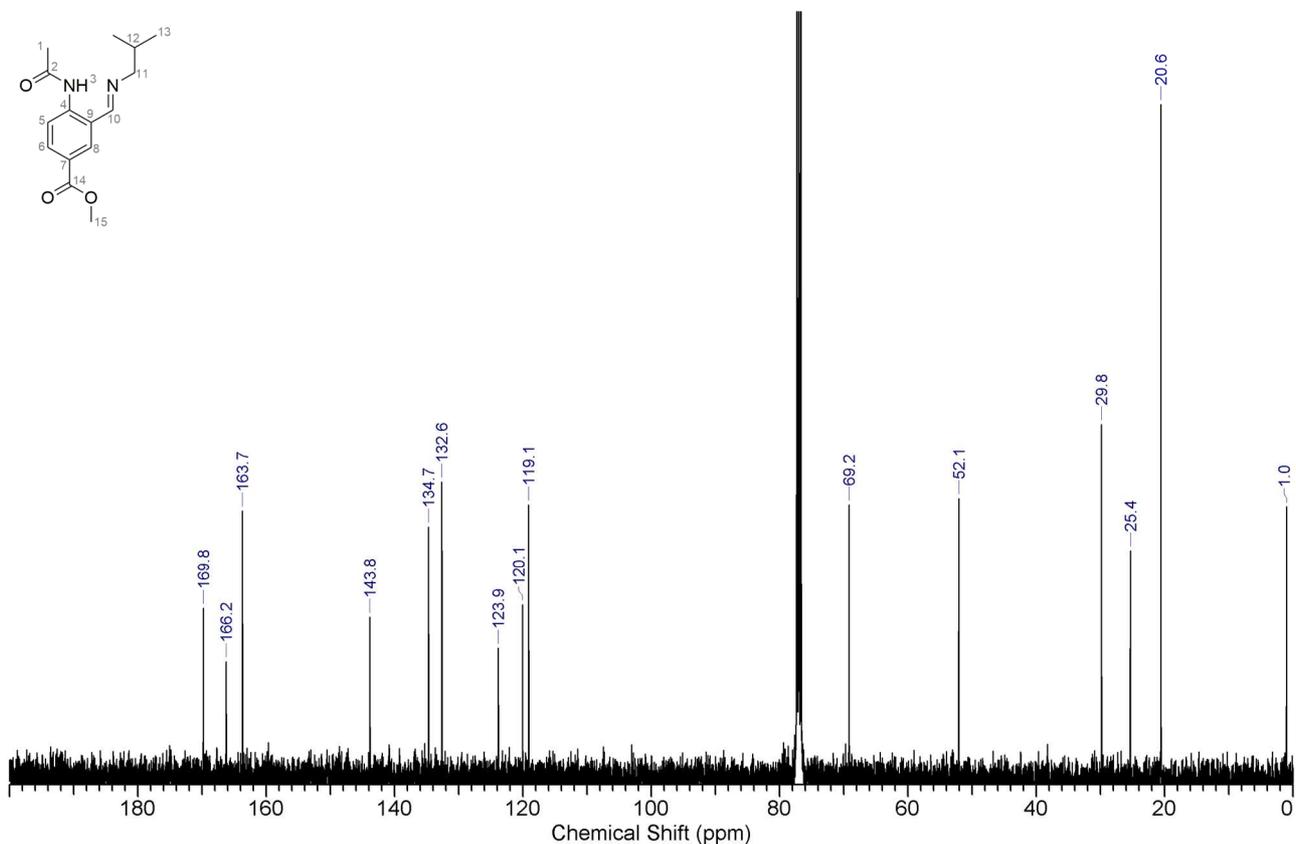
**Figure S9.** <sup>1</sup>H NMR spectrum (400 MHz) of compound **4a** in DMSO-*d*<sub>6</sub> at 298 K.



**Figure S10.** <sup>13</sup>C NMR spectrum (101 MHz) of compound **4a** in DMSO-*d*<sub>6</sub> at 298 K.



**Figure S11.**  $^1\text{H}$  NMR spectrum (400 MHz) of compound **7a** in  $\text{CDCl}_3$  at 298 K.



**Figure S12.**  $^{13}\text{C}$  NMR spectrum (101 MHz) of compound **7a** in  $\text{CDCl}_3$  at 298 K.

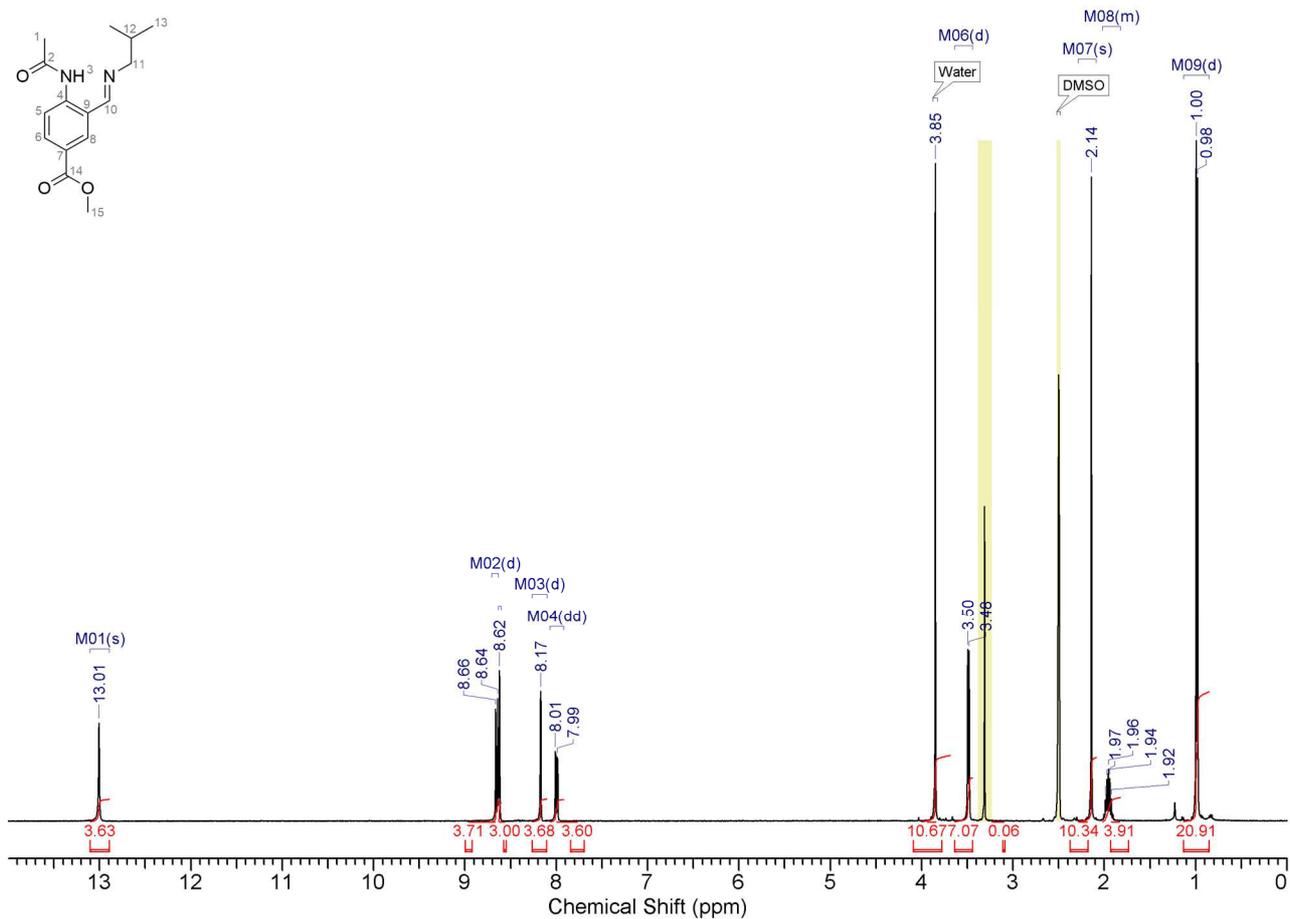


Figure S13. <sup>1</sup>H NMR spectrum (400 MHz) of compound **7a** in DMSO-*d*<sub>6</sub> at 298 K.

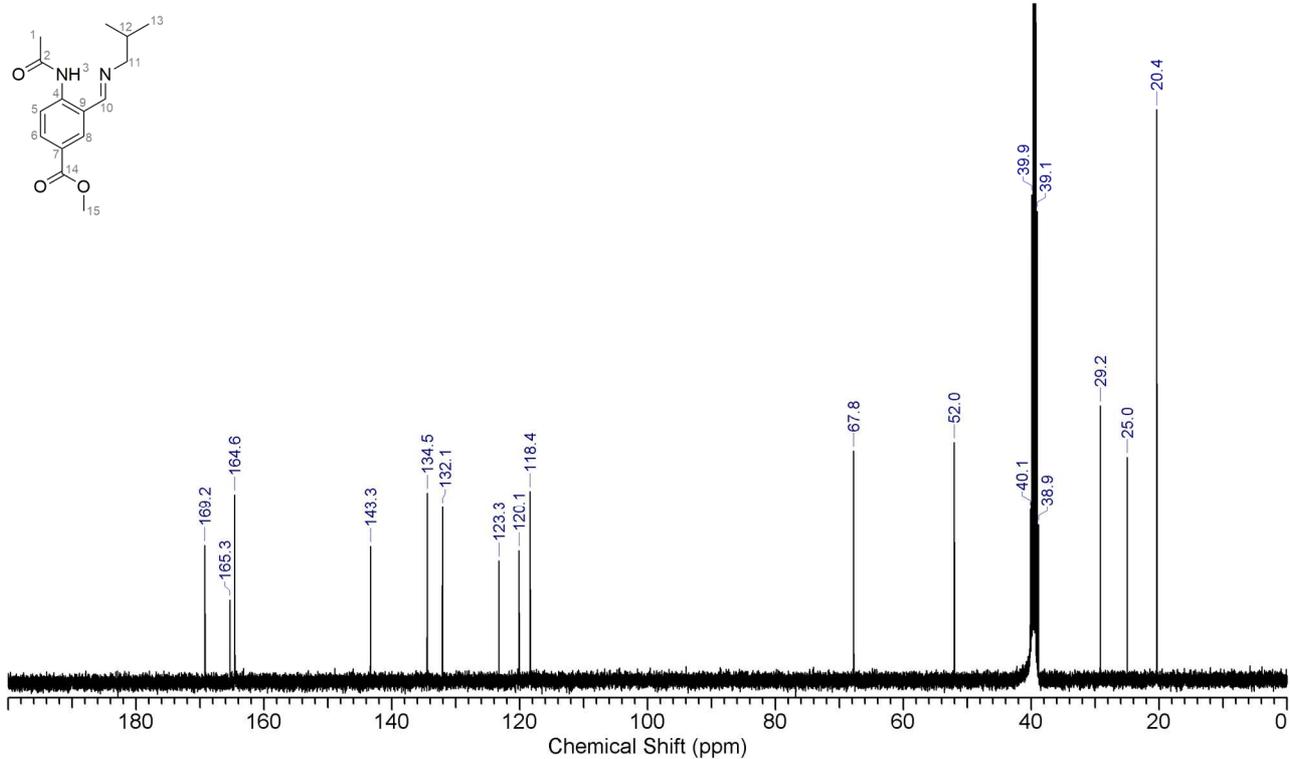
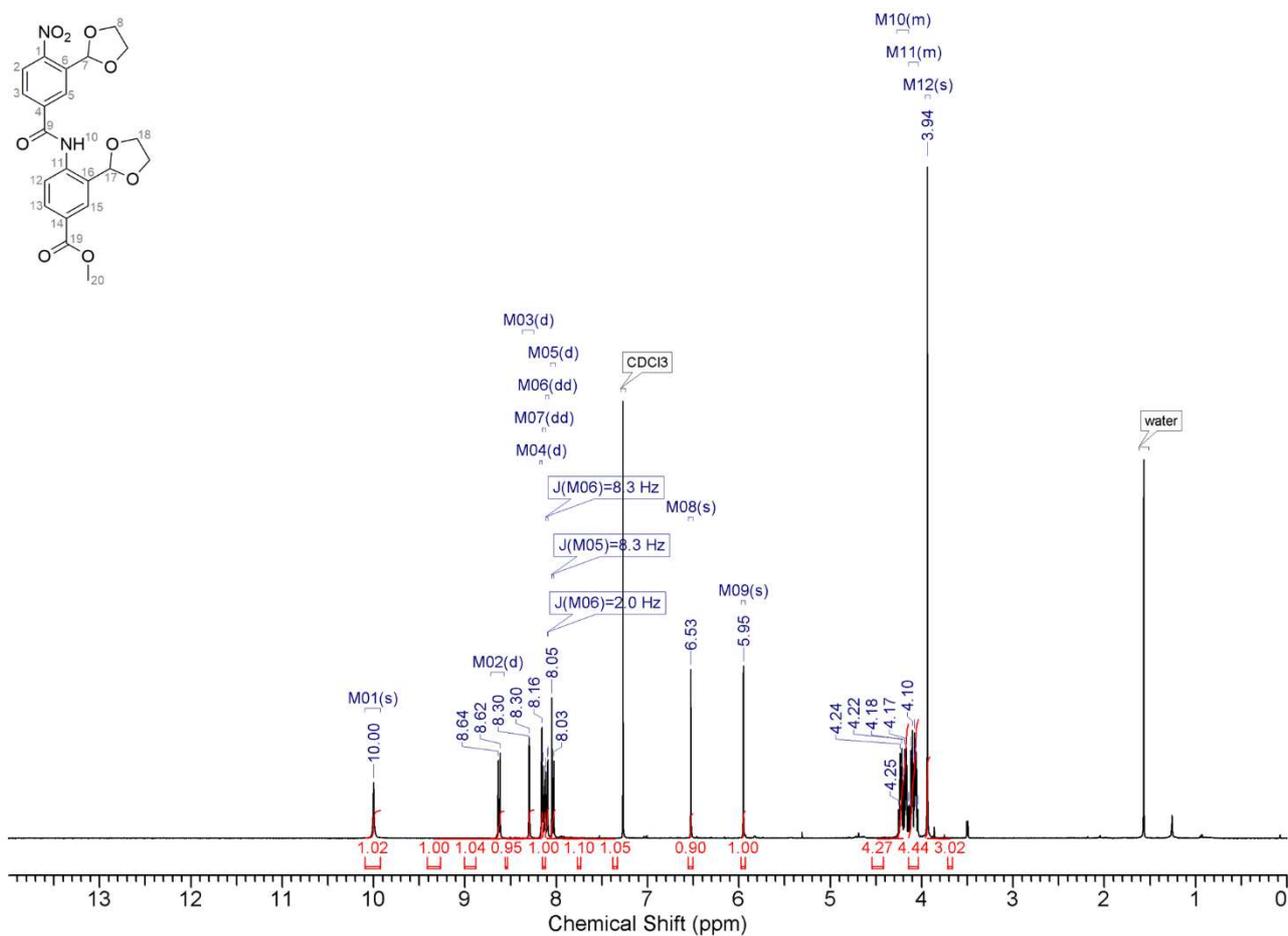
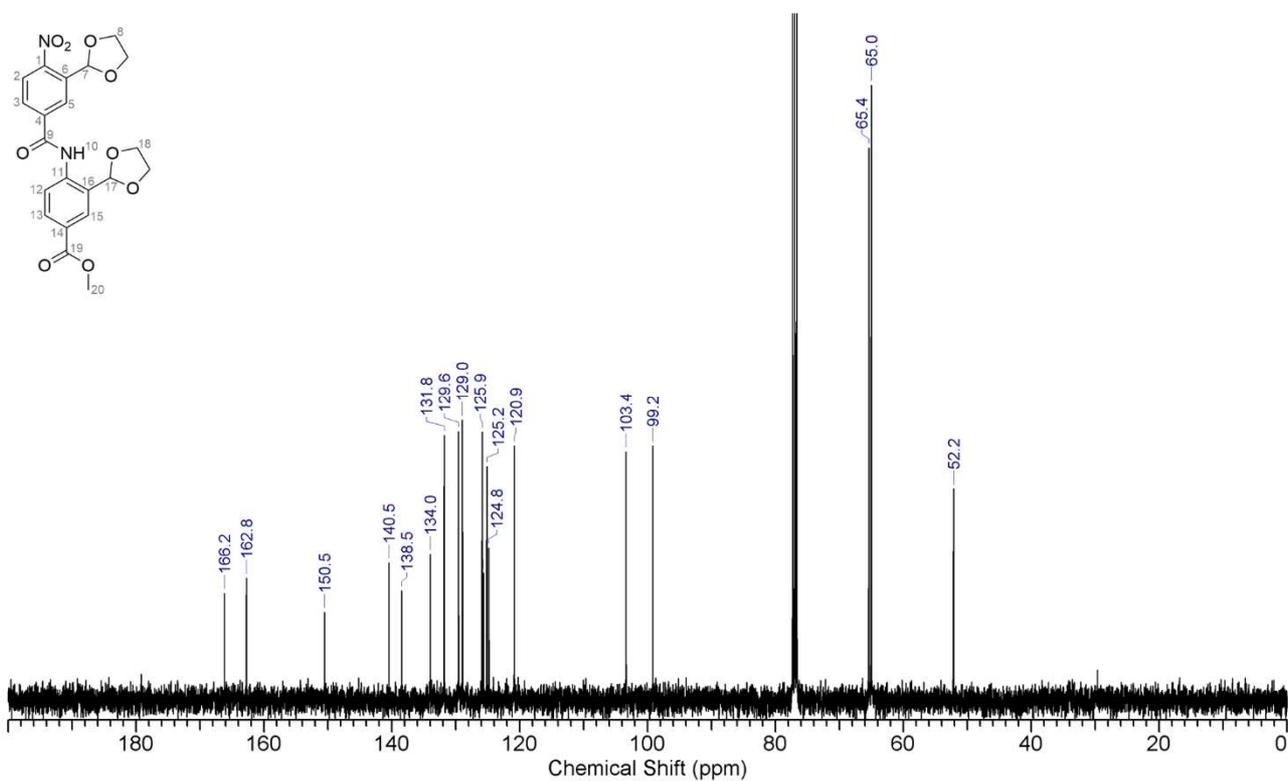


Figure S14. <sup>13</sup>C NMR spectrum (101 MHz) of compound **7a** in DMSO-*d*<sub>6</sub> at 298 K.



**Figure S15.** <sup>1</sup>H NMR spectrum (400 MHz) of compound **12** in CDCl<sub>3</sub> at 298 K.



**Figure S16.** <sup>13</sup>C NMR spectrum (101 MHz) of compound **12** in CDCl<sub>3</sub> at 298 K.

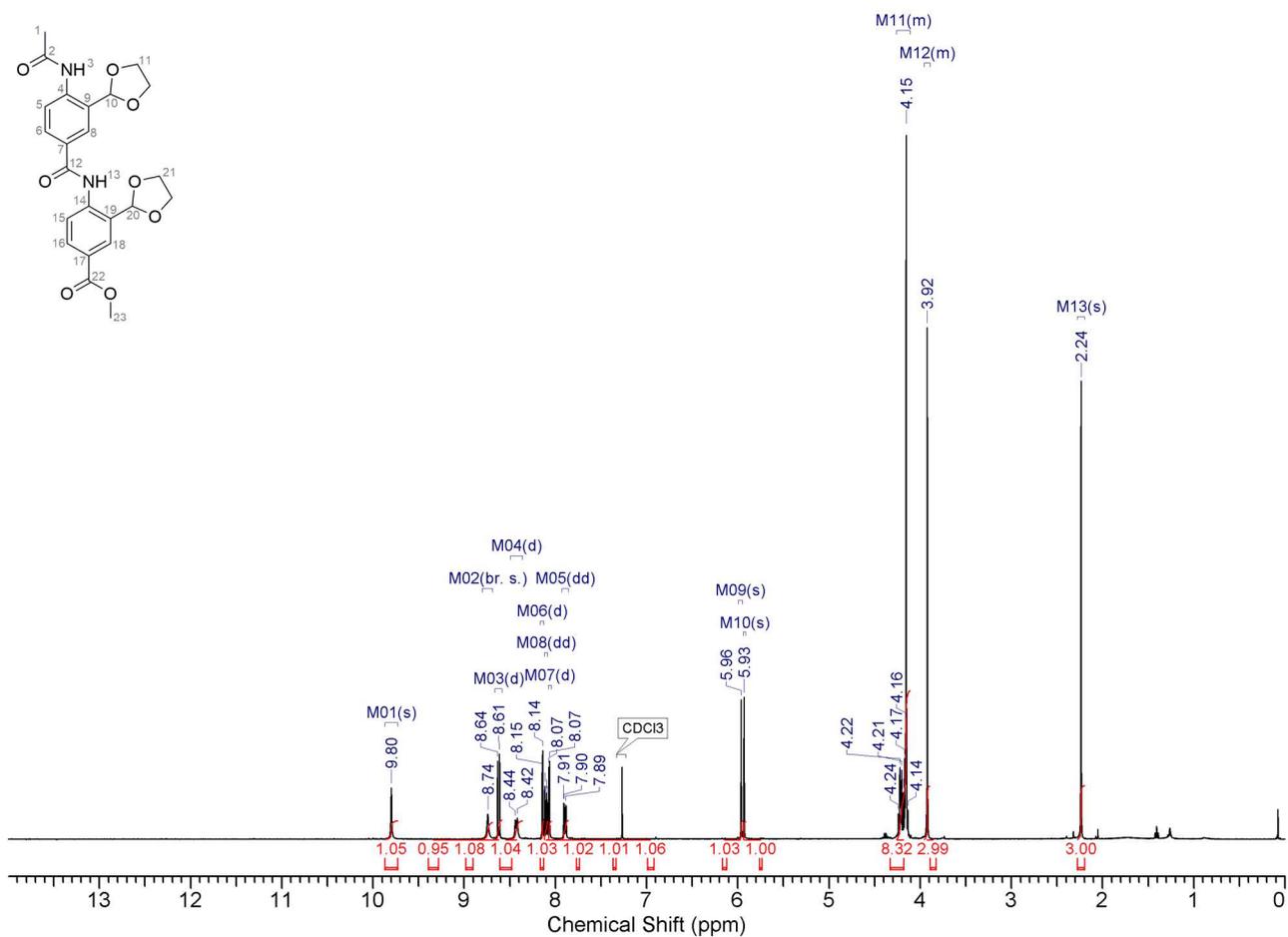


Figure S17. <sup>1</sup>H NMR spectrum (400 MHz) of compound **2a** in CDCl<sub>3</sub> at 298 K.

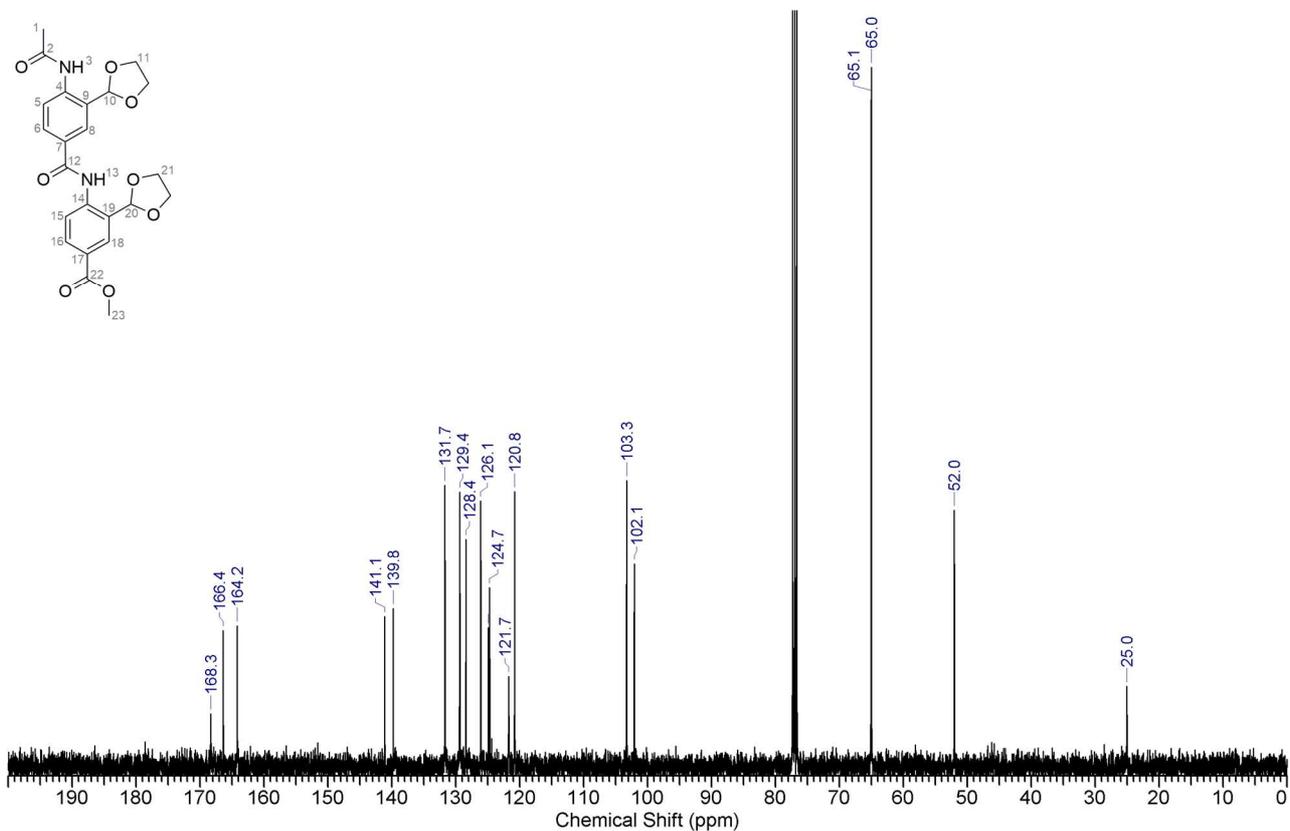
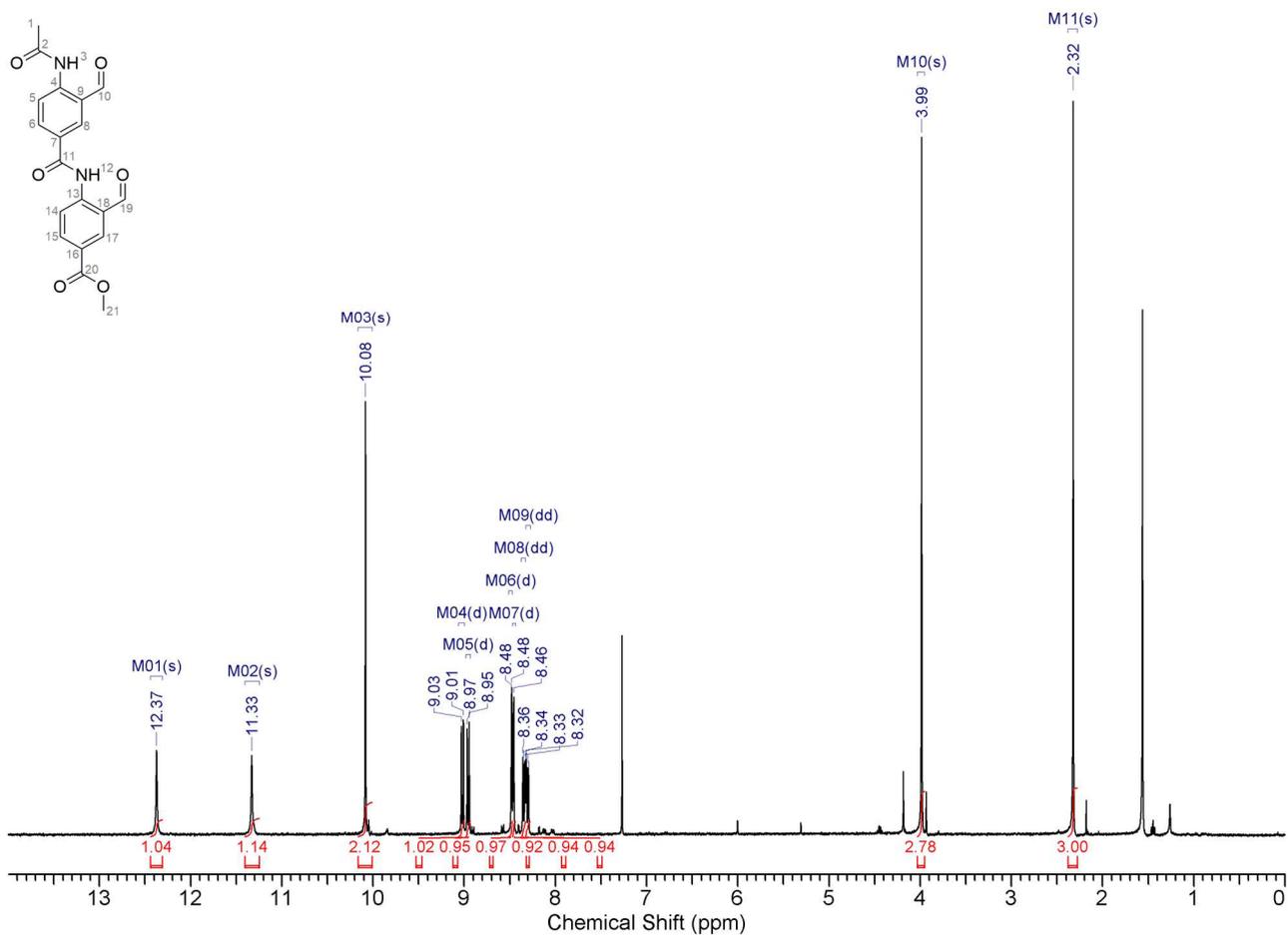
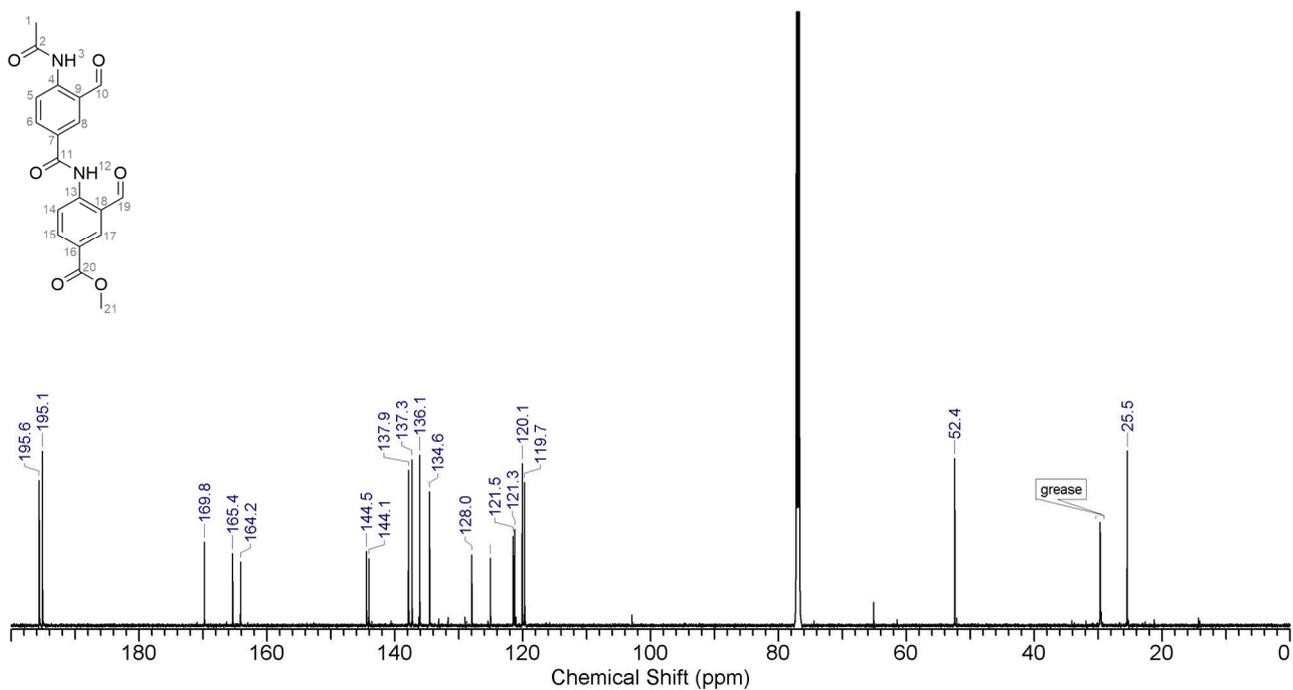


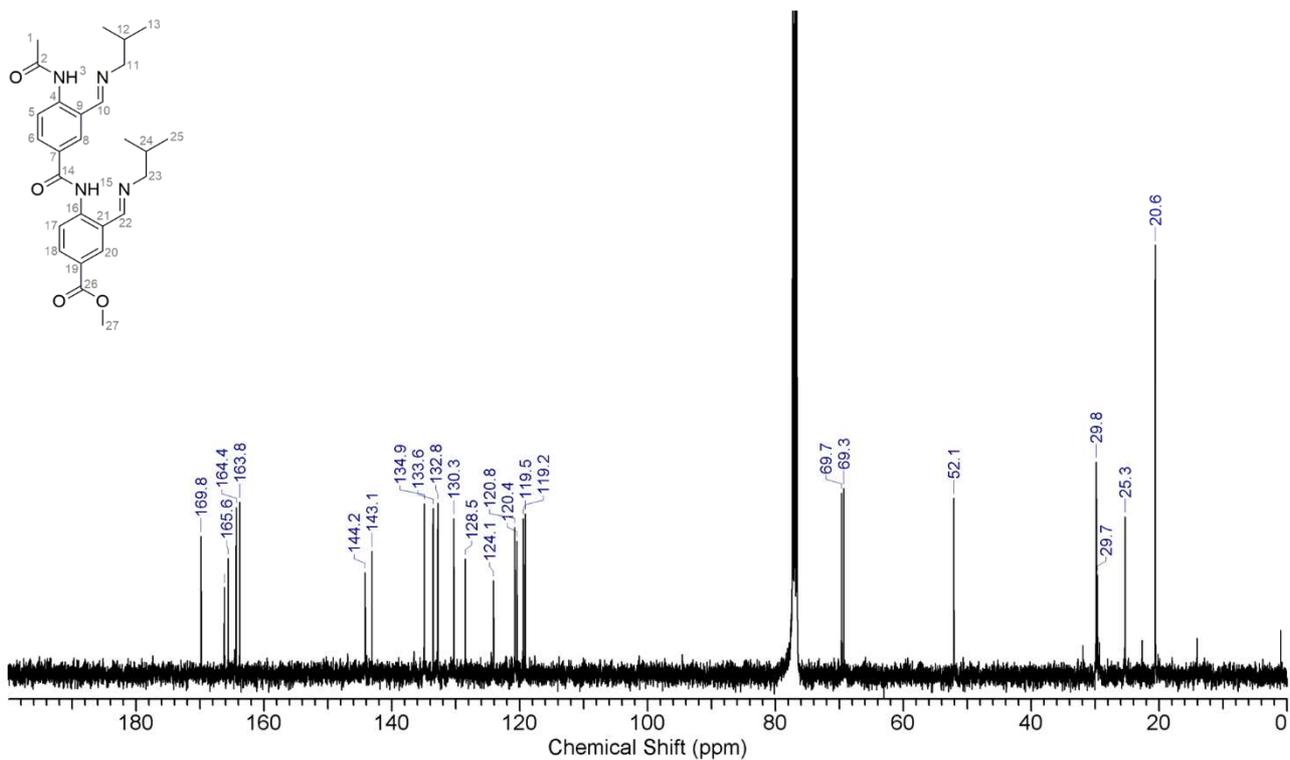
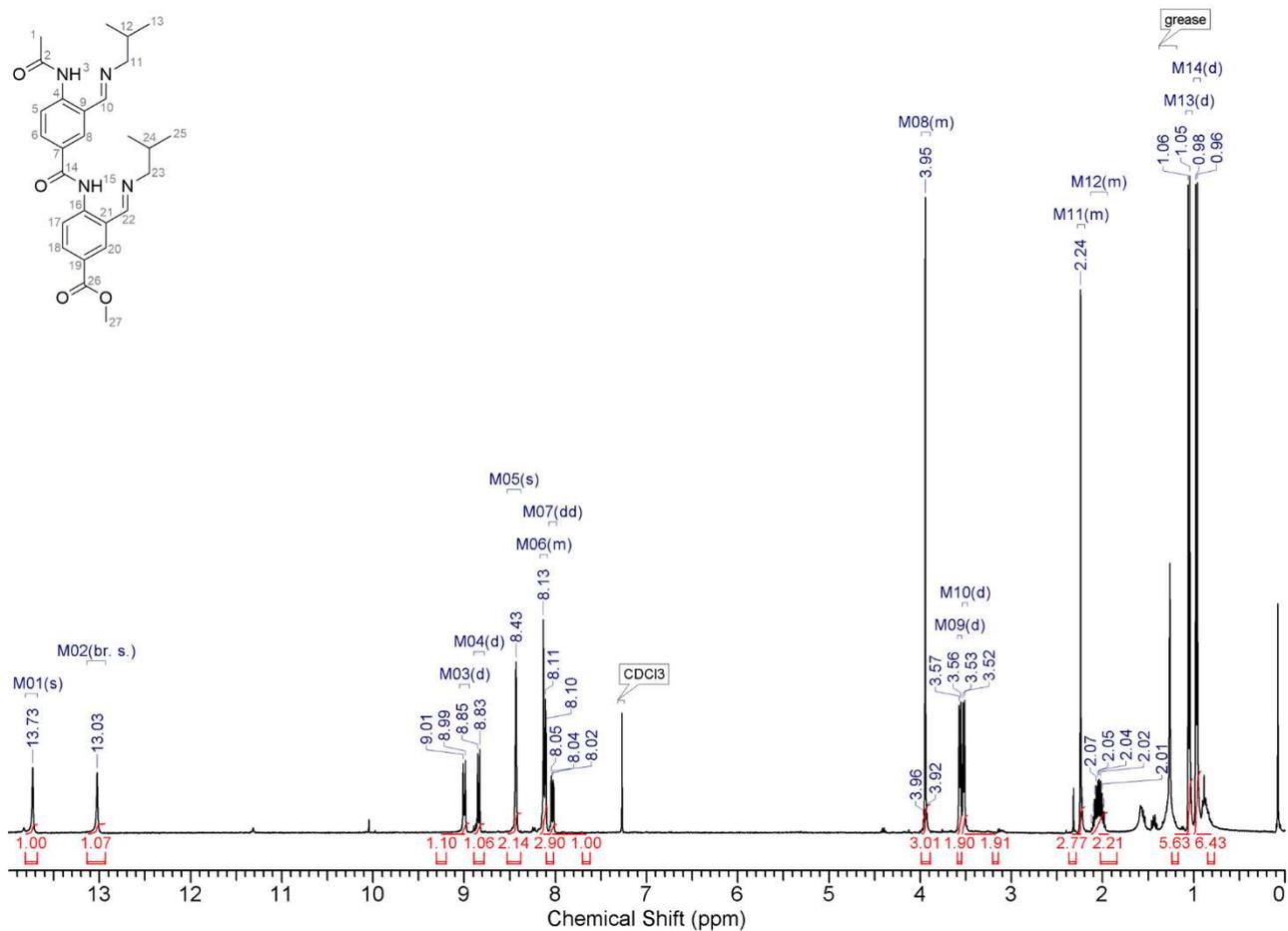
Figure S18. <sup>13</sup>C NMR spectrum (101 MHz) of compound **2a** in CDCl<sub>3</sub> at 298 K.



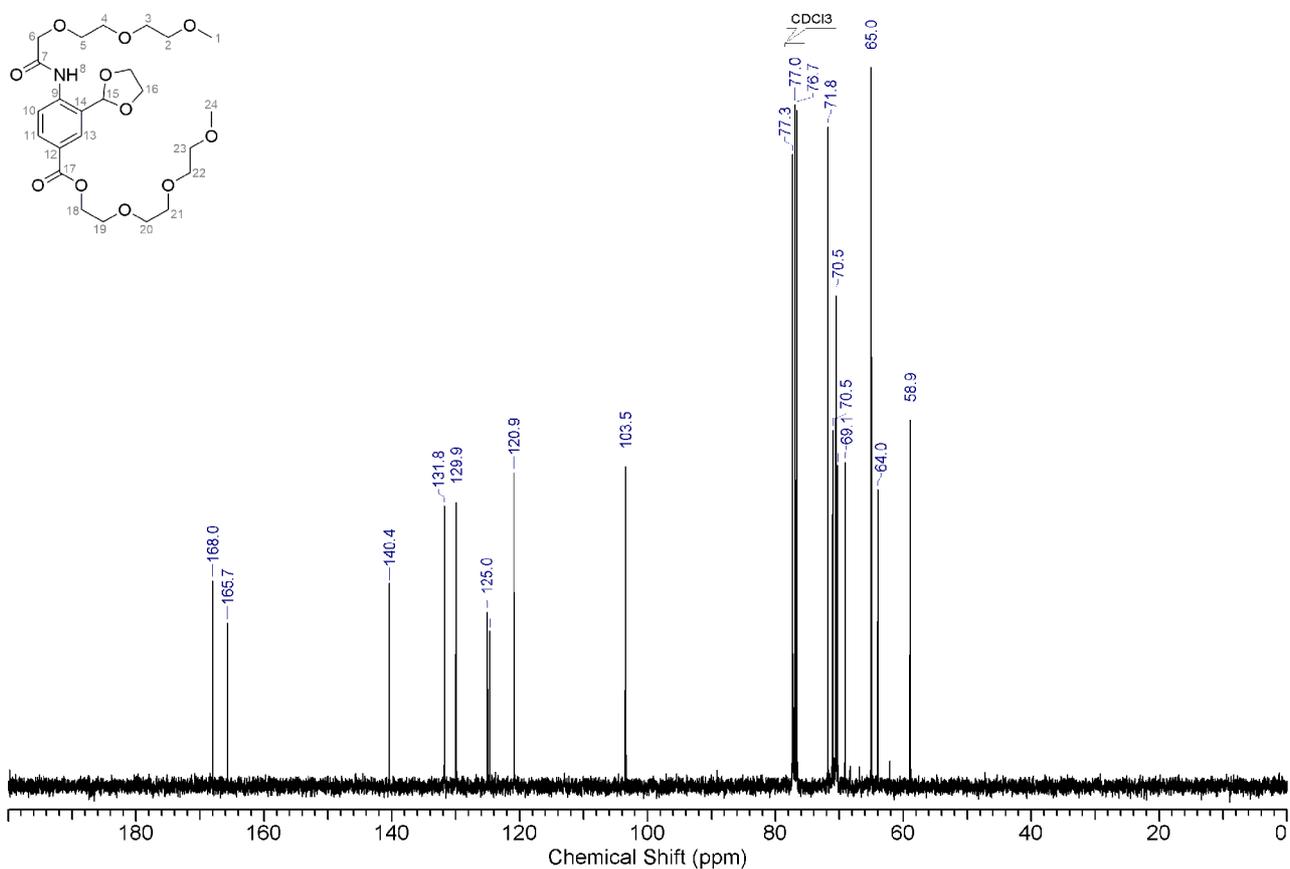
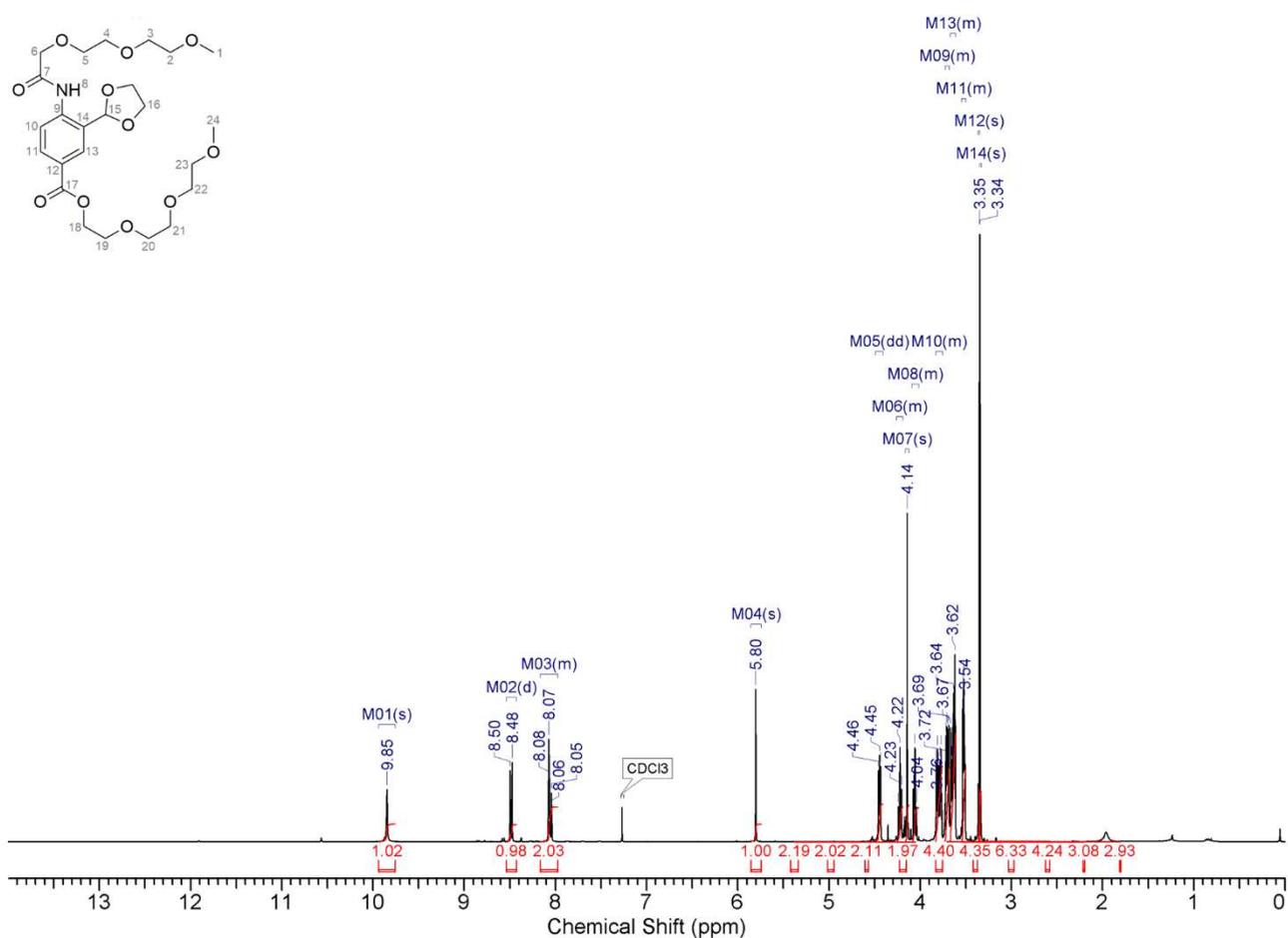
**Figure S19.** <sup>1</sup>H NMR spectrum (400 MHz) of compound **5a** in CDCl<sub>3</sub> at 298 K.

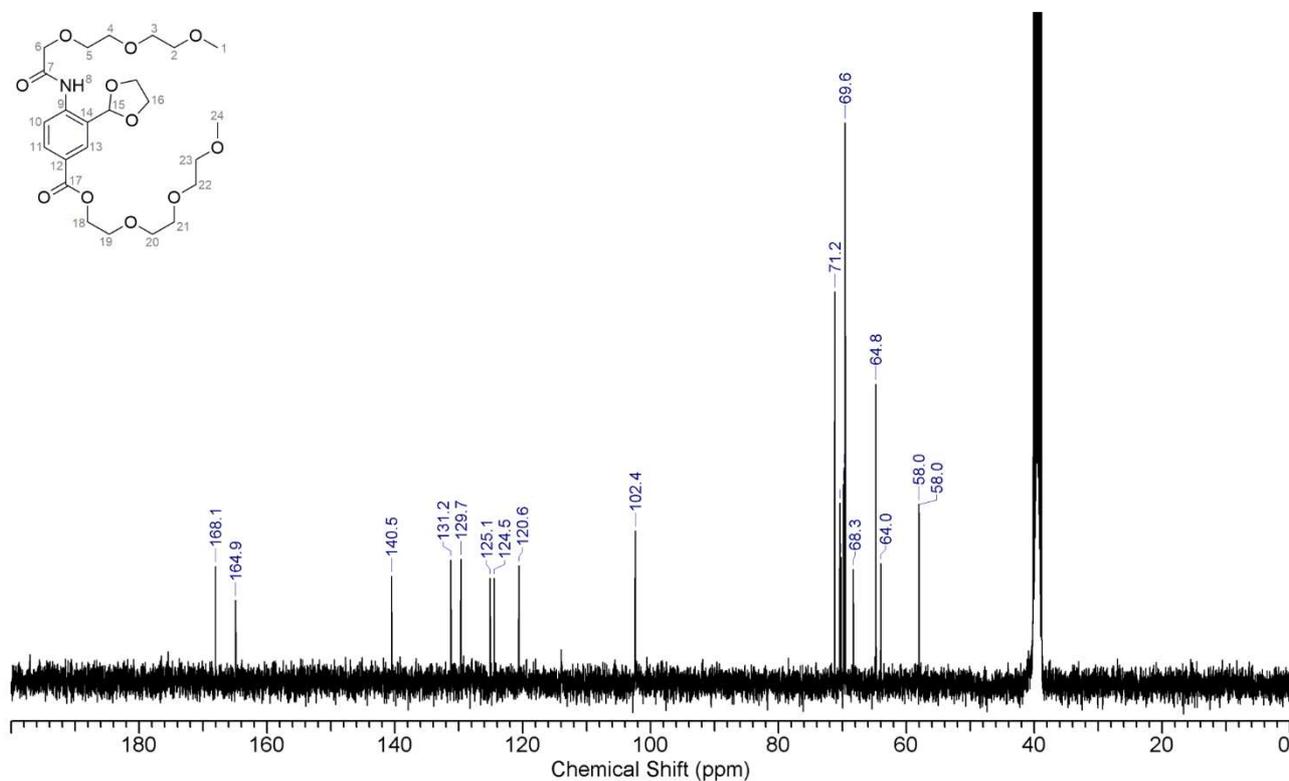
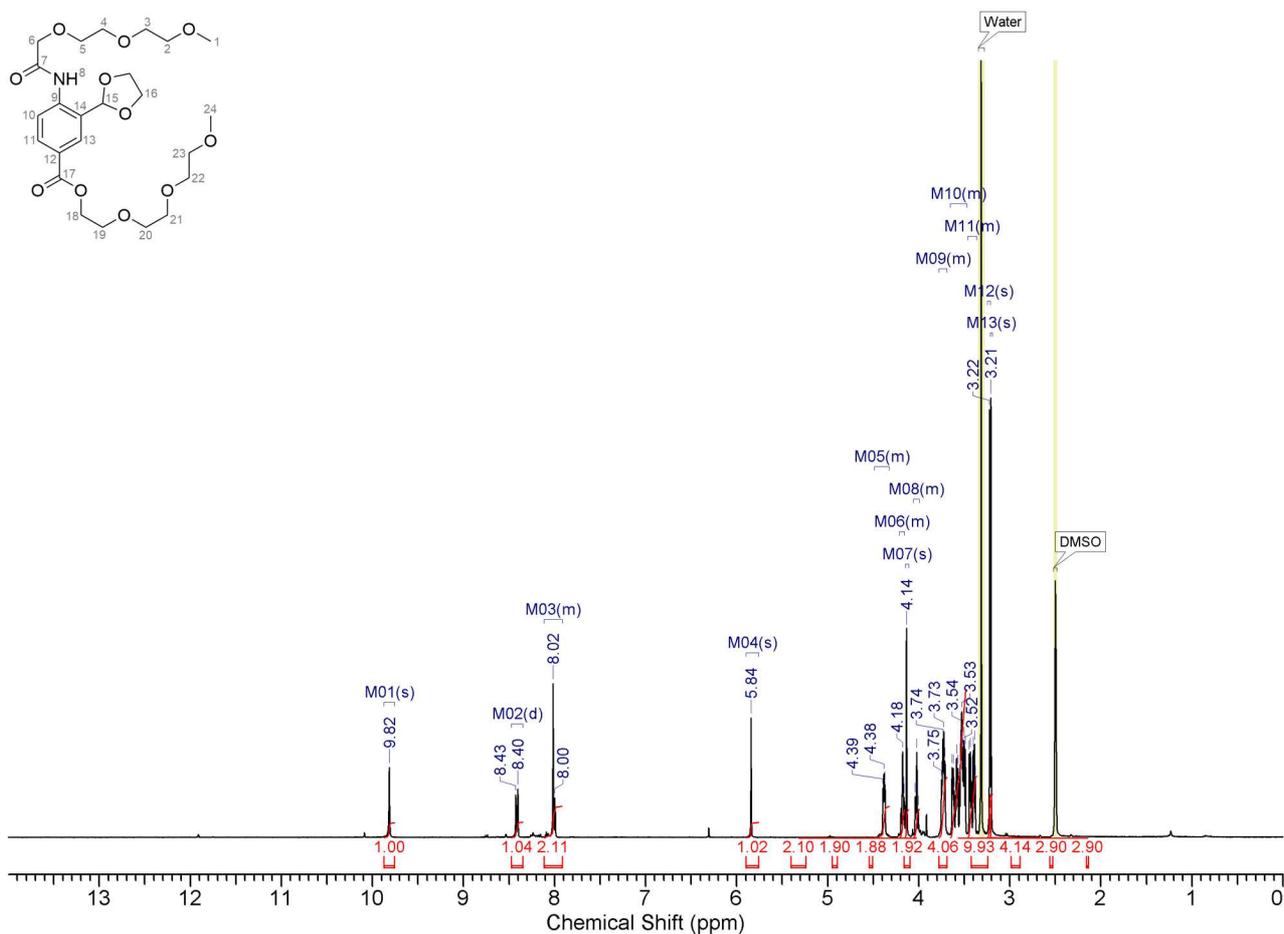


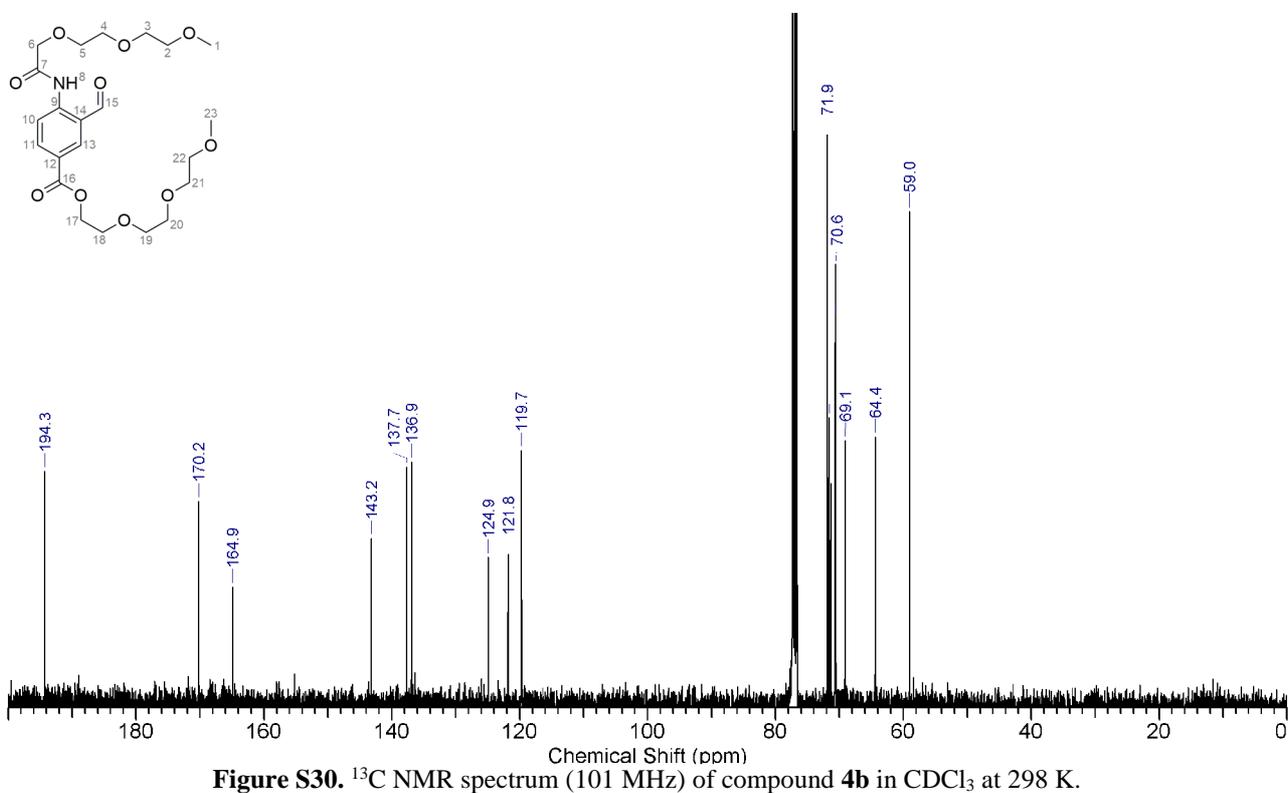
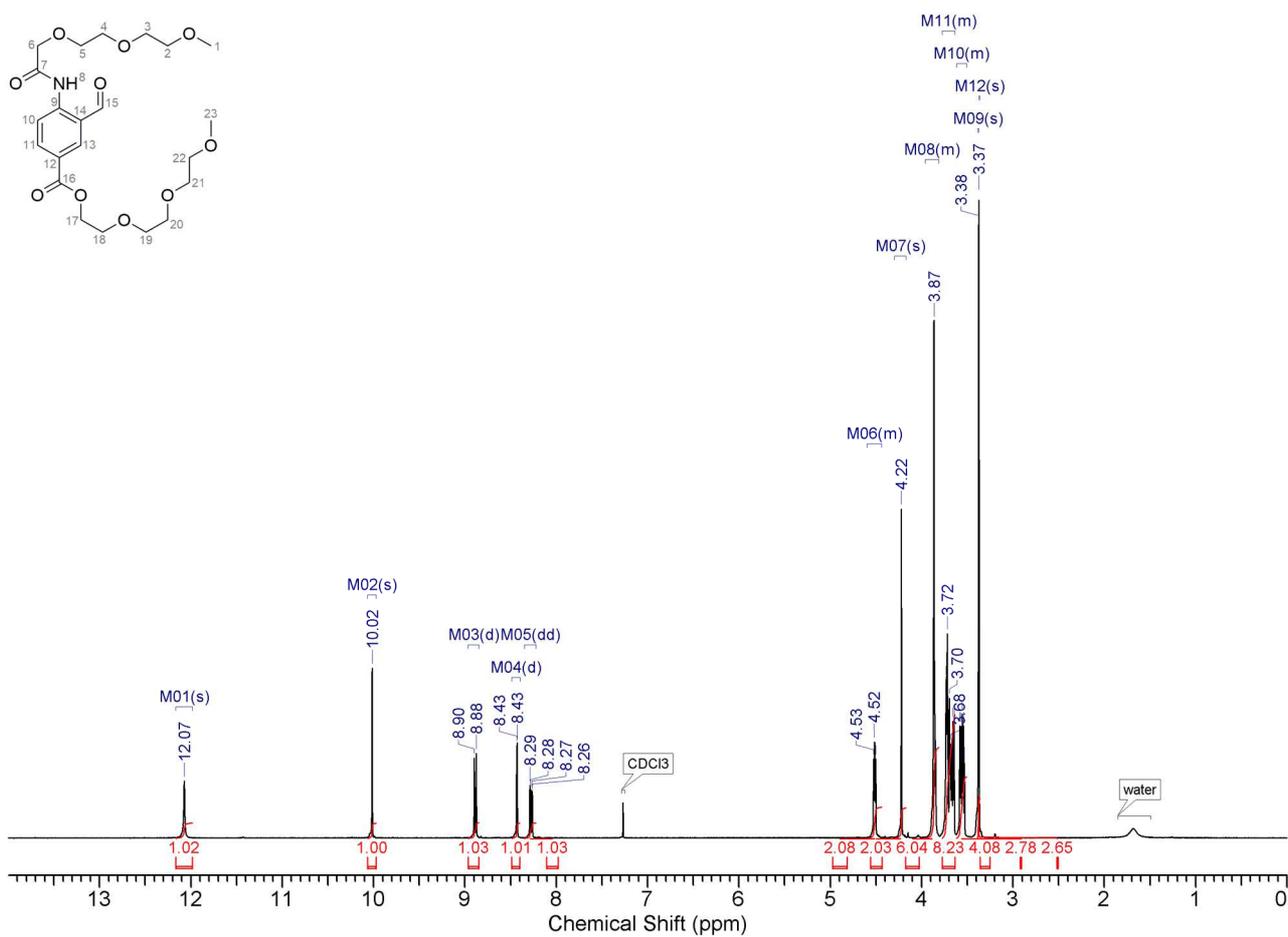
**Figure S20.** <sup>13</sup>C NMR spectrum (151 MHz) of compound **5a** in CDCl<sub>3</sub> at 298 K.

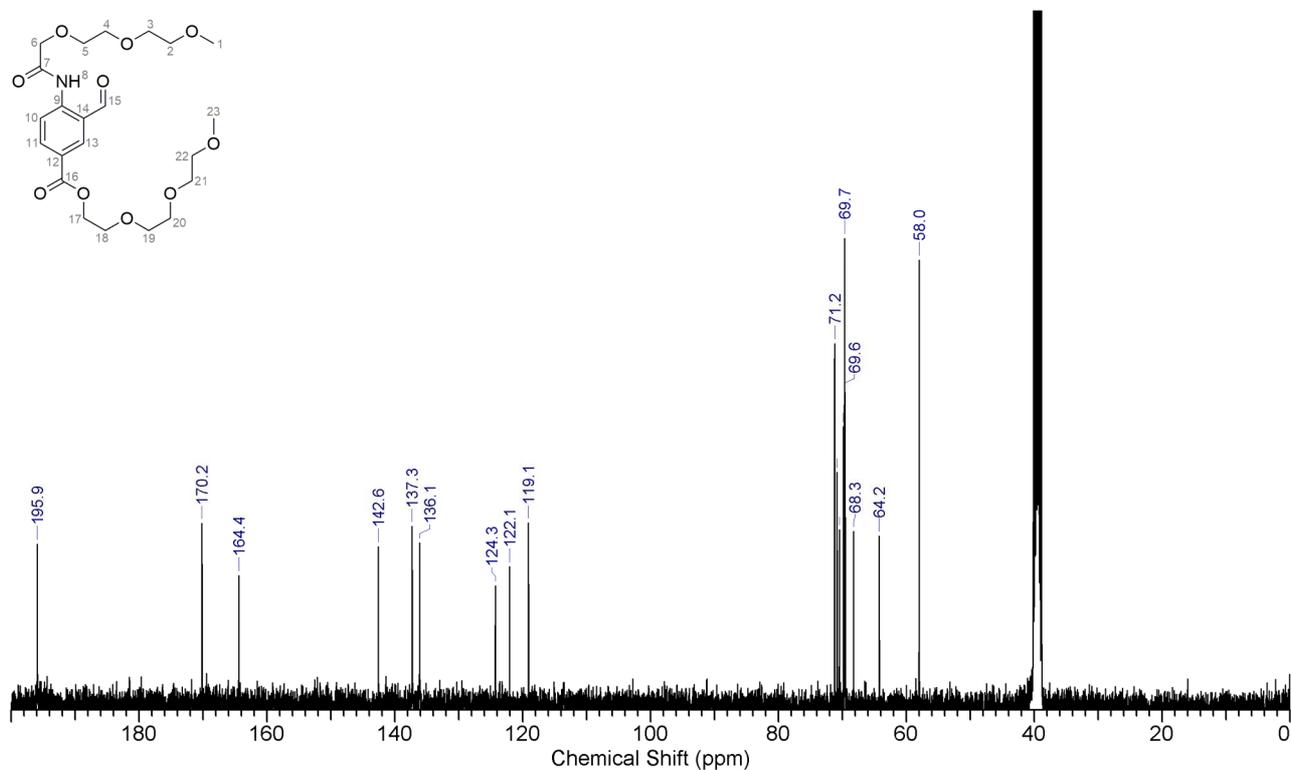
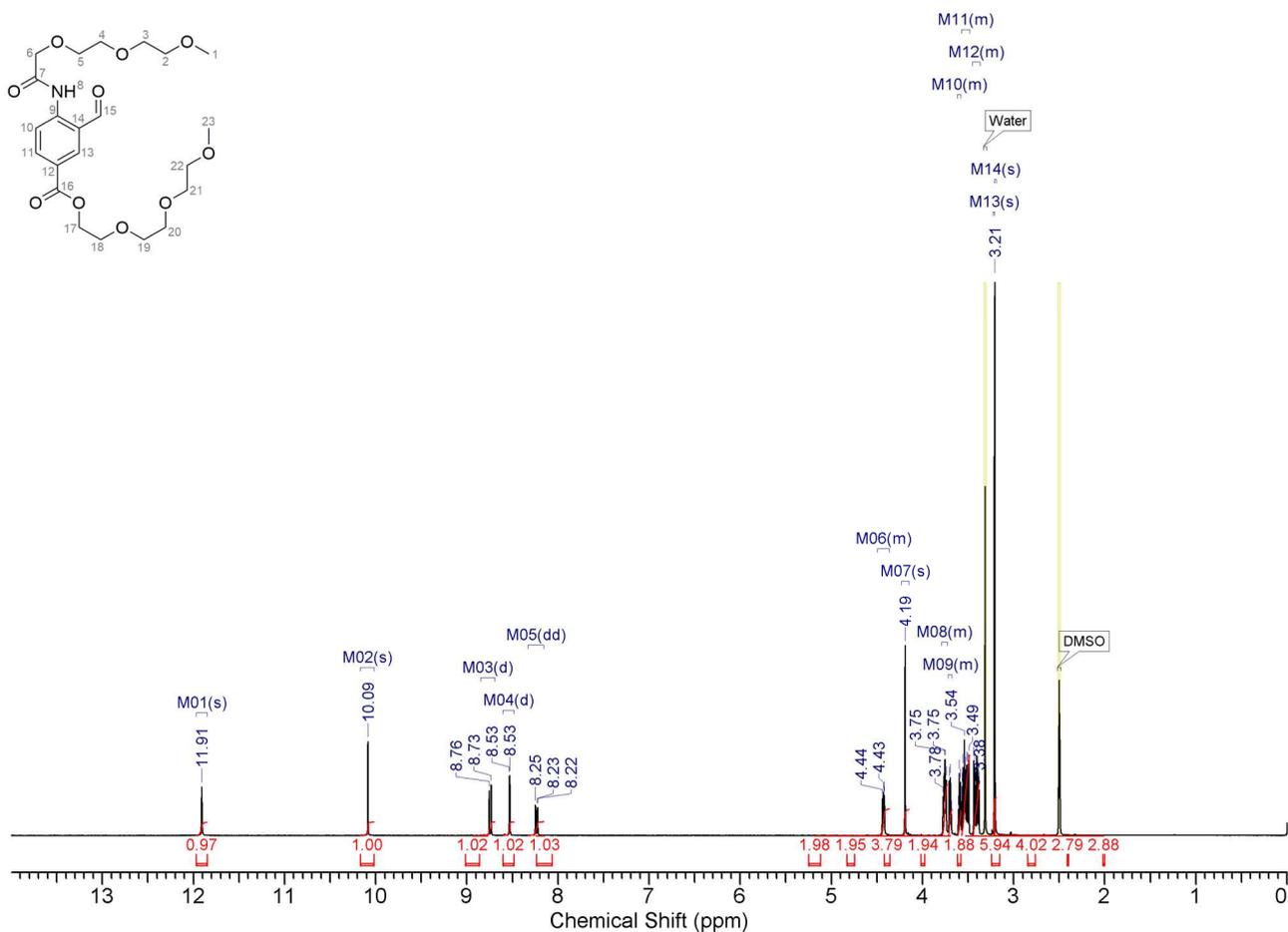












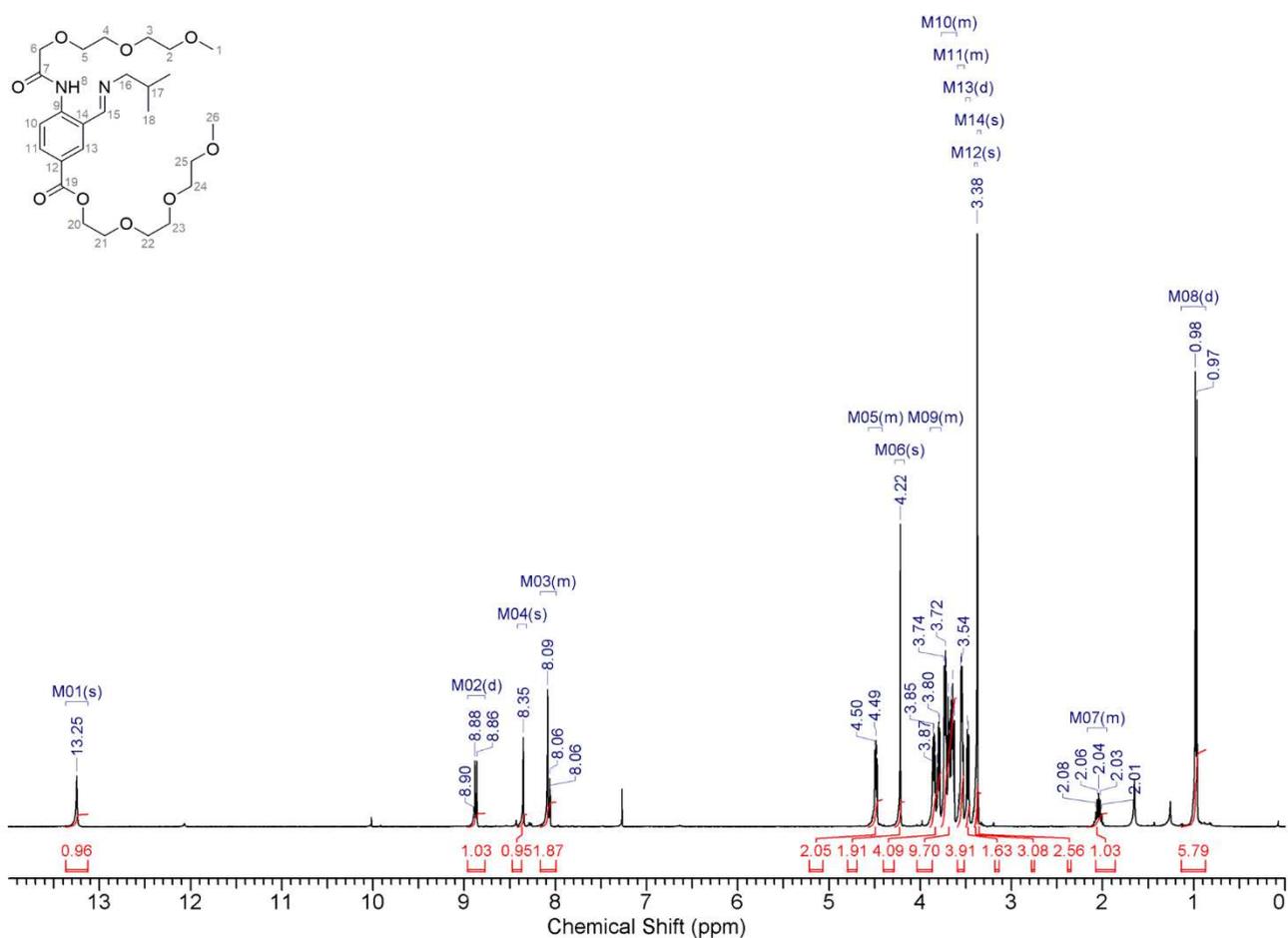


Figure S33.  $^1\text{H}$  NMR spectrum (400 MHz) of compound **7b** in  $\text{CDCl}_3$  at 298 K.

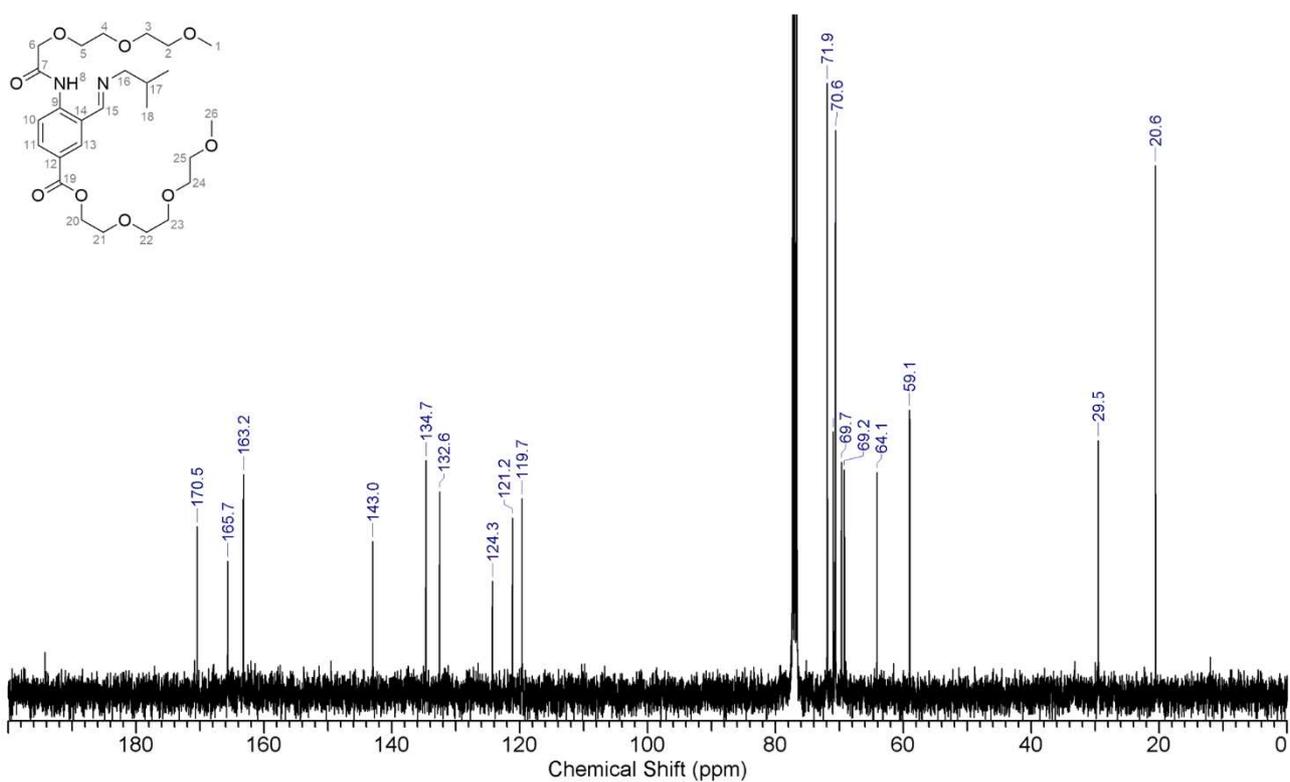
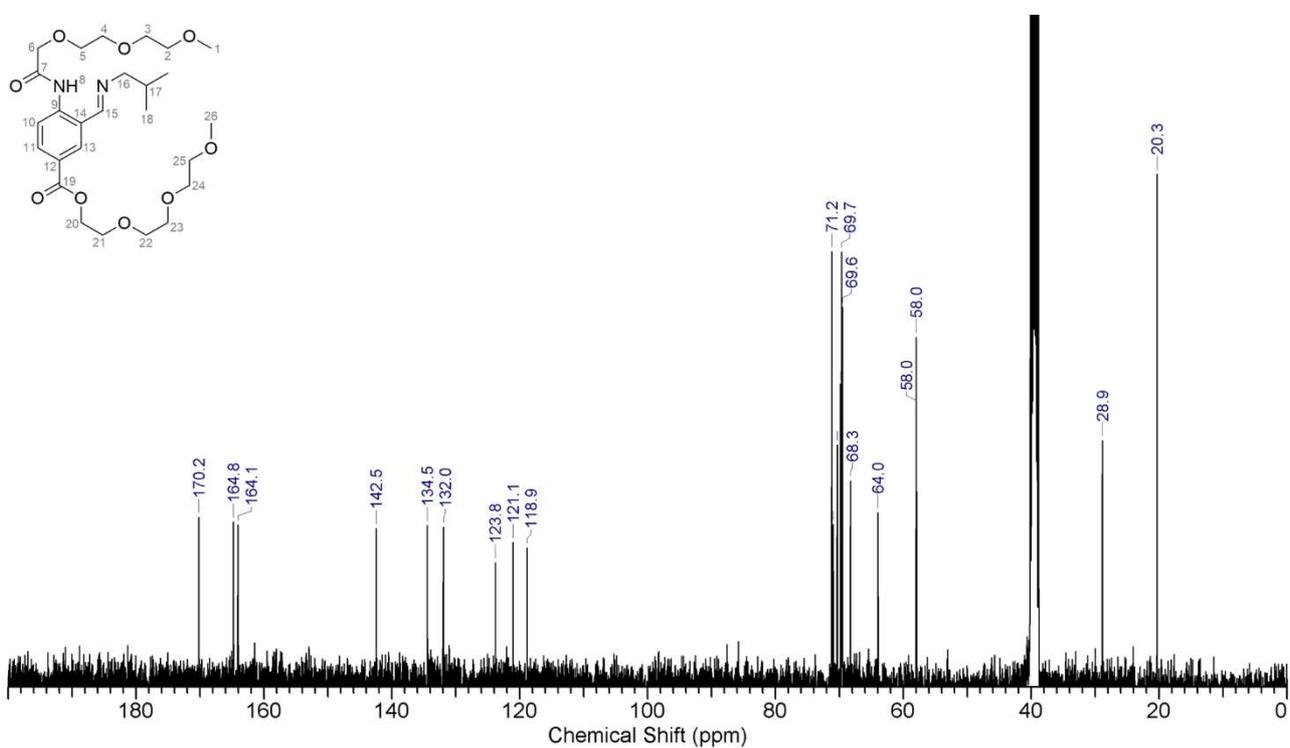
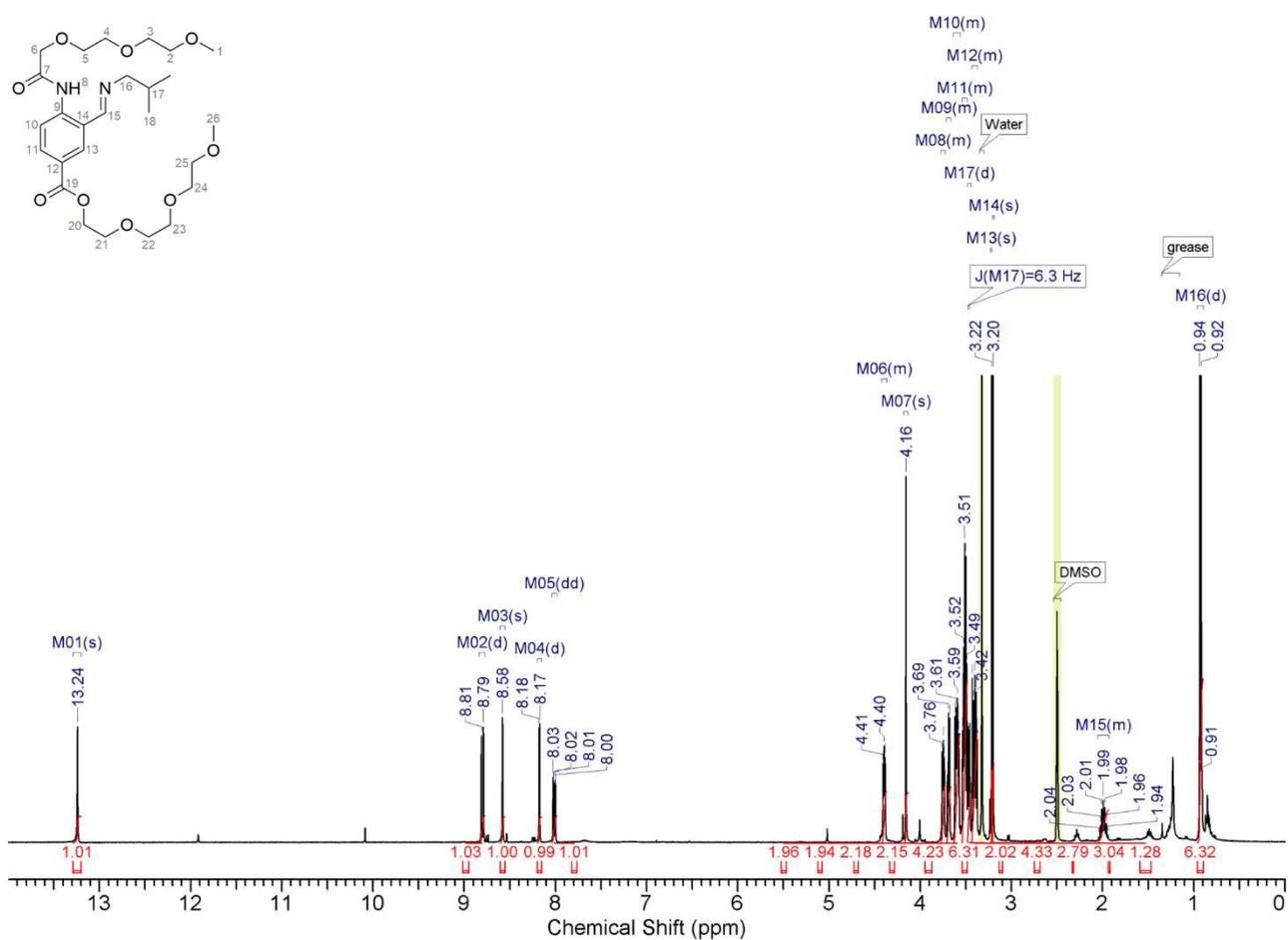


Figure S34.  $^{13}\text{C}$  NMR spectrum (101 MHz) of compound **7b** in  $\text{CDCl}_3$  at 298 K.



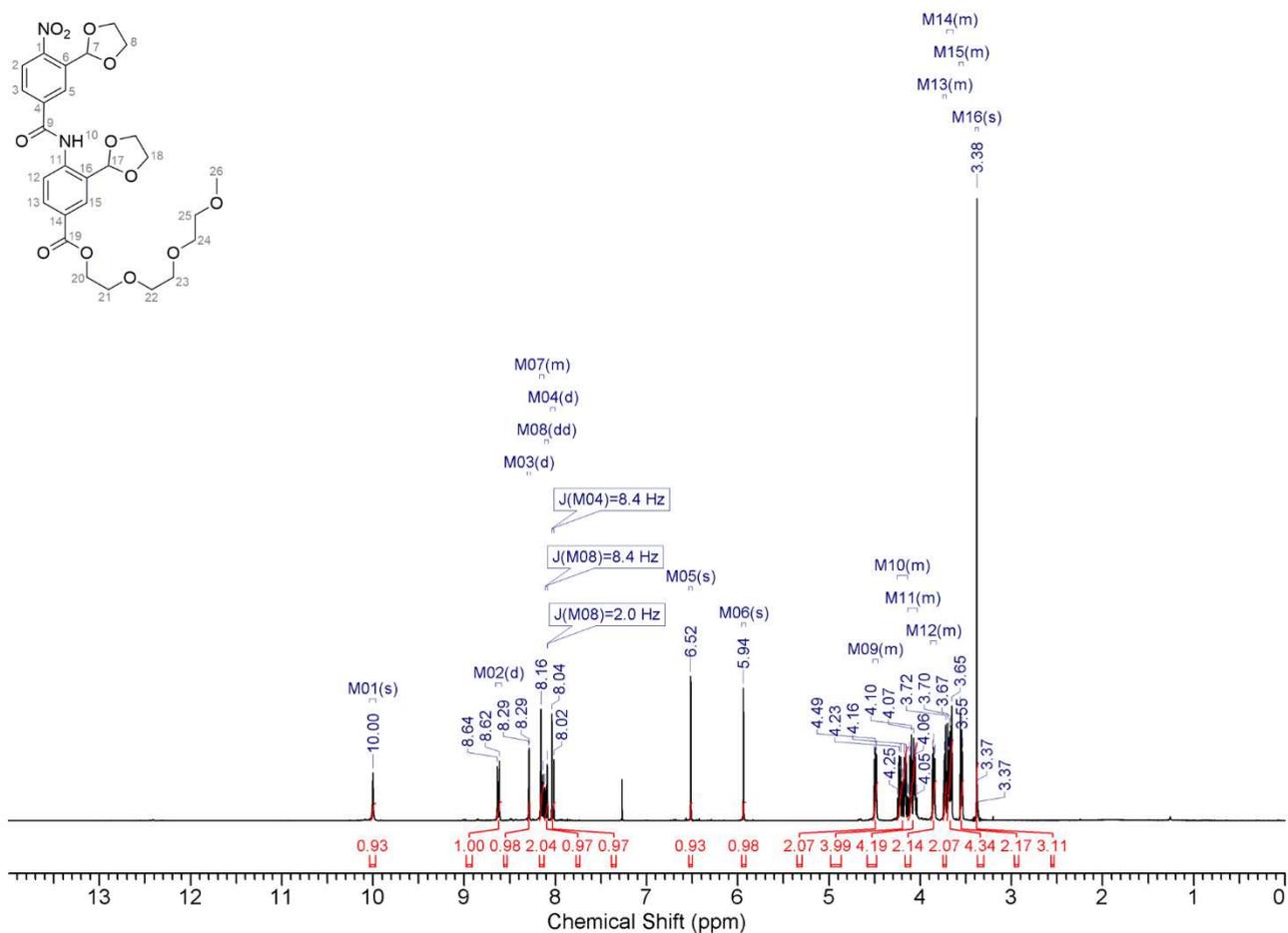


Figure S37. <sup>1</sup>H NMR spectrum (400 MHz) of compound **14** in CDCl<sub>3</sub> at 298 K.

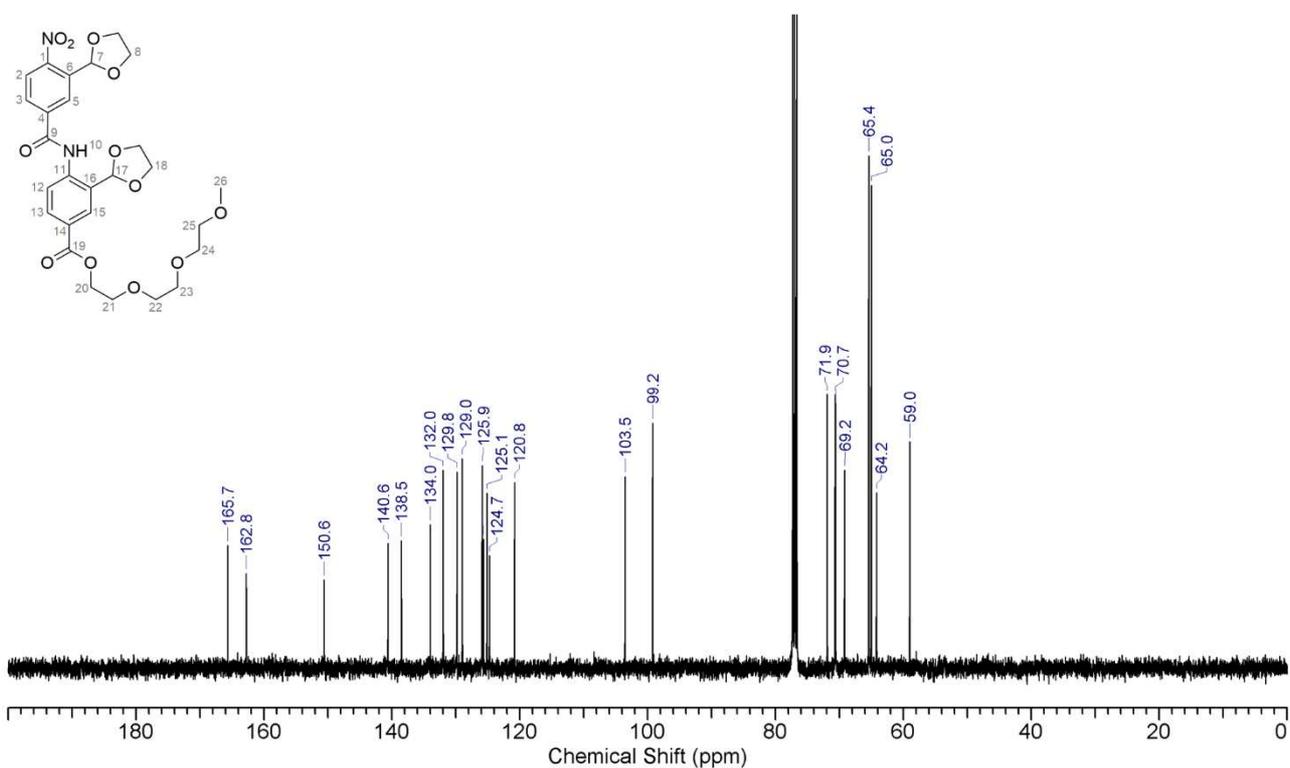
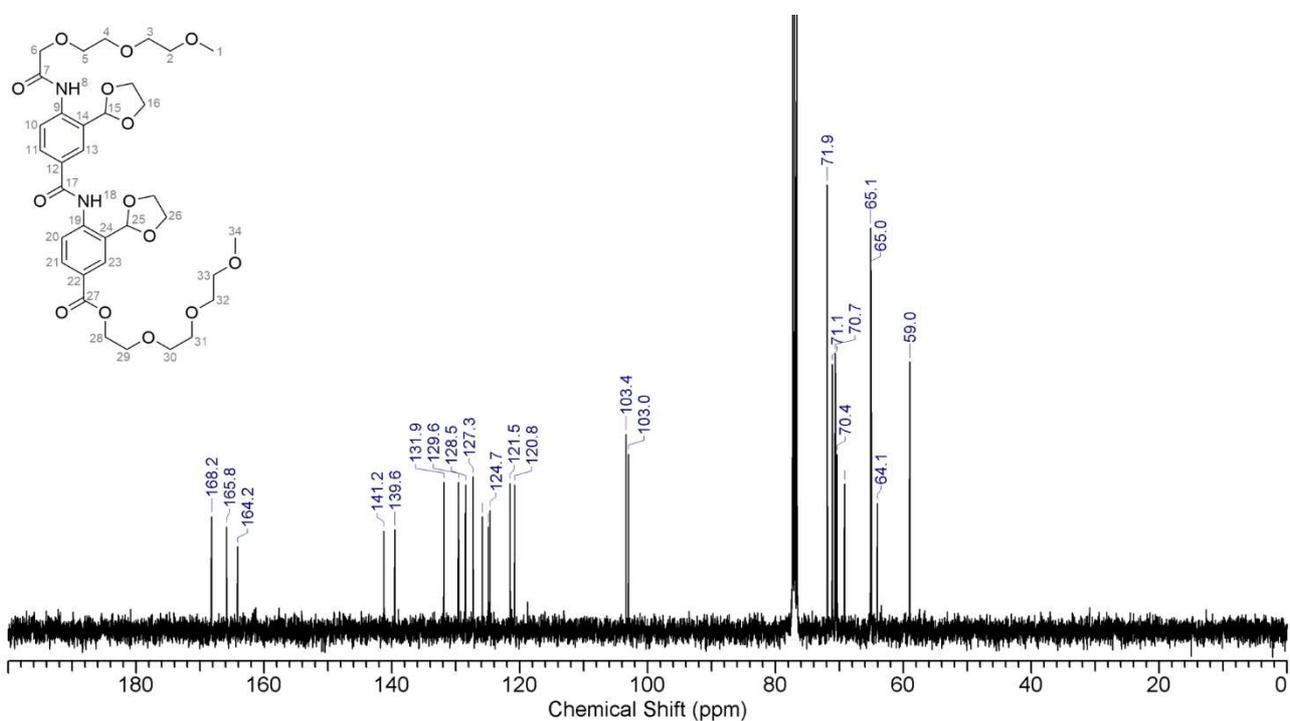
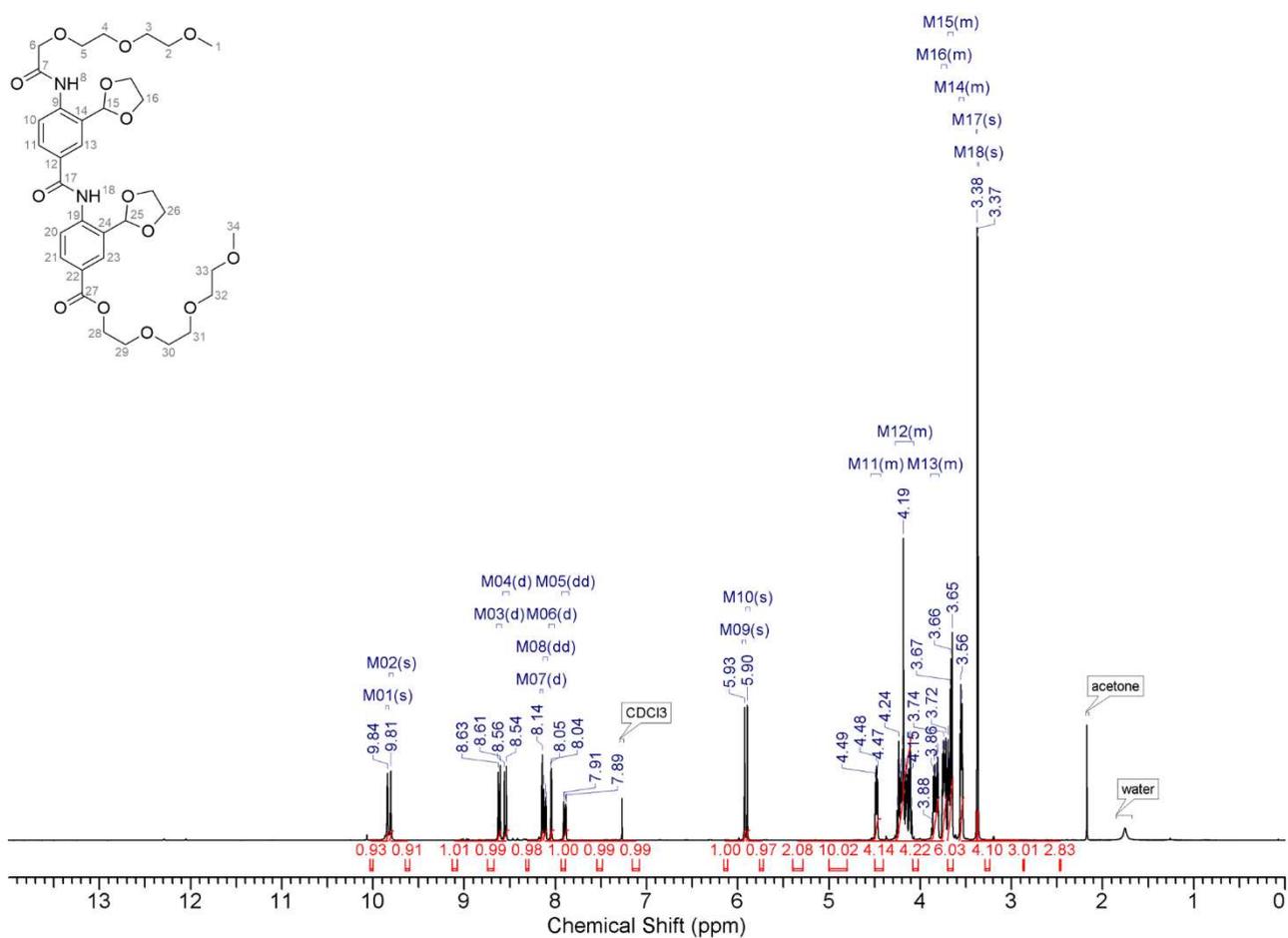


Figure S38. <sup>13</sup>C NMR spectrum (101 MHz) of compound **14** in CDCl<sub>3</sub> at 298 K.



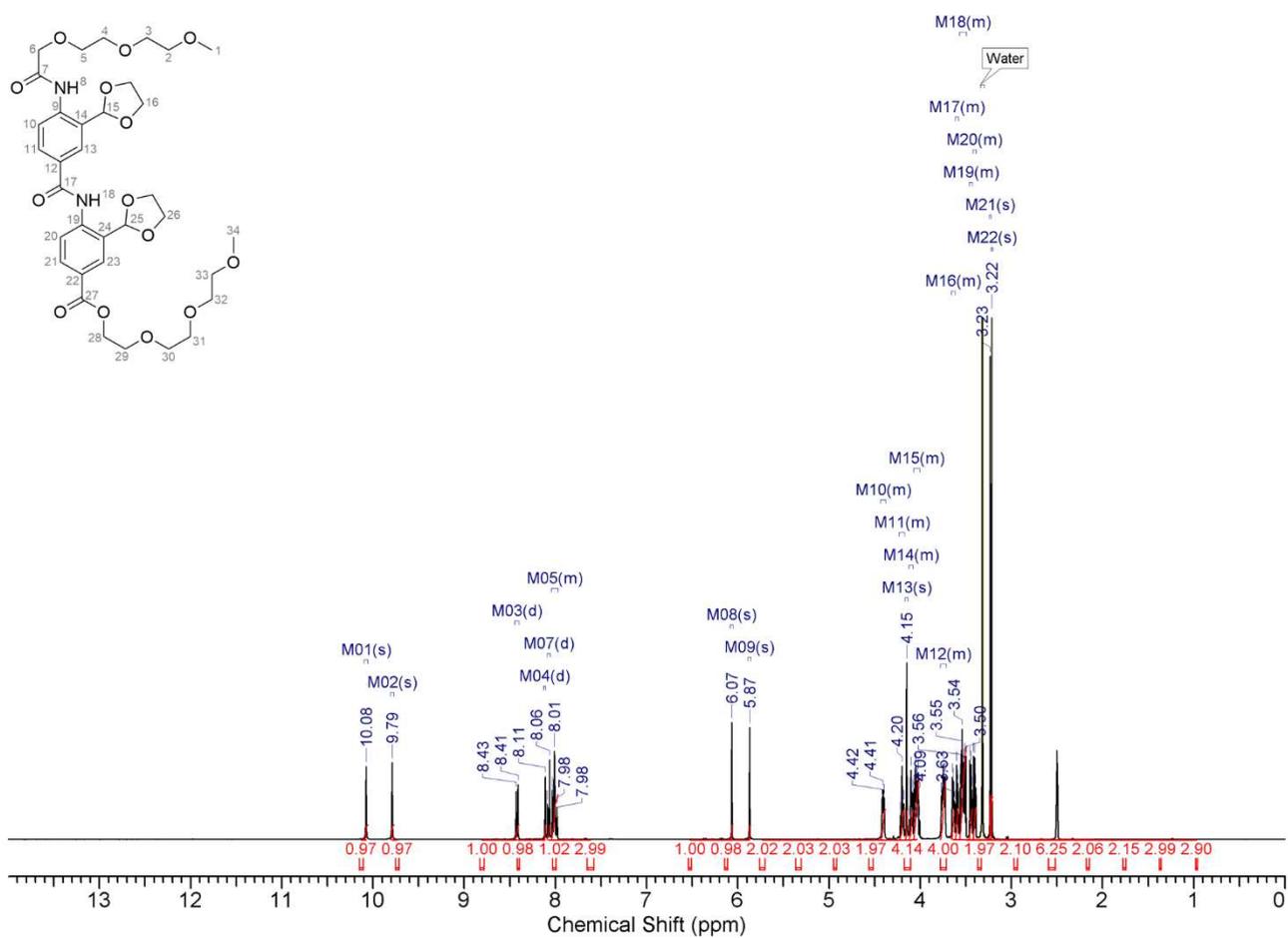


Figure S41.  $^1\text{H}$  NMR spectrum (400 MHz) of compound **2b** in  $\text{DMSO-}d_6$  at 298 K.

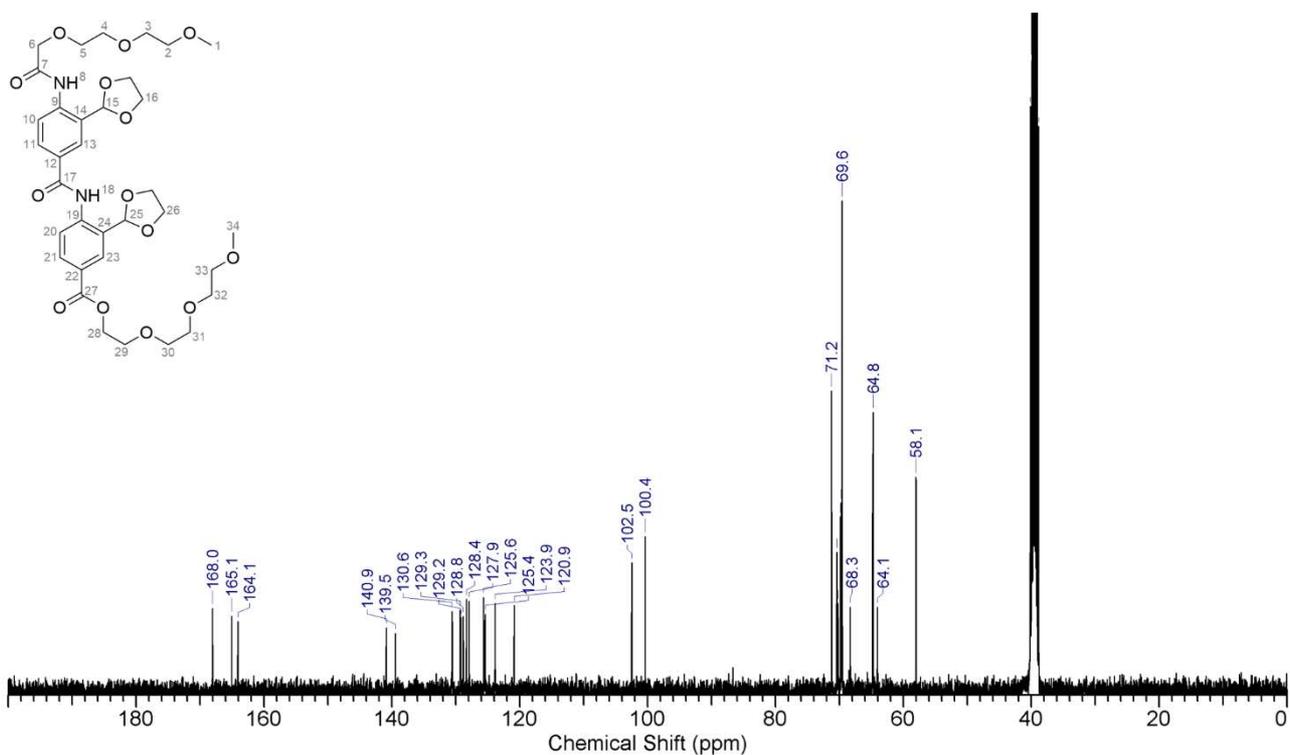
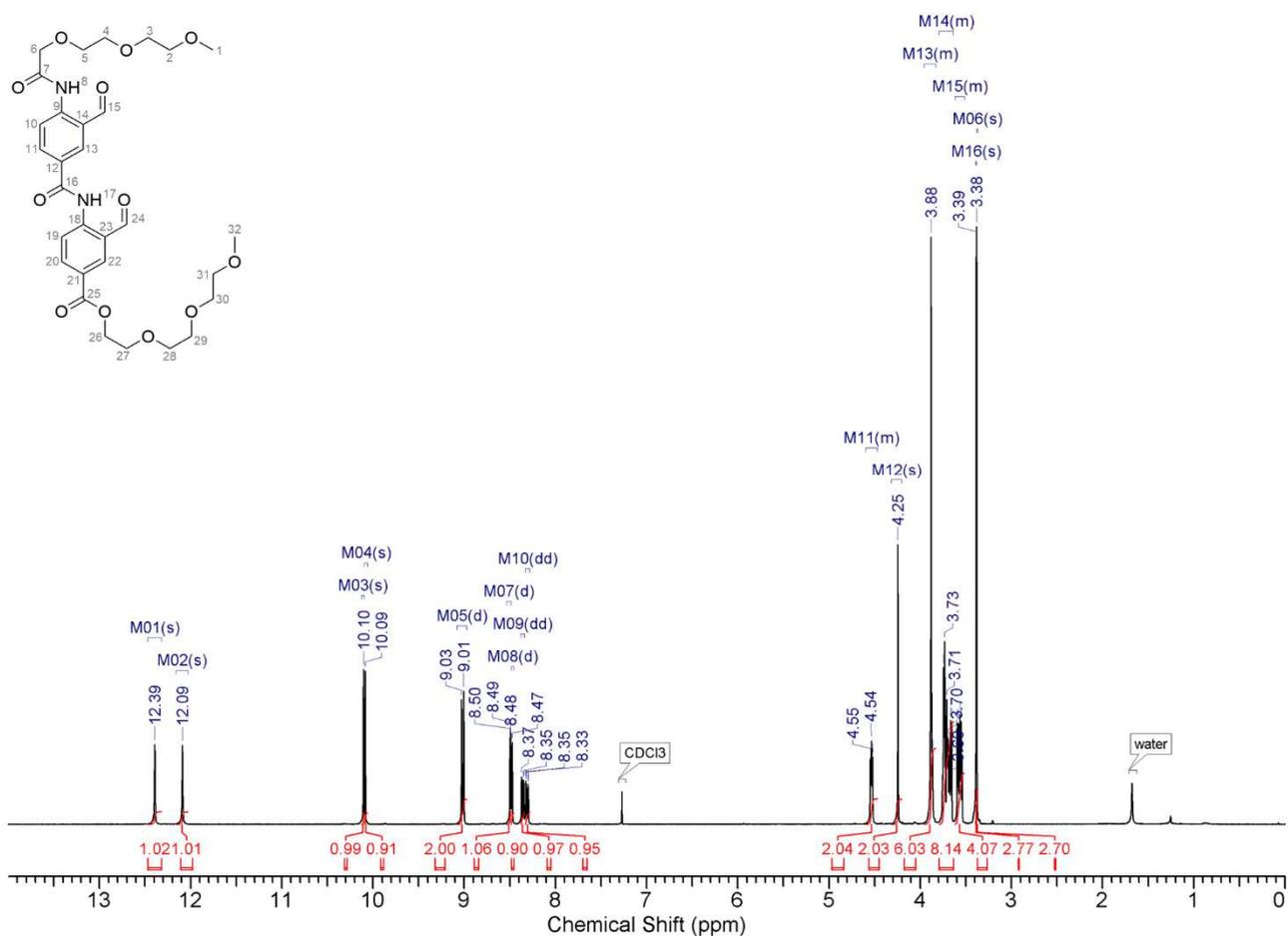
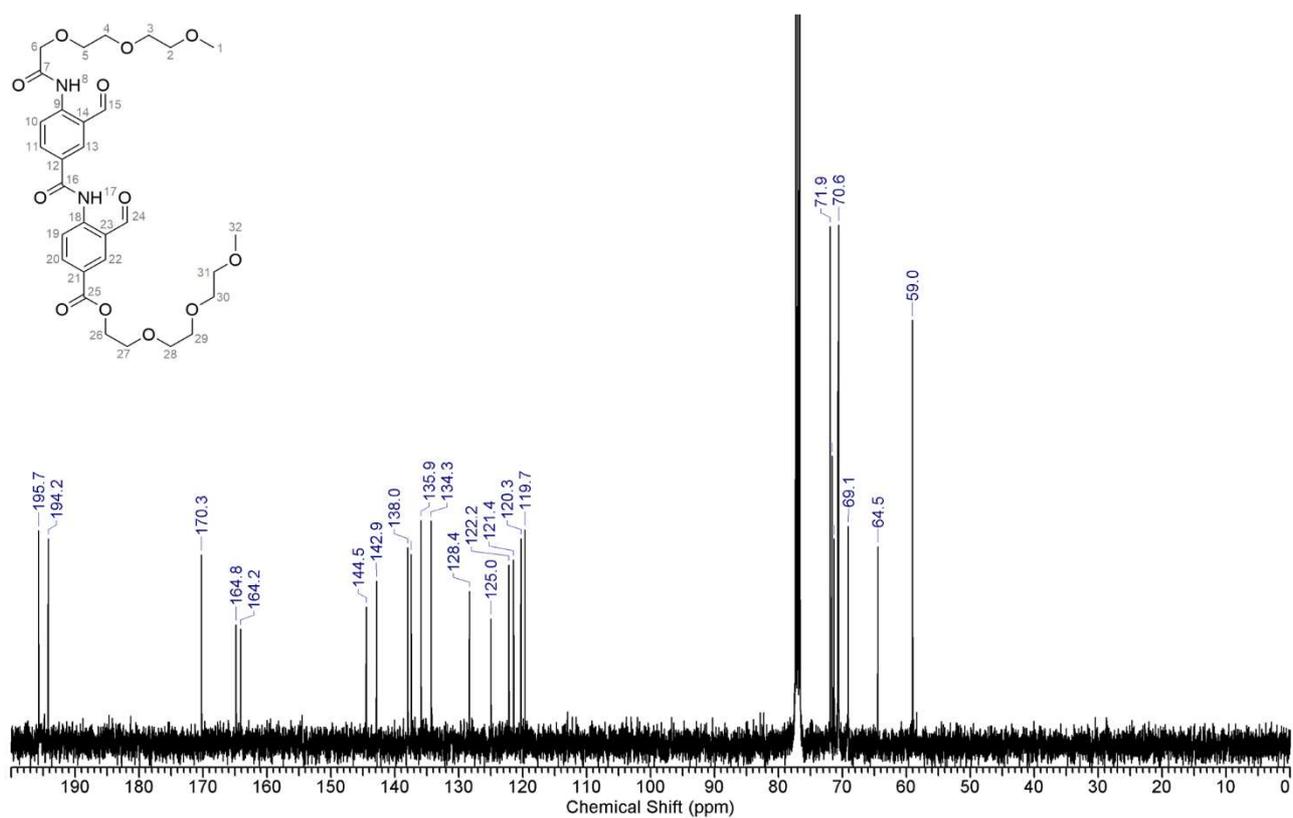


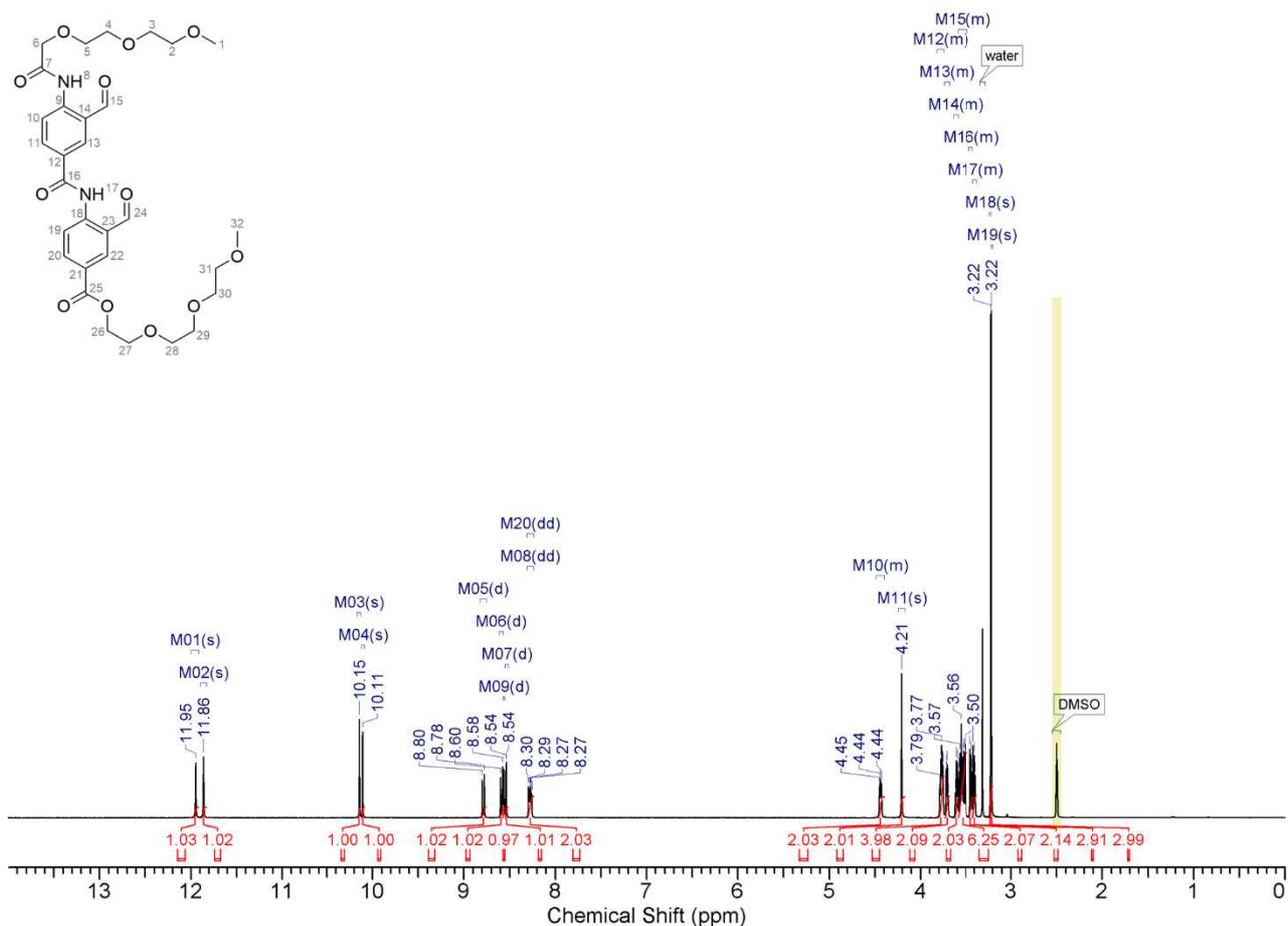
Figure S42.  $^{13}\text{C}$  NMR spectrum (101 MHz) of compound **2b** in  $\text{DMSO-}d_6$  at 298 K.



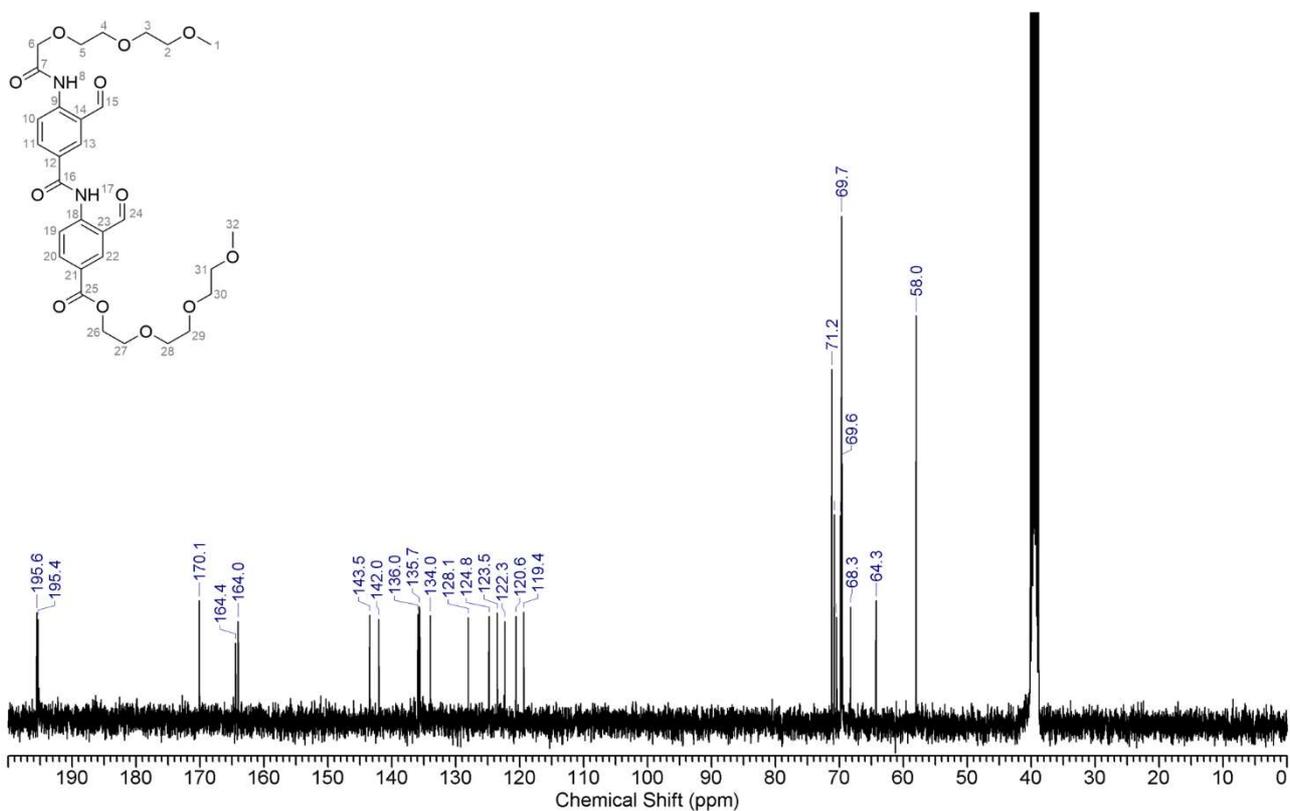
**Figure S43.** <sup>1</sup>H NMR spectrum (400 MHz) of compound **5b** in CDCl<sub>3</sub> at 298 K.



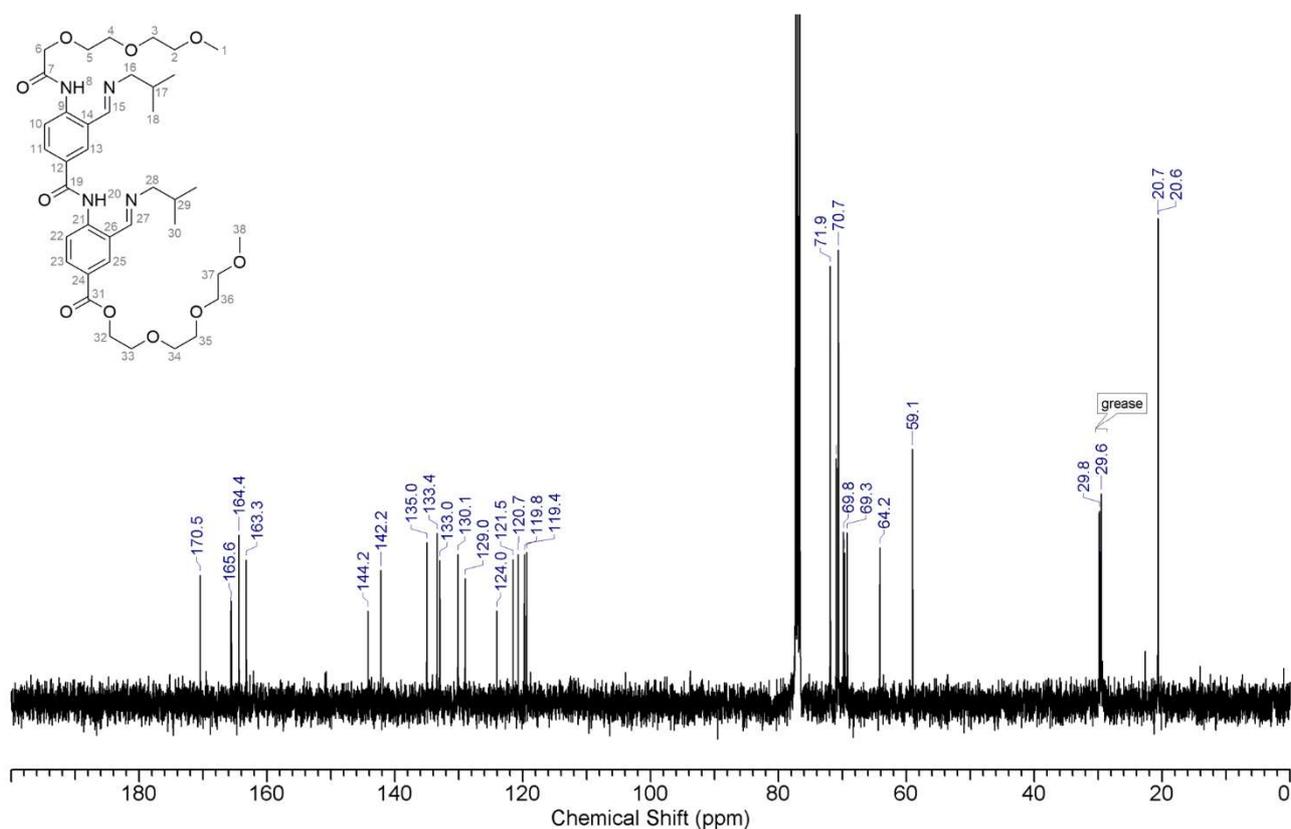
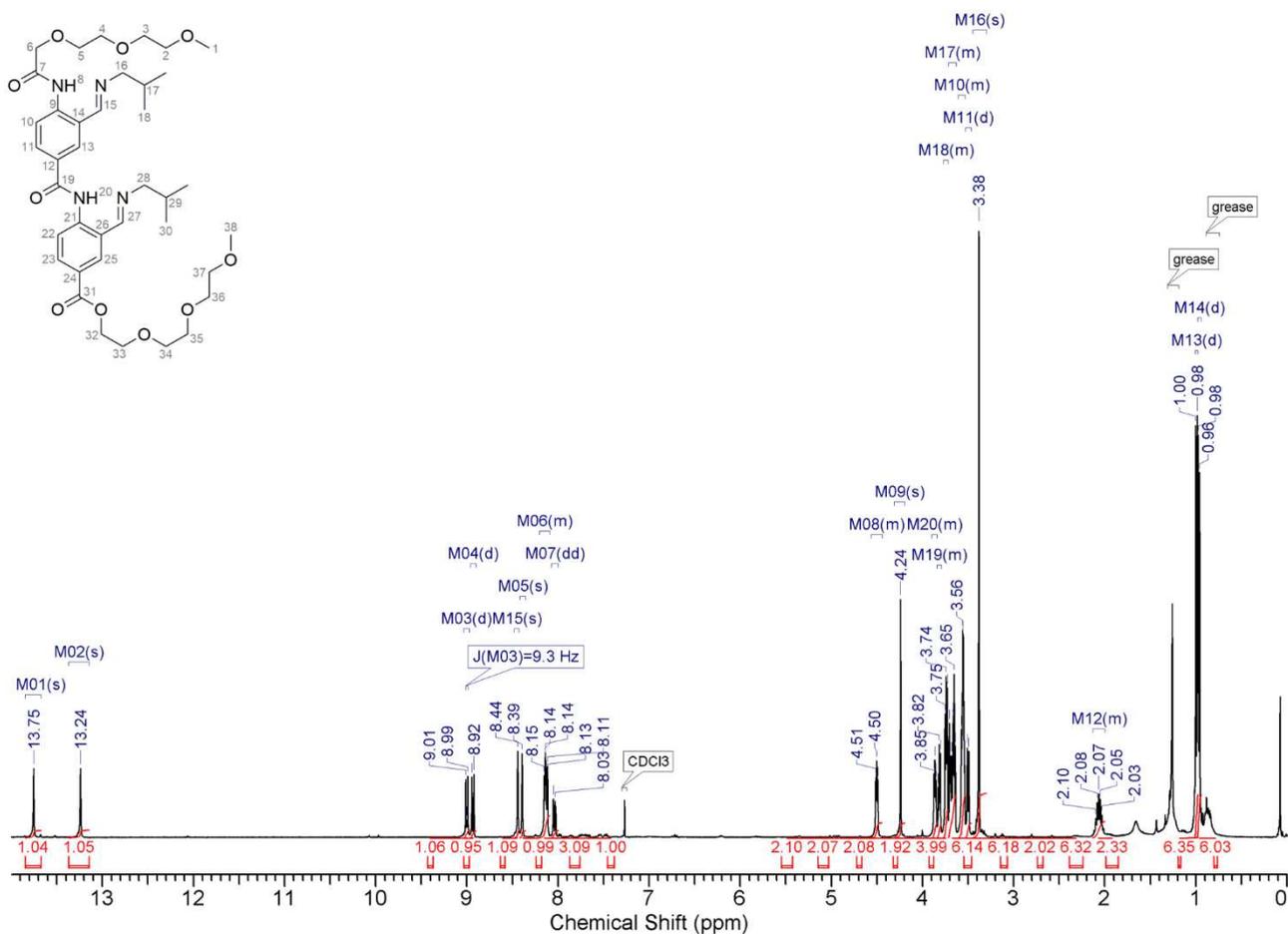
**Figure S44.** <sup>13</sup>C NMR spectrum (101 MHz) of compound **5b** in CDCl<sub>3</sub> at 298 K.

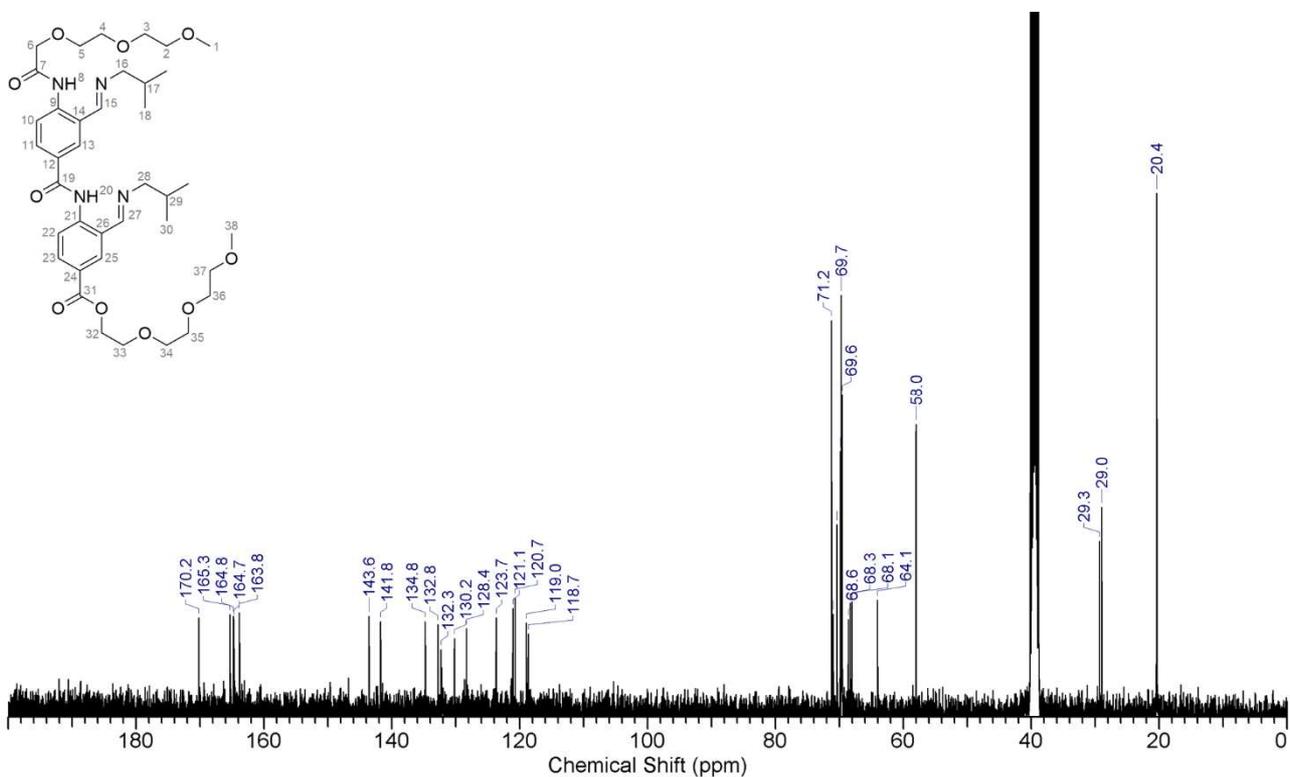
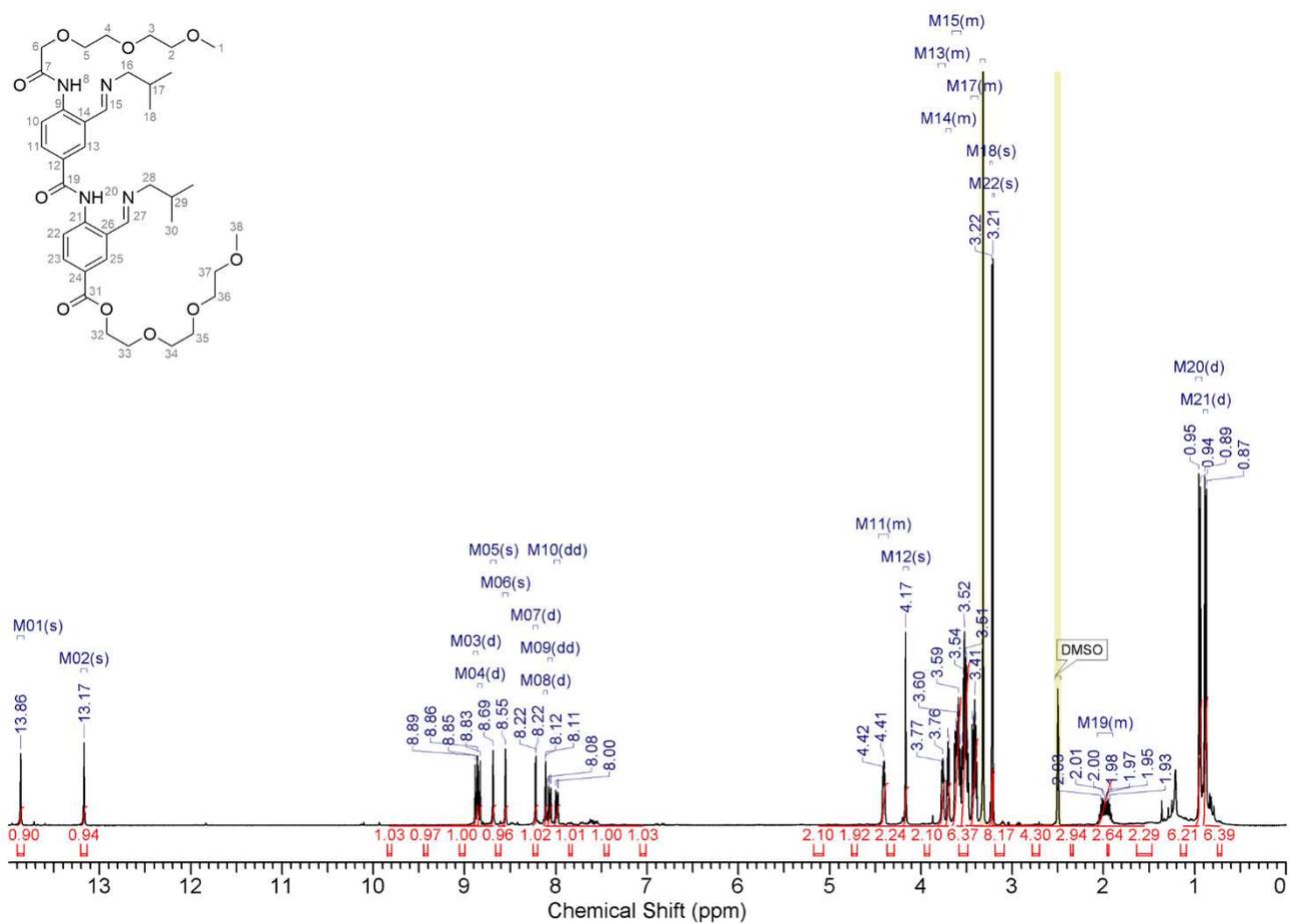


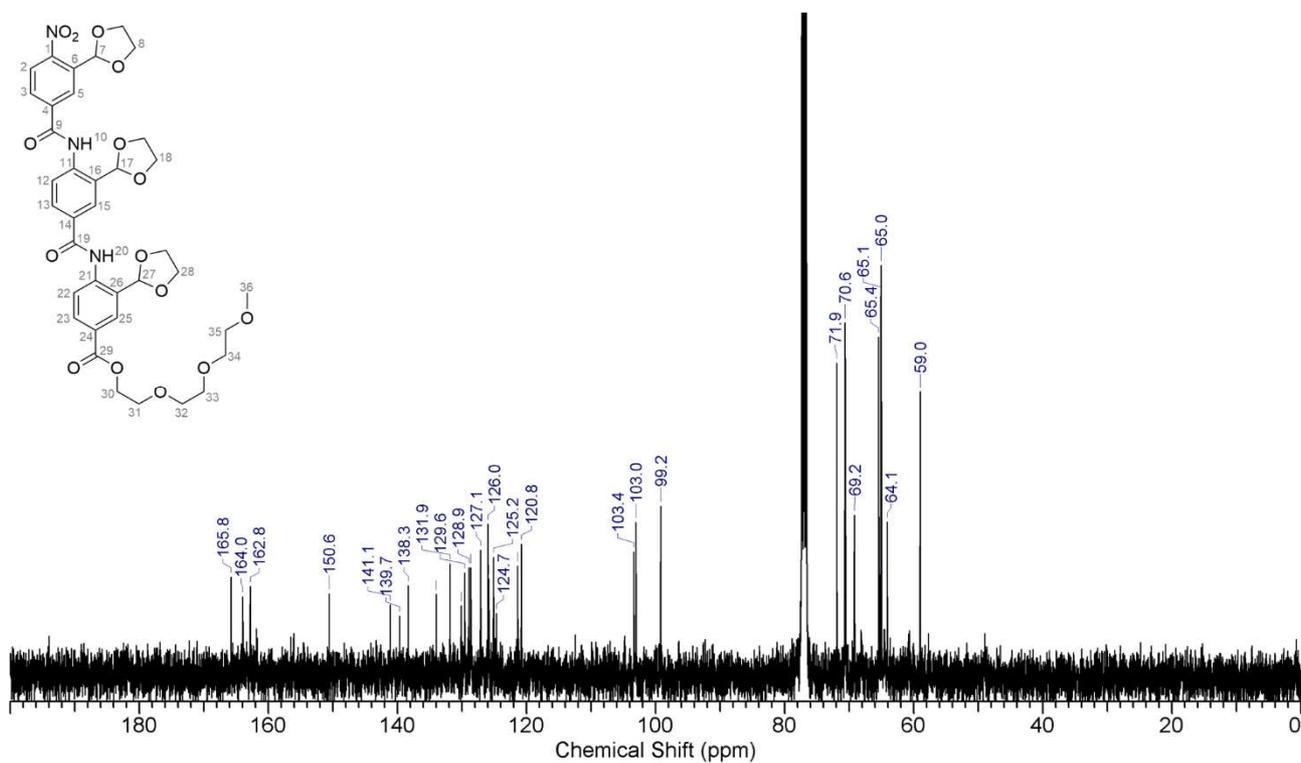
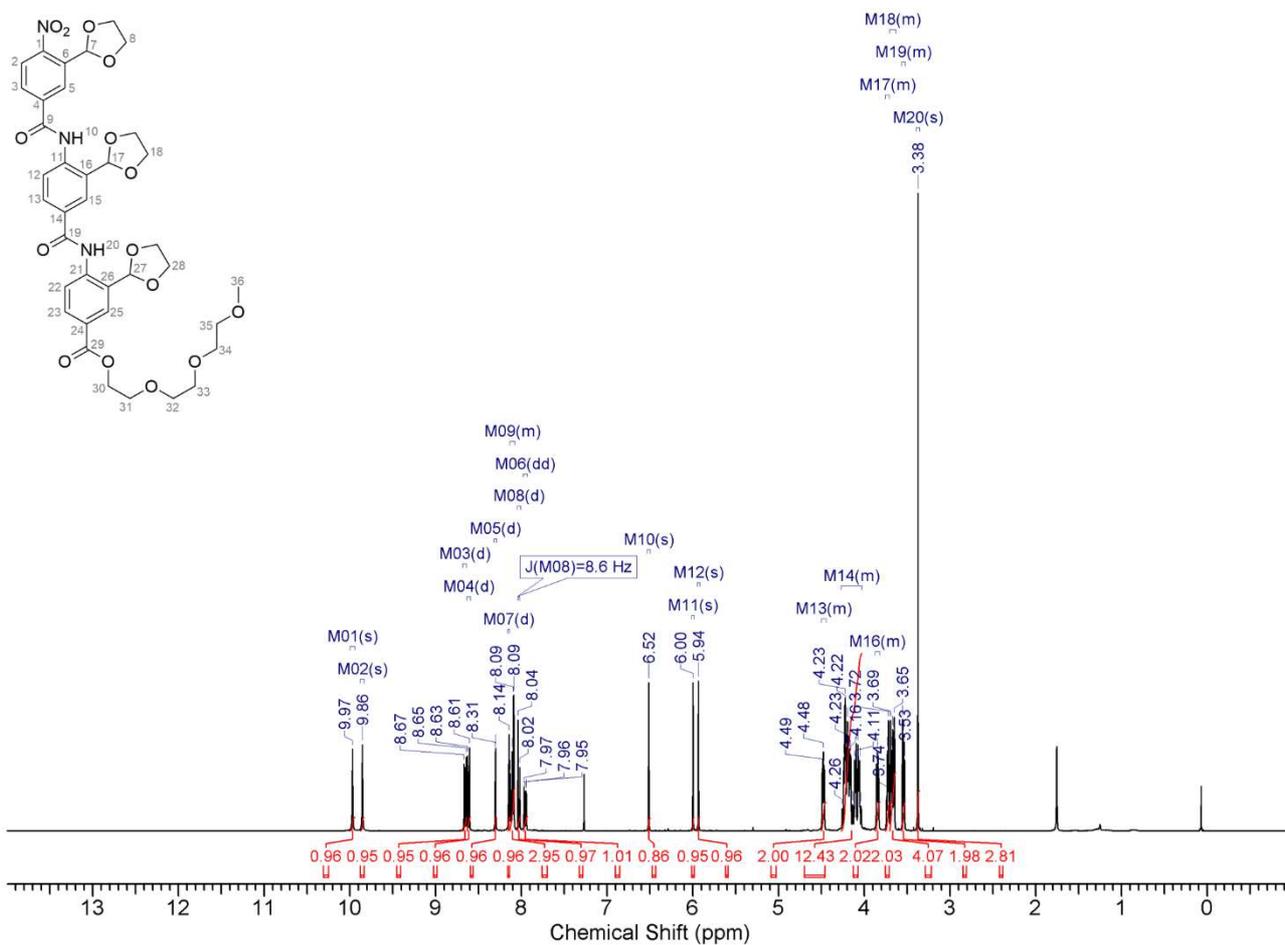
**Figure S45.** <sup>1</sup>H NMR spectrum (400 MHz) of compound **5b** in DMSO-*d*<sub>6</sub> at 298 K.



**Figure S46.** <sup>13</sup>C NMR spectrum (101 MHz) of compound **5b** in DMSO-*d*<sub>6</sub> at 298 K.







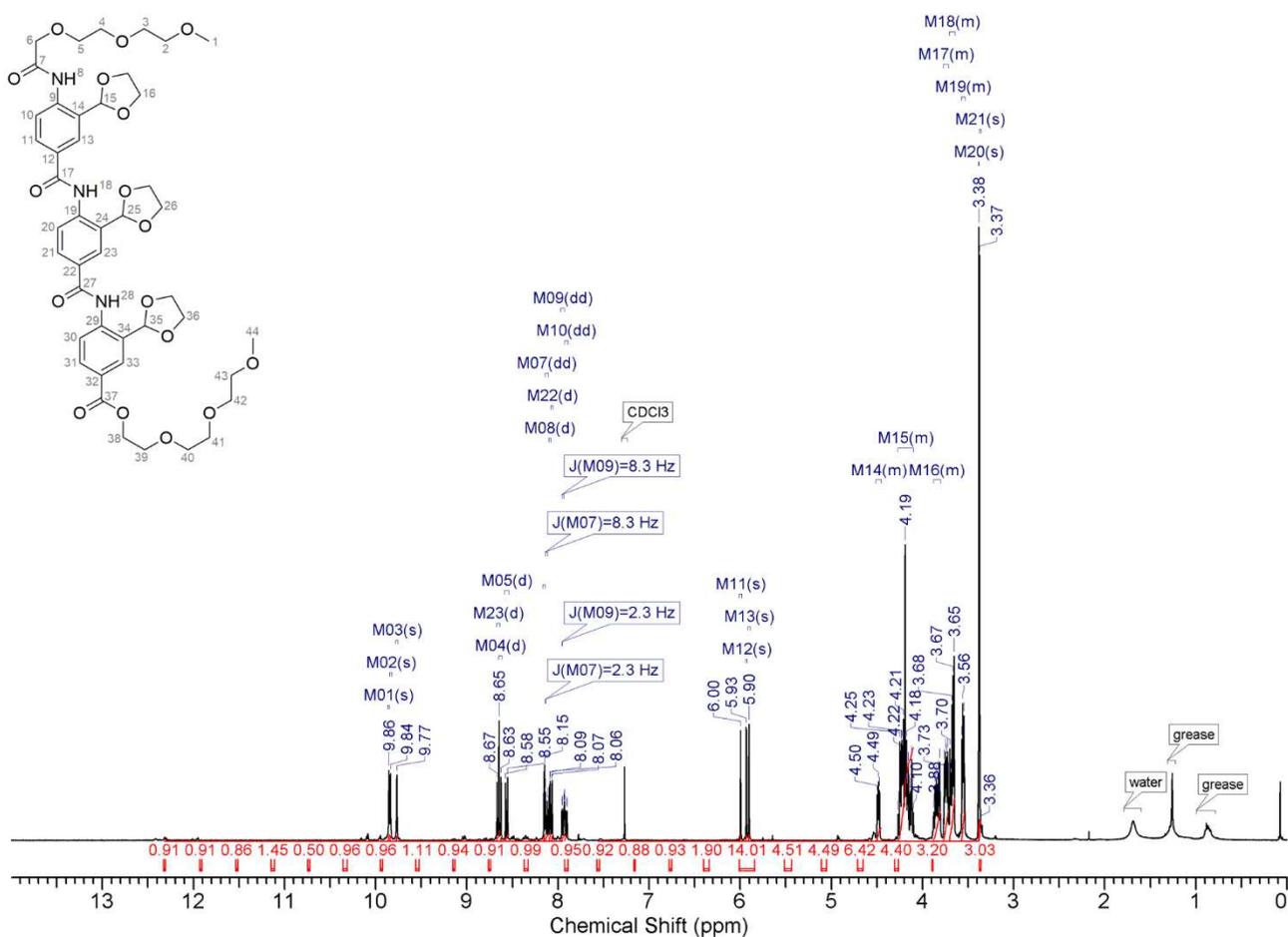


Figure S53. <sup>1</sup>H NMR spectrum (400 MHz) of compound **3b** in CDCl<sub>3</sub> at 298 K.

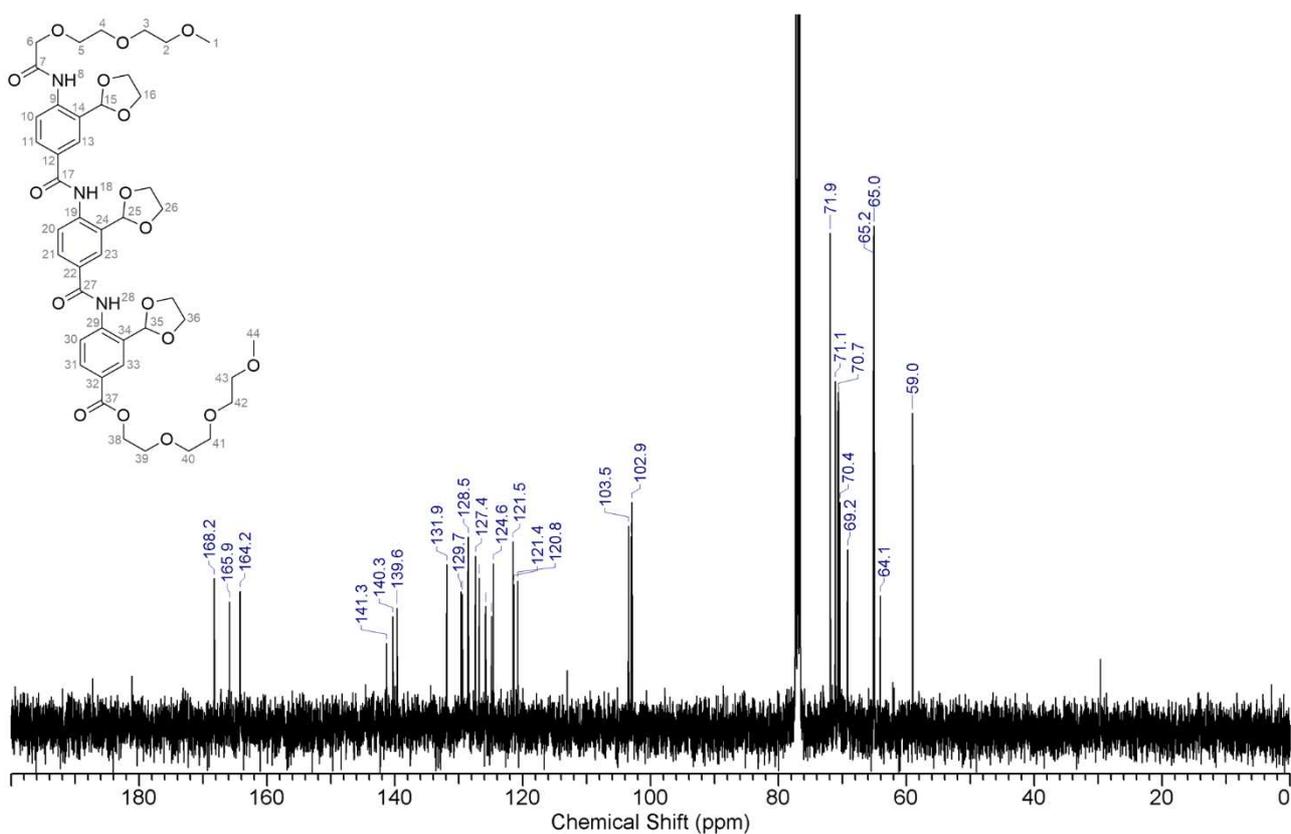


Figure S54. <sup>13</sup>C NMR spectrum (101 MHz) of compound **3b** in CDCl<sub>3</sub> at 298 K.

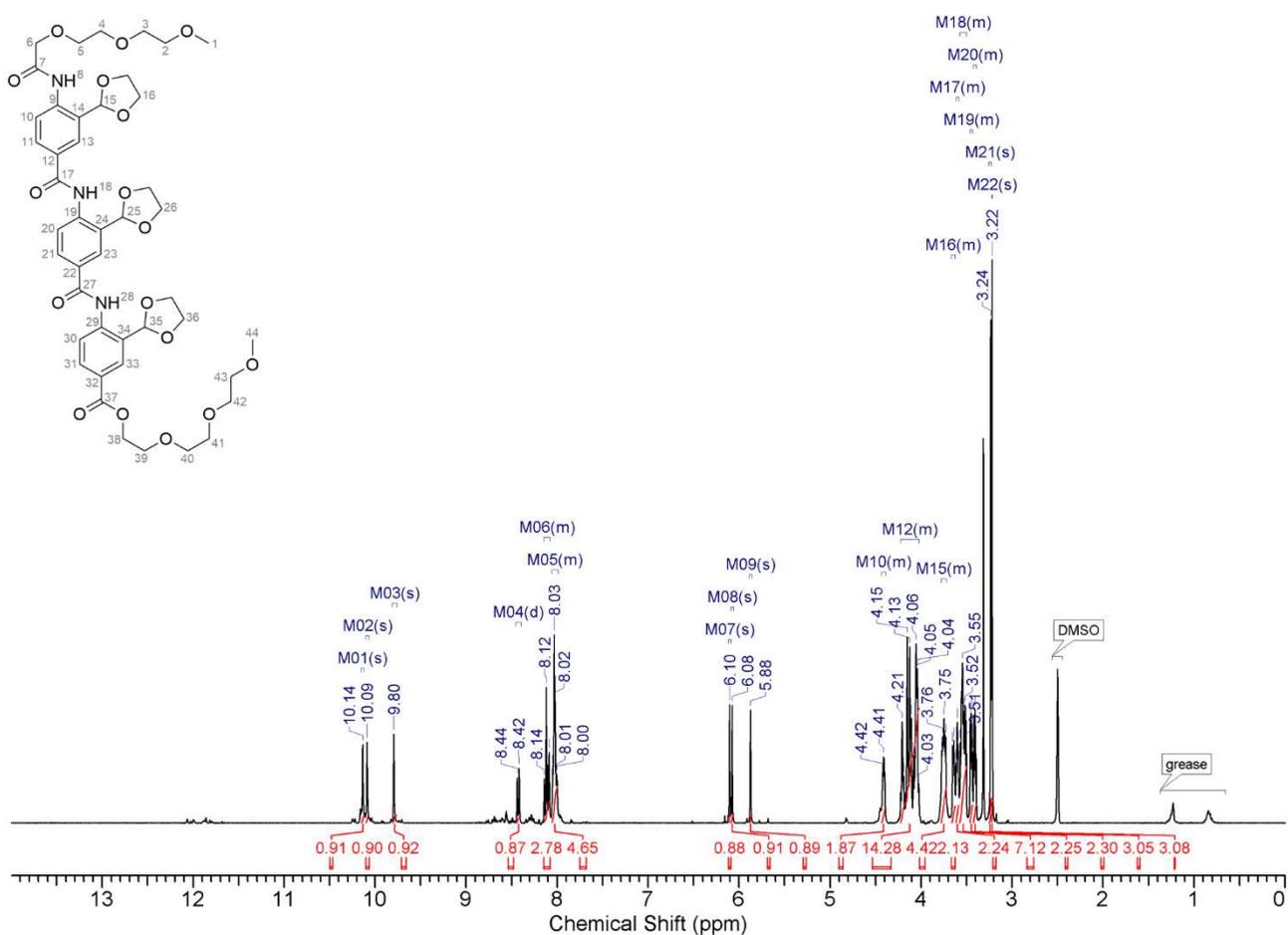


Figure S55. <sup>1</sup>H NMR spectrum (400 MHz) of compound **3b** in DMSO-*d*<sub>6</sub> at 298 K.

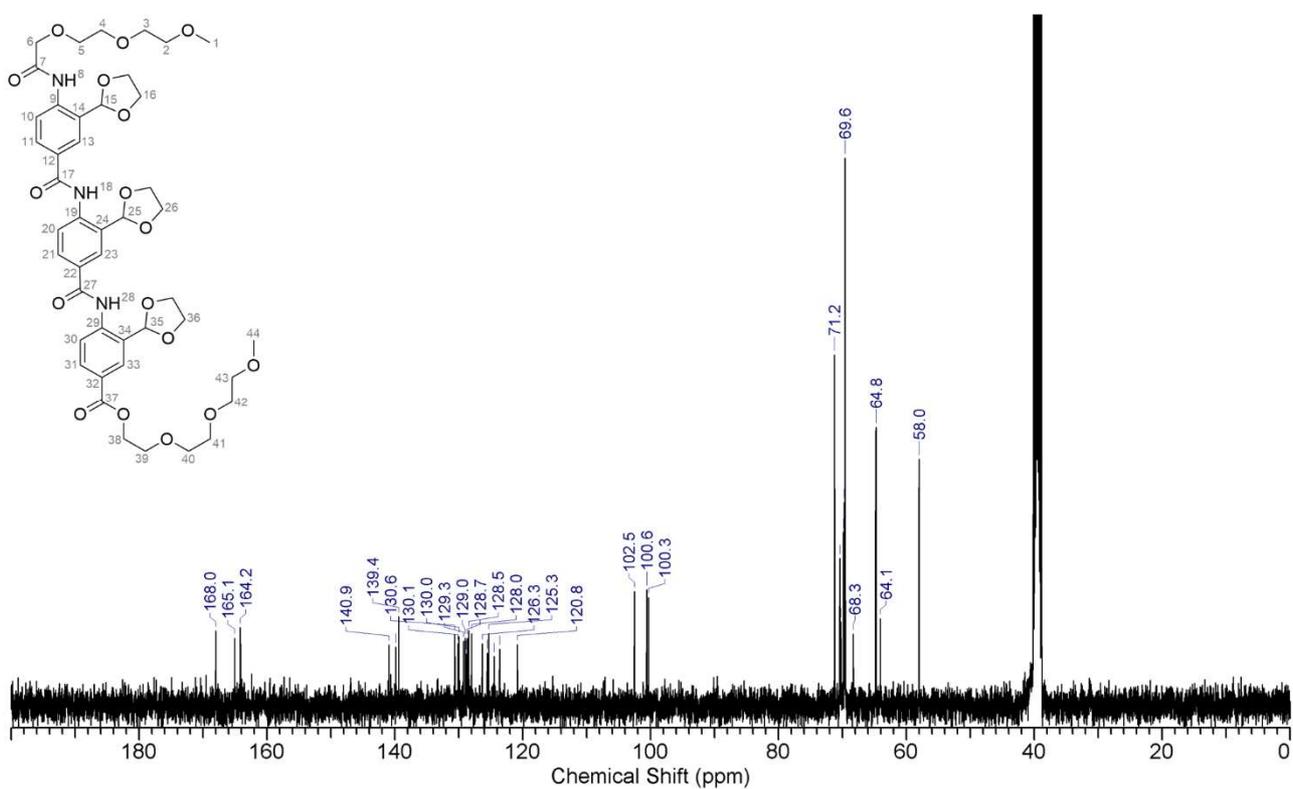


Figure S56. <sup>13</sup>C NMR spectrum (101 MHz) of compound **3b** in DMSO-*d*<sub>6</sub> at 298 K.

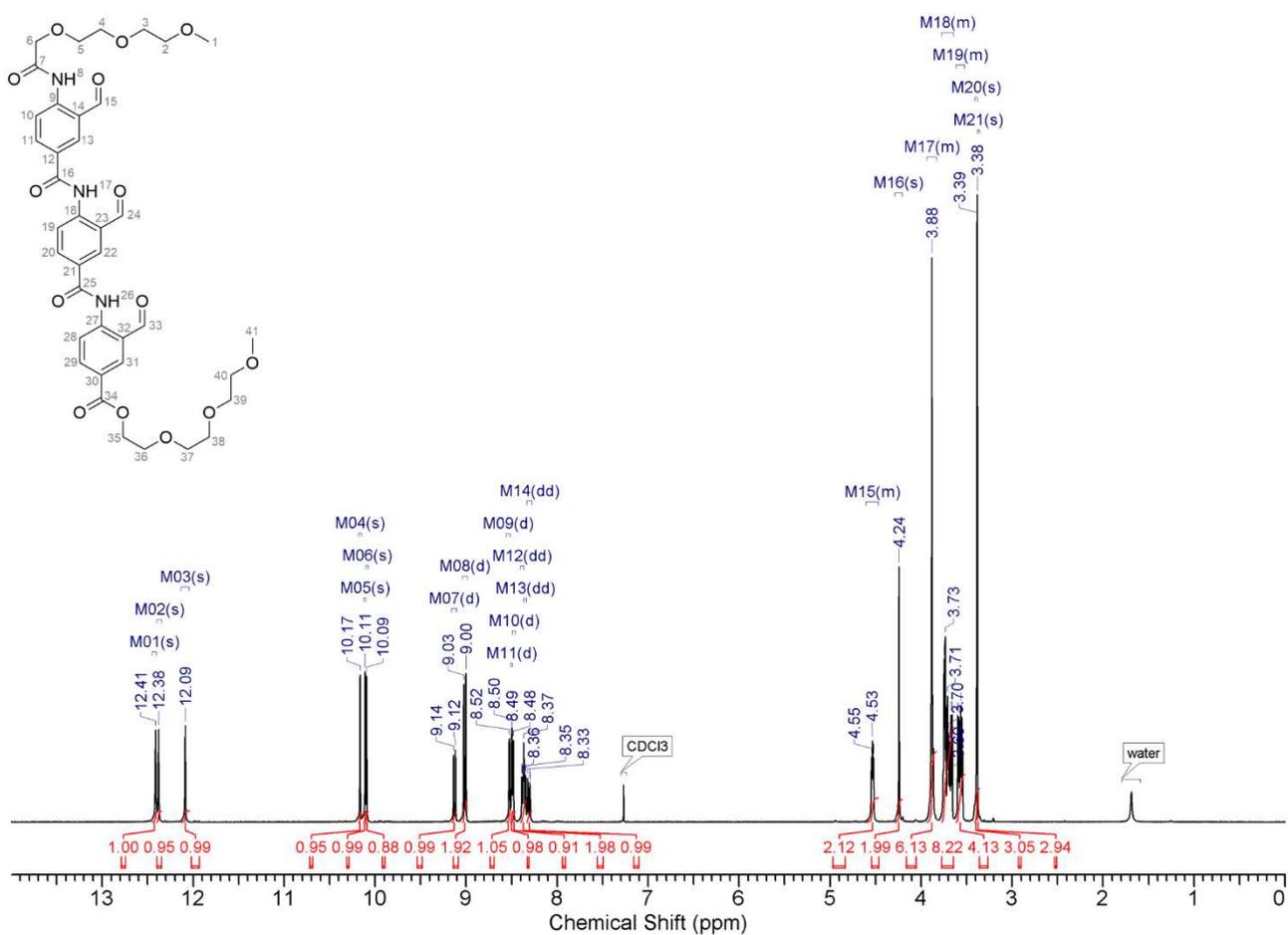


Figure S57. <sup>1</sup>H NMR spectrum (400 MHz) of compound **6b** in CDCl<sub>3</sub> at 298 K.

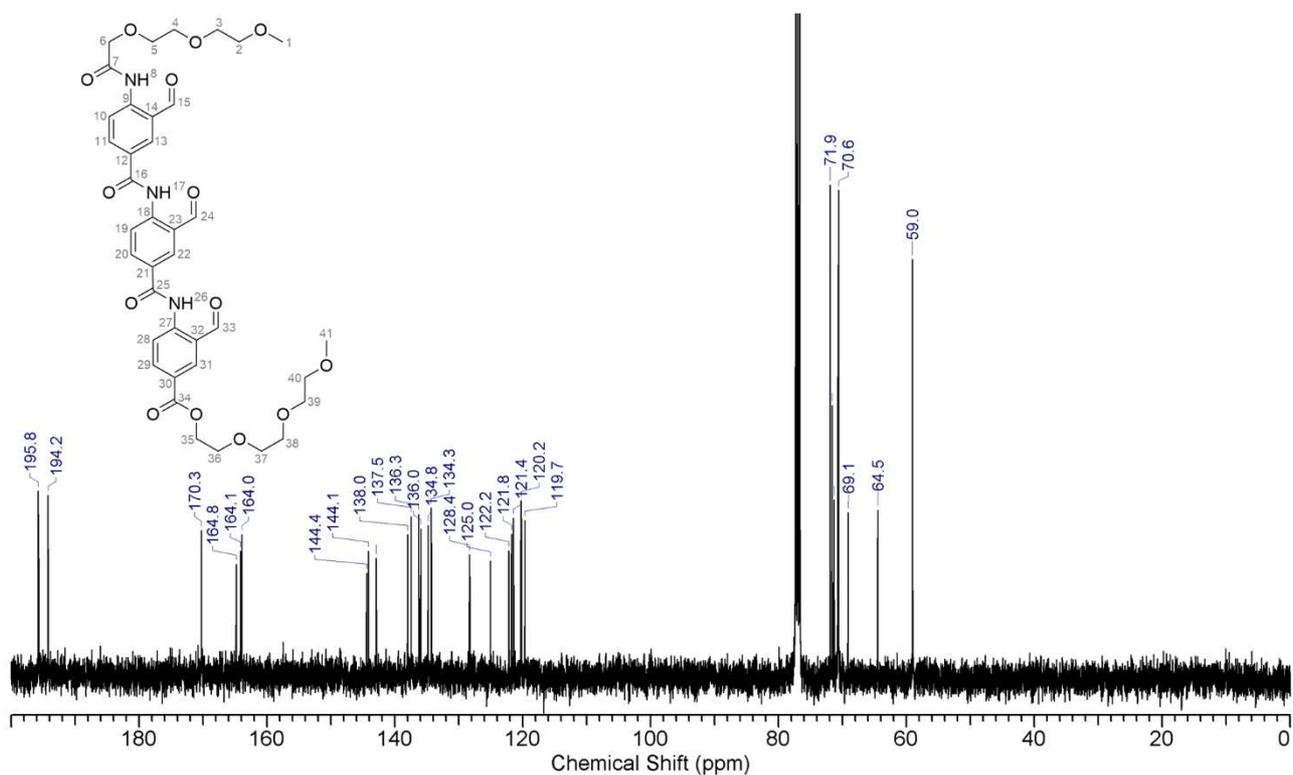


Figure S58. <sup>13</sup>C NMR spectrum (101 MHz) of compound **6b** in CDCl<sub>3</sub> at 298 K.

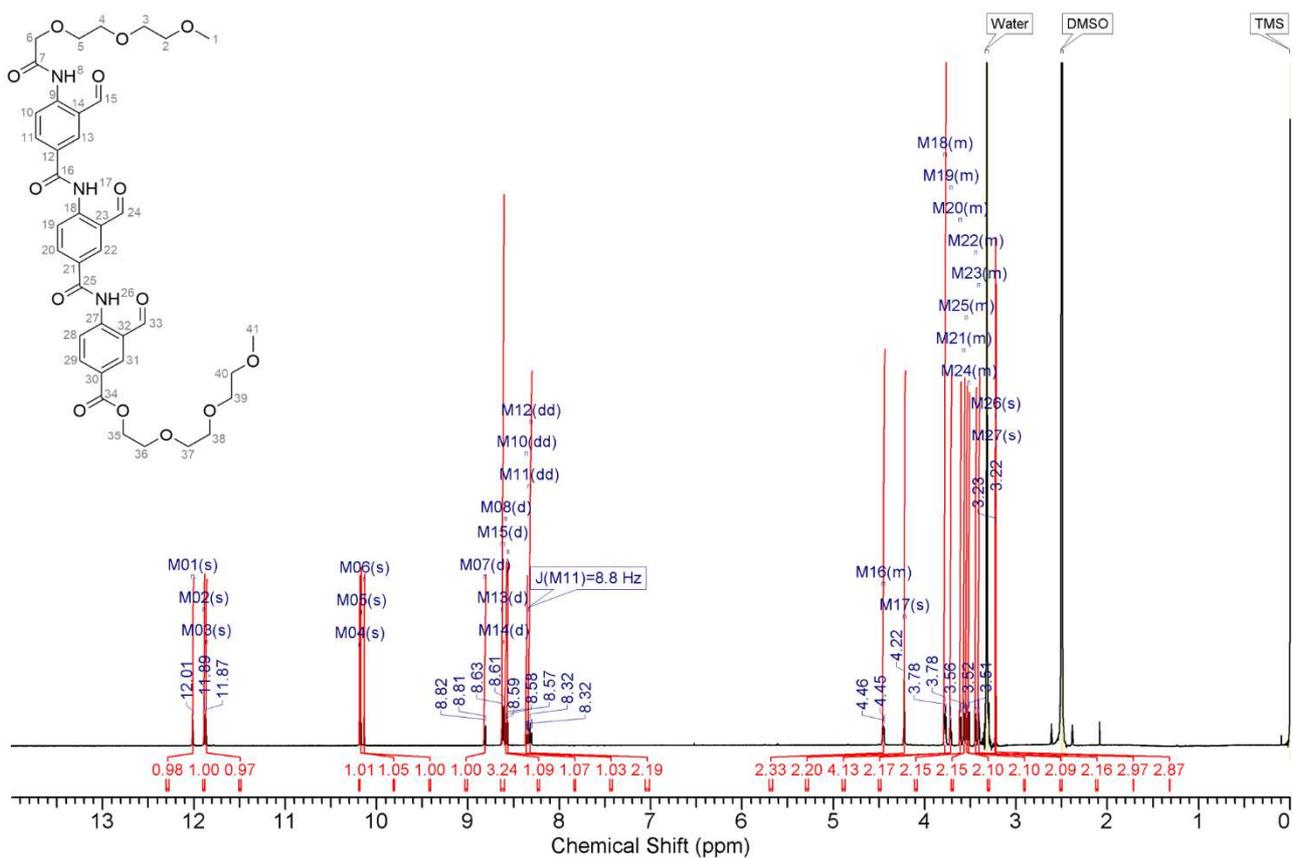


Figure S59. <sup>1</sup>H NMR spectrum (600 MHz) of compound **6b** in DMSO-*d*<sub>6</sub> at 298 K.

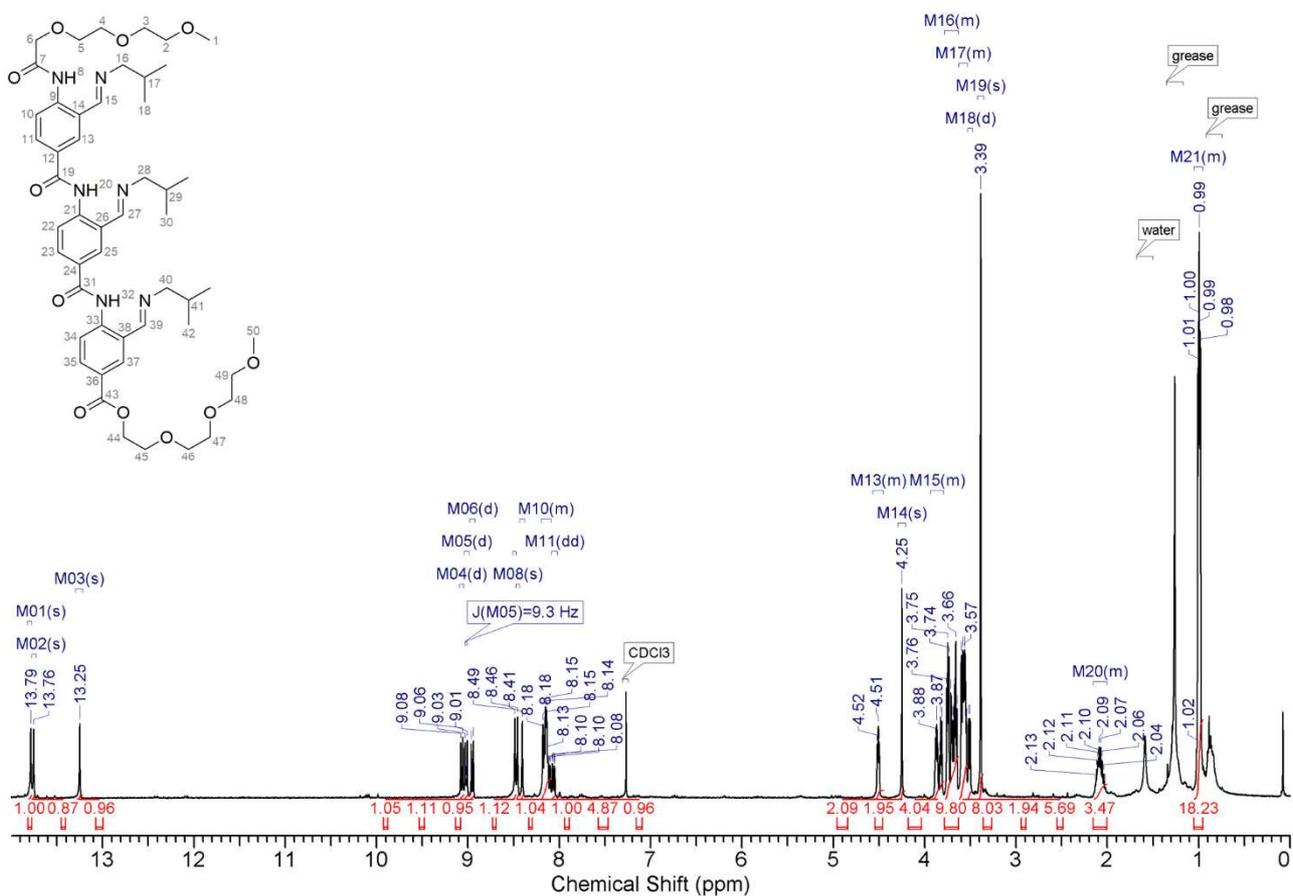


Figure S60. <sup>1</sup>H NMR spectrum (400 MHz) of compound **9b** in CDCl<sub>3</sub> at 298 K.

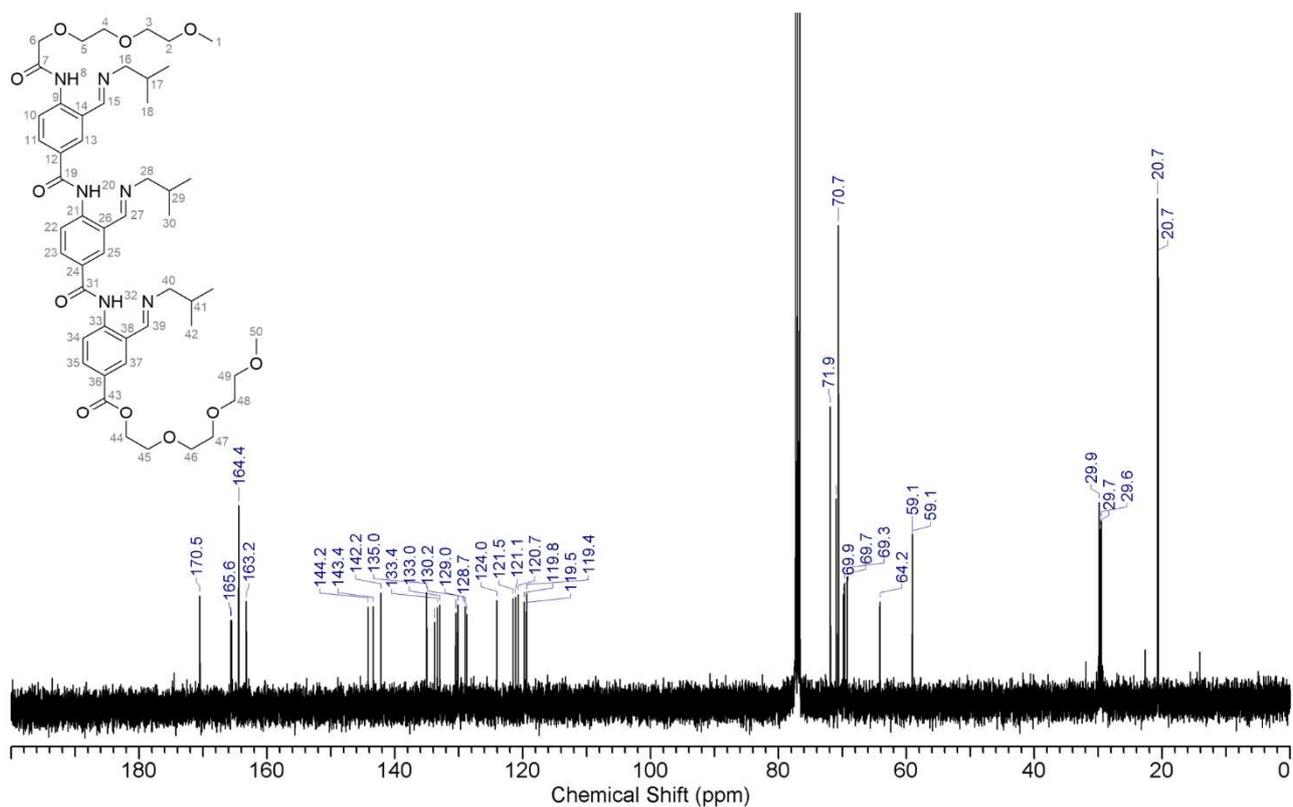


Figure S61.  $^{13}\text{C}$  NMR spectrum (101 MHz) of compound **9b** in  $\text{CDCl}_3$  at 298 K.

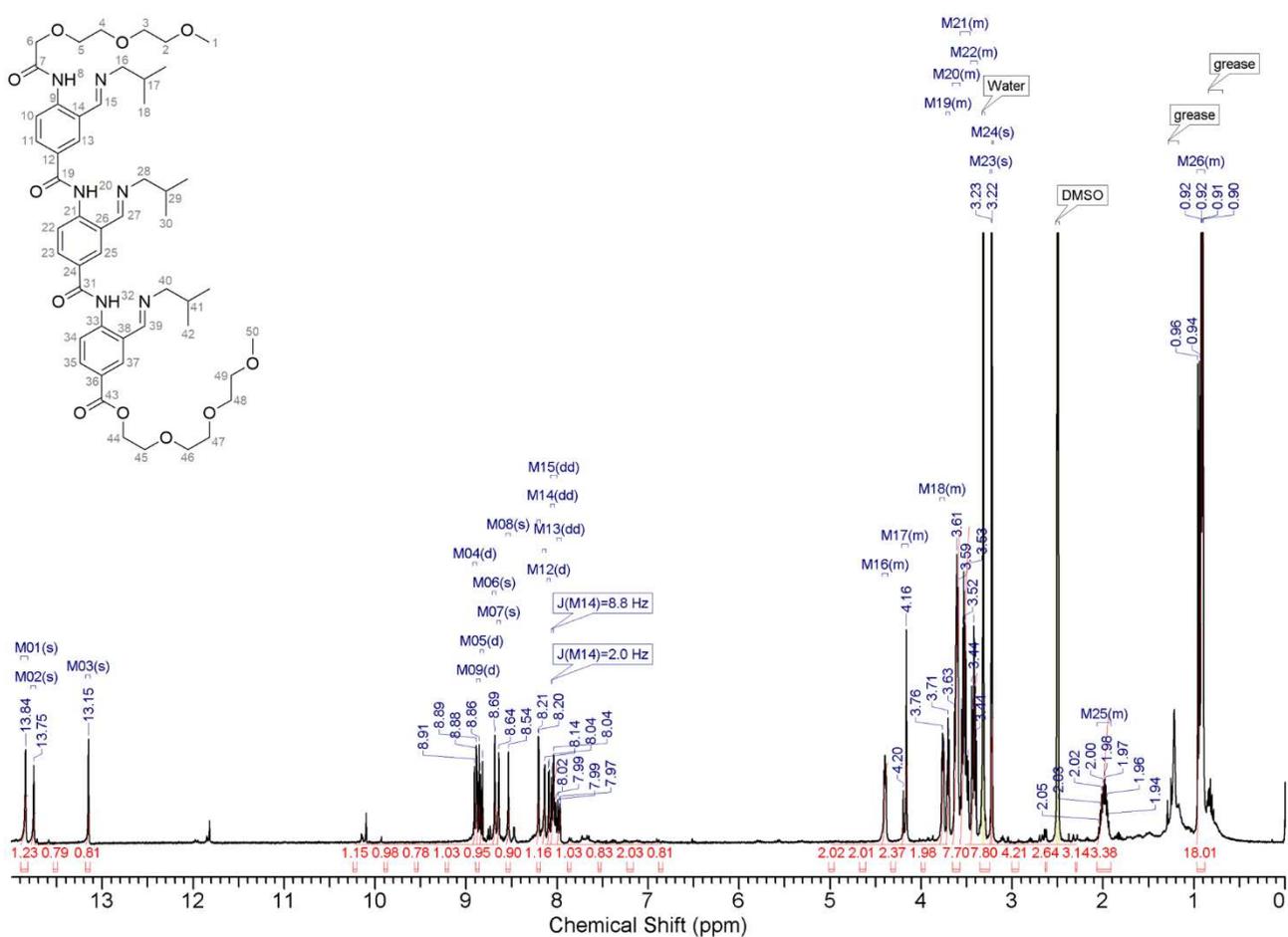
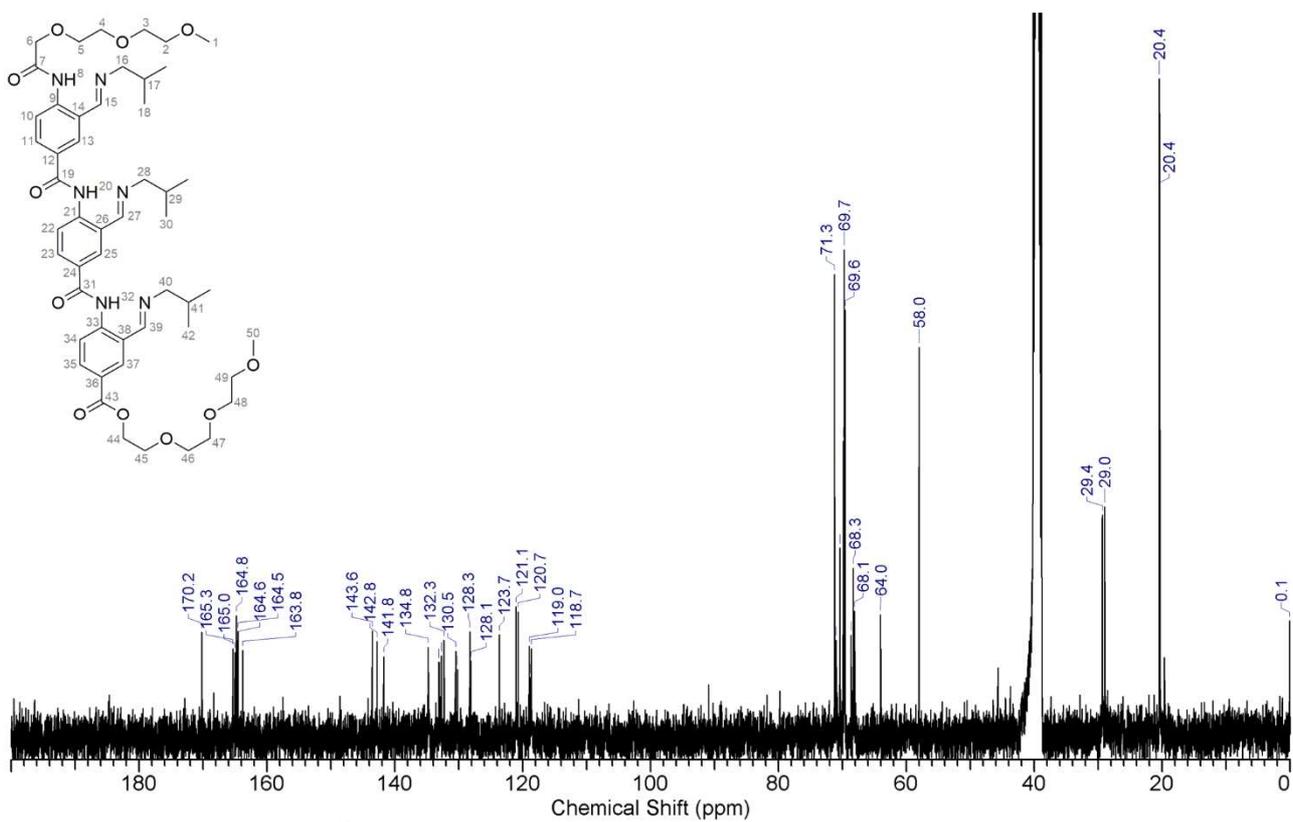


Figure S62.  $^1\text{H}$  NMR spectrum (400 MHz) of compound **9b** in  $\text{DMSO}-d_6$  at 298 K.



**Figure S63.**  $^{13}\text{C}$  NMR spectrum (101 MHz) of compound **9b** in  $\text{DMSO-}d_6$  at 298 K.

### 3. Conformational analysis

#### 3.1. Single crystal X-ray diffraction (solid state)

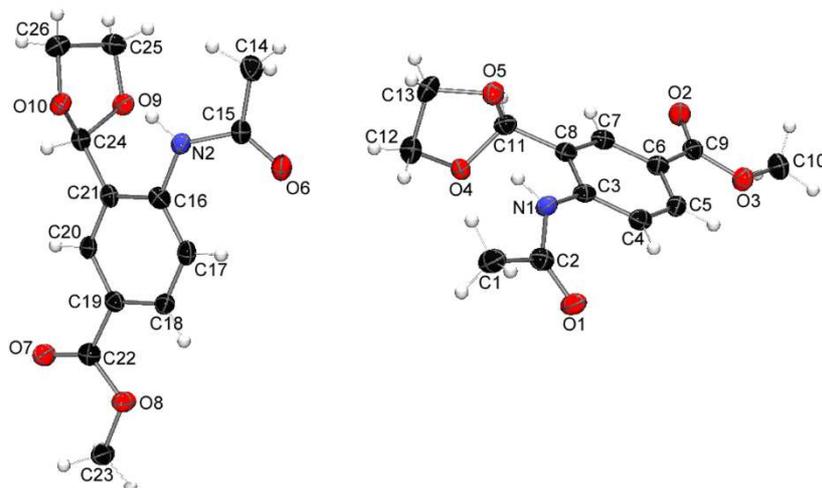
Low temperature single crystal X-ray diffraction was performed on crystals of monomers **1a**, **4a** and **7a**, and on dimers **8a**, **8b** and **14** at the X-ray facilities in New York University. Data collection was performed with a Bruker SMART APEXII CCD area detector on a D8 goniometer at 100 K. The temperature during data collection was controlled to 100 K with an Oxford Cryosystems Series 700+ Cooler. Preliminary lattice parameters and orientation matrices were obtained from three sets of frames. Data were collected using graphite-monochromated and 0.5 mm-MonoCap-collimated MoK $\alpha$  radiation with the  $\omega$  scan method.<sup>2</sup> Data were processed with the INTEGRATE program of the APEX2 software for reduction and cell refinement.<sup>2</sup> Multi-scan absorption corrections were applied using the SCALE program for area detector. The structure was solved using SHELXS-97 and refined using SHELXL-2014/7.<sup>3</sup> In all cases non-hydrogen atoms were refined anisotropically until convergence. Hydrogen atoms were stereochemically fixed at idealized positions (unless where they are involved in hydrogen bonding) and then refined isotropically. Hydrogen bonds are calculated using the HTAB command in SHELXL-2014/7. Graphics were generated using ORTEP-III, MERCURY 3.0 or ViewerLite and Pov-Ray. Crystallographic data have been deposited with the Cambridge Crystallographic Data Centre (CCDC 1487433-1487438) and copies can be obtained free of charge via [www.ccdc.cam.ac.uk/data\\_request/cif](http://www.ccdc.cam.ac.uk/data_request/cif).

#### X-ray data for compound **1a**, CCDC 1487433

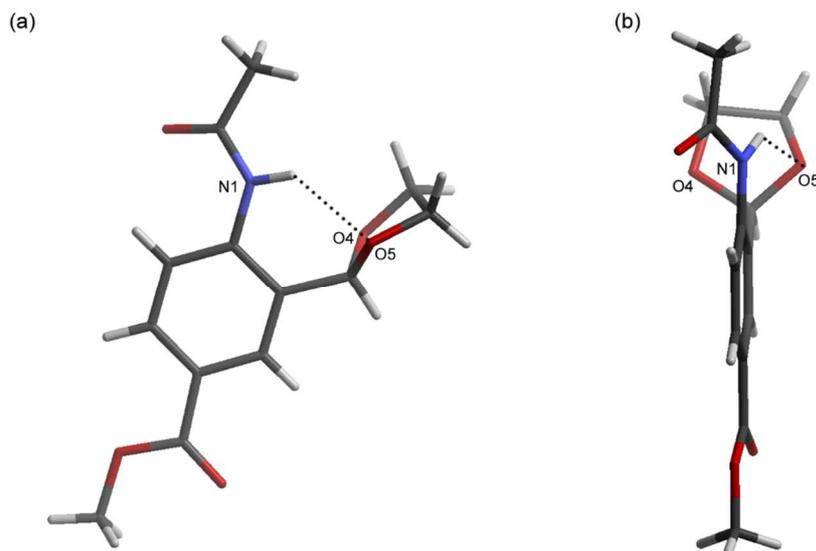
Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **1a** in 1:1 chloroform:dichloromethane. Crystal data for compound **1a**: C<sub>13</sub>H<sub>15</sub>NO<sub>5</sub>,  $M_r$  = 265.26 g/mol, crystal size = 0.54 x 0.09 x 0.06 mm<sup>3</sup>, colourless needle, monoclinic, space group  $P21/n$ ,  $a$  = 24.202(5) Å,  $b$  = 4.5207(10) Å,  $c$  = 24.218(5) Å,  $\alpha$  = 90°,  $\beta$  = 112.835(3)°,  $\gamma$  = 90°,  $V$  = 2442.0(9) Å<sup>3</sup>,  $Z$  = 8,  $\rho_c$  = 1.443 g cm<sup>-3</sup>,  $\mu$  = 0.112 mm<sup>-1</sup>, radiation and wavelength = MoK $\alpha$  (0.71073),  $T$  = 100(2) K,  $\theta_{max}$  = 26.420, reflections collected: 32397, independent reflections: 5023 ( $R_{int}$  = 0.0621), 352 parameters,  $R$  indices (all data):  $R_1$  = 0.0677,  $wR_2$  = 0.1430, final  $R$  indices [ $I > 2\sigma I$ ]:  $R_1$  = 0.0585,  $wR_2$  = 0.1376,  $GOOF$  = 1.086, largest diff. peak and hole = 0.492 and -0.242 e Å<sup>-3</sup>. The crystal was found to be a 2-component twin and a twin matrix [0 0 -1 0 -1 0 -1 0 0] was applied during refinement.

**Table S1.** Hydrogen bond properties for **1a**.

Donor-H...Acceptor	D-H (Å)	H...A (Å)	D...A (Å)	∠D-H...A (°)
N1-H1N...O5	0.89(4)	2.23(4)	2.905(4)	132(3)
N2-H2N...O9	0.87(4)	2.53(4)	3.259(3)	141(3)
N2-H2N...O10	0.87(4)	2.22(4)	2.913(3)	136(3)



**Figure S64.** ORTEP diagram of the unit cell of **1a** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.



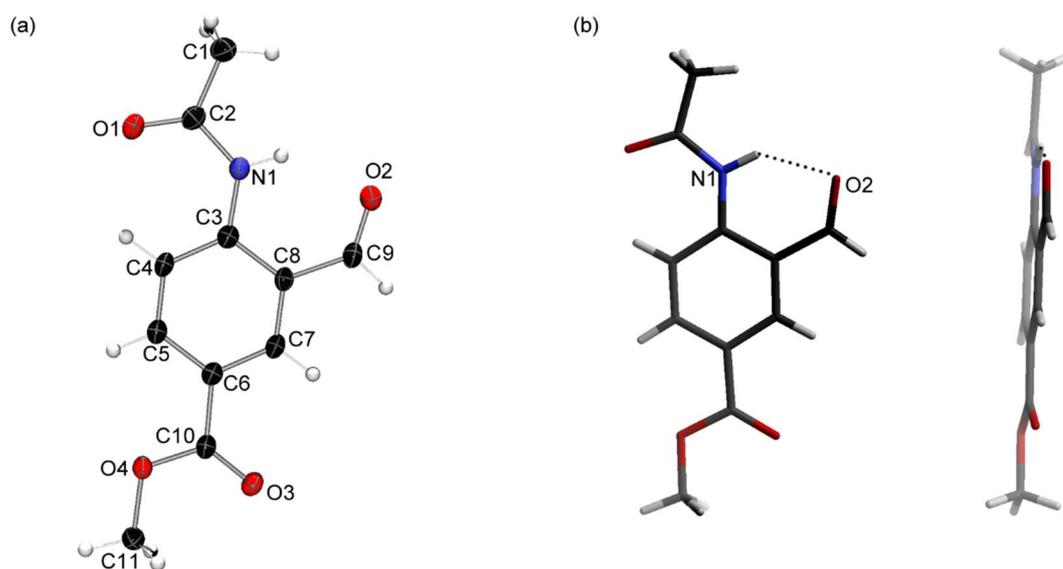
**Figure S65.** Two views of the crystal structure of **1a** showing intramolecular hydrogen bonding (dashed lines). Only atoms involved in hydrogen bonding are labelled for clarity. The amide NH (N1) is not bifurcated and only shows a hydrogen bond with one of the oxygen atoms of the acetal group (O5).

## X-ray data for compound **4a**, CCDC 1487438

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **4a** in diethyl ether. Crystal data for compound **4a**: C<sub>11</sub>H<sub>11</sub>NO<sub>4</sub>, *M<sub>r</sub>* = 221.21 g/mol, crystal size = 0.32 x 0.19 x 0.02 mm<sup>3</sup>, colourless block, triclinic, space group *P*-1, *a* = 3.8385(6) Å, *b* = 10.3480(17) Å, *c* = 12.844(2) Å,  $\alpha$  = 93.967(3) °,  $\beta$  = 90.145(3) °,  $\gamma$  = 95.188(3) °, *V* = 506.86(14) Å<sup>3</sup>, *Z* = 2,  $\rho_c$  = 1.449 g cm<sup>-3</sup>,  $\mu$  = 0.112 mm<sup>-1</sup>, radiation and wavelength = MoK $\alpha$  (0.71073), *T* = 100(2) K,  $\theta_{max}$  = 27.902, reflections collected: 9353, independent reflections: 2393 (*R*<sub>int</sub> = 0.0278), 149 parameters, *R* indices (all data): *R*<sub>1</sub> = 0.0624, *wR*<sub>2</sub> = 0.1662, final *R* indices [*I* > 2 $\sigma$ *I*]: *R*<sub>1</sub> = 0.0551, *wR*<sub>2</sub> = 0.1581, *GOOF* = 1.066, largest diff. peak and hole = 0.520 and -0.318 e Å<sup>3</sup>.

**Table S2.** Hydrogen bond properties for **4a**.

Donor-H...Acceptor	D-H (Å)	H...A (Å)	D...A (Å)	$\angle$ D-H...A (°)
N1-H1N...O2	0.87(2)	2.00(2)	2.7123(19)	137.7(18)



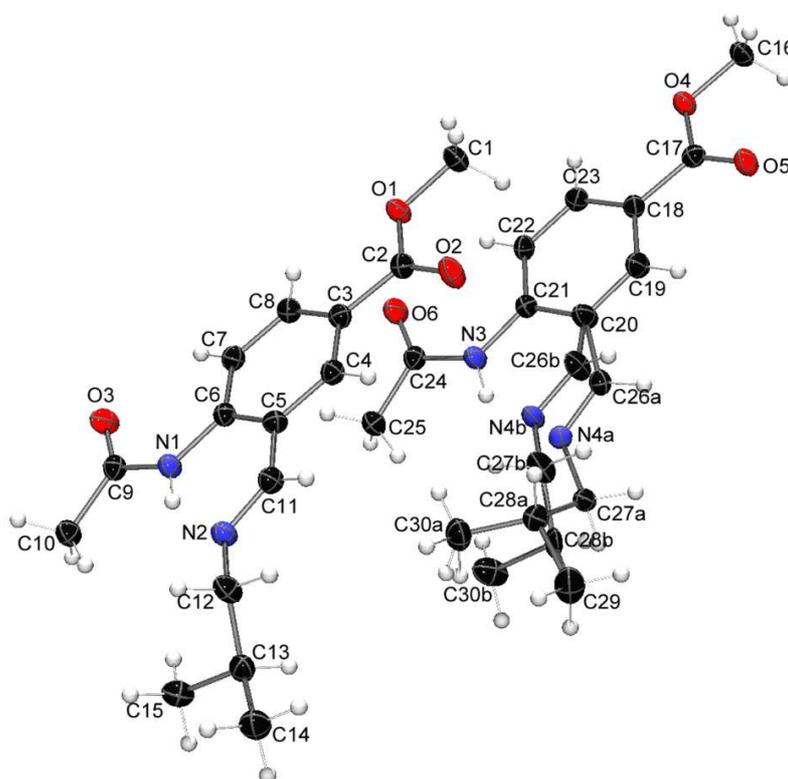
**Figure S66.** (a) ORTEP diagram of the unit cell of **4a** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Two views of the crystal structure of **4a** showing intramolecular hydrogen bonding (dashed lines). Only atoms involved in hydrogen bonding are labelled for clarity.

## X-ray data for compound **7a**, CCDC 1487434

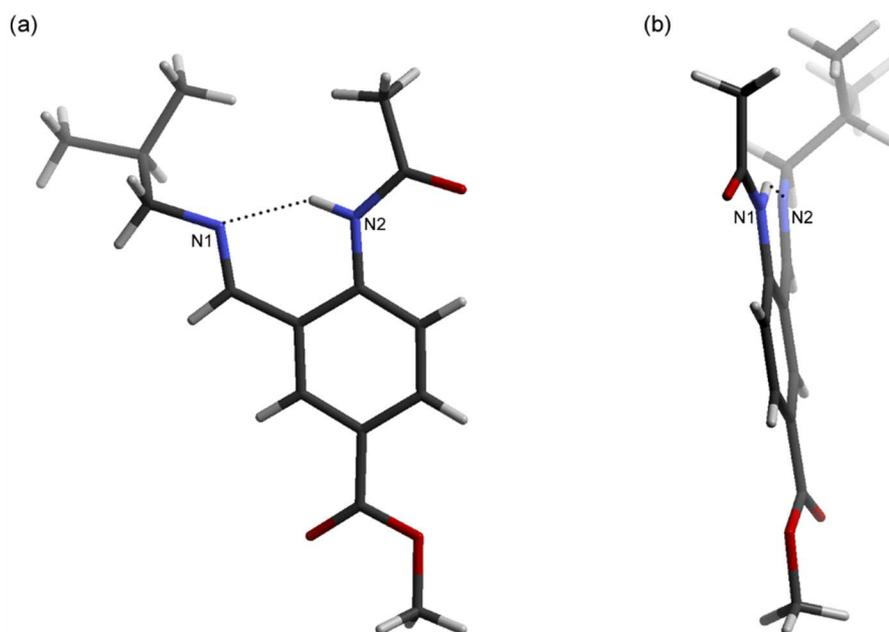
Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **7a** in 1:1 chloroform:dichloromethane. Crystal data for compound **7a**:  $C_{15}H_{20}N_2O_3$ ,  $M_r = 276.33$  g/mol, crystal size = 0.40 x 0.13 x 0.06 mm<sup>3</sup>, pale yellow block, triclinic, space group  $P-1$ ,  $a = 10.9082(7)$  Å,  $b = 10.9750(7)$  Å,  $c = 12.5933(8)$  Å,  $\alpha = 79.4910(10)^\circ$ ,  $\beta = 80.0990(10)^\circ$ ,  $\gamma = 89.4990(10)^\circ$ ,  $V = 1459.87(16)$  Å<sup>3</sup>,  $Z = 4$ ,  $\rho_c = 1.257$  g cm<sup>-3</sup>,  $\mu = 0.088$  mm<sup>-1</sup>, radiation and wavelength = MoK $\alpha$  (0.71073),  $T = 100(2)$  K,  $\theta_{max} = 26.405$ , reflections collected: 19880, independent reflections: 5979 ( $R_{int} = 0.0255$ ), 415 parameters,  $R$  indices (all data):  $R_1 = 0.0636$ ,  $wR_2 = 0.1471$ , final  $R$  indices [ $I > 2\sigma I$ ]:  $R_1 = 0.0497$ ,  $wR_2 = 0.1358$ ,  $GOOF = 1.044$ , largest diff. peak and hole = 0.360 and -0.351 e Å<sup>3</sup>. Disorder was present in one of the isobutyl chains, which was split over two parts with 33% and 77% abundance.

**Table S3.** Hydrogen bond properties for **7a**.

Donor-H...Acceptor	D-H (Å)	H...A (Å)	D...A (Å)	$\angle D-H...A$ (°)
N1-H1N...N2	0.89(2)	1.95(2)	2.7157(18)	143.8(17)
N3-H3N...N4A	0.93(2)	1.91(2)	2.720(2)	143.8(17)
N3-H3N...N4B	0.93(2)	1.99(2)	2.773(6)	140.1(17)



**Figure S67.** ORTEP diagram of the unit cell of **7a** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.



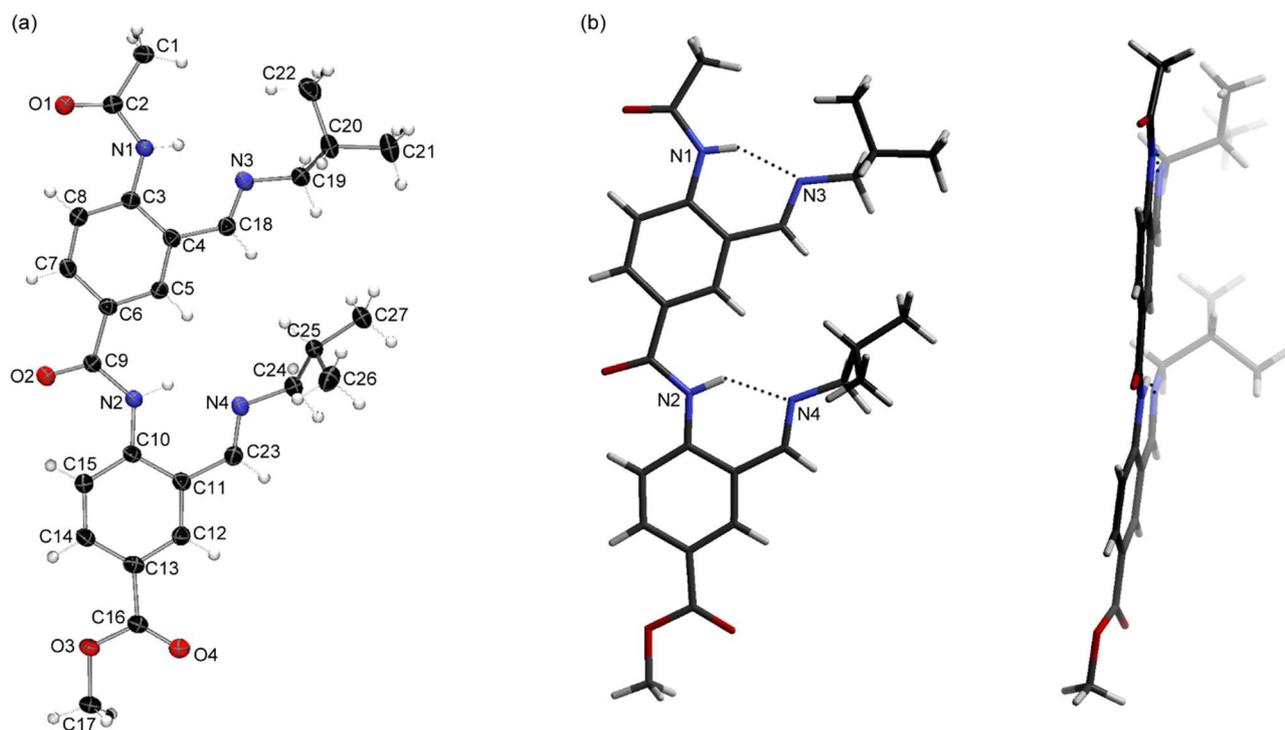
**Figure S68.** Two views of the crystal structure of **7a** showing intramolecular hydrogen bonding (dashed lines). Only atoms involved in hydrogen bonding are labelled for clarity.

#### X-ray data for compound **8a**, CCDC 1487435

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **8a** in dichloromethane. Crystal data for compound **8a**:  $C_{27}H_{34}N_4O_4$ ,  $M_r = 478.58$  g/mol, crystal size = 0.39 x 0.13 x 0.10 mm<sup>3</sup>, colourless block, triclinic, space group  $P-1$ ,  $a = 7.1080(14)$  Å,  $b = 13.001(3)$  Å,  $c = 15.084(3)$  Å,  $\alpha = 112.002(3)^\circ$ ,  $\beta = 102.265(3)^\circ$ ,  $\gamma = 90.523(3)^\circ$ ,  $V = 1256.8(4)$  Å<sup>3</sup>,  $Z = 2$ ,  $\rho_c = 1.265$  g cm<sup>-3</sup>,  $\mu = 0.086$  mm<sup>-1</sup>, radiation and wavelength = MoK $\alpha$  (0.71073),  $T = 100(2)$  K,  $\theta_{max} = 27.150$ , reflections collected: 21850, independent reflections: 5526 ( $R_{int} = 0.0342$ ), 324 parameters,  $R$  indices (all data):  $R_1 = 0.0788$ ,  $wR_2 = 0.1939$ , final  $R$  indices [ $I > 2\sigma I$ ]:  $R_1 = 0.0623$ ,  $wR_2 = 0.1825$ ,  $GOOF = 1.092$ , largest diff. peak and hole = 0.464 and -0.386 e Å<sup>3</sup>.

**Table S4.** Hydrogen bond properties for **8a**.

Donor-H...Acceptor	D-H (Å)	H...A (Å)	D...A (Å)	$\angle D-H...A$ (°)
N1-H1N...N3	0.86(3)	1.93(3)	2.694(3)	147(3)
N2-H2N...N4	0.92(3)	1.93(3)	2.701(3)	140(3)



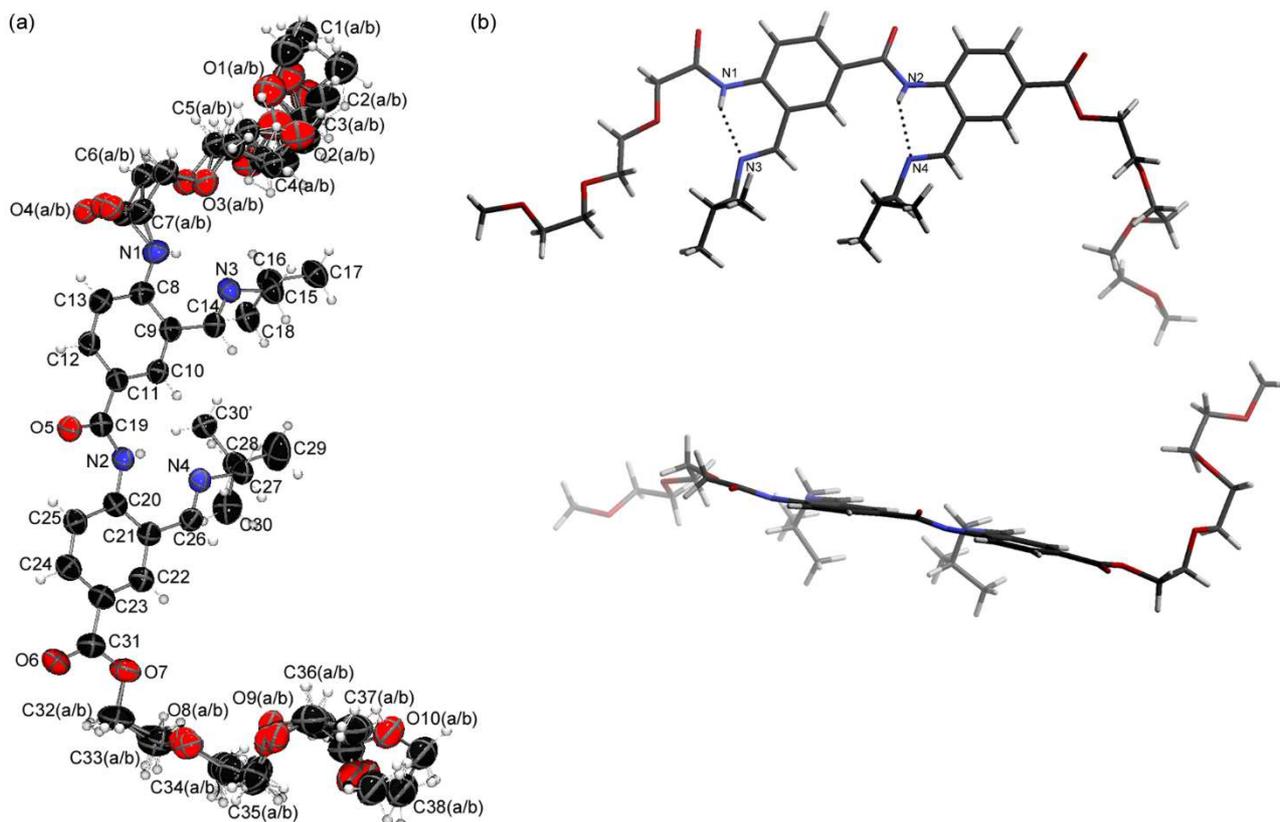
**Figure S69.** (a) ORTEP diagram of the unit cell of **8a** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Two views of the crystal structure of **8a** showing intramolecular hydrogen bonding (dashed lines). Only atoms involved in hydrogen bonding are labelled for clarity.

### X-ray data for compound **8b**, CCDC 1487437

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **8b** in 1:1 chloroform:dichloromethane. Crystal data for compound **8b**:  $C_{38}H_{56}N_4O_{10}$ ,  $M_r = 728.86$  g/mol, crystal size = 0.38 x 0.24 x 0.10 mm<sup>3</sup>, colourless block, triclinic, space group *P*-1,  $a = 10.0905(6)$  Å,  $b = 10.9163(6)$  Å,  $c = 18.1284(10)$  Å,  $\alpha = 96.7110(10)^\circ$ ,  $\beta = 93.4450(10)^\circ$ ,  $\gamma = 101.6850(10)^\circ$ ,  $V = 1934.86(19)$  Å<sup>3</sup>,  $Z = 2$ ,  $\rho_c = 1.251$  g cm<sup>-3</sup>,  $\mu = 0.090$  mm<sup>-1</sup>, radiation and wavelength = MoK $\alpha$  (0.71073),  $T = 100(2)$  K,  $\theta_{max} = 25.054$ , reflections collected: 33355, independent reflections: 6827 ( $R_{int} = 0.0274$ ), 564 parameters, 99 restraints, *R* indices (all data):  $R_1 = 0.1308$ ,  $wR_2 = 0.3546$ , final *R* indices [ $I > 2\sigma I$ ]:  $R_1 = 0.1036$ ,  $wR_2 = 0.3158$ ,  $GOOF = 1.464$ , largest diff. peak and hole = 0.254 and -0.159 e Å<sup>3</sup>. Significant disorder was present in the appending ethyleneglycol chains and isobutyl side arms, resulting in lower crystal quality. The disordered atoms were split over various positions and a number of restraint were used to obtain a reasonable fit (EADP, DFIX, DANG, FLAT, SADI, SIMU and ISOR restraints, as well as a DAMP command to achieve good refinement). A SWAT command was also employed in SHELX to account for residual solvent and side chain disorder.

**Table S5.** Hydrogen bond properties for **8b**.

Donor-H...Acceptor	D-H (Å)	H...A (Å)	D...A (Å)	∠D-H...A (°)
N1-H1N...O3	0.71(3)	2.33(3)	2.668(4)	111(3)
N1-H1N...N3	0.71(3)	2.06(3)	2.692(3)	148(3)
N2-H2N...N4	0.60(2)	2.21(2)	2.710(2)	143(3)



**Figure S70.** (a) ORTEP diagram of the unit cell of **8b** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Two views of the crystal structure of **8b** showing intramolecular hydrogen bonding (dashed lines). Only atoms involved in hydrogen bonding are labelled for clarity.

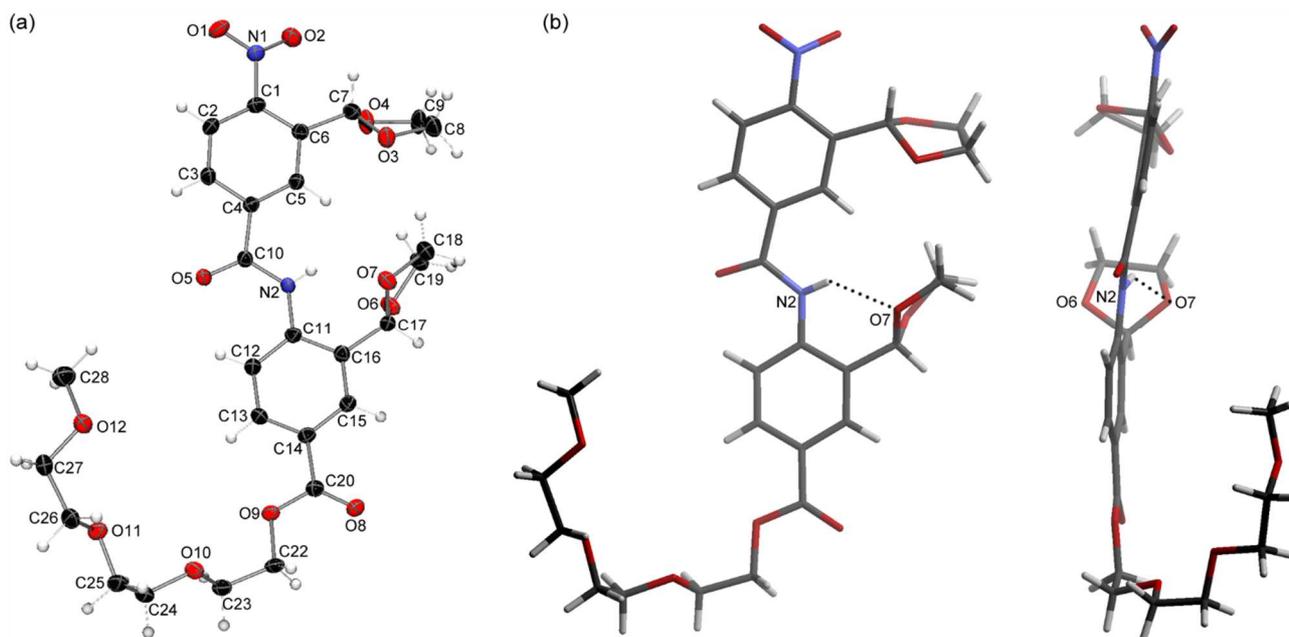
### X-ray data for compound **14**, CCDC 1487436

Single crystals suitable for X-ray diffraction were formed spontaneously during the synthesis of compound **14**. Crystal data for compound **14**:  $C_{27}H_{32}N_2O_{12}$ ,  $M_r = 576.54$  g/mol, crystal size = 0.34 x 0.18 x 0.15 mm<sup>3</sup>, yellow fragment, monoclinic, space group  $P21/n$ ,  $a = 14.5173(5)$  Å,  $b = 12.8034(4)$  Å,  $c = 14.6624(5)$  Å,  $\alpha = 90^\circ$ ,  $\beta = 105.9480(10)^\circ$ ,  $\gamma = 90^\circ$ ,  $V = 2620.42(15)$  Å<sup>3</sup>,  $Z = 4$ ,  $\rho_c = 1.461$  g cm<sup>-3</sup>,  $\mu = 0.116$  mm<sup>-1</sup>, radiation and wavelength = MoK $\alpha$  (0.71073),  $T = 100(2)$  K,  $\theta_{max} = 26.394$ , reflections collected: 46815, independent reflections: 5361 ( $R_{int} = 0.0326$ ), 374

parameters,  $R$  indices (all data):  $R_1 = 0.0564$ ,  $wR_2 = 0.1360$ , final  $R$  indices [ $I > 2\sigma I$ ]:  $R_1 = 0.0466$ ,  $wR_2 = 0.1280$ ,  $GOOF = 1.032$ , largest diff. peak and hole = 0.523 and -0.280 e  $\text{\AA}^3$ .

**Table S6.** Hydrogen bond properties for **14**.

Donor-H $\cdots$ Acceptor	D-H ( $\text{\AA}$ )	H $\cdots$ A ( $\text{\AA}$ )	D $\cdots$ A ( $\text{\AA}$ )	$\angle$ D-H $\cdots$ A ( $^\circ$ )
N2-H1N $\cdots$ O6	0.75(2)	2.64(2)	3.0363(19)	115.5(18)
N2-H1N $\cdots$ O7	0.75(2)	2.26(2)	2.8620(19)	138(2)

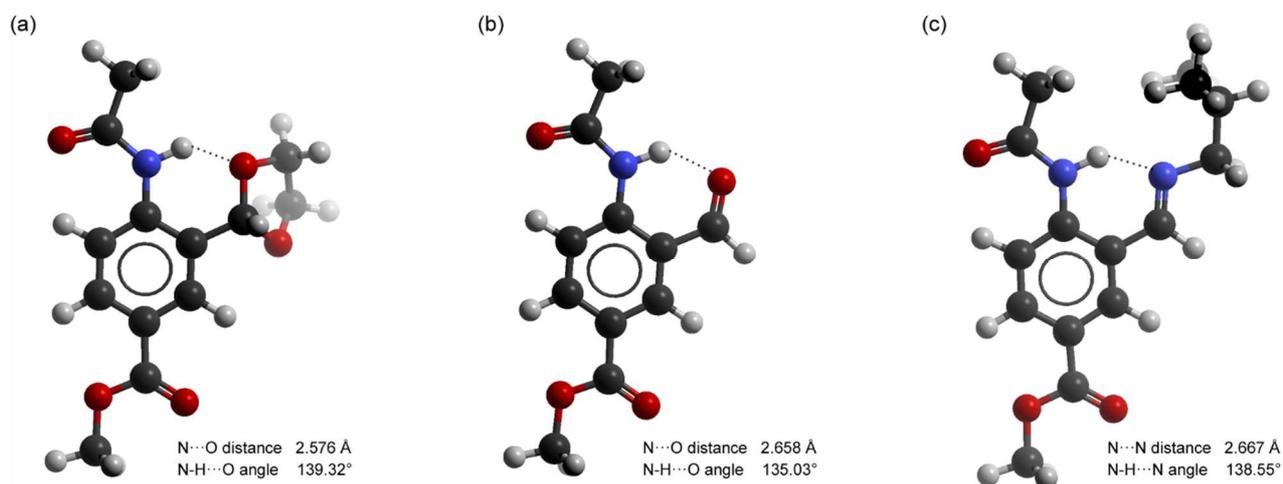


**Figure S71.** (a) ORTEP diagram of the unit cell of **14** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Two views of the crystal structure of **14** showing intramolecular hydrogen bonding (dashed lines). Only atoms involved in hydrogen bonding are labelled for clarity. The amide NH (N2) is not bifurcated and only shows a hydrogen bond with one of the oxygen atoms of the acetal group (O7).

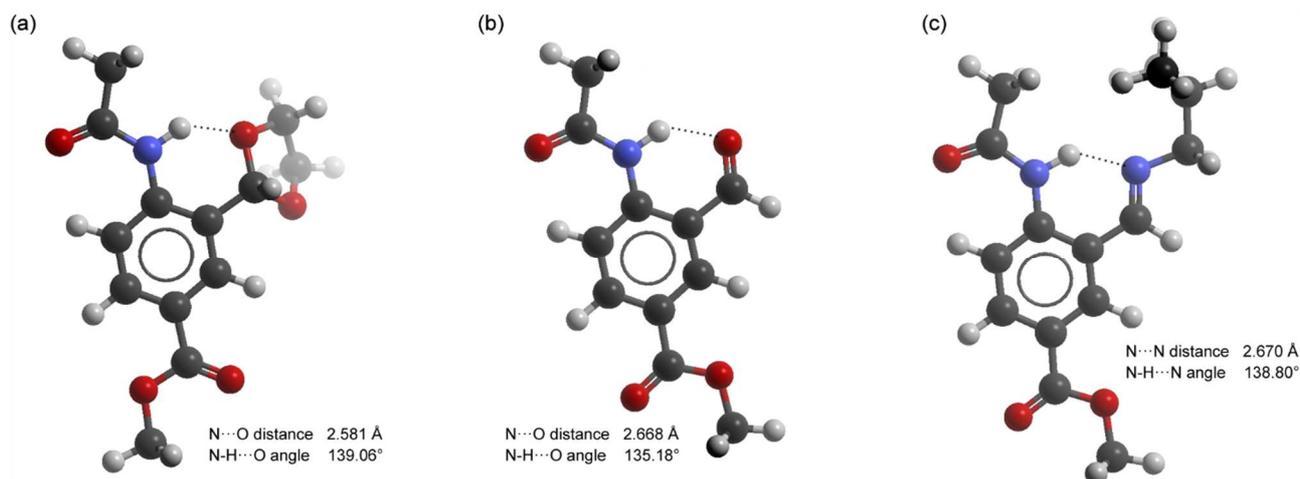
### 3.2. Computational Modelling

All computer modelling was performed using Molecular Operating Environment<sup>TM</sup> version 2014:09 (MOE) and conformational analysis employed an MMFF94x force field with an implicit chloroform or water solvent model (generalized Born solvation model, dielectric constant exterior 4.81 for chloroform and 80.1 for water). Structures were first generally energy minimized, followed by a LowModeMD conformational search of various possible low-energy conformers. In general, the lowest energy conformer found in chloroform and water were very similar. The modelling was performed for the acetal, aldehyde and imine containing structures. Only the structures capped with methyl groups were modelled, because the oligoethylene chains are expected to be too flexible and can adopt many conformers, which will render the conformational search more difficult. The

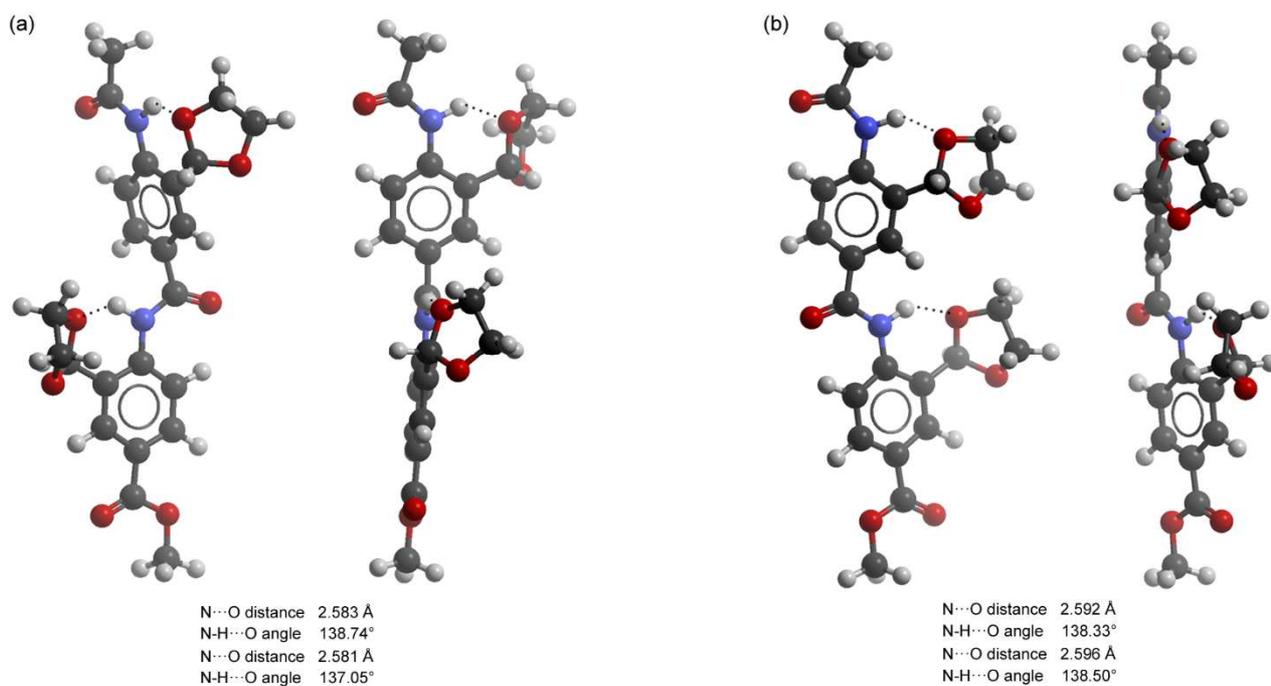
resulting lowest energy conformers are shown in Figures S72-S79. All structures indicate the formation of an intramolecular hydrogen bond with the amide NH as the donor and the oxygen atom in the acetal and aldehyde or the nitrogen atom in the imine functionality as acceptor. In most cases the functional groups are displayed on the same face of the helix mimetic in the lowest energy conformer.



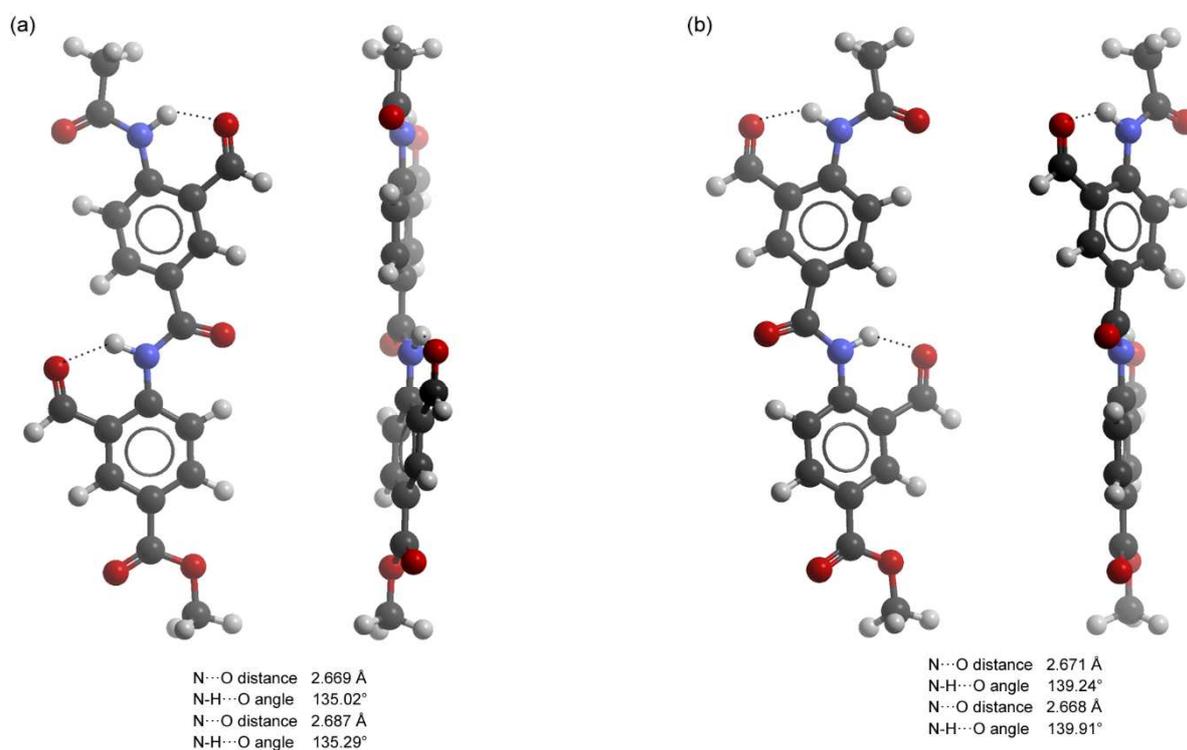
**Figure S72.** Energy-minimized structure of monomers **1a**, **4a** and **7a** in chloroform. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) acetal **1a**, (b) aldehyde **4a**, (c) imine **7a**.



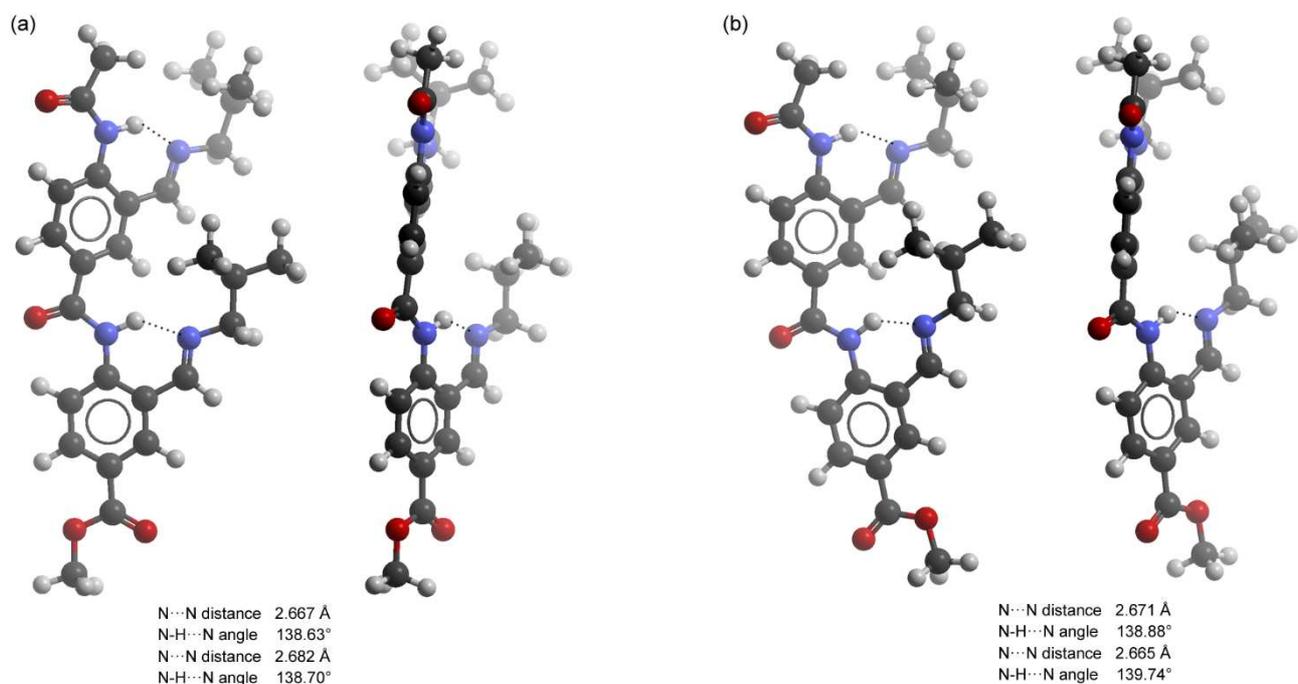
**Figure S73.** Energy-minimized structure of monomers **1a**, **4a** and **7a** in water. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) acetal **1a**, (b) aldehyde **4a**, (c) imine **7a**.



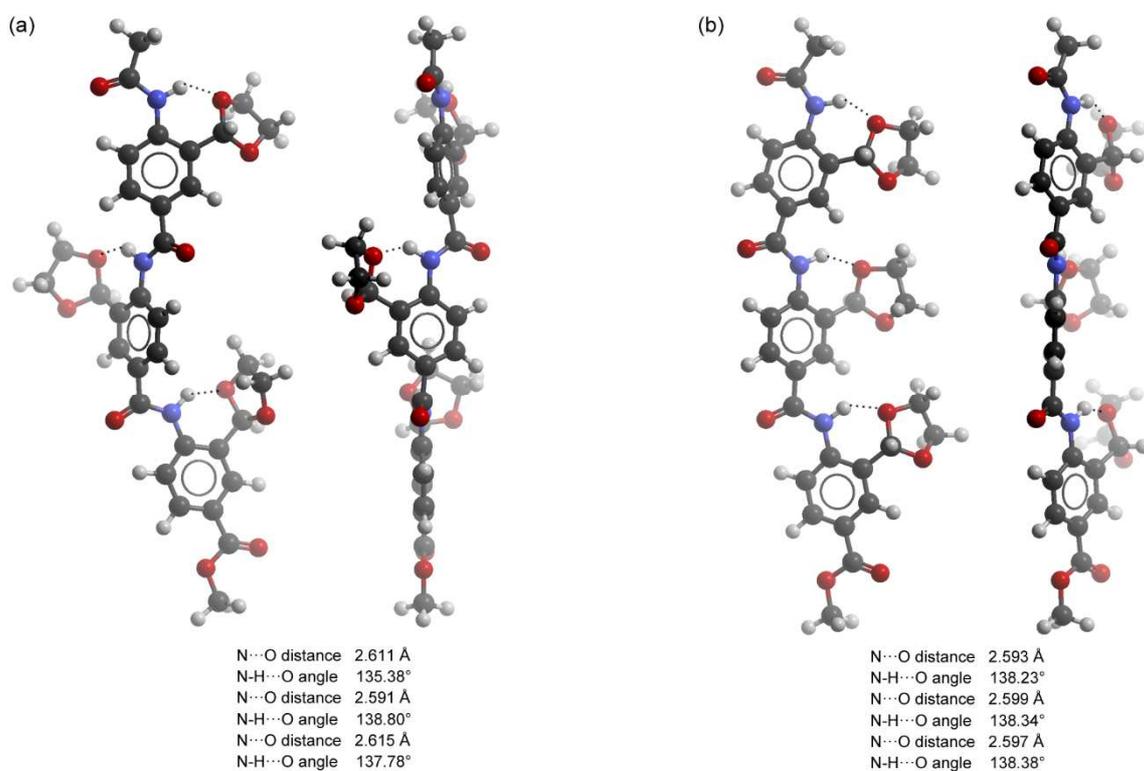
**Figure S74.** Two views on the energy-minimized structure of dimeric acetal **2a** in chloroform and water. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) **2a** in chloroform, (b) **2a** in water.



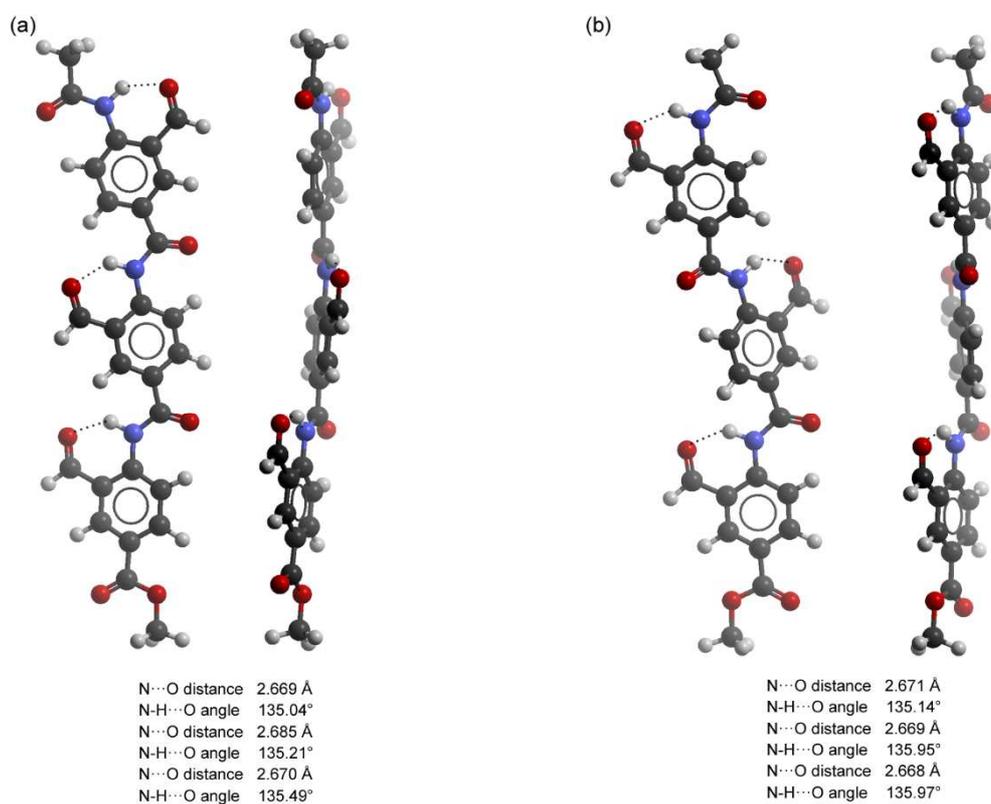
**Figure S75.** Two views on the energy-minimized structure of dimeric aldehyde **5a** in chloroform and water. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) **5a** in chloroform, (b) **5a** in water.



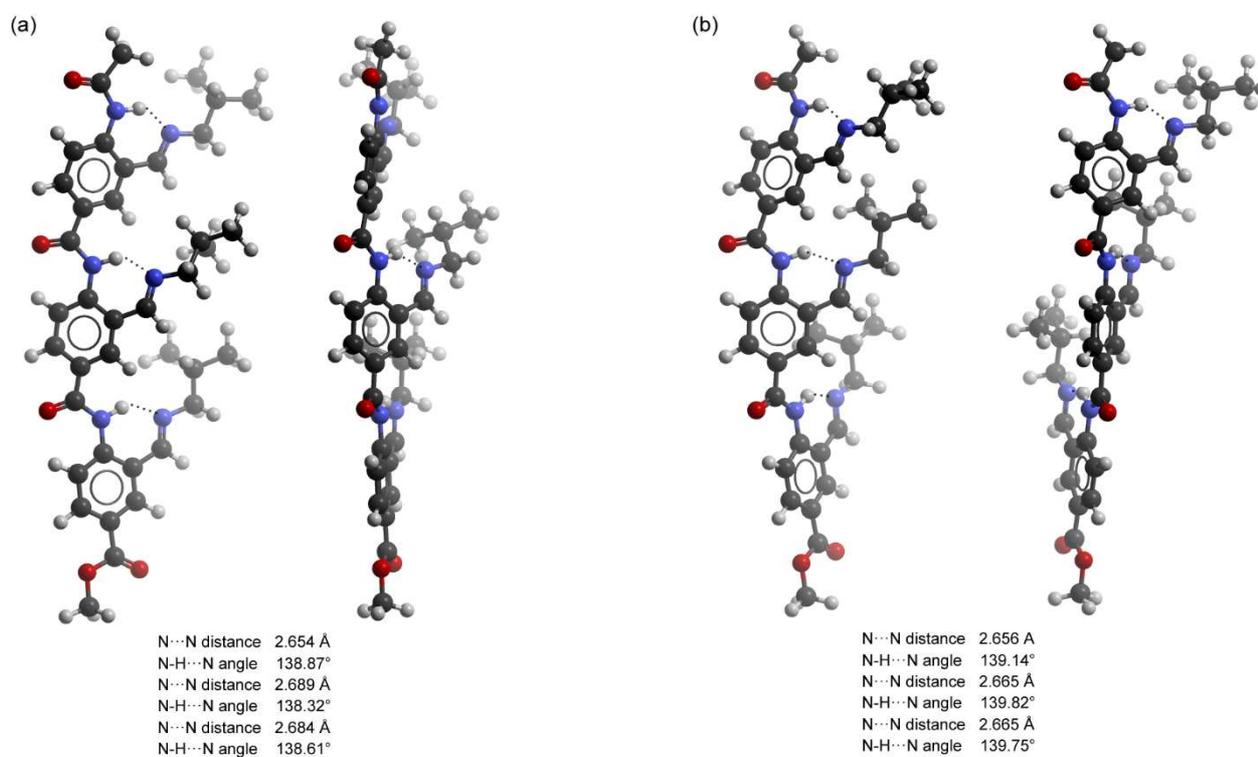
**Figure S76.** Two views on the energy-minimized structure of dimeric imine **8a** in chloroform and water. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) **8a** in chloroform, (b) **8a** in water.



**Figure S77.** Two views on the energy-minimized structure of trimeric acetal **3a** in chloroform and water. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) **3a** in chloroform, (b) **3a** in water.

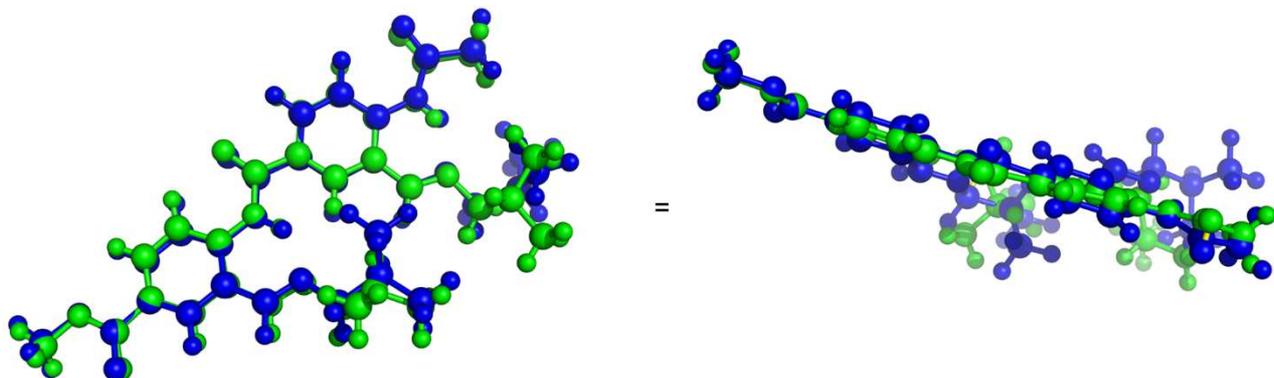


**Figure S78.** Two views on the energy-minimized structure of trimeric aldehyde **6a** in chloroform and water. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) **6a** in chloroform, (b) **6a** in water.

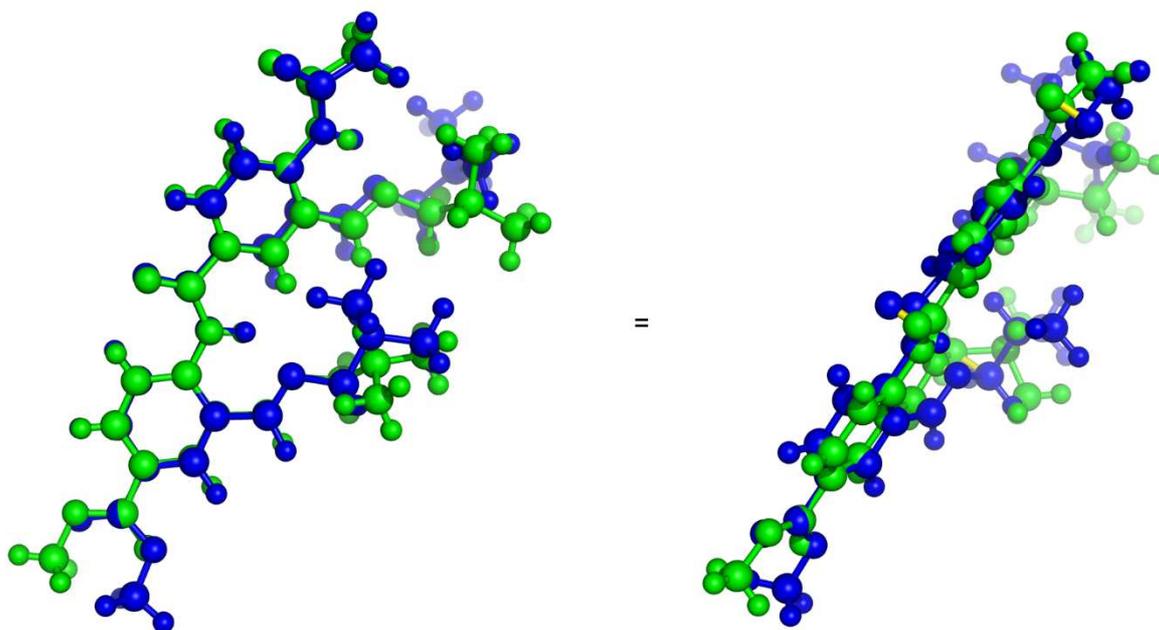


**Figure S79.** Two views on the energy-minimized structure of trimeric imine **9a** in chloroform and water. Hydrogen bonds are represented by dashed lines and the donor-acceptor distances and the hydrogen bond angles are also given (calculated using Mercury 3.5.1). (a) **9a** in chloroform, (b) **9a** in water.

In order to check the validity of the MOE modelling, the lowest energy conformer of compound **8a** in chloroform and water was overlaid with the crystal structure of compound **8a** using Pymol v0.99 (Figure S80-S81). This resulted in a good agreement between the computer models and the experimental crystal structure with RMSDs of 0.5-0.7 Å.



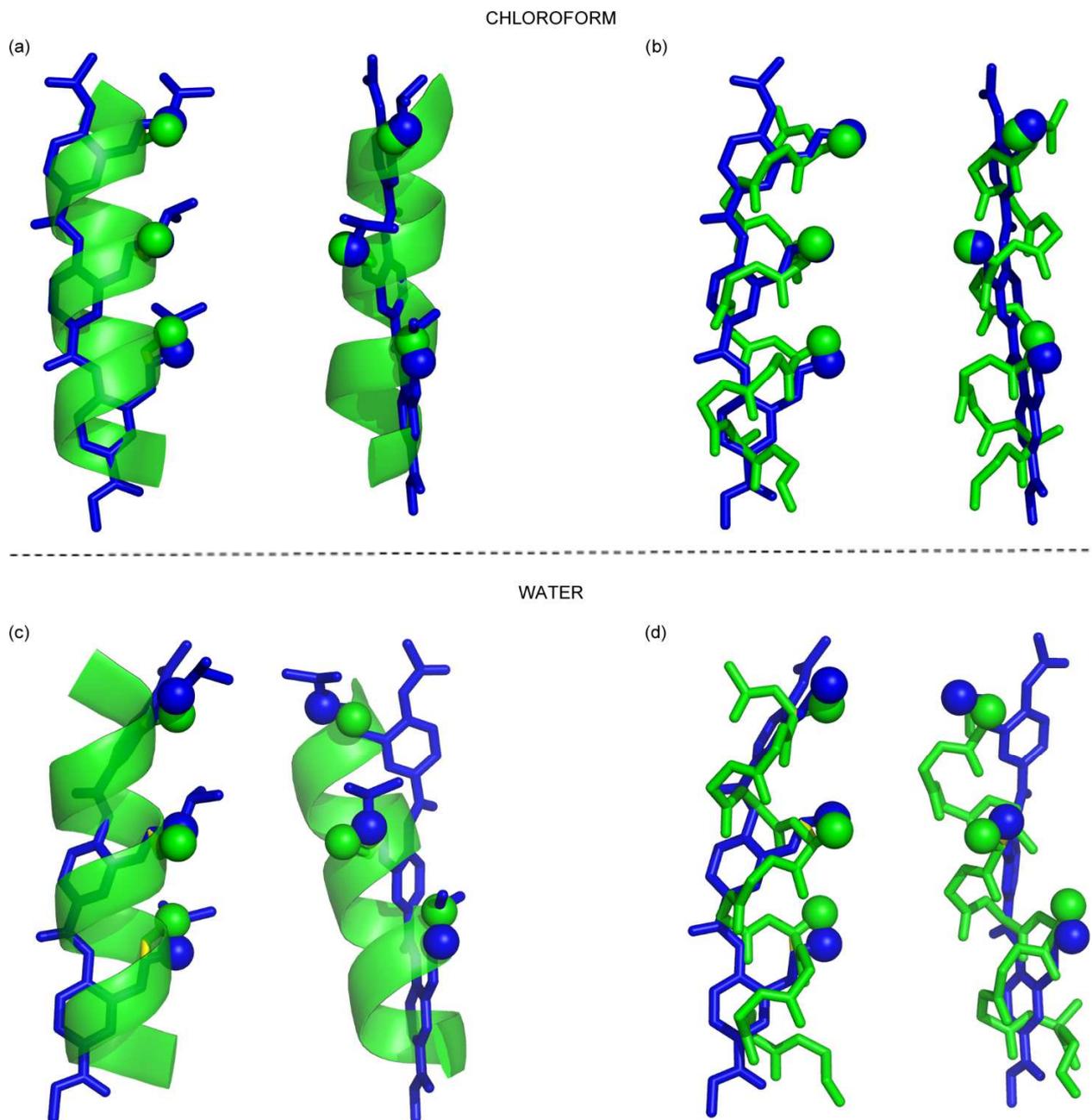
**Figure S80.** Two different views on the Pymol v0.99 overlay between the lowest energy conformer of compound **8a** in chloroform calculated by MOE (blue) and the experimental crystal structure obtained for compound **8a** (green). RMSD = 0.562 Å.



**Figure S81.** Two different views on the Pymol v0.99 overlay between the lowest energy conformer of compound **8a** in water calculated by MOE (blue) and the experimental crystal structure obtained for compound **8a** (green). RMSD = 0.700 Å.

In order to check if the imine containing compounds are good  $\alpha$ -helix mimetics, the lowest energy conformers of compound **9a** in chloroform and water were overlaid with the  $i$ ,  $i+4$  and  $i+7$  residues in a natural  $\alpha$ -helix extracted from PDB entry 2P32,<sup>4</sup> residues 537-546 (Figures S82-S83). Six point RMSD values (using the  $\alpha$ - and  $\beta$ -carbon atoms with respective imine nitrogen and

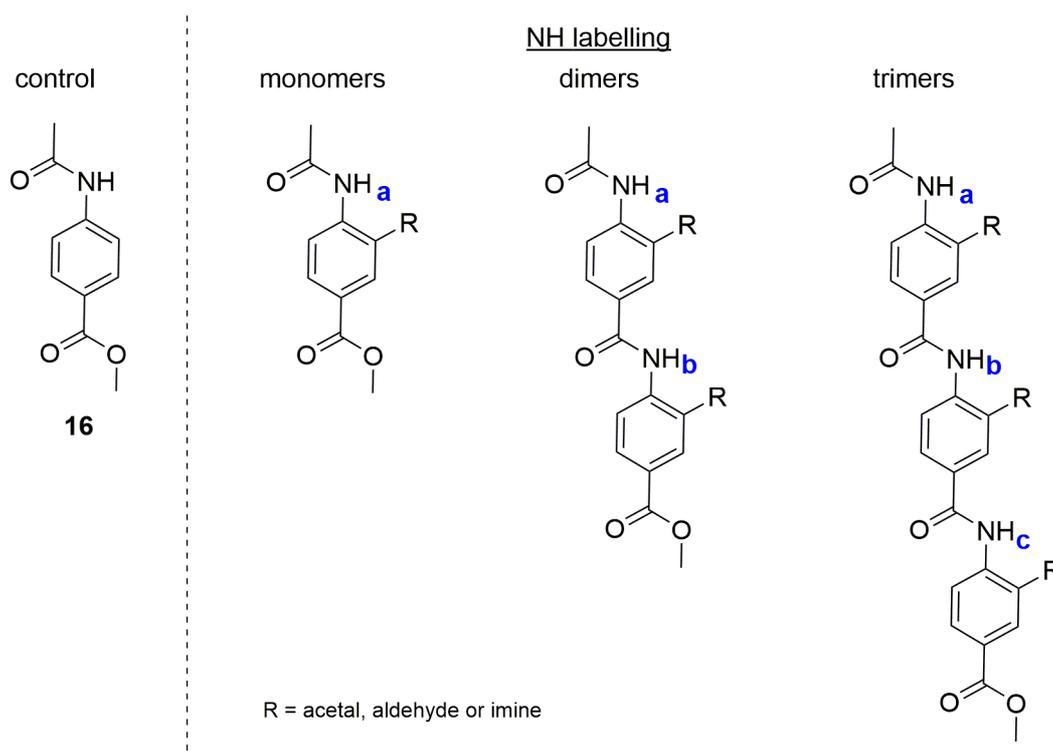
isobutyl CH<sub>2</sub> group) were also calculated and confirmed that these structures represent possible  $\alpha$ -helix mimetics.



**Figure S82.** Pymol v0.99 overlay between the lowest energy conformer of compound **9a** in chloroform (a+b) and water (c+d) calculated by MOE (blue) and natural  $\alpha$ -helix extracted from PDB entry 2P32 (green). RMSD = 0.809 Å (chloroform) and 1.261 Å (water). (a+c) The  $\alpha$ -helix is shown as a cartoon, with the  $\beta$ -carbons overlaid with the isobutyl CH<sub>2</sub> groups shown as spheres for clarity. (b+d) The  $\alpha$ -helix is shown as sticks (backbone only, side chain are omitted for clarity), with the  $\beta$ -carbons overlaid with the isobutyl CH<sub>2</sub> groups shown as spheres for clarity.

### 3.3. Dilution studies (solution)

$^1\text{H}$  NMR spectra of various concentrations of compounds **1b-9b**, as well as **1a**, **4a**, **7a** and control compound **16** (which cannot form intramolecular hydrogen bonds), were obtained in both  $\text{CDCl}_3$  and  $\text{DMSO-}d_6$ . The chemical shifts of the amide NH signals are given in Table S7 ( $\text{CDCl}_3$ ) and Table S8 ( $\text{DMSO-}d_6$ ) and do not depend significantly on the concentration of the solution. This can indicate that no aggregation takes place at high concentration and could also suggest that the amide NH protons are involved in intramolecular hydrogen bonds that are not broken upon dilution. Only compound **16** shows a strong dependence on concentration in  $\text{CDCl}_3$ , which indicates that it might be hydrogen bonding to another molecule (intermolecular hydrogen bond), which is broken upon dilution. All the acetals, aldehyde and imine containing mimetics on the other hand suggest the existence of intramolecular hydrogen bonds, as also observed in the solid state (X-ray diffraction, Section 3.1) and during modelling (Section 3.2). The labelling of the amide NHs used in Tables S7-S8 and in the other NMR studies is given in Figure S83.



**Figure S83.** Structure of control compound **16** and labelling of the NHs during the various conformational analysis studies in solution.

**Table S7.** Chemical shift (ppm) of the amide NH protons in CDCl<sub>3</sub> at 298 K at various concentrations.

<i>imine</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>7a</b>	<b>a</b>	13.03	13.03	13.03	13.03	13.03
<b>7b</b>	<b>a</b>	13.25	13.25	13.25	13.25	13.25
<b>8b</b>	<b>a</b>	13.24	13.24	13.24	13.24	13.24
	<b>b</b>	13.76	13.76	13.76	13.76	13.75
<b>9b</b>	<b>a</b>	13.25	13.25	13.25	13.25	n/a
	<b>b</b>	13.76	13.76	13.76	13.76	n/a
	<b>c</b>	13.79	13.79	13.79	13.79	n/a
<i>aldehyde</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>4a</b>	<b>a</b>	11.32	11.32	11.32	11.32	11.32
<b>4b</b>	<b>a</b>	12.08	12.08	12.08	12.08	12.07
<b>5b</b>	<b>a</b>	12.10	12.10	12.10	12.09	12.09
	<b>b</b>	12.41	12.40	12.40	12.40	12.39
<b>6b</b>	<b>a</b>	12.11	12.11	12.11	12.10	12.09
	<b>b</b>	12.41	12.41	12.41	12.40	12.38
	<b>c</b>	12.45	12.45	12.44	12.43	12.41
<i>acetal</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>1a</b>	<b>a</b>	8.76	8.76	8.76	8.76	8.76
<b>1b</b>	<b>a</b>	9.86	9.86	9.86	9.86	9.86
<b>2b</b>	<b>a</b>	9.85	9.85	9.85	9.85	9.84
	<b>b</b>	9.81	9.81	9.81	9.81	9.81
<b>3b</b>	<b>a</b>	9.86	9.86	9.86	9.86	9.85
	<b>b</b>	9.77	9.77	9.77	9.77	9.77
	<b>c</b>	9.84	9.84	9.84	9.84	9.84
<i>control</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>16</b>	<b>a</b>	<b>7.25</b>	<b>7.25</b>	<b>7.28</b>	<b>7.30</b>	<b>7.37</b>

**Table S8.** Chemical shift (ppm) of the amide NH protons in DMSO-*d*<sub>6</sub> at 298 K at various concentrations.

<i>imine</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>7a</b>	<b>a</b>	13.01	13.01	13.01	13.01	13.01
<b>7b</b>	<b>a</b>	13.24	13.24	13.24	13.24	13.24
<b>8b</b>	<b>a</b>	13.19	13.19	13.19	13.18	13.17
	<b>b</b>	13.91	13.91	13.90	13.89	13.86
<b>9b</b>	<b>a</b>	13.19	13.19	13.17	13.15	n/a
	<b>b</b>	13.86	13.85	13.80	13.75	n/a
	<b>c</b>	13.95	13.94	13.89	13.84	n/a
<i>aldehyde</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>4a</b>	<b>a</b>	11.00	11.00	11.00	11.00	11.00
<b>4b</b>	<b>a</b>	11.91	11.91	11.91	11.91	11.91
<b>5b</b>	<b>a</b>	11.88	11.88	11.87	11.87	11.86
	<b>b</b>	11.96	11.96	11.96	11.95	11.95
<b>6b</b>	<b>a</b>	11.89	11.89	n/a	n/a	n/a
	<b>b</b>	11.87	11.87	n/a	n/a	n/a
	<b>c</b>	12.01	12.01	n/a	n/a	n/a
<i>acetal</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>1a</b>	<b>a</b>	9.42	9.43	9.43	9.42	9.43
<b>1b</b>	<b>a</b>	9.82	9.82	9.82	9.82	9.82
<b>2b</b>	<b>a</b>	9.79	9.79	9.79	9.79	9.79
	<b>b</b>	10.08	10.08	10.08	10.08	10.08
<b>3b</b>	<b>a</b>	9.79	9.79	9.79	9.79	9.80
	<b>b</b>	10.09	10.09	10.09	10.09	10.09
	<b>c</b>	10.14	10.14	10.14	10.14	10.14
<i>control</i>						
	NH	0.5 mM	1 mM	5 mM	10 mM	25 mM
<b>16</b>	<b>a</b>	10.28	10.28	10.28	10.28	10.28

### 3.4. DMSO-*d*<sub>6</sub> and CDCl<sub>3</sub> comparison (solution)

Intramolecular hydrogen bonding can also be shown by comparing the chemical shift of the donor hydrogen atom in various solvents. Abraham *et al.* developed the A-value, given by  $A = 0.0065 + 0.133 * (\delta_{DMSO} - \delta_{CDCl_3})$ , which can be used to classify intramolecular hydrogen bonds.<sup>5</sup> For NH donors,  $A < 0.05$  suggest intramolecular hydrogen bonding and  $A > 0.16$  indicates that there is no hydrogen bonding. The results are shown in Table S9 and suggest intramolecular hydrogen

bonding in both the imine and aldehyde containing compounds, while the acetals are an intermediate case (A-values are intermediate and no conclusion can be drawn). Control compound **16** has a large difference in chemical shift between CDCl<sub>3</sub> and DMSO-*d*<sub>6</sub> and clearly shows no evidence of intramolecular hydrogen bonding.

**Table S9.** Chemical shift (ppm) of the amide NH protons in CDCl<sub>3</sub> and DMSO-*d*<sub>6</sub> at 298 K using a 5 mM concentration for monomers and dimers and a 1 mM concentration for trimers (values are the average of 3 independent repeats) and derived A-values. A-values *outside* the range of intramolecular hydrogen bonding are shown in bold.

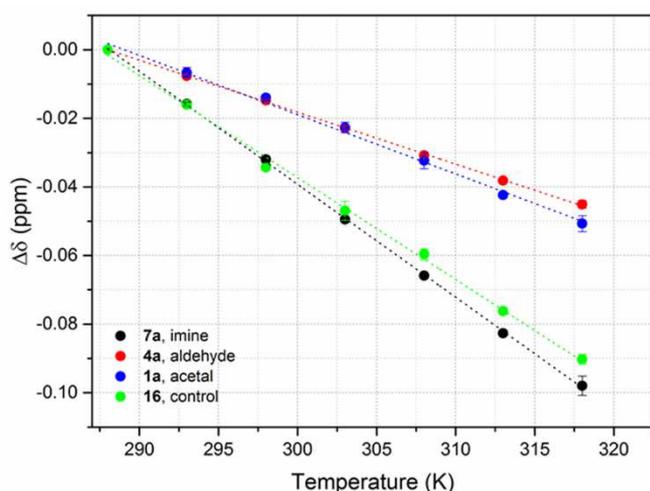
<i>imine</i>					
	NH	$\delta$ CDCl <sub>3</sub>	$\delta$ DMSO	$\Delta\delta$	A-value
<b>7a</b>	<b>a</b>	13.0335	13.0214	-0.0121	0.0049
<b>7b</b>	<b>a</b>	13.2525	13.2426	-0.0099	0.0052
<b>8b</b>	<b>a</b>	13.2433	13.1898	-0.0535	-0.0006
	<b>b</b>	13.7569	13.9052	0.1483	0.0262
<b>9b</b>	<b>a</b>	13.2505	13.1964	-0.0541	-0.0007
	<b>b</b>	13.7560	13.8541	0.0981	0.0195
	<b>c</b>	13.7879	13.9452	0.1573	0.0274
<i>aldehyde</i>					
	NH	$\delta$ CDCl <sub>3</sub>	$\delta$ DMSO	$\Delta\delta$	A-value
<b>4a</b>	<b>a</b>	11.3142	11.0035	-0.3107	-0.0348
<b>4b</b>	<b>a</b>	12.0703	11.9172	-0.1531	-0.0139
<b>5b</b>	<b>a</b>	12.0901	11.8752	-0.2149	-0.0221
	<b>b</b>	12.3969	11.9570	-0.4399	-0.0520
<b>6b</b>	<b>a</b>	12.1010	11.8872	-0.2138	-0.0219
	<b>b</b>	12.4067	11.8674	-0.5393	-0.0652
	<b>c</b>	12.4389	12.0116	-0.4273	-0.0503
<i>acetal</i>					
	NH	$\delta$ CDCl <sub>3</sub>	$\delta$ DMSO	$\Delta\delta$	A-value
<b>1a</b>	<b>a</b>	8.7623	9.4359	0.6736	<b>0.0961</b>
<b>1b</b>	<b>a</b>	9.8612	9.8148	-0.0464	0.0003
<b>2b</b>	<b>a</b>	9.8455	9.7874	-0.0581	-0.0012
	<b>b</b>	9.8091	10.0723	0.2632	0.0415
<b>3b</b>	<b>a</b>	9.8590	9.7931	-0.0659	-0.0023
	<b>b</b>	9.7746	10.0869	0.3123	0.0480
	<b>c</b>	9.8419	10.1360	0.2941	0.0456
<i>control</i>					
	NH	$\delta$ CDCl <sub>3</sub>	$\delta$ DMSO	$\Delta\delta$	A-value
<b>16</b>	<b>a</b>	7.2789	10.2632	2.9843	<b>0.4034</b>

### 3.5. Variable temperature NMR (solution)

<sup>1</sup>H NMR spectra were collected for the various acetal, aldehyde and imine containing helix mimetics at various temperatures in the range 288-318 K (CDCl<sub>3</sub>) or 298-353 K (DMSO-*d*<sub>6</sub>). The chemical shift of the amide NH signals was plotted against the temperature and subjected to a linear fit (Figures S84-S91). This allows the calculation of temperature coefficients  $\Delta\delta/\Delta T$  given by the slope of the plot. Normally a value less negative than -0.004 ppm/K (-4 ppb/K) is taken as evidence for intramolecular hydrogen bonding, but many exceptions exist to this rule.<sup>6,7</sup> An overview of the calculated temperature coefficients is given in Table S10. For all of the compounds temperature coefficients less negative than -0.004 ppm/K were observed in chloroform, including for control compound **16**. In DMSO-*d*<sub>6</sub> on the other hand, the data is in agreement with the findings of the A-values (Section 3.4), with acetal **1a** and control compound **16** displaying temperature coefficients more negative than -4 ppb/K and thus not involved in intramolecular hydrogen bonding. The aldehyde and imine bearing compounds have temperature coefficients less negative than -4 ppb/K and the temperature coefficients are also largely independent of the solvent used. This suggests that the imines and aldehydes are involved in intramolecular hydrogen bonding to the amide NHs.

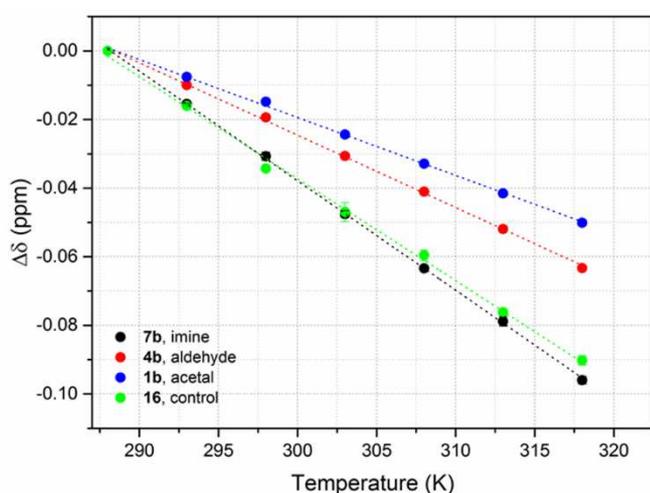
**Table S10.** Temperature coefficients  $\Delta\delta/\Delta T$  (in ppb/K) of the amide NH protons in CDCl<sub>3</sub> and DMSO-*d*<sub>6</sub> at 298 K using a 5 mM concentration for monomers and dimers and a 1 mM concentration for trimers (values are the average of 3 independent repeats). Values *outside* the range of intramolecular hydrogen bonding are shown in bold.

	$\Delta\delta/\Delta T$ in CDCl <sub>3</sub> (ppb/K)			$\Delta\delta/\Delta T$ in DMSO- <i>d</i> <sub>6</sub> (ppb/K)		
	<i>NHa</i>	<i>NHb</i>	<i>NHc</i>	<i>NHa</i>	<i>NHb</i>	<i>NHc</i>
<i>imine</i>						
<b>7a</b>	-3.30	n/a	n/a	-2.83	n/a	n/a
<b>7b</b>	-3.20	n/a	n/a	-3.24	n/a	n/a
<b>8b</b>	-3.18	-3.52	n/a	-2.84	-3.32	n/a
<b>9b</b>	-3.18	-3.53	-3.51	-2.81	-3.23	-3.27
<i>aldehyde</i>						
<b>4a</b>	-1.52	n/a	n/a	-1.95	n/a	n/a
<b>4b</b>	-2.11	n/a	n/a	-3.02	n/a	n/a
<b>5b</b>	-2.07	-1.85	v	-2.61	-1.79	n/a
<b>6b</b>	-2.05	-1.79	-1.83	-2.65	-1.30	-1.95
<i>acetal</i>						
<b>1a</b>	-1.73	n/a	n/a	<b>-4.53</b>	n/a	n/a
<b>1b</b>	-1.69	n/a	n/a	-1.56	n/a	n/a
<b>2b</b>	-1.66	-1.67	n/a	-1.55	-3.00	n/a
<b>3b</b>	-1.68	-1.46	-1.73	-1.58	-3.27	-3.19
<i>control</i>						
<b>16</b>	-2.97	n/a	n/a	<b>-4.41</b>	n/a	n/a



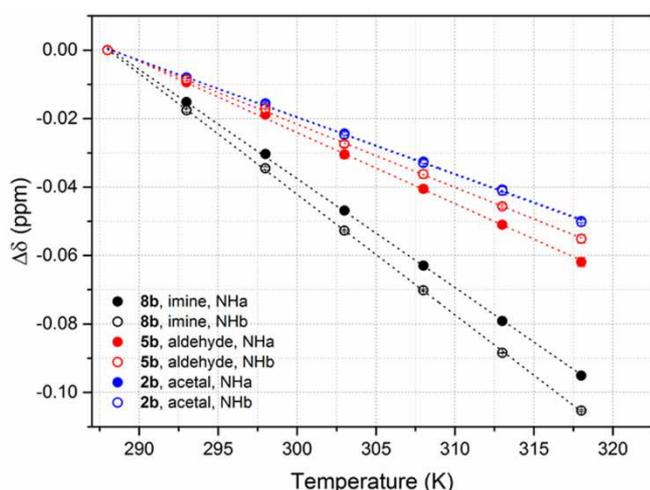
Equation	y = a + b*x			
Weight	No Weighting			
Residual Sum of Squares	1.57107E-6	6.06825E-7	9.0119E-6	1.45795E-5
Pearson's r	-0.9999	-0.99981	-0.99785	-0.99882
Adj. R-Square	0.99975	0.99955	0.99484	0.99718
		Value	Standard Error	
7a	Intercept	0.94943	0.00642	
	Slope	-0.0033	2.11867E-5	
4a	Intercept	0.43651	0.00399	
	Slope	-0.00152	1.31673E-5	
1a	Intercept	0.49894	0.01538	
	Slope	-0.00173	5.07428E-5	
16	Intercept	0.85488	0.01957	
	Slope	-0.00297	6.45413E-5	

**Figure S84.** Plot of the amide NH chemical shift in CDCl<sub>3</sub> versus temperature for the acetyl-capped monomers **7a**, **4a**, **1a** and control **16**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.



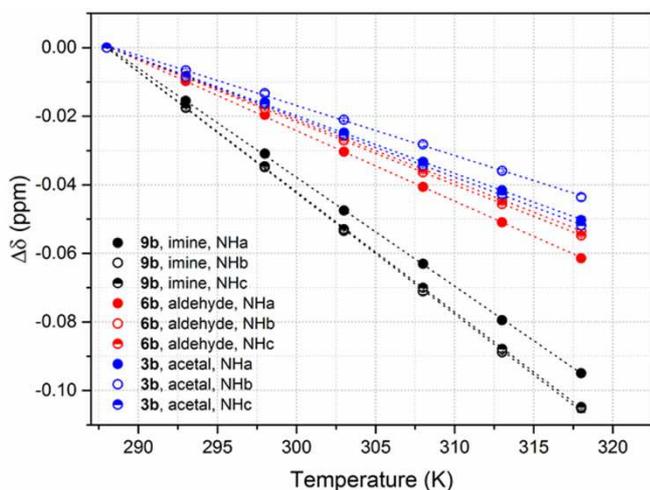
Equation	y = a + b*x			
Weight	No Weighting			
Residual Sum of Squares	1.47202E-6	2.43693E-6	2.35857E-6	1.45795E-5
Pearson's r	-0.9999	-0.99961	-0.99941	-0.99882
Adj. R-Square	0.99975	0.99906	0.99858	0.99718
		Value	Standard Error	
7b	Intercept	0.92081	0.00622	
	Slope	-0.0032	2.0508E-5	
4b	Intercept	0.60833	0.008	
	Slope	-0.00211	2.63868E-5	
1b	Intercept	0.48673	0.00787	
	Slope	-0.00169	2.59592E-5	
16	Intercept	0.85488	0.01957	
	Slope	-0.00297	6.45413E-5	

**Figure S85.** Plot of the amide NH chemical shift in CDCl<sub>3</sub> versus temperature for the PEG-capped monomers **7b**, **4b**, **1b** and control **16**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.



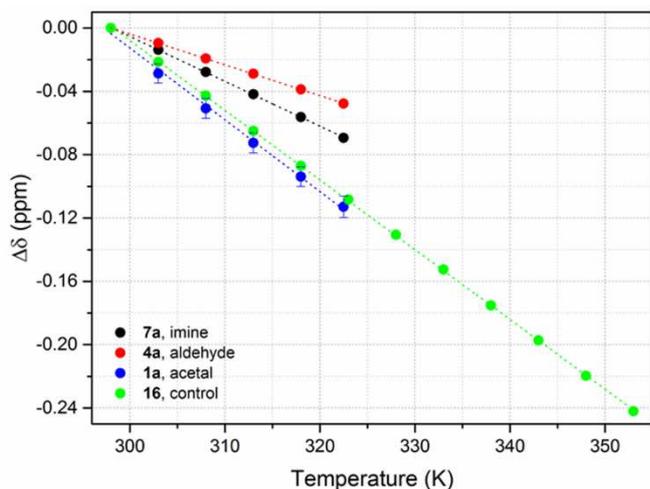
Equation	y = a + b*x			
Weight	No Weighting			
Residual Sum of Squares	1.25571E-6	6.59643E-7	2.17234E-6	
Pearson's r	1.11143E-6	9.90476E-7	6.70754E-7	
Adj. R-Square	-0.99991	-0.99996	-0.99964	
	-0.99977	-0.99974	-0.99983	
	0.99979	0.99991	0.99914	
	0.99944	0.99938	0.99959	
		Value	Standard Error	
8b-NHa	Intercept	0.91775	0.00574	
	Slope	-0.00318	1.89414E-5	
8b-NHb	Intercept	1.01496	0.00416	
	Slope	-0.00352	1.37284E-5	
5b-NHa	Intercept	0.59842	0.00755	
	Slope	-0.00207	2.49132E-5	
5b-NHb	Intercept	0.53209	0.0054	
	Slope	-0.00185	1.782E-5	
2b-NHa	Intercept	0.47822	0.0051	
	Slope	-0.00166	1.68224E-5	
2b-NHb	Intercept	0.48021	0.0042	
	Slope	-0.00167	1.38436E-5	

**Figure S86.** Plot of the amide NH chemical shift in CDCl<sub>3</sub> versus temperature for the PEG-capped dimers **8b**, **5b** and **2b**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.



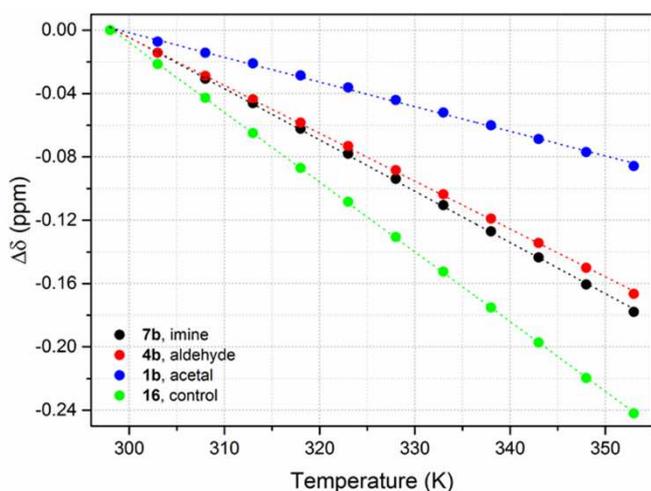
Equation	$y = a + b \cdot x$		
Weight	No Weighting		
Residual Sum of Squares	5.81587E-7	1.55698E-6	5.31587E-7
Pearson's r	5.33929E-7	3.84762E-7	2.73492E-7
	4.68849E-7	1.17571E-6	5.3746E-7
	-0.99996	-0.99991	-0.99997
	-0.99991	-0.99991	-0.99994
	-0.99988	-0.99961	-0.99987
Adj. R-Square	0.9999	0.99979	0.99993
	0.99978	0.99979	0.99986
	0.99971	0.99905	0.99969
	<b>Value</b>	<b>Standard Error</b>	
9b-NHa	Intercept 0.91562	0.00391	
	Slope -0.00318	1.28906E-5	
9b-NHb	Intercept 1.01687	0.00639	
	Slope -0.00353	2.10915E-5	
9b-NHc	Intercept 1.00954	0.00374	
	Slope -0.00351	1.2324E-5	
6b-NHa	Intercept 0.59201	0.00374	
	Slope -0.00205	1.23512E-5	
6b-NHb	Intercept 0.51663	0.00318	
	Slope -0.00179	1.04848E-5	
6b-NHc	Intercept 0.52878	0.00268	
	Slope -0.00183	8.83972E-6	
3b-NHa	Intercept 0.48381	0.00351	
	Slope -0.00168	1.1574E-5	
3b-NHb	Intercept 0.4207	0.00556	
	Slope -0.00146	1.83281E-5	
3b-NHc	Intercept 0.4987	0.00376	
	Slope -0.00173	1.23919E-5	

**Figure S87.** Plot of the amide NH chemical shift in  $\text{CDCl}_3$  versus temperature for the PEG-capped trimers **9b**, **6b** and **3b**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.



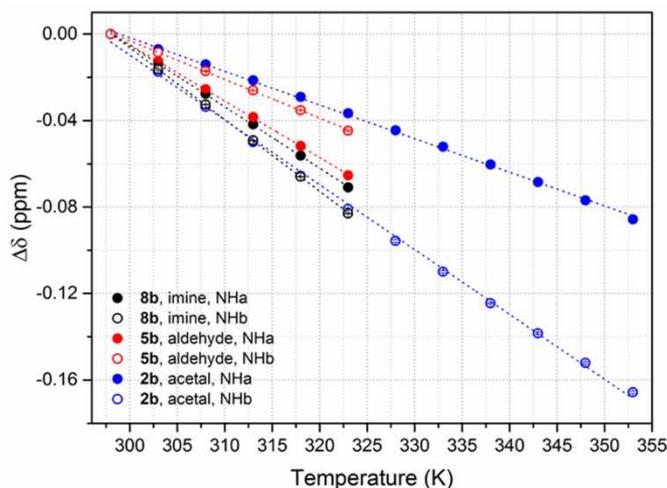
Equation	$y = a + b \cdot x$		
Weight	No Weighting		
Residual Sum of Squares	4.29262E-7	1.13025E-7	2.60614E-5
Pearson's r	-0.99994	-0.99996	-0.99851
Adj. R-Square	0.99984	0.99991	0.99628
		<b>Value</b>	<b>Standard Error</b>
7a	Intercept 0.84415	0.00493	
	Slope -0.00283	1.58866E-5	
4a	Intercept 0.58031	0.00253	
	Slope -0.00195	8.15185E-6	
1a	Intercept 1.34733	0.03844	
	Slope -0.00453	1.23785E-4	
16	Intercept 1.31382	0.00322	
	Slope -0.00441	9.86457E-6	

**Figure S88.** Plot of the amide NH chemical shift in  $\text{DMSO}-d_6$  versus temperature for the acetyl-capped monomers **7a**, **4a**, **1a** and control **16**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.



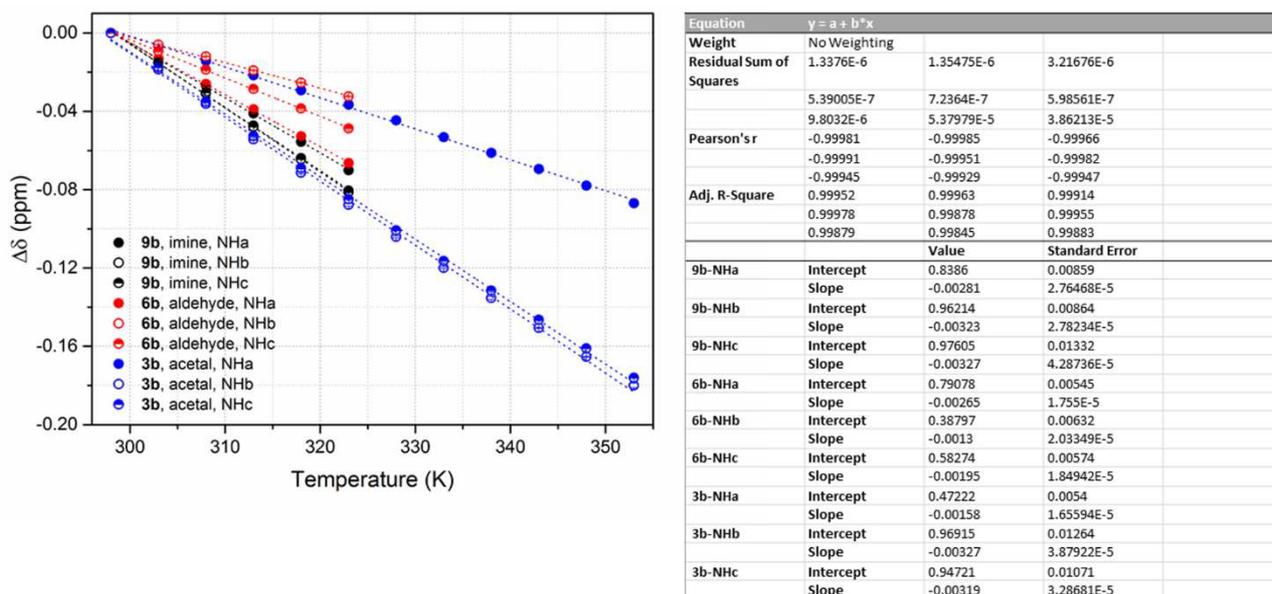
Equation	$y = a + b \cdot x$			
Weight	No Weighting			
Residual Sum of Squares	1.23923E-5	8.40728E-6	1.22317E-5	3.47882E-6
Pearson's r	-0.99983	-0.99987	-0.99929	-0.99997
Adj. R-Square	0.99964	0.99972	0.99845	0.99994
		Value	Standard Error	
7b	Intercept	0.96697	0.00607	
	Slope	-0.00324	1.86182E-5	
4b	Intercept	0.90191	0.005	
	Slope	-0.00302	1.53352E-5	
1b	Intercept	0.46538	0.00603	
	Slope	-0.00156	1.84972E-5	
16	Intercept	1.31382	0.00322	
	Slope	-0.00441	9.86457E-6	

**Figure S89.** Plot of the amide NH chemical shift in DMSO- $d_6$  versus temperature for the PEG-capped monomers **7b**, **4b**, **1b** and control **16**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.



Equation	$y = a + b \cdot x$		
Weight	No Weighting		
Residual Sum of Squares	6.27259E-7	3.90116E-7	5.35196E-7
Pearson's r	6.4091E-7	8.19264E-6	4.31858E-5
	-0.99991	-0.99996	-0.99991
	-0.99977	-0.99953	-0.99933
Adj. R-Square	0.99978	0.9999	0.99978
	0.99943	0.99896	0.99852
		Value	Standard Error
8b-NHa	Intercept	0.84555	0.00588
	Slope	-0.00284	1.89323E-5
8b-NHb	Intercept	0.9883	0.00464
	Slope	-0.00332	1.49306E-5
5b-NHa	Intercept	0.77864	0.00543
	Slope	-0.00261	1.74879E-5
5b-NHb	Intercept	0.53292	0.00594
	Slope	-0.00179	1.91372E-5
2b-NHa	Intercept	0.46471	0.00493
	Slope	-0.00155	1.51382E-5
2b-NHb	Intercept	0.88992	0.01133
	Slope	-0.003	3.47562E-5

**Figure S90.** Plot of the amide NH chemical shift in DMSO- $d_6$  versus temperature for the PEG-capped dimers **8b**, **5b** and **2b**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.



**Figure S91.** Plot of the amide NH chemical shift in DMSO- $d_6$  versus temperature for the PEG-capped trimers **9b**, **6b** and **3b**. Data points are the average of three independent repeats (error bars represent standard deviations). Dotted lines are the result of a linear fit, with the obtained values shown in the right-hand table.

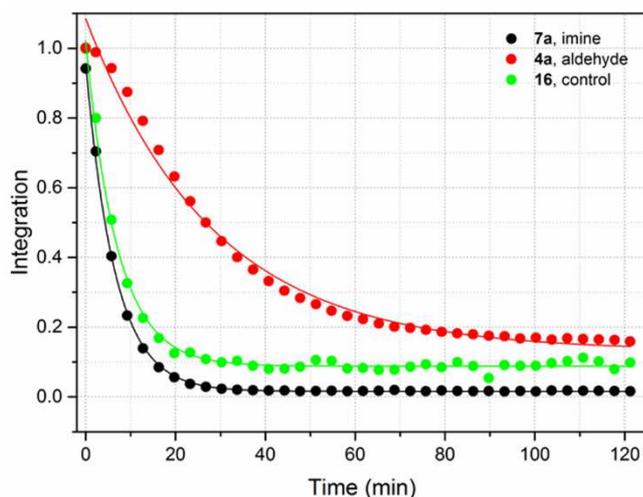
### 3.6. Deuterium exchange (solution)

When an exchangeable proton is involved in hydrogen bonding, the proton-deuterium exchange is expected to be slower than when it is not involved in hydrogen bonding. Therefore, experiments were conducted where either 5  $\mu\text{L}$  methanol- $d_4$  was added to 750  $\mu\text{L}$  of a 5 mM solution of the mimetics in  $\text{CDCl}_3$ , or 10  $\mu\text{L}$   $\text{D}_2\text{O}$  was added to 750  $\mu\text{L}$  of a 5 mM solution in DMSO- $d_6$  (1 mM was used for the less soluble trimers). The integration of the amide NH signals was plotted against time and fitted to an exponential decay function to calculate the ‘half-life’ of the exchange (Figures S92-S99). The calculated half-lives are given in Table S11. The acetal-containing compounds degraded quickly upon the addition of methanol to chloroform (presumably due to the presence of HCl in chloroform). The aldehyde-containing compounds did not always fit to the exponential decay curve and the deuterium exchange data can therefore also not be used as proof of hydrogen bonding. Furthermore, the imine-containing compounds have similar half-lives to control compound **16**. It is possible that the aggregation of the control compound (as shown by dilution studies, Section S3.3) leads to a slower than expected exchange. In brief, deuterium exchange could hint towards intramolecular hydrogen bonding in the imines, but the data is inconclusive.

**Table S11.** Half-lives (in minutes) of the amide NH protons in a deuterium exchange experiment in CDCl<sub>3</sub> or DMSO-*d*<sub>6</sub> at 298 K using a 5 mM concentration for monomers and dimers and a 1 mM concentration for trimers.

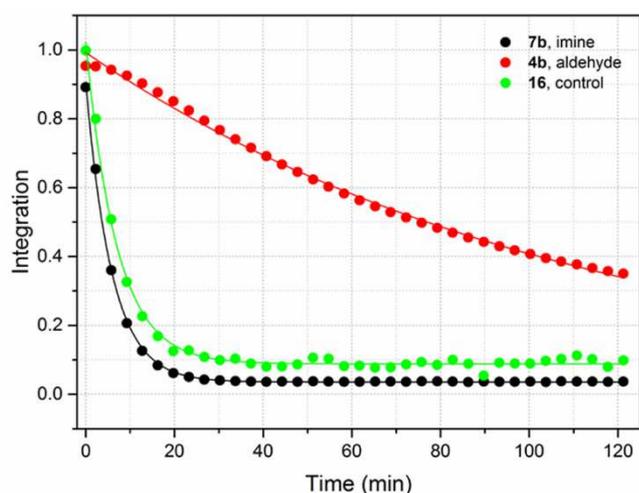
	half-life in CDCl <sub>3</sub> (min)			half-life in DMSO- <i>d</i> <sub>6</sub> (min)		
	NHa	NHb	NHc	NHa	NHb	NHc
<i>imine</i>						
<b>7a</b>	4.47	n/a	n/a	27.2	n/a	n/a
<b>7b</b>	4.05	n/a	n/a	29.7	n/a	n/a
<b>8b</b>	3.73	5.83	n/a	46.6	23.9	n/a
<b>9b</b>	2.30	4.01	3.30	86	45.2	33.2
<i>aldehyde</i>						
<b>4a</b>	19.5*	n/a	n/a	37.6	n/a	n/a
<b>4b</b>	77*	n/a	n/a	586	n/a	n/a
<b>5b</b>	21.0*	2.8*	n/a	1319	15.1	n/a
<b>6b</b>	16.3*	1.83*	1.82*	>1000	38.5	44.9
<i>acetal</i>						
<b>1a</b>	degradation	n/a	n/a	117.3	n/a	n/a
<b>1b</b>	degradation	n/a	n/a	3.45	n/a	n/a
<b>2b</b>	degradation	degradation	n/a	160**	23.4**	n/a
<b>3b</b>	degradation	degradation	degradation	898	119	105
<i>control</i>						
<b>16</b>	4.8***	n/a	n/a	23.9	n/a	n/a

\*Does not fit exponential decay; \*\*Some peaks split in two, possible degradation; \*\*\*On top of exchange, addition of methanol-*d*<sub>4</sub> causes shift of NH signal.



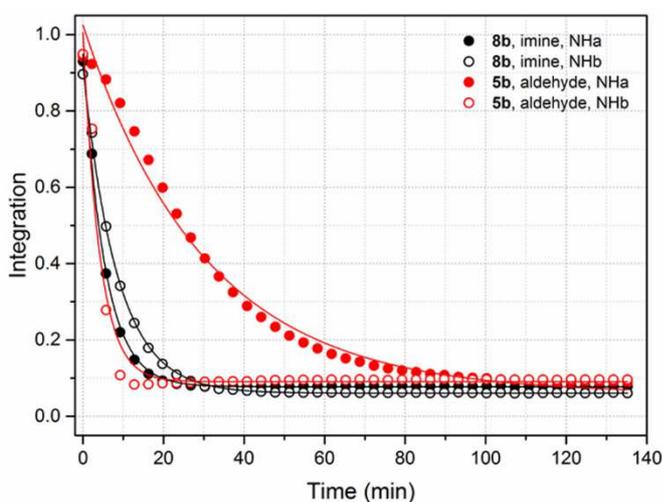
Model	ExpDec1		
Equation	$y = A1 \cdot \exp(-x/t1) + y0$		
Reduced Chi-Sqr	3.24414E-5	7.18686E-4	1.92195E-4
Adj. R-Square	0.99917	0.98992	0.99521
		Value	Standard Error
<b>7a</b>	y0	0.01601	0.00107
	A1	0.94033	0.00509
	t1	6.4518	0.07158
	k	0.155	0.00172
	tau	4.47205	0.04961
<b>4a</b>	y0	0.13172	0.00994
	A1	0.95291	0.01647
	t1	28.14727	1.22016
	k	0.03553	0.00154
	tau	19.5102	0.84575
<b>16</b>	y0	0.08815	0.00264
	A1	0.93307	0.0122
	t1	6.98587	0.18521
	k	0.14315	0.0038
	tau	4.84224	0.12838

**Figure S92.** Plot of the integration of the amide NH proton NMR signal in CDCl<sub>3</sub> versus time (after the addition of methanol-*d*<sub>4</sub> at *t* = 0 min) for the acetyl-capped monomers **7a**, **4a** and control **16**. Acetal monomer **1a** degraded during the experiment. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table (tau = half-life).



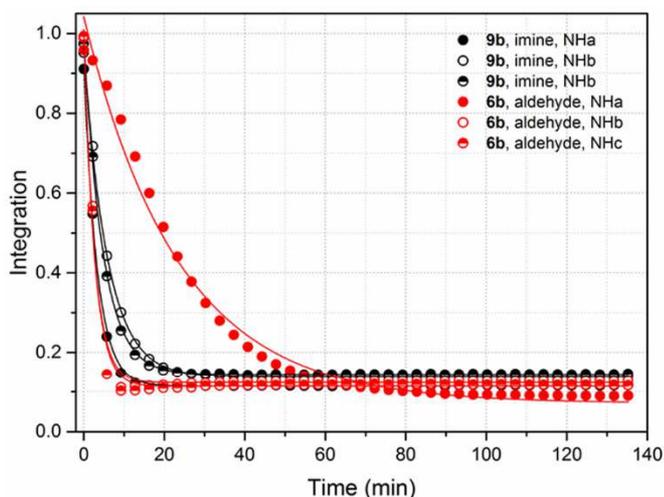
Model	ExpDec1			
Equation	$y = A1 \cdot \exp(-x/t1) + y0$			
Reduced Chi-Sqr	3.34969E-5	1.39548E-4	1.92195E-4	
Adj. R-Square	0.99896	0.99645	0.99521	
		Value	Standard Error	
7b	y0	0.03559	0.00108	
	A1	0.86983	0.00526	
	t1	5.84872	0.07364	
	k	0.17098	0.00215	
	tau	4.05403	0.05104	
4b	y0	0.00644	0.04303	
	A1	0.98576	0.03969	
	t1	111.33228	7.99201	
	k	0.00898	6.44783E-4	
	tau	77.16965	5.53964	
16	y0	0.08815	0.00264	
	A1	0.93307	0.0122	
	t1	6.98587	0.18521	
	k	0.14315	0.0038	
	tau	4.84224	0.12838	

**Figure S93.** Plot of the integration of the amide NH proton NMR signal in CDCl<sub>3</sub> versus time (after the addition of methanol-d<sub>4</sub> at t = 0 min) for the PEG-capped monomers **7b**, **4b** and control **16**. Acetal monomer **1b** degraded during the experiment. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table (tau = half-life).



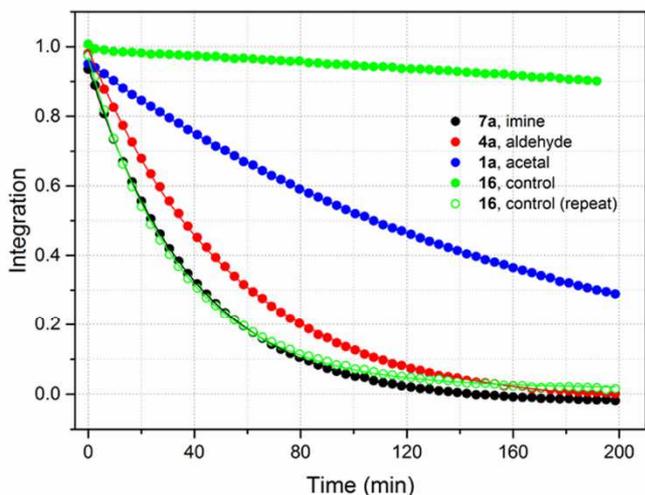
Model	ExpDec1			
Equation	$y = A1 \cdot \exp(-x/t1) + y0$			
Reduced Chi-Sqr	6.21638E-5	3.74858E-5	6.87749E-4	8.79919E-4
Adj. R-Square	0.99781	0.99888	0.99037	0.96957
		Value	Standard Error	
8b-NHa	y0	0.07757	0.00137	
	A1	0.87144	0.00726	
	t1	5.37949	0.09415	
	k	0.18589	0.00325	
	tau	3.72878	0.06526	
8b-NHb	y0	0.06033	0.00112	
	A1	0.85582	0.00518	
	t1	8.41151	0.10016	
	k	0.11888	0.00142	
	tau	5.83041	0.06942	
5b-NHa	y0	0.05879	0.00889	
	A1	0.96536	0.01557	
	t1	30.30077	1.19607	
	k	0.033	0.0013	
	tau	21.00289	0.82905	
5b-NHb	y0	0.09063	0.00504	
	A1	0.91434	0.0284	
	t1	4.12412	0.28312	
	k	0.24248	0.01665	
	tau	2.85863	0.19624	

**Figure S94.** Plot of the integration of the amide NH proton NMR signal in CDCl<sub>3</sub> versus time (after the addition of methanol-d<sub>4</sub> at t = 0 min) for the PEG-capped dimers **8b** and **5b**. Acetal monomer **2b** degraded during the experiment. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table (tau = half-life).



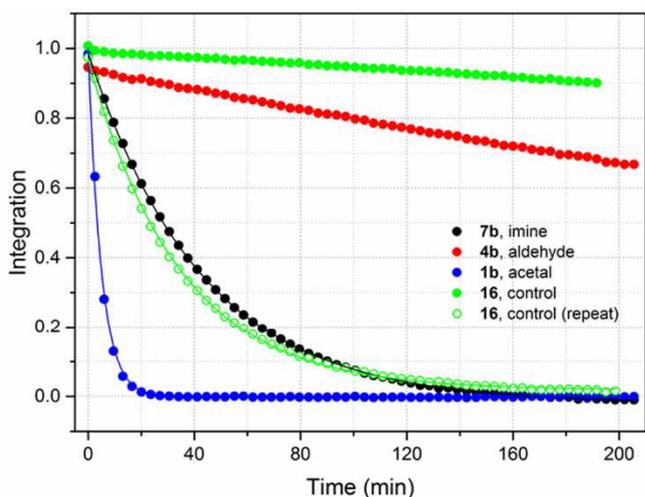
Model	ExpDec1		
Equation	$y = A1 * \exp(-x/t1) + y0$		
Reduced Chi-Sqr	3.86845E-5	1.8371E-5	2.31664E-5
	7.02723E-4	3.62029E-4	3.04931E-4
Adj. R-Square	0.99809	0.9993	0.99908
	0.98964	0.98431	0.98706
		Value	Standard Error
9b-NHa	y0	0.1143	0.00104
	A1	0.80561	0.0061
	t1	3.31164	0.05803
	k	0.30197	0.00529
	tau	2.29545	0.04023
9b-NHb	y0	0.1383	7.48211E-4
	A1	0.82441	0.0039
	t1	5.78939	0.05678
	k	0.17273	0.00169
	tau	4.0129	0.03936
9b-NHc	y0	0.14433	8.25809E-4
	A1	0.84092	0.00452
	t1	4.76003	0.05495
	k	0.21008	0.00243
	tau	3.2994	0.03809
6b-NHa	y0	0.07092	0.00713
	A1	0.97036	0.01692
	t1	23.49341	0.8864
	k	0.04257	0.00161
	tau	16.28439	0.61441
6b-NHb	y0	0.12325	0.00316
	A1	0.88448	0.01897
	t1	2.64065	0.13803
	k	0.37869	0.01979
	tau	1.83036	0.09567
6b-NHc	y0	0.11414	0.0029
	A1	0.8954	0.01741
	t1	2.63158	0.12483
	k	0.38	0.01803
	tau	1.82407	0.08653

**Figure S95.** Plot of the integration of the amide NH proton NMR signal in  $CDCl_3$  versus time (after the addition of methanol- $d_4$  at  $t = 0$  min) for the PEG-capped trimers **9b** and **6b**. Acetal monomer **3b** degraded during the experiment. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table ( $\tau =$  half-life).



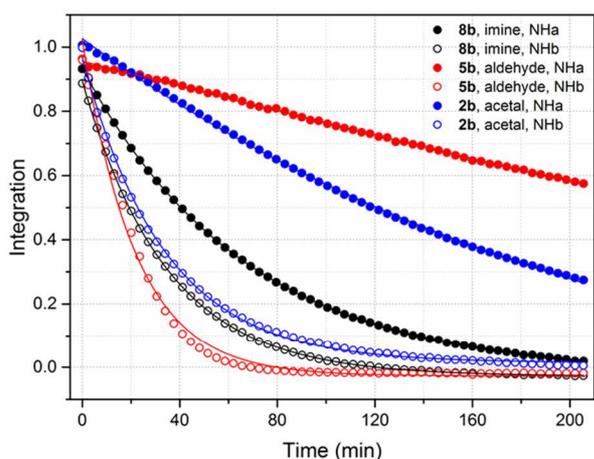
Model	ExpDec1		
Equation	$y = A1 * \exp(-x/t1) + y0$		
Reduced Chi-Sqr	3.30275E-6	1.78787E-5	1.50752E-6
	5.46214E-6	2.2354E-5	
Adj. R-Square	0.99995	0.99974	0.99996
	0.99263	0.99962	
		Value	Standard Error
7a	y0	-0.02365	4.55967E-4
	A1	0.96675	9.72487E-4
	t1	39.36519	0.08859
	k	0.0254	5.71679E-5
	tau	27.28587	0.0614
4a	y0	-0.03122	0.00155
	A1	1.02342	0.00223
	t1	54.27821	0.31545
	k	0.01842	1.07075E-4
	tau	37.62279	0.21866
1a	y0	-0.0095	0.00313
	A1	0.96315	0.00284
	t1	169.29097	0.93893
	k	0.00591	3.27618E-5
	tau	117.34356	0.65082
16	y0	3.14201	4.6598
	A1	-2.14855	4.65915
	t1	-4614.63669	9802.93344
	k	-2.16702E-4	4.60342E-4
	tau	-3198.62241	6794.87567
16 (repeat)	y0	0.01822	0.00106
	A1	0.95076	0.00265
	t1	34.44222	0.20132
	k	0.02903	1.69711E-4
	tau	23.87353	0.13955

**Figure S96.** Plot of the integration of the amide NH proton NMR signal in  $DMSO-d_6$  versus time (after the addition of  $D_2O$  at  $t = 0$  min) for the acetyl-capped monomers **7a**, **4a**, **1a** and control **16**. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table ( $\tau =$  half-life). The experiment was repeated for control **16** and indicated a high uncertainty in the experiments.



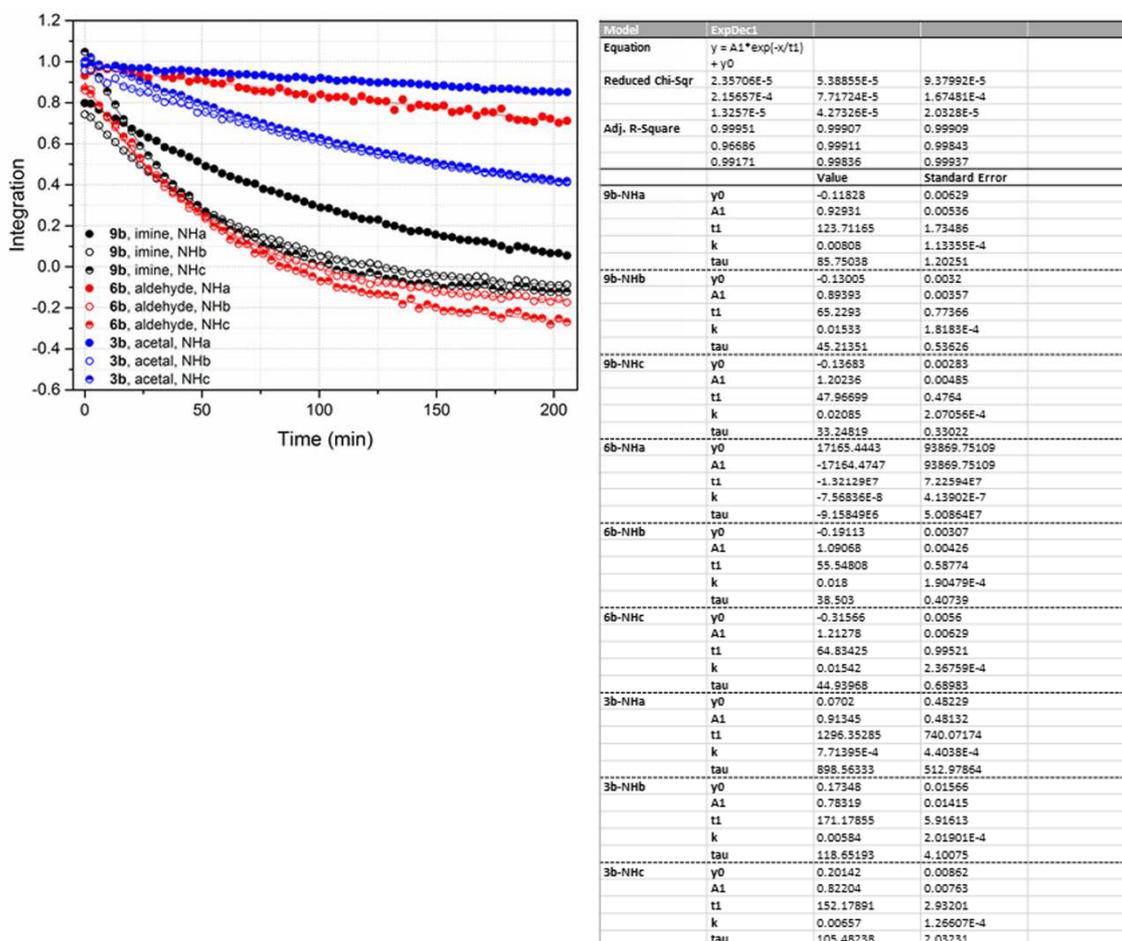
Model	ExpDec1		
Equation	$y = A1 * \exp(-x/t1) + y0$		
Reduced Chi-Sqr	5.79044E-6	3.8408E-6	4.13668E-5
	5.46214E-6	2.2354E-5	
Adj. R-Square	0.99992	0.99943	0.99825
	0.99263	0.99962	
	Value	Standard Error	
7b	y0	-0.01986	6.22749E-4
	A1	1.00798	0.00125
	t1	42.82204	0.12183
	k	0.02335	6.64388E-5
	tau	29.68197	0.08445
4b	y0	-0.33787	0.11048
	A1	1.27875	0.10995
	t1	845.29947	82.26359
	k	0.00118	1.15129E-4
	tau	585.91695	57.02077
1b	y0	-0.00207	8.76657E-4
	A1	1.00169	0.00608
	t1	4.98031	0.06227
	k	0.20079	0.00251
	tau	3.45209	0.04316
16	y0	3.14201	4.6598
	A1	-2.14855	4.65915
	t1	-4614.63669	9802.93344
	k	-2.16702E-4	4.60342E-4
	tau	-3198.62241	6794.87567
16 (repeat)	y0	0.01822	0.00106
	A1	0.95076	0.00265
	t1	34.44222	0.20132
	k	0.02903	1.69711E-4
	tau	23.87353	0.13955

**Figure S97.** Plot of the integration of the amide NH proton NMR signal in DMSO- $d_6$  versus time (after the addition of D<sub>2</sub>O at t = 0 min) for the PEG-capped monomers **7b**, **4b**, **1b** and control **16**. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table (tau = half-life).



Model	ExpDec1		
Equation	$y = A1 * \exp(-x/t1) + y0$		
Reduced Chi-Sqr	1.43999E-6	3.49466E-6	1.39427E-5
	3.79317E-4	3.93812E-5	5.89691E-5
Adj. R-Square	0.99998	0.99994	0.99894
	0.99341	0.99922	0.99899
	Value	Standard Error	
8b-NHa	y0	-0.02455	5.47505E-4
	A1	0.95837	5.8893E-4
	t1	67.18938	0.12401
	k	0.01488	2.74699E-5
	tau	46.57213	0.08596
8b-NHb	y0	-0.02766	4.01846E-4
	A1	0.92266	0.00105
	t1	34.49843	0.08073
	k	0.02899	6.78363E-5
	tau	23.91249	0.05596
5b-NHa	y0	-2.80144	1.06553
	A1	3.75771	1.06453
	t1	1903.06836	569.23089
	k	5.25467E-4	1.57174E-4
	tau	1319.10647	394.56079
5b-NHb	y0	-0.02695	0.0033
	A1	1.04947	0.01284
	t1	21.85396	0.48993
	k	0.04576	0.00103
	tau	15.14801	0.3396
2b-NHa	y0	-0.26402	0.02692
	A1	1.29018	0.02536
	t1	230.99516	7.30268
	k	0.00433	1.3686E-4
	tau	160.11365	5.06183
2b-NHb	y0	0.01416	0.00163
	A1	0.953	0.00434
	t1	33.80962	0.3144
	k	0.02958	2.75044E-4
	tau	23.43504	0.21793

**Figure S98.** Plot of the integration of the amide NH proton NMR signal in DMSO- $d_6$  versus time (after the addition of D<sub>2</sub>O at t = 0 min) for the PEG-capped dimers **8b**, **5b** and **2b**. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table (tau = half-life).



**Figure S99.** Plot of the integration of the amide NH proton NMR signal in DMSO- $d_6$  versus time (after the addition of D<sub>2</sub>O at t = 0 min) for the PEG-capped trimers **9b**, **6b** and **3b**. The solid lines are the result of an exponential decay fit, with the obtained values shown in the right-hand table (tau = half-life).

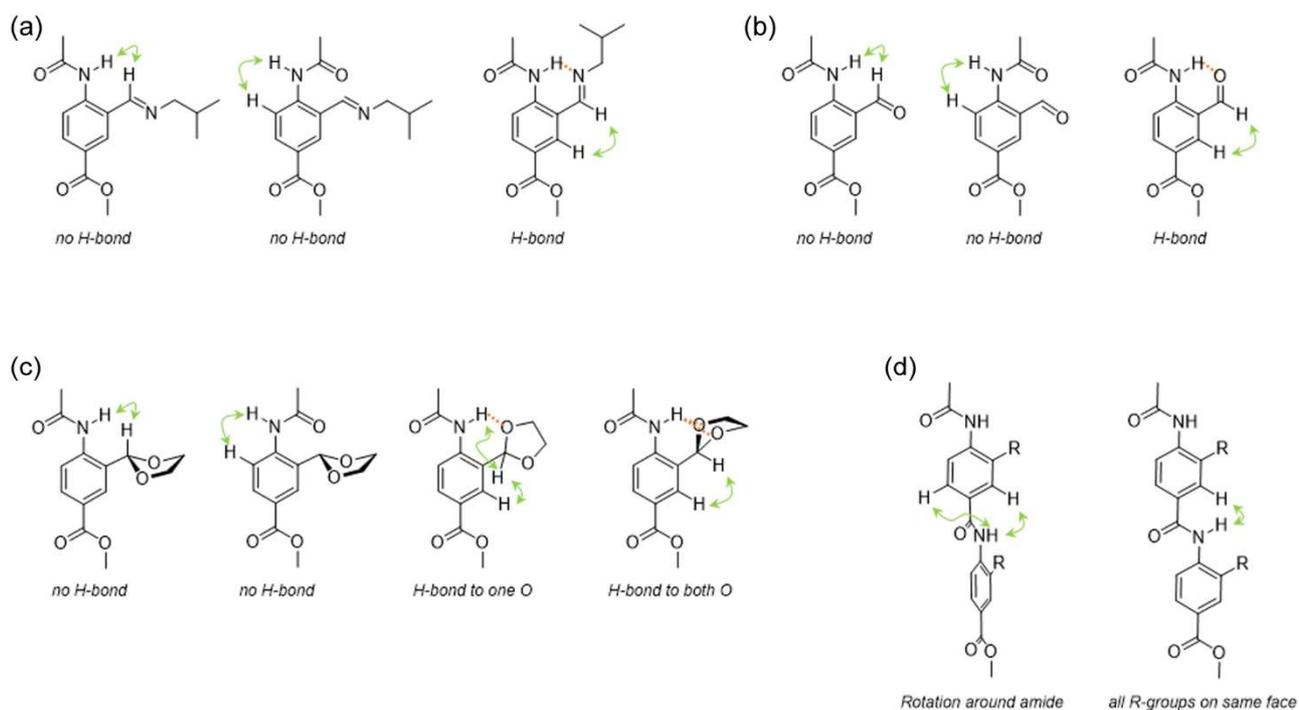
### 3.7. NOESY/ROESY NMR (solution)

<sup>1</sup>H-<sup>1</sup>H NOESY or ROESY NMR (ROESY for compounds with MW > 700 g/mol, i.e. trimers **3b**, **6b** and **9b**) could potentially be used to determine the most preferred conformation of the  $\alpha$ -helix mimetics in solution (CDCl<sub>3</sub> and DMSO- $d_6$  were used as solvents and the spectra were collected at 298 K). The obtained NOESY spectra are shown in Figures S101-S124.

The intramolecular hydrogen bonding can be confirmed by nOe, as shown in Figure S100. When there is no hydrogen bond present the amide NH proton could give nOe cross peaks with the *ortho* CH or with the imine/aldehyde CH group, while these nOe signals will not be present when the compounds are strongly hydrogen bonded (see Figure S100). The case of the acetal containing compounds is more difficult, because hydrogen bonds could potential form with only one or with both of the oxygen atoms of the acetal group (Figure S100).

<sup>1</sup>H-<sup>1</sup>H NOESY experiments can also be used to see if the side chains are on the same face of the mimetic or whether free rotation occurs around the amide bonds that allows the side chains to be

displayed in various relative positions. When rotation occurs (or in the case of a staggered conformation), the amide NH proton can have cross peaks with both of the aromatic signals of the neighbouring aromatic ring (Figure S100), while if all of the side chains are on the same face of the helix mimetic, the nOe cross peak will be more pronounced for one of the two aromatic protons.



**Figure S100.** Possible conformations of the imines (a), aldehyde (b) and acetals (c) that can be the result of the absence/presence of hydrogen bonding, as well as possible rotation around the amide bonds (d). Hydrogen bonds are represented by dashed orange lines. The expected strong nOe signals are shown for each conformer by a green arrow.

### Imine containing compounds (7a, 7b, 8b and 9b): Figures S101-S108

The  $^1\text{H}$ - $^1\text{H}$  NOESY (or ROESY) NMR spectra in  $\text{CDCl}_3$  and  $\text{DMSO-}d_6$  show similar cross peaks, and the conformation of the imines is expected to be the same in both solvents. There are no significant cross peaks between the amide protons and the *ortho* CH protons (H8-H10, H20-H22, H32-H34), or between the amide NH protons and the imine CH protons (H8-H15, H20-H27, H32-H39). On the other hand, there are strong cross peaks between the imine CH and the aromatic region (H13-H15, H25-H27, H37-H39). These two observations suggest the existence of a hydrogen bond between the amide NH and the imine nitrogen atom. The amide NH signals show cross peaks of roughly equal intensity with the two neighbouring aromatic protons (H20-H11/H20-H13 and H32-H23/H32-H25), indicating free rotation around the bond between the aromatic ring and the amide carbonyl function. This free rotation was also confirmed by the fact that in  $\text{DMSO-}d_6$

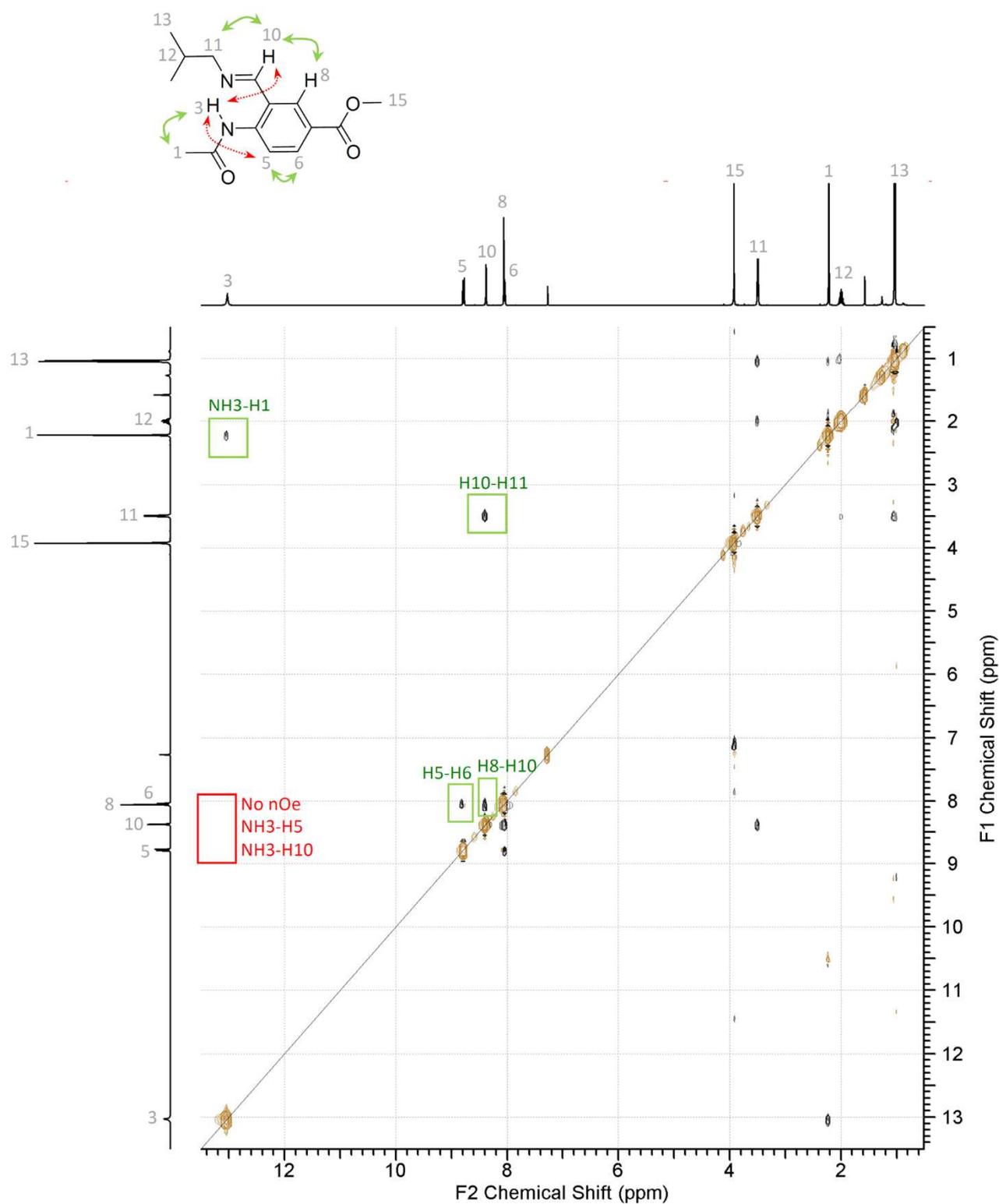
some of the cross peaks have the same phase as the diagonal, suggesting rotational exchange (this is presumably not observed in CDCl<sub>3</sub> because the rotation will be faster in this less viscous solvent).

#### Aldehyde containing compounds (4a, 4b, 5b and 6b): Figures S109-S116

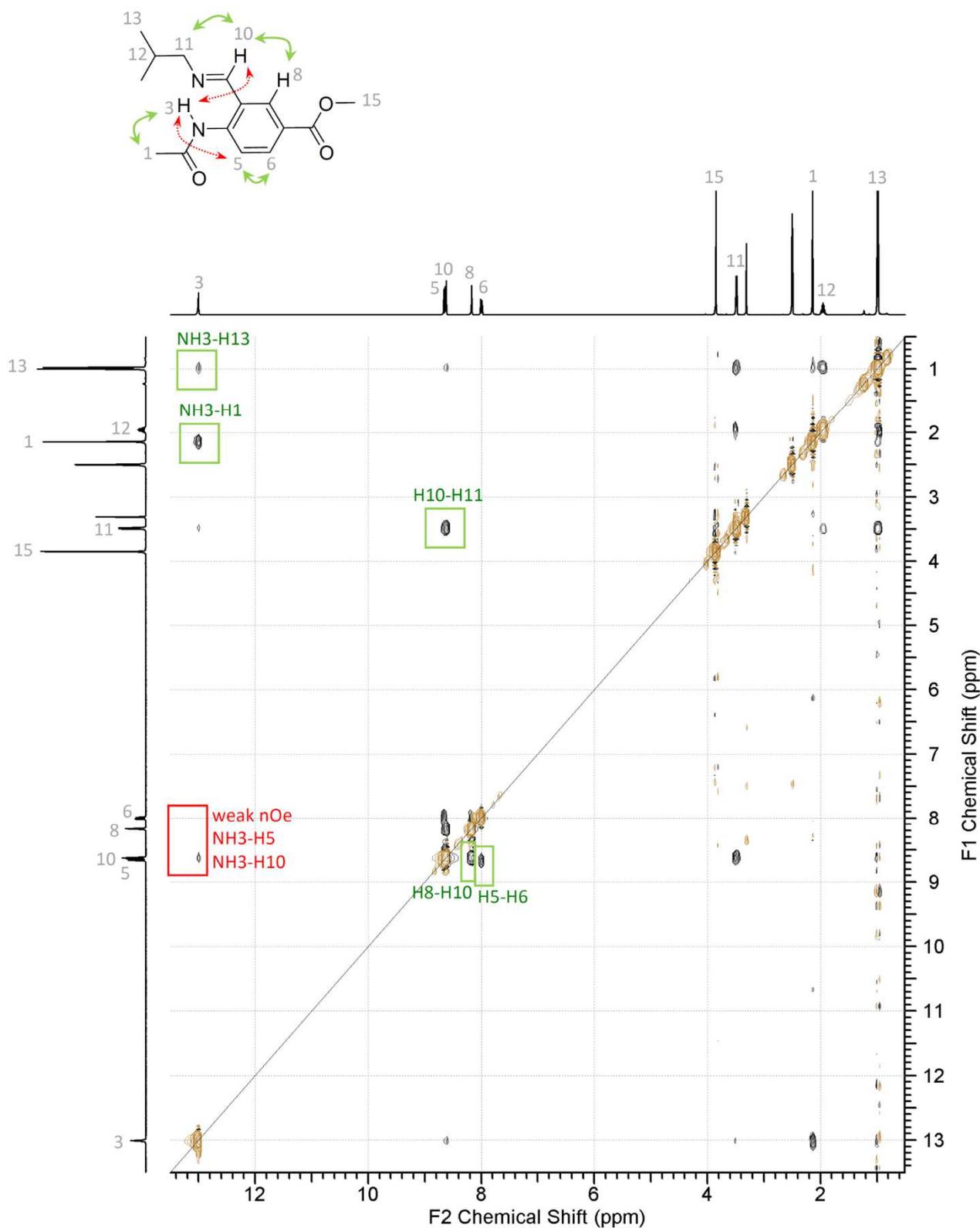
The <sup>1</sup>H-<sup>1</sup>H NOESY (or ROESY) NMR spectra in CDCl<sub>3</sub> and DMSO-*d*<sub>6</sub> show similar cross peaks. There are no significant cross peaks between the amide protons and the *ortho* CH protons (H8-H10, H17-H19, H26-H28), or between the amide NH protons and the aldehyde CH protons (H8-H15, H17-H24, H26-H33), while there are strong cross peaks between the aldehyde CH and the aromatic region (H13-H15, H22-H24, H31-H33), suggesting the existence of a hydrogen bond between the amide NH and the aldehyde oxygen atom. The amide NH signals show cross peaks of roughly equal intensity with the two neighbouring aromatic protons (H17-H11/H17-H13 and H26-H20/H26-H22), indicating free rotation around the amide bond. This free rotation was also confirmed by the fact that in DMSO-*d*<sub>6</sub> some of the cross peaks have the same phase as the diagonal.

#### Acetal containing compounds (1a, 1b, 2b and 3b): Figures S117-S124

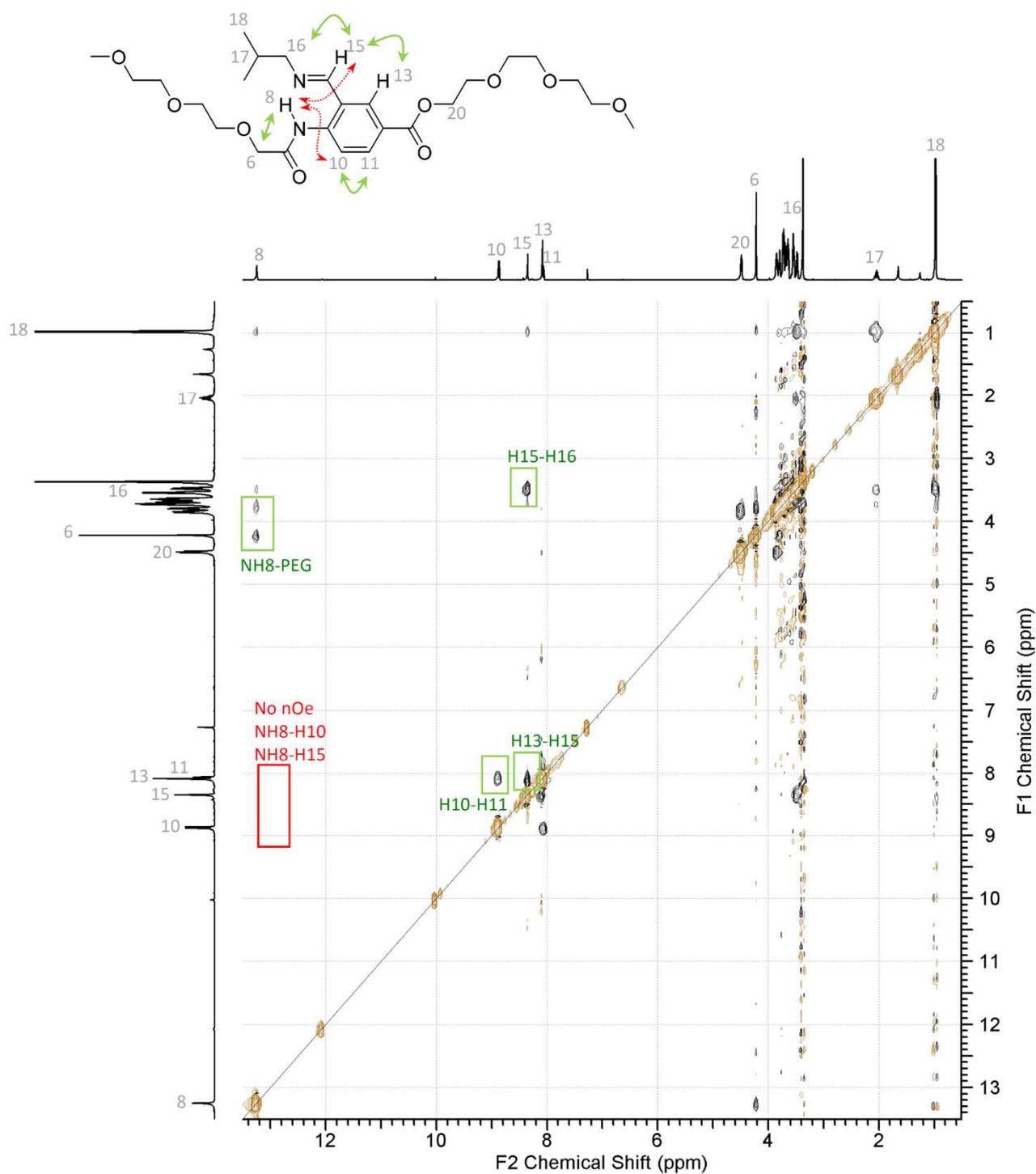
The <sup>1</sup>H-<sup>1</sup>H NOESY (or ROESY) NMR spectra in CDCl<sub>3</sub> and DMSO-*d*<sub>6</sub> show similar cross peaks. There are no significant cross peaks between the amide protons and the *ortho* CH protons (H8-H10, H18-H20, H28-H30), there are no or weak cross peaks (depending on the compound) between the amide NH protons and the acetal CH protons (H8-H15, H18-H25, H28-H35), and there are strong cross peaks between the acetal CH and the aromatic region (H13-H15, H23-H25, H33-H35). This suggests the existence of a hydrogen bond between the amide NH and an acetal oxygen atom, but it cannot be concluded whether the hydrogen bond is with one or both of the acetal oxygen atoms. The amide NH signals show cross peaks of roughly equal intensity with the two neighbouring aromatic protons (H18-H11/H18-H13 and H28-H21/H28-H23), indicating free rotation around the amide bond. This free rotation was also confirmed by the fact that in DMSO-*d*<sub>6</sub> some of the cross peaks have the same phase as the diagonal.



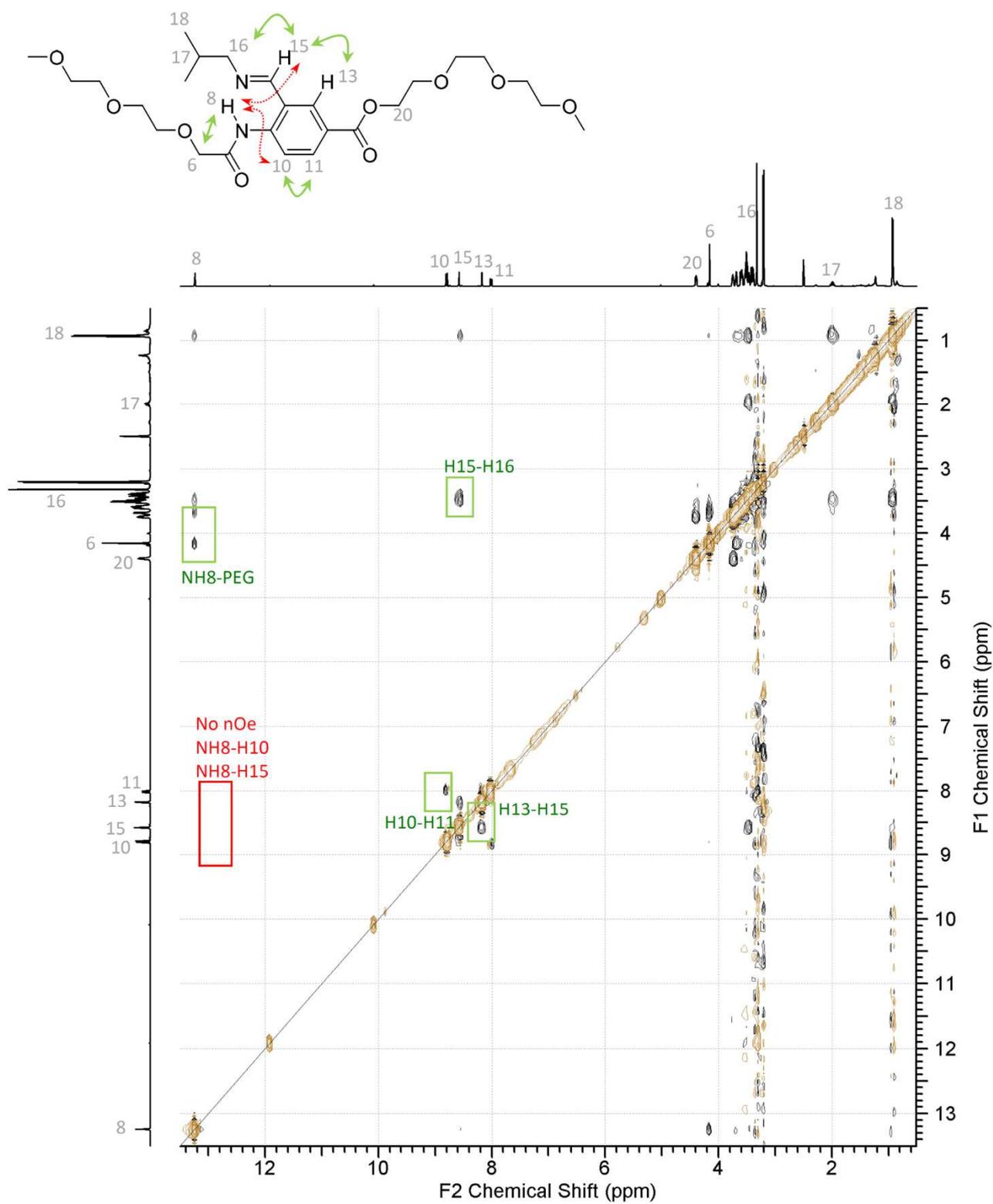
**Figure S101.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **7a** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **7a**.



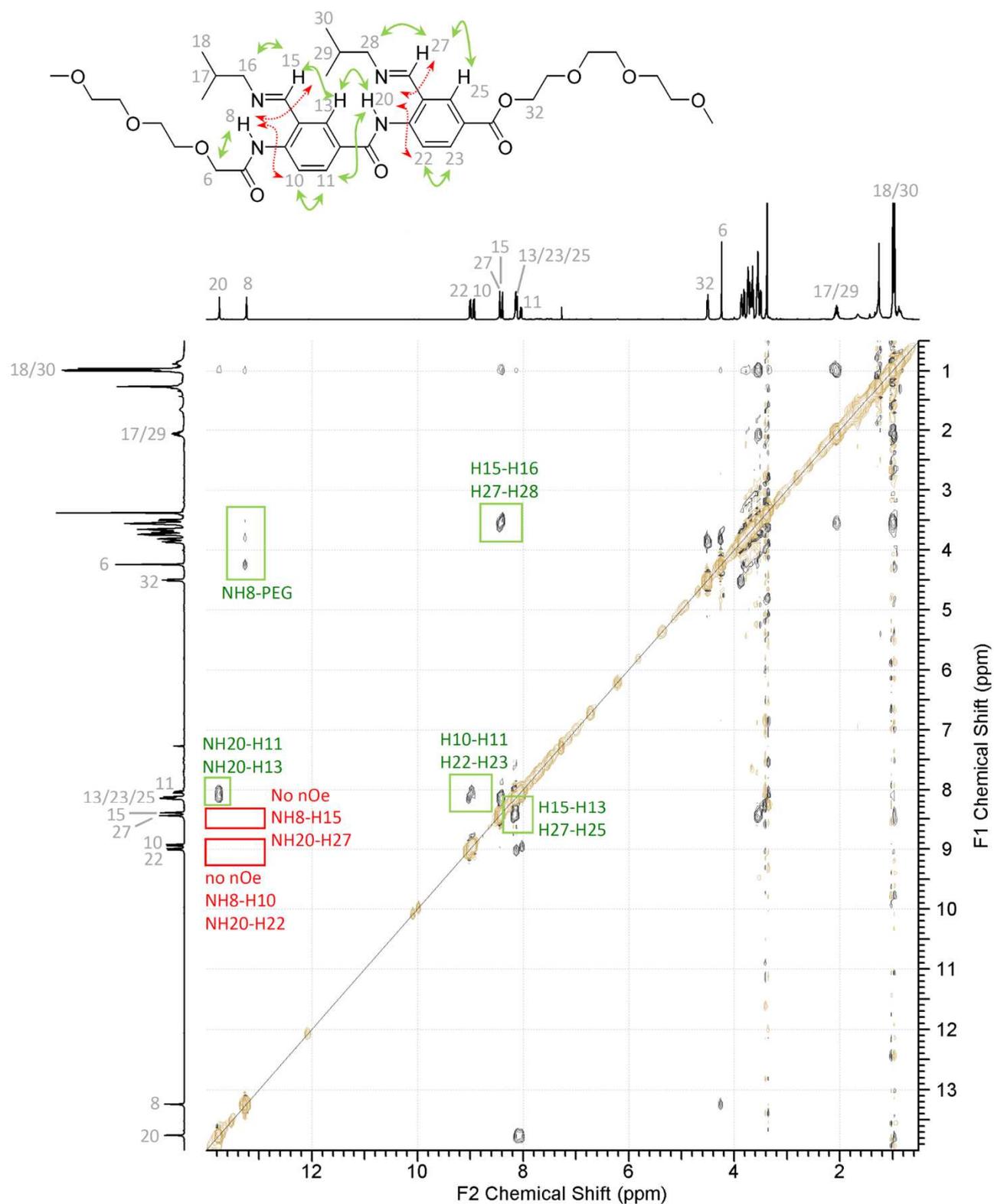
**Figure S102.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **7a** in  $\text{DMSO}-d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **7a**.



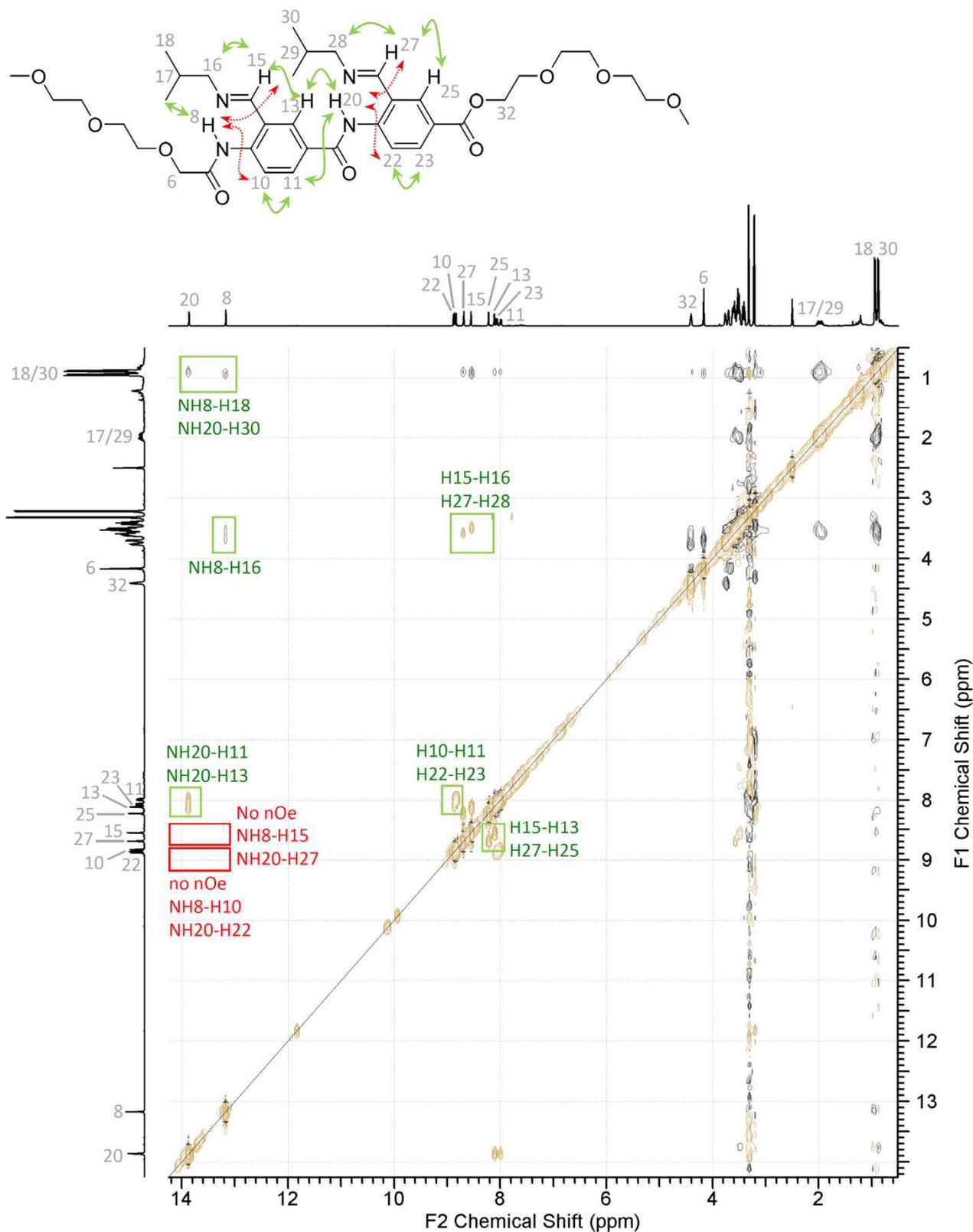
**Figure S103.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **7b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **7b**.



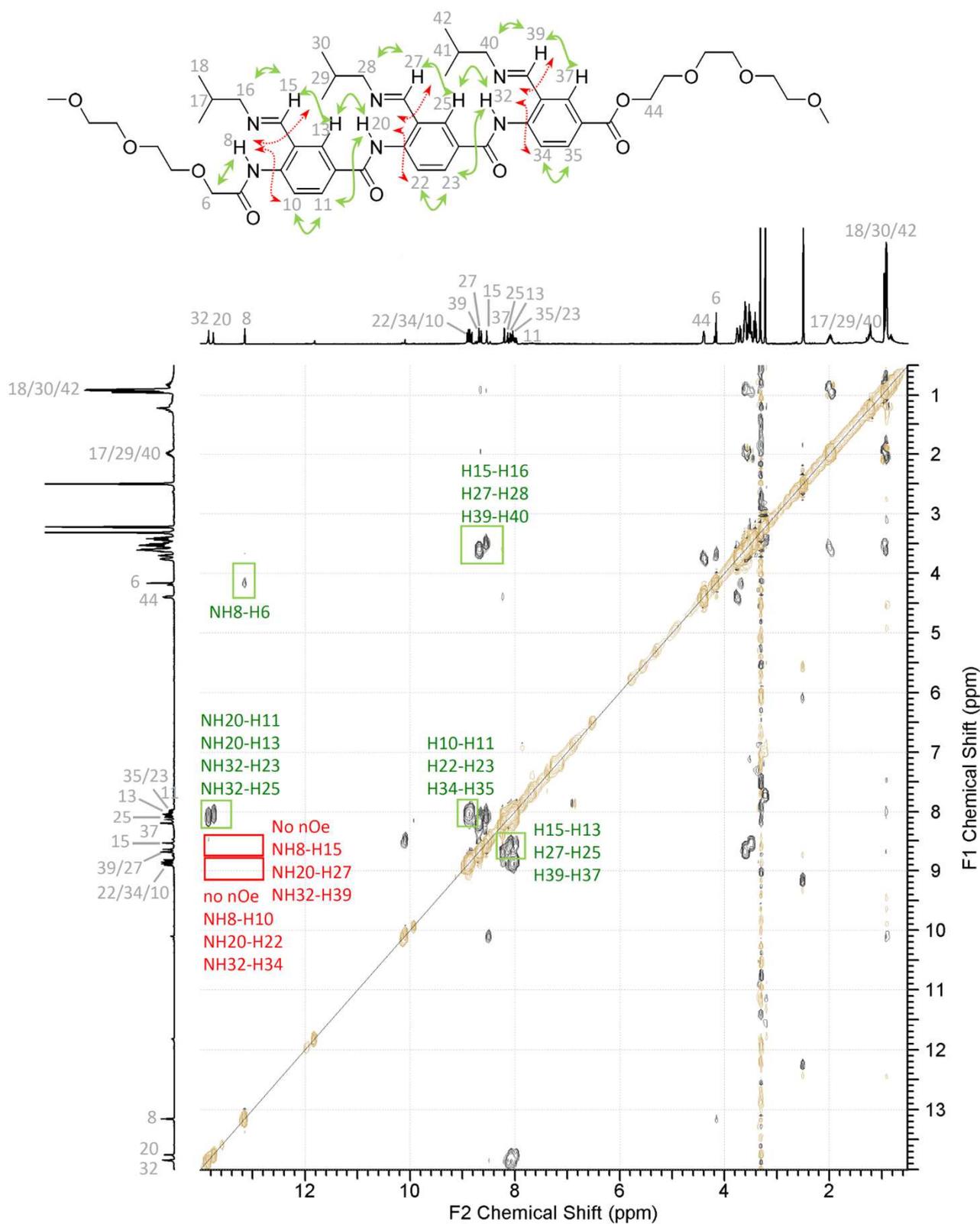
**Figure S104.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **7b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **7b**.



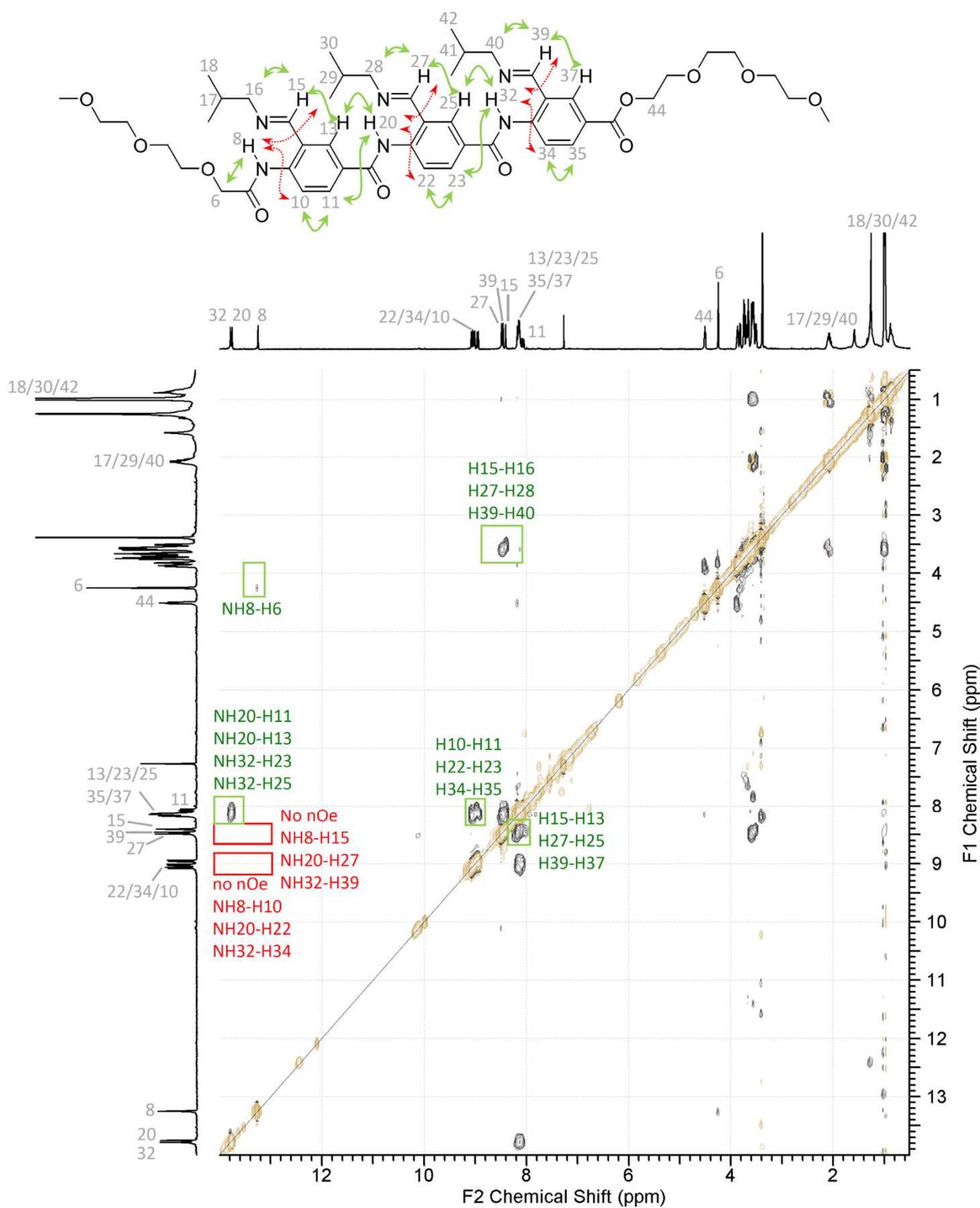
**Figure S105.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **8b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **8b**. Equal intensity cross peaks for H20-H11 and H20-H13 suggests rotation around the amide bond.



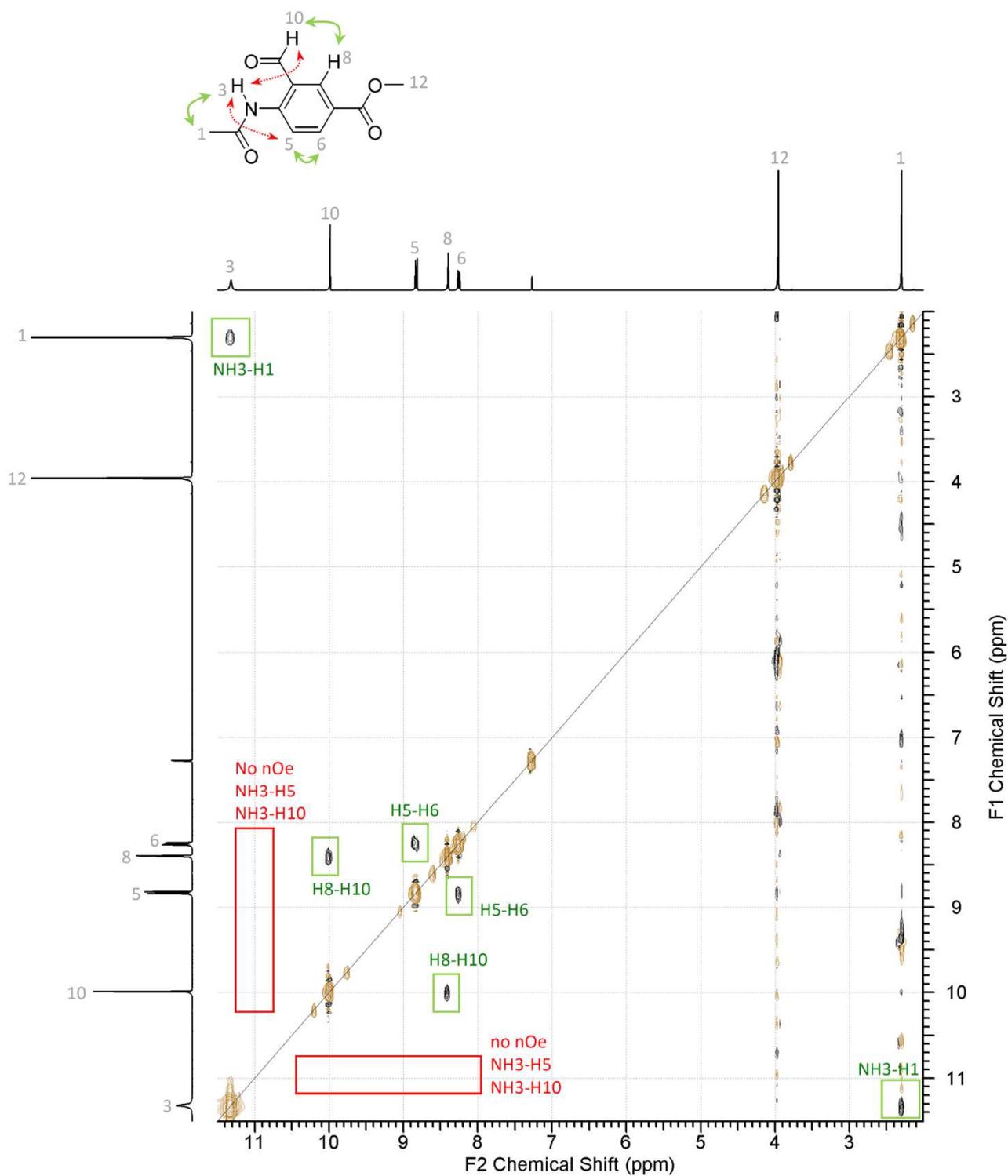
**Figure S106.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **8b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **8b**. Rotation around the amide bond is suggested by the appearance of cross peaks with the same phase as the diagonal due to rotational exchange.



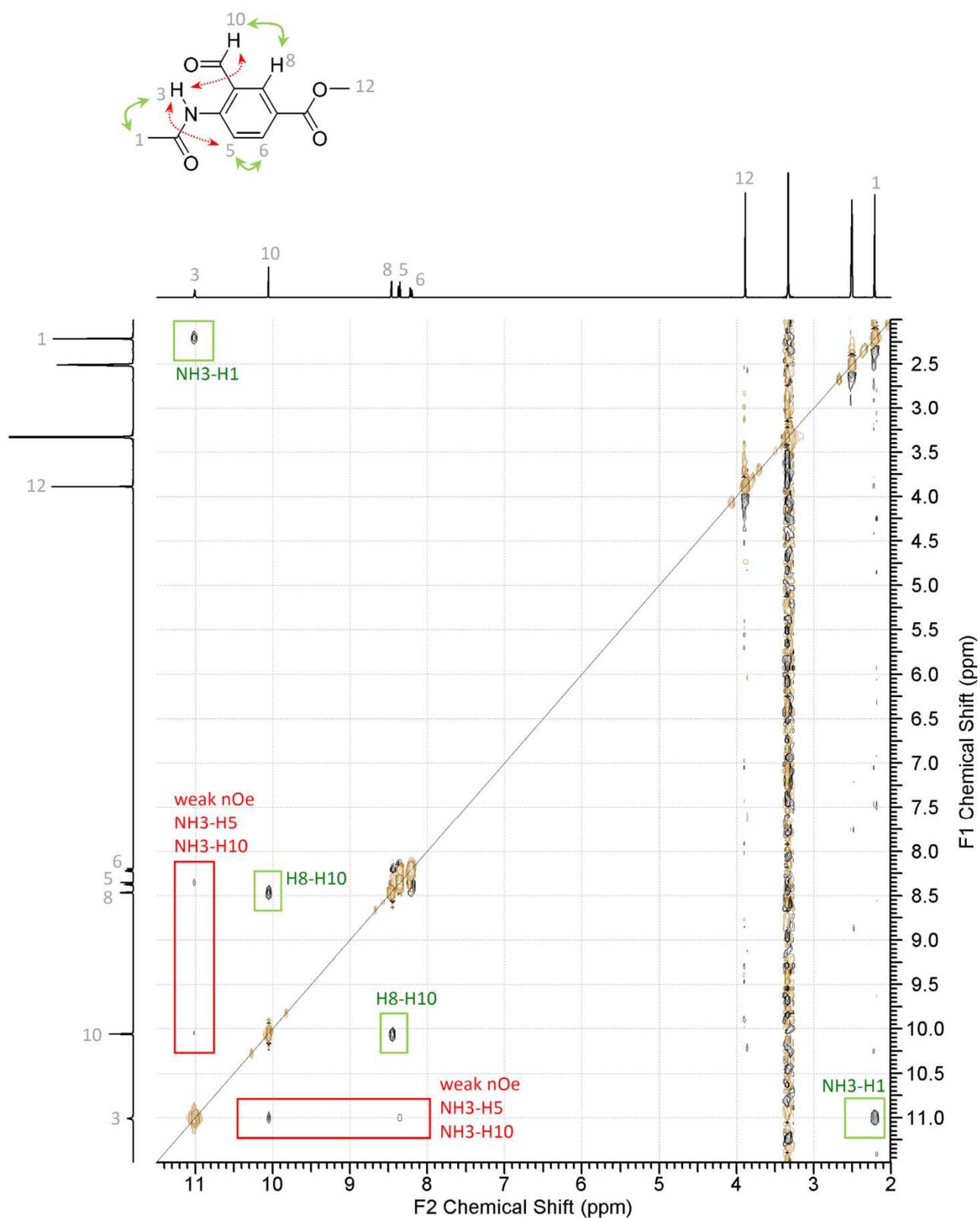
**Figure S107.**  $^1\text{H}$ - $^1\text{H}$  ROESY NMR spectrum (400, 400 MHz) of compound **9b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **9b**. Equal intensity cross peaks for H20-H11 and H20-H13, and for H32-H23 and H32-H25, suggests rotation around the amide bonds.



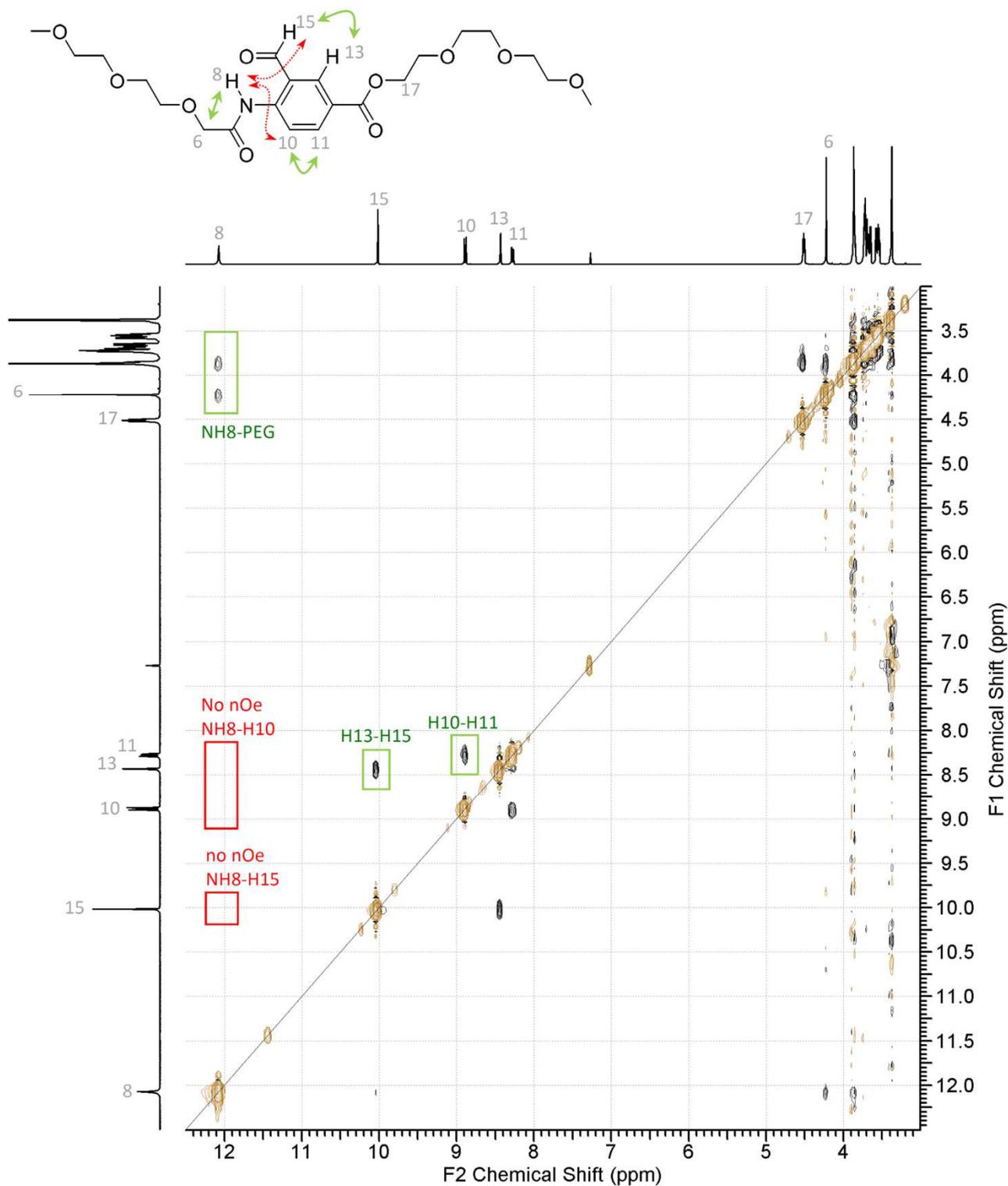
**Figure S108.**  $^1\text{H}$ - $^1\text{H}$  ROESY NMR spectrum (400, 400 MHz) of compound **9b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **9b**. Rotation is expected, but cannot be confirmed due to significant overlap.



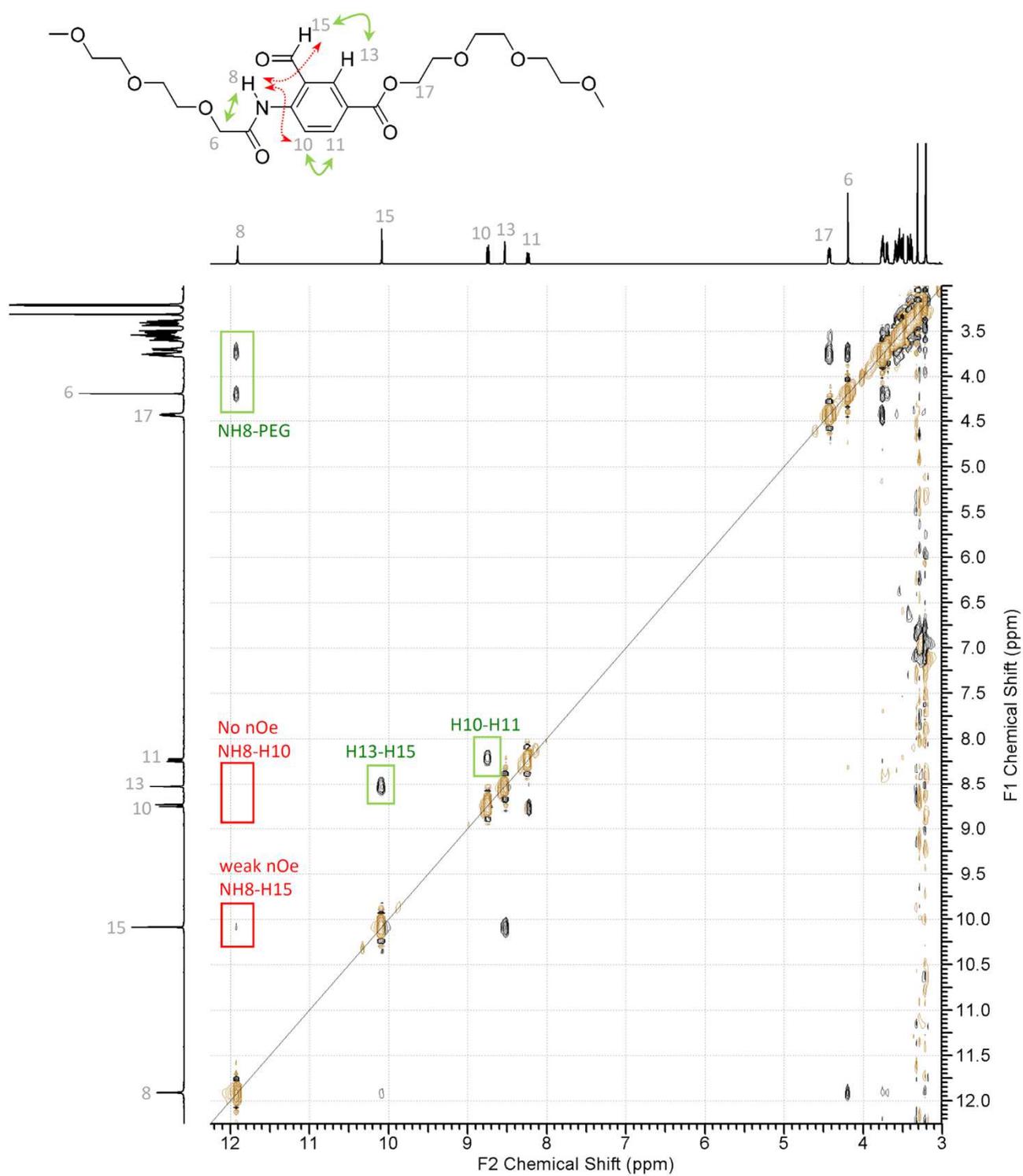
**Figure S109.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **4a** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **4a**.



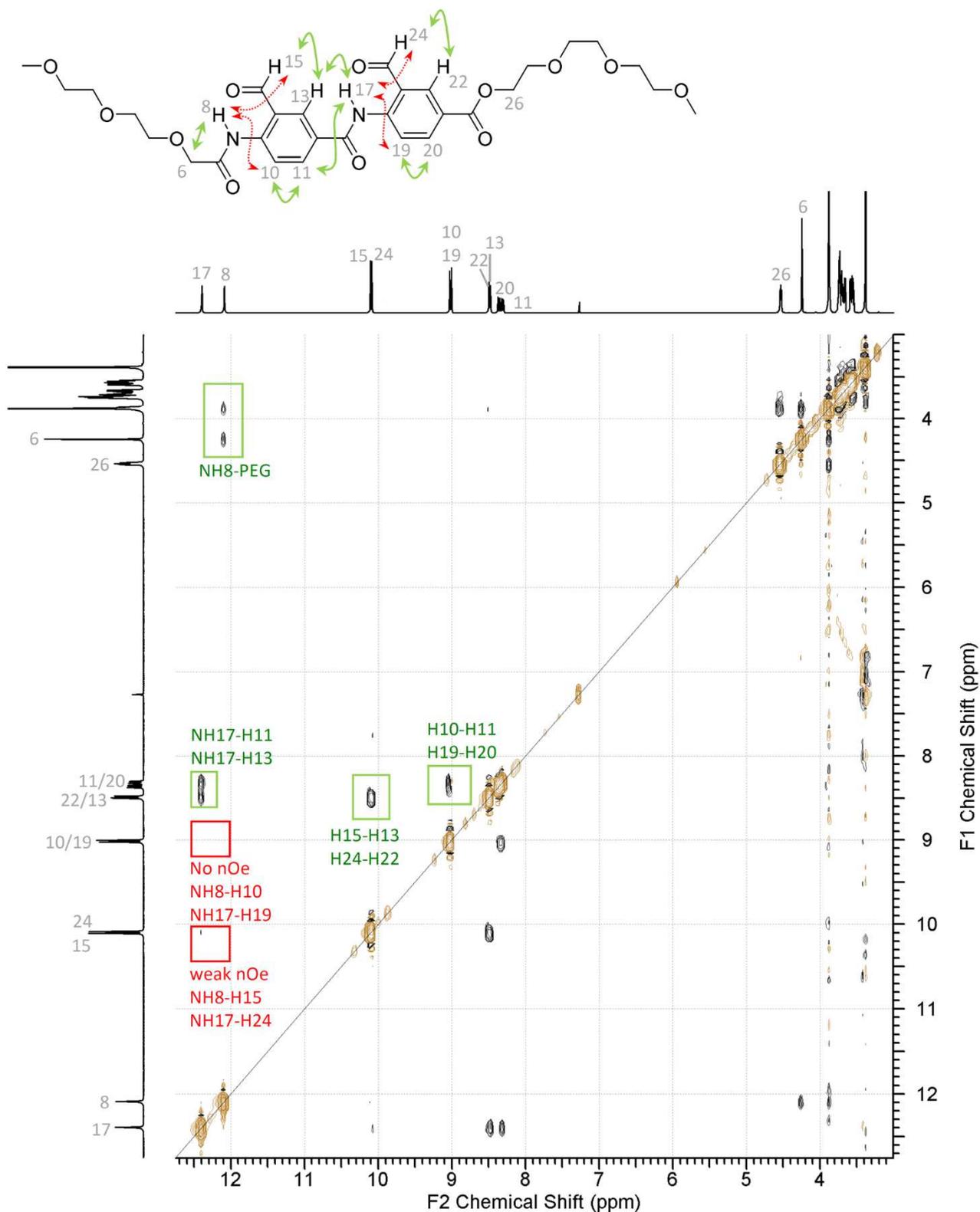
**Figure S110.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **4a** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **4a**.



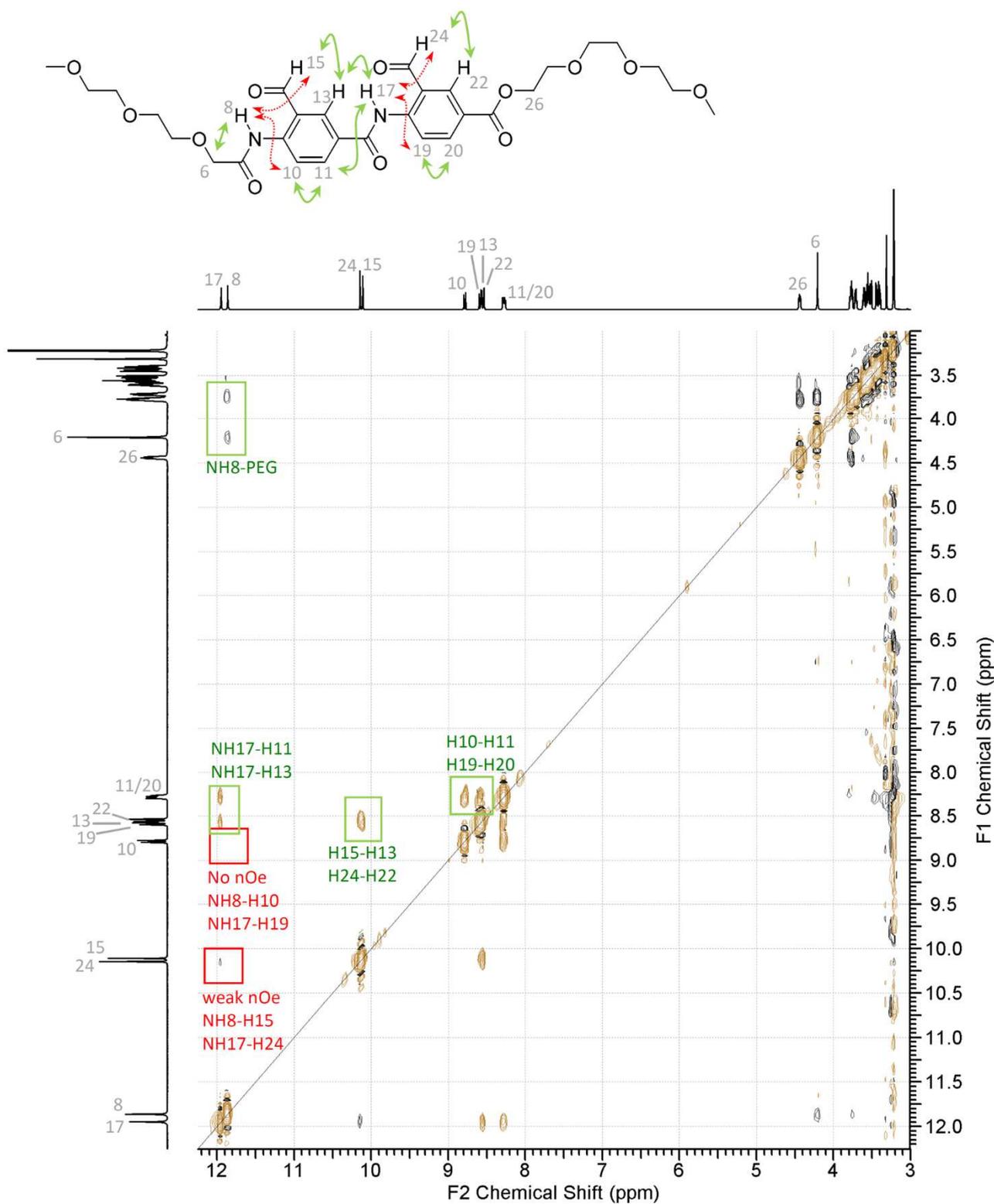
**Figure S111.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **4b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **4b**.



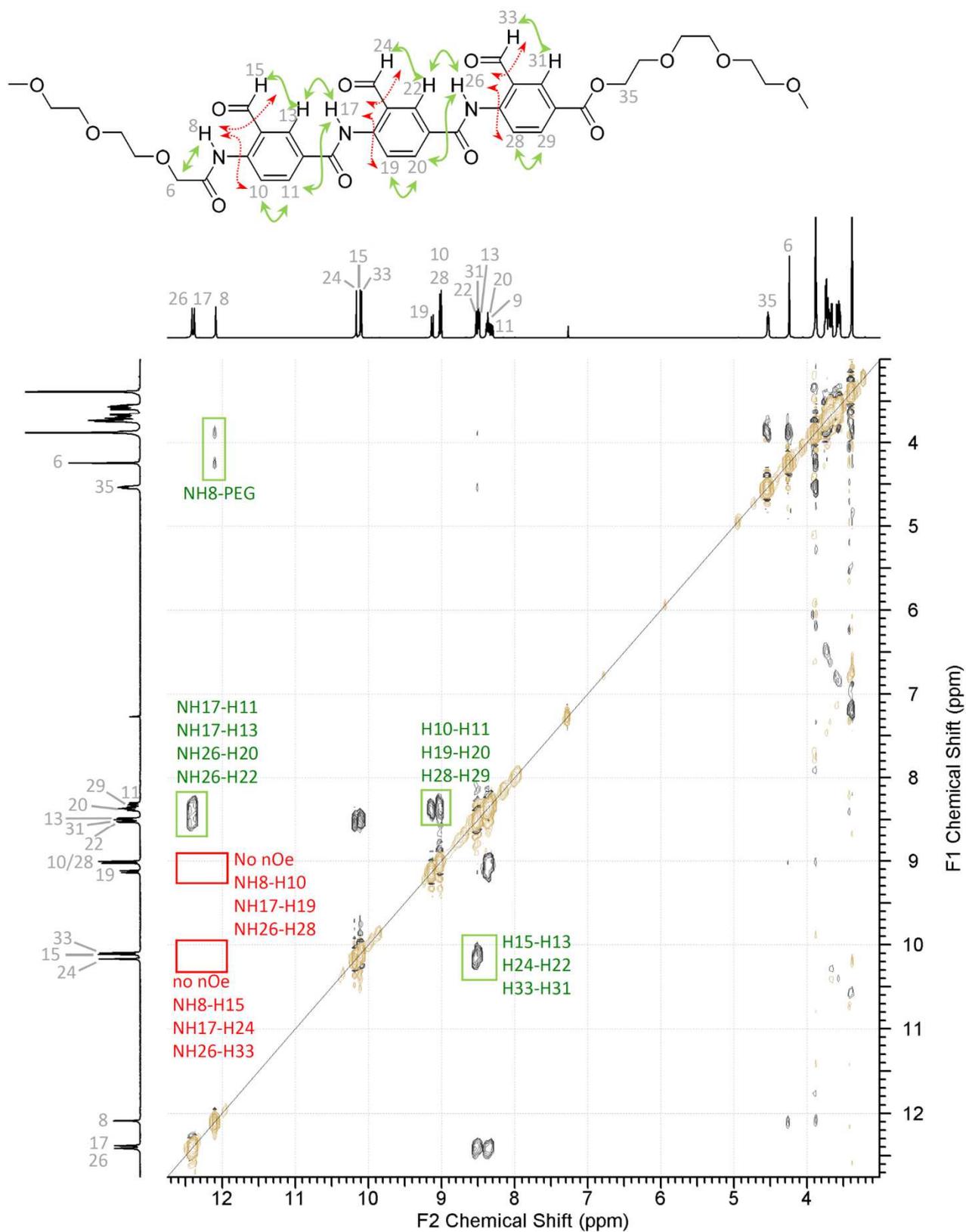
**Figure S112.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **4b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **4b**.



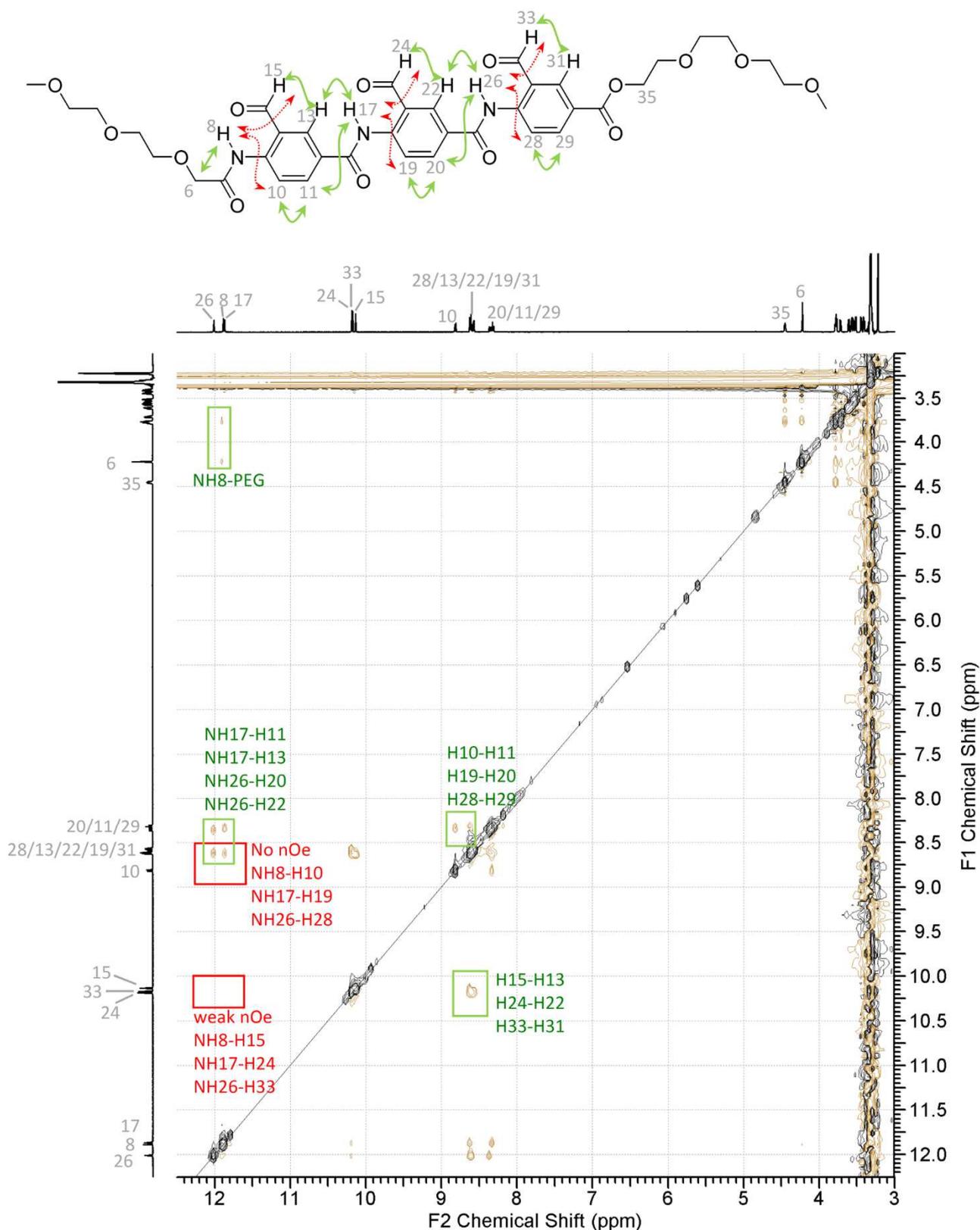
**Figure S113.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **5b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **5b**. Equal intensity cross peaks for H17-H11 and H17-H13 suggests rotation around the amide bond.



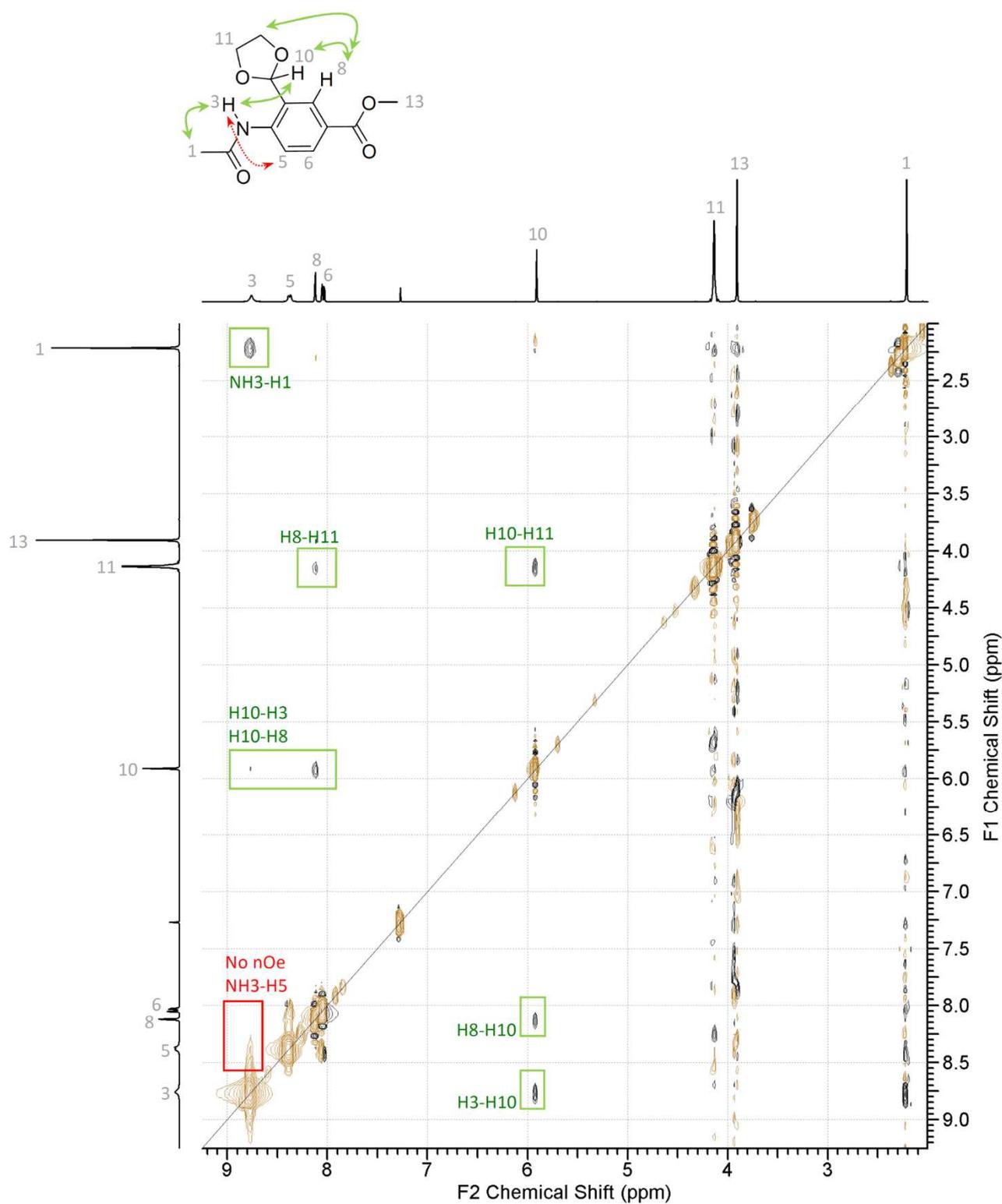
**Figure S114.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **5b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **5b**. Rotation around the amide bond is suggested by the appearance of cross peaks with the same phase as the diagonal due to rotational exchange (all cross peaks also appear in the equivalent ROESY spectrum).



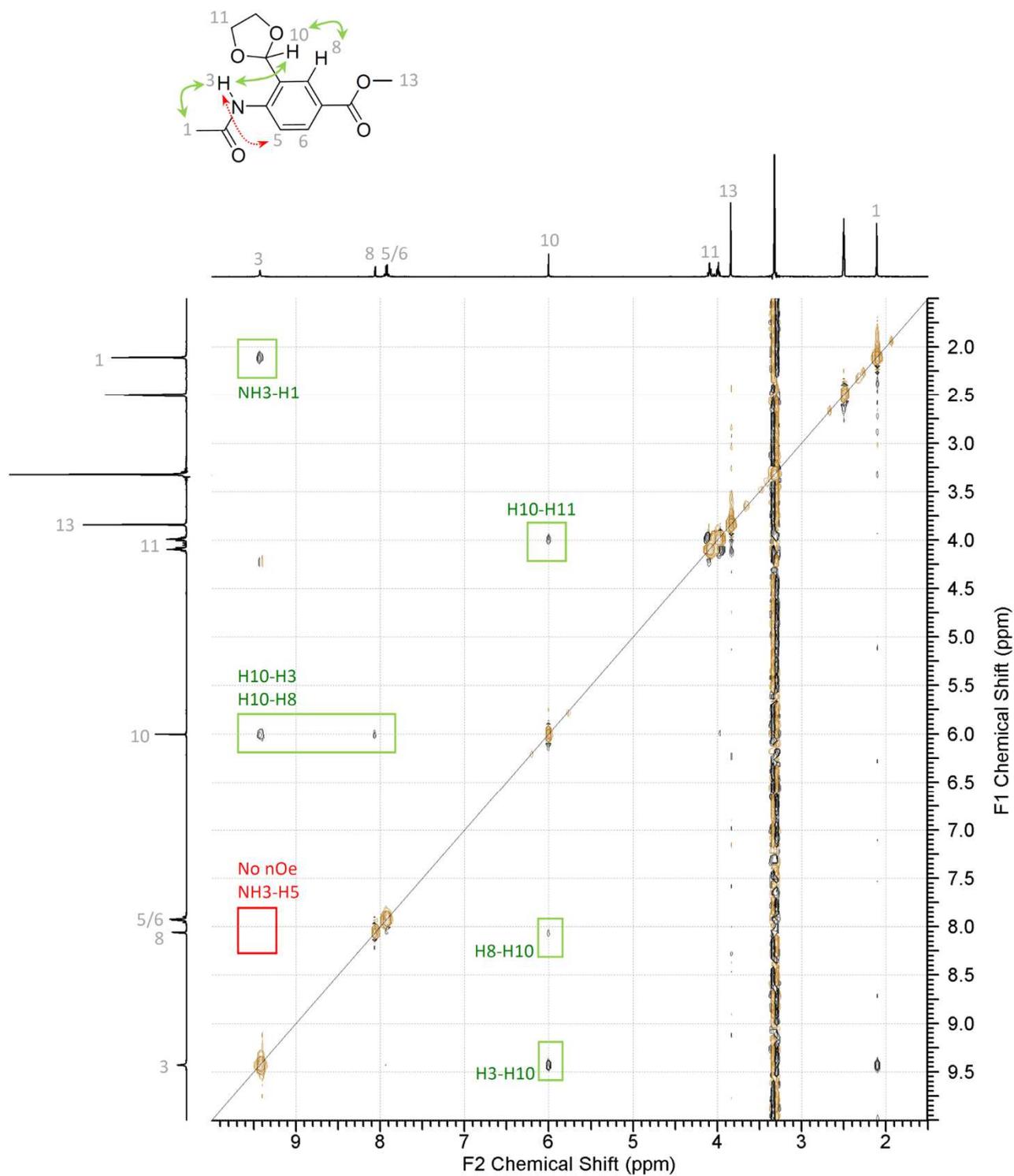
**Figure S115.**  $^1\text{H}$ - $^1\text{H}$  ROESY NMR spectrum (400, 400 MHz) of compound **6b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **6b**. Rotation of the amide bond is expected, but cannot be fully confirmed due to overlapping signals.



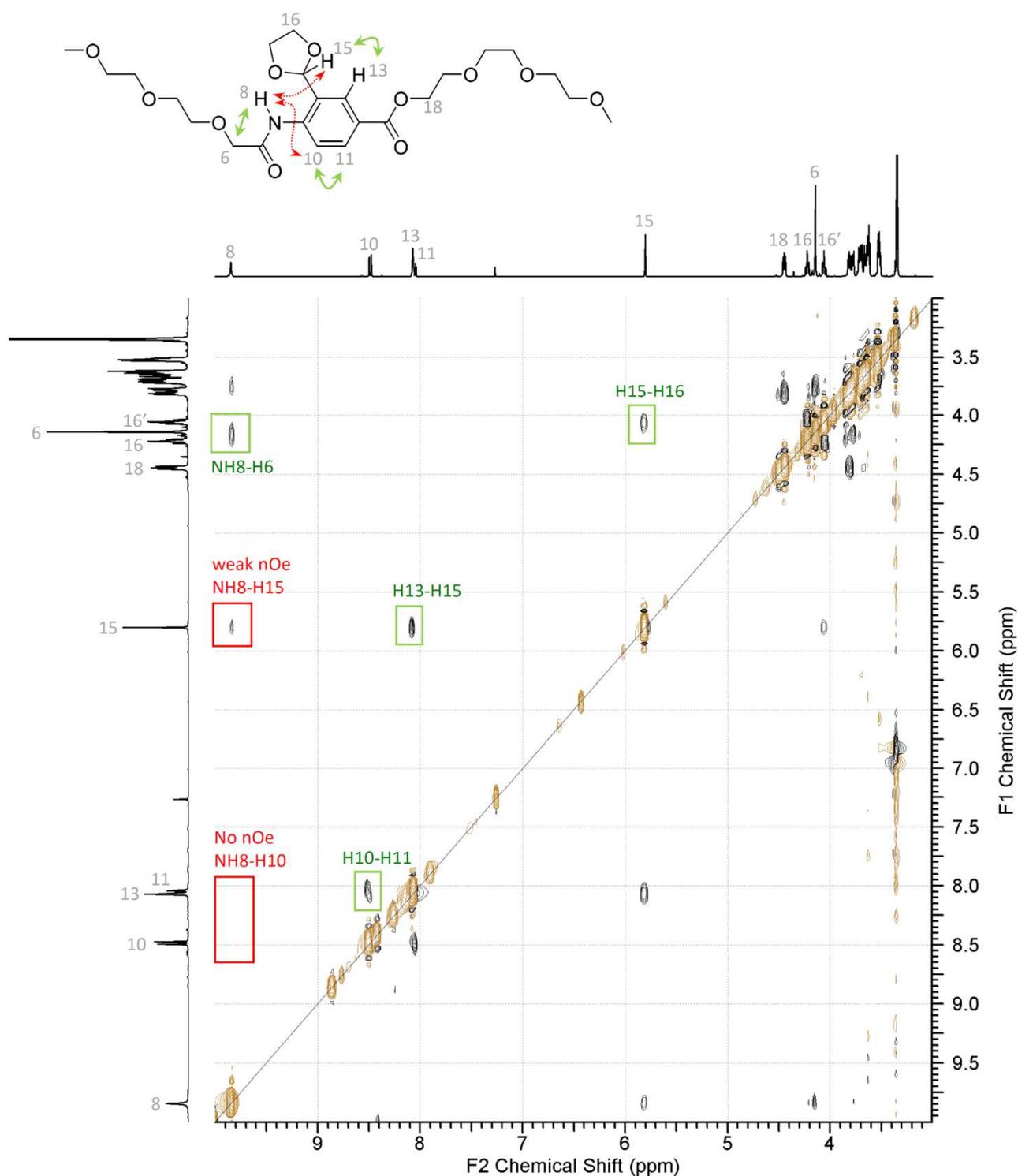
**Figure S116.**  $^1\text{H}$ - $^1\text{H}$  ROESY NMR spectrum (600, 600 MHz) of compound **6b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **6b**. Equal intensity cross peaks for H17-H11 and H17-H13, and for H26-H20 and H26-H22, suggests rotation around the amide bonds.



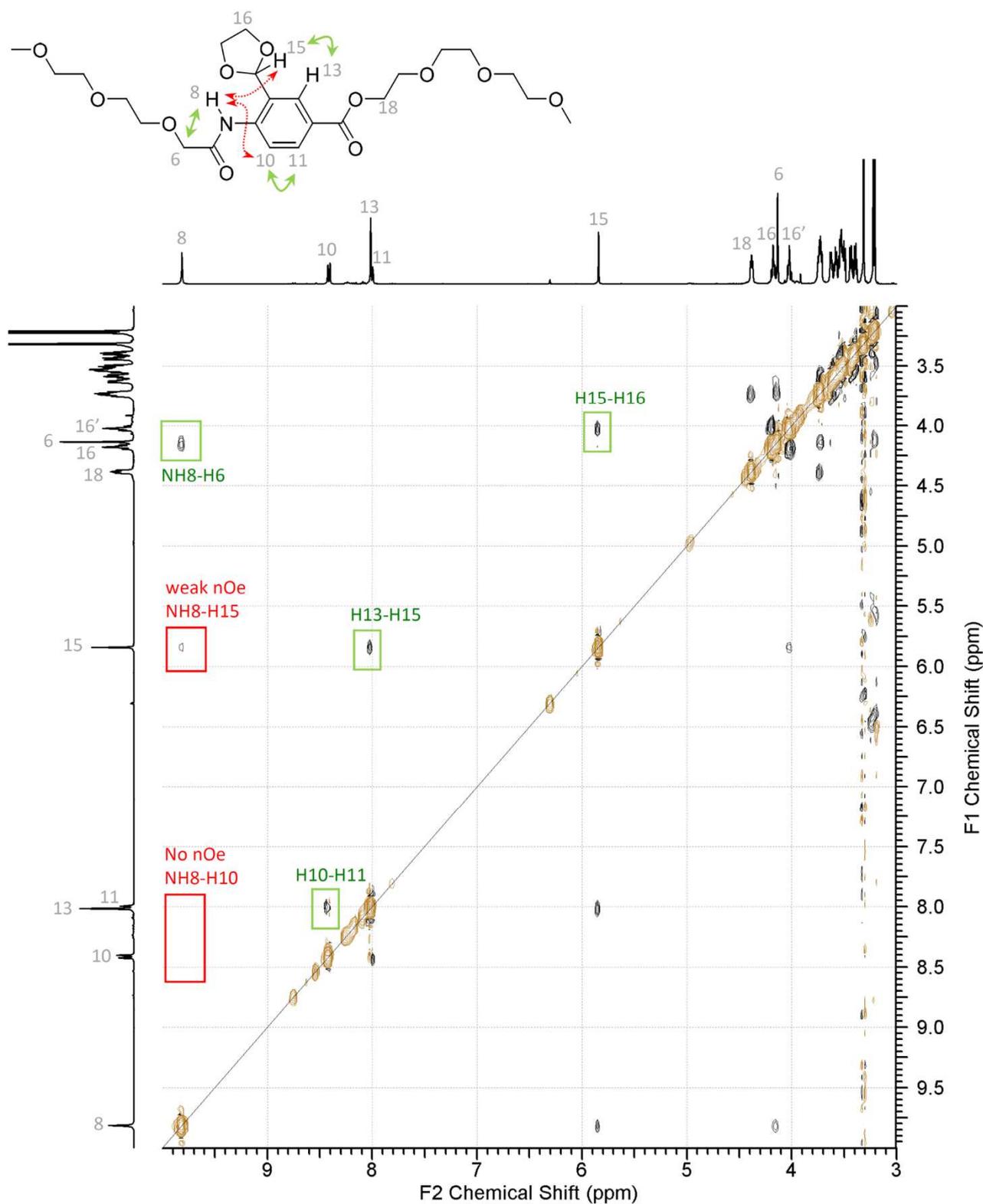
**Figure S117.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **1a** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **1a**. Cross peaks of H10 are observed with both H3 and H8 (but stronger with H8), suggesting hydrogen bonding to one or two of the oxygen atoms in the acetal group.



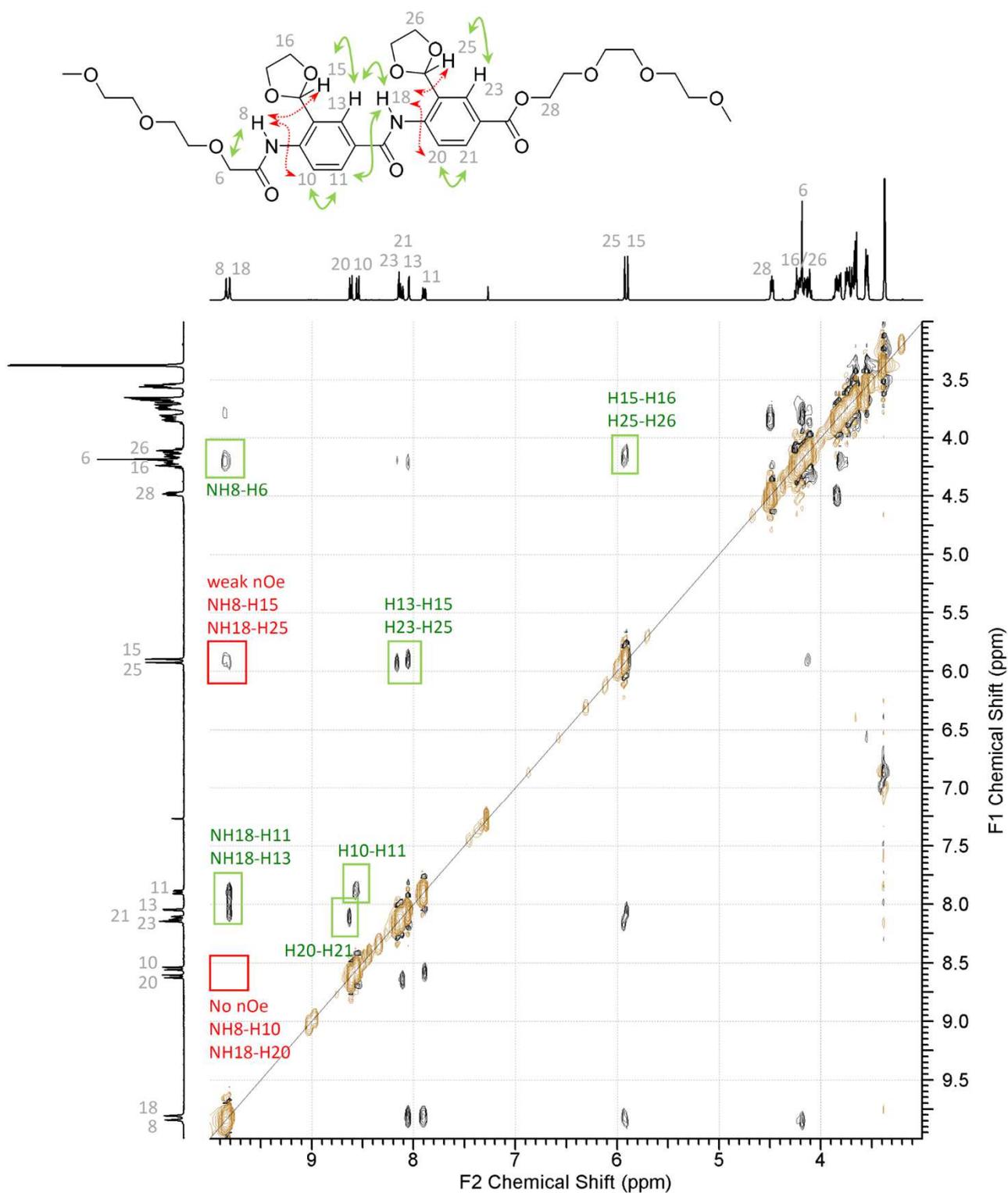
**Figure S118.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **1a** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **1a**. Cross peaks of H10 are observed with both H3 and H8 (a bit stronger with H3), suggesting hydrogen bonding to only one of the oxygen atoms in the acetal group.



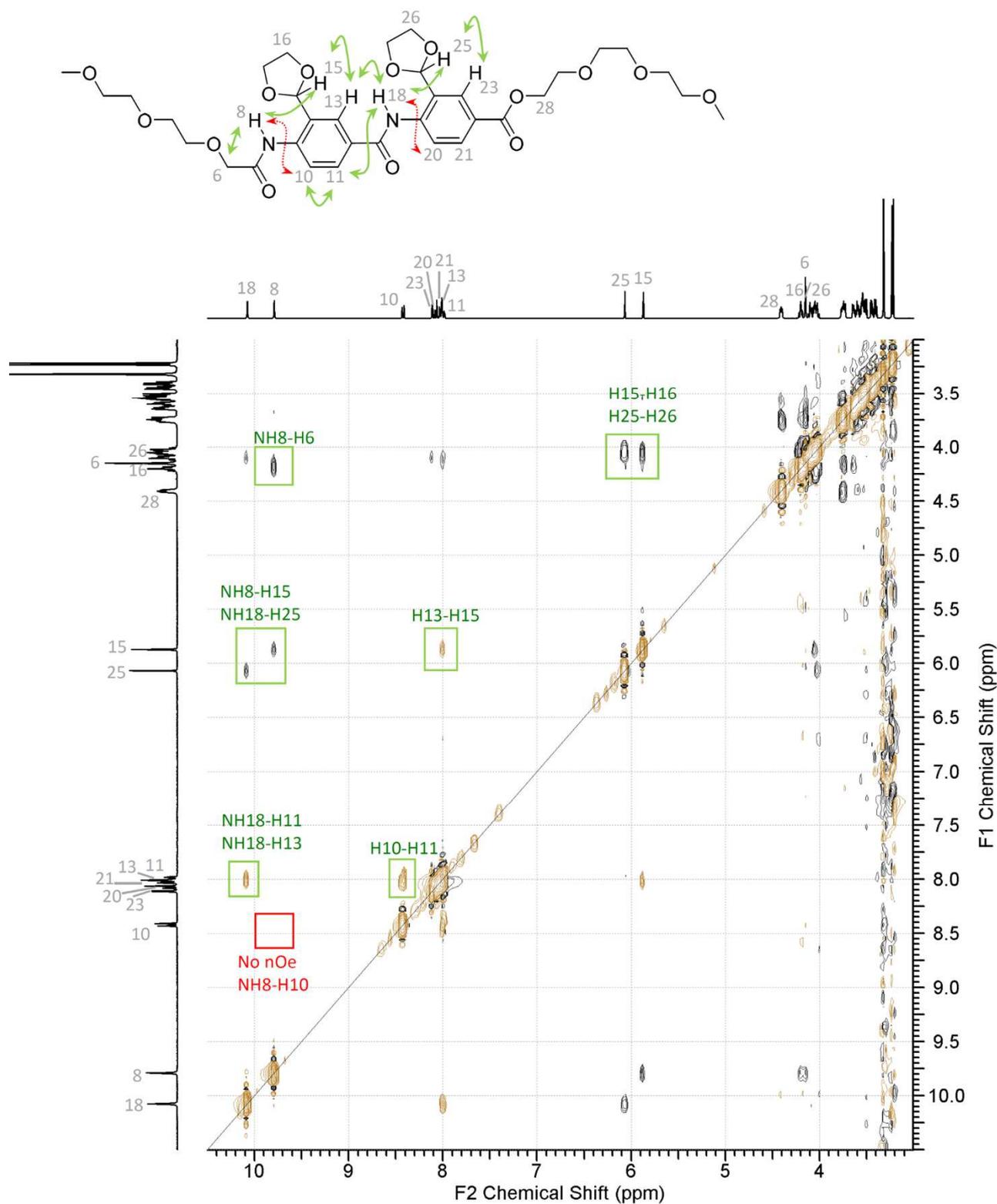
**Figure S119.** <sup>1</sup>H-<sup>1</sup>H NOESY NMR spectrum (400, 400 MHz) of compound **1b** in CDCl<sub>3</sub> at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **1b**. The nOe cross peak between H15 and H13 is much stronger than between H15 and H8, suggesting that the amide NH is hydrogen bonded to both acetal oxygen atoms.



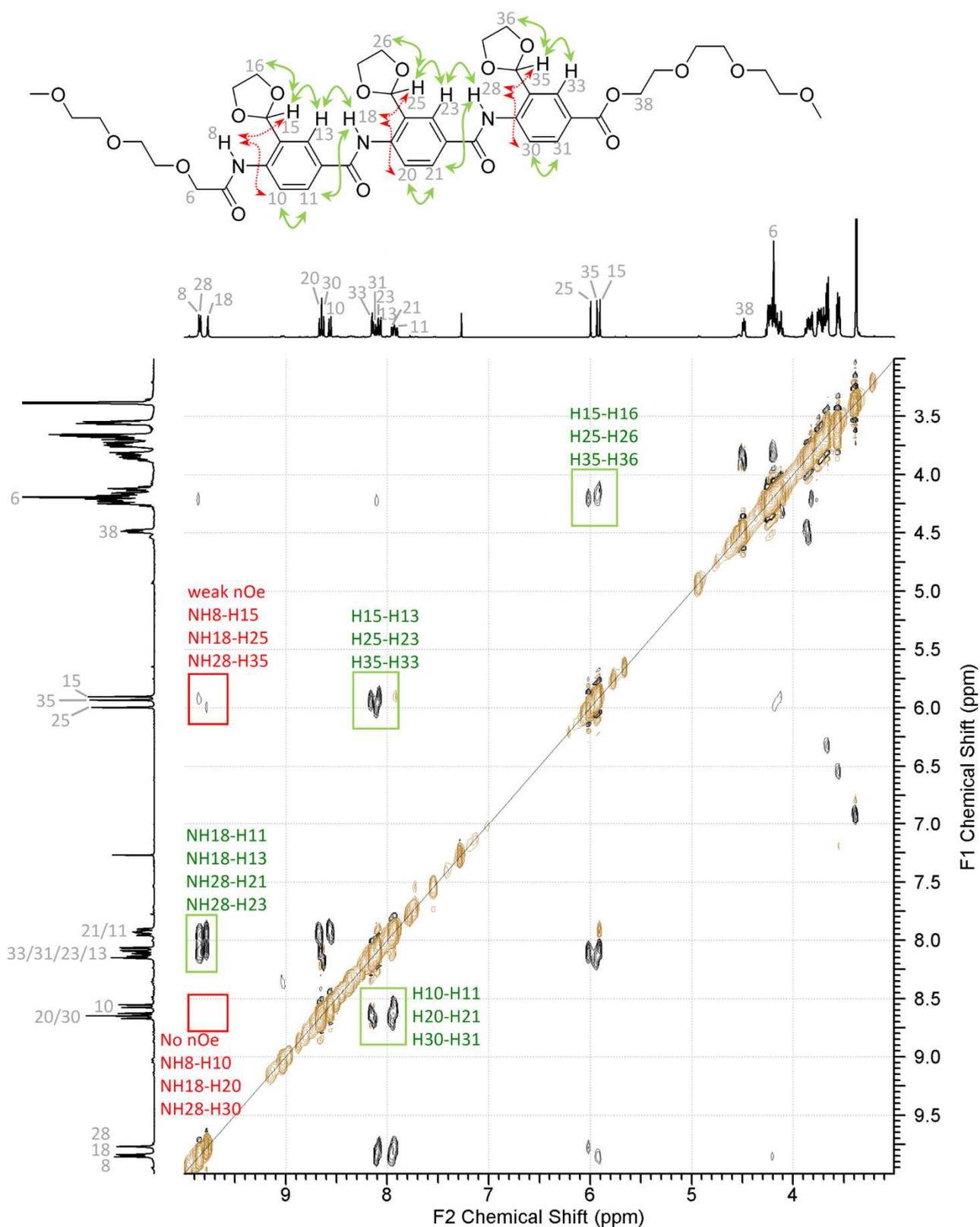
**Figure S120.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **1b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **1b**. The nOe cross peak between H15 and H13 is much stronger than between H15 and H8, suggesting that the amide NH is hydrogen bonded to both acetal oxygen atoms.



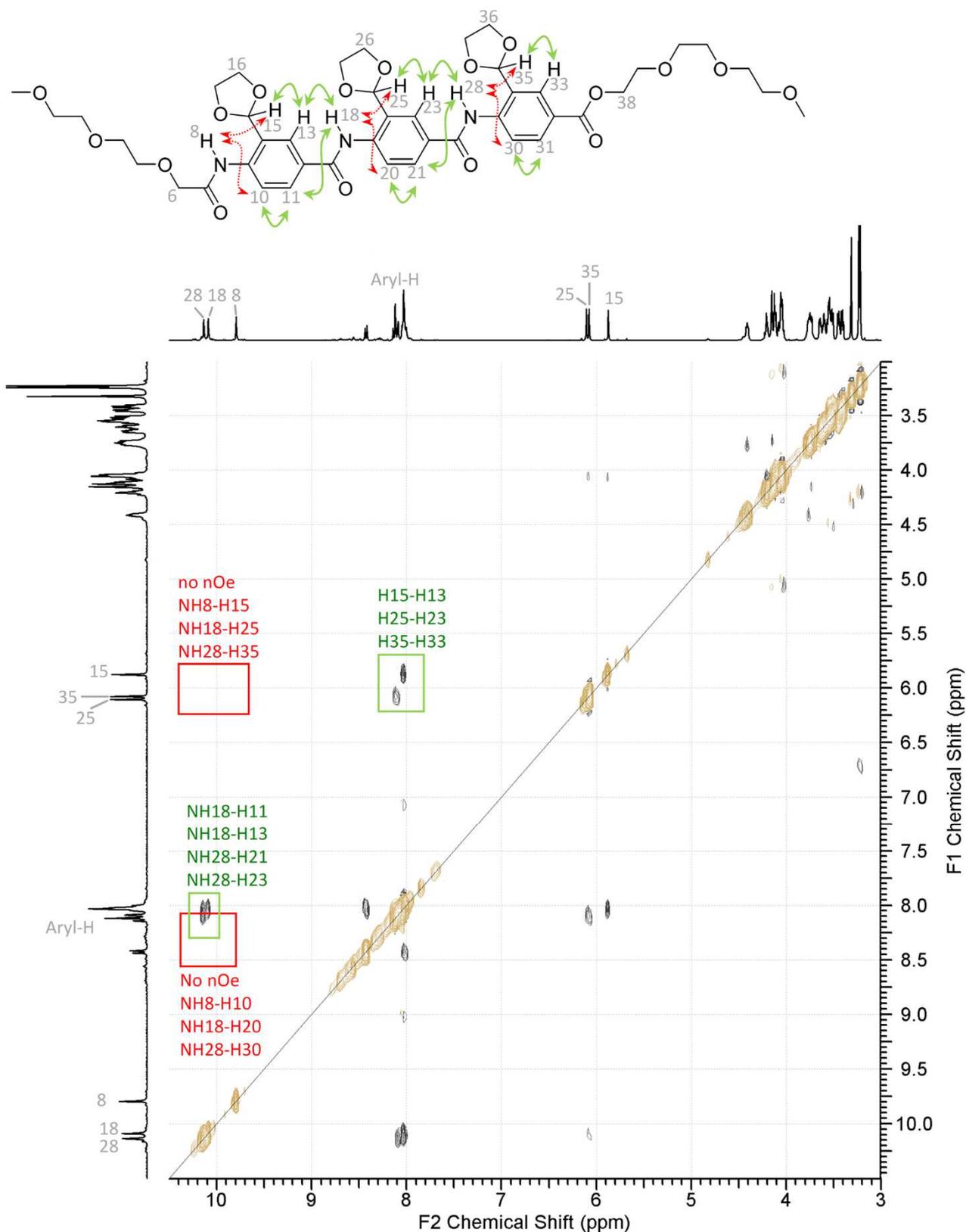
**Figure S121.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **2b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **2b**. The nOe cross peaks between H15-H13 or H25-H23 are much stronger than between H15-H8 or H25-H18, suggesting that the amide NHs are hydrogen bonded to both acetal oxygen atoms. Equal intensity cross peaks for H18-H11 and H18-H13 suggests rotation around the amide bond.



**Figure S122.**  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectrum (400, 400 MHz) of compound **2b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **2b**. Overlap of the aromatic signals does not allow the full assignment of the cross peaks. Cross peaks of H15 and H25 are observed with both H8 and H18, and H13 and H23, suggesting hydrogen bonding to only one of the oxygen atoms in the acetal group. Rotation around the amide bond is suggested by the appearance of cross peaks with the same phase as the diagonal due to rotational exchange (all cross peaks also appear in the equivalent ROESY spectrum).



**Figure S123.**  $^1\text{H}$ - $^1\text{H}$  ROESY NMR spectrum (400, 400 MHz) of compound **3b** in  $\text{CDCl}_3$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **3b**. The nOe cross peaks between H15-H13, H25-H23 or H35-H33 are much stronger than between H15-H8, H25-H18 or H35-H28, suggesting that the amide NHs are hydrogen bonded to both acetal oxygen atoms. Equal intensity cross peaks for H18-H11 and H18-H13, and for H28-H21 and H28-H23, suggests rotation around the amide bonds.



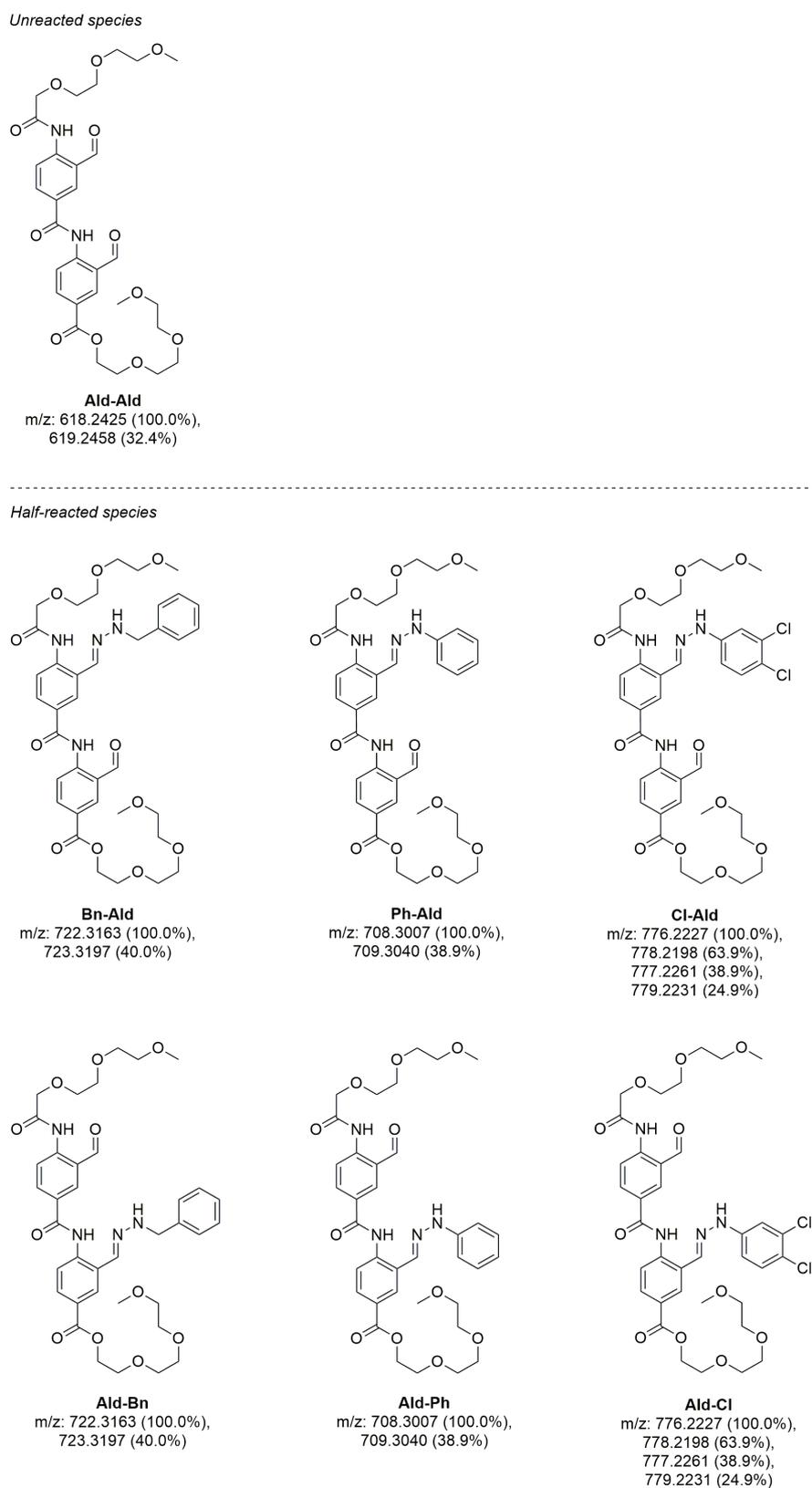
**Figure S124.**  $^1\text{H}$ - $^1\text{H}$  ROESY NMR spectrum (400, 400 MHz) of compound **3b** in  $\text{DMSO-}d_6$  at 298 K, with selected nOe cross peaks highlighted in green, and selected absent (or weak) nOe cross peaks highlighted in red, in both the spectrum and the chemical structure of **3b**. The peaks could not be fully assigned due to significant overlap and the assignment is therefore partly based on analogy with the monomer and dimer. The nOe cross peaks between H15-H13, H25-H23 or H35-H33 are much stronger than between H15-H8, H25-H18 or H35-H28, suggesting that the amide NHs are hydrogen bonded to both acetal oxygen atoms. Equal intensity cross peaks for H18-H11 and H18-H13, and for H28-H21 and H28-H23, suggests rotation around the amide bonds.

## 4. Library synthesis

A library of hydrazones was created by stirring a mixture of dimeric aldehyde **5b** and benzyl-hydrazine, phenylhydrazine and 3,4-dichlorophenylhydrazine (2 equivalents in total) in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2) for 24 hours at room temperature. The resulting mixture was then analysed using reversed-phase HPLC (Thermo Scientific Dionex Ultimate 3000) and the peaks were further analysed using ESI mass spectrometry (Agilent 1100 series LC/MSD Trap XCT). The HPLC solvents were 'A: 90% water, 10% acetonitrile, 0.1% HCOOH' and 'B: 5% water, 95% acetonitrile, 0.1% HCOOH'. An 80 min HPLC run was used, where the first 5 min consisted of 25% B, followed by a linear gradient to 100% B over 60 min, and finally flushing 10 min at 100% B and 10 min at 100% A. HPLC traces were analysed using Origin 9.1.0. An overview of all of the potential hydrazones (half-reacted, homodimers and heterodimers) that could be formed in the library and their respective molecular weight are shown in Figures S125-S126. Some of the half-reacted species and heterodimers constitute structural isomers with the same exact mass and these species can therefore not be differentiated using mass spectrometry. When two separate HPLC peaks are found with identical corresponding ESI+  $m/z$  values, they will correspond to the two structural isomers.

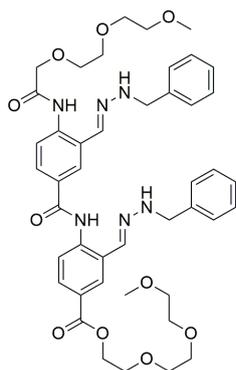
Initially, the starting materials were run to see where they appear during a HPLC run. The starting materials were run both fresh (directly injected into the HPLC) or after stirring for 24 hours in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2) at room temperature to check for degradation. The HPLC traces are shown in Figures S127-S130 and show that some of the hydrazines degrade under the experimental conditions. This might lower the yield of the library synthesis but should not be a problem if the hydrazone formation is quicker than hydrazine degradation. In a second step, the hydrazone formation was executed with each of the hydrazine individually (e.g. stirring a mixture of **5b** and 2 equivalents benzylhydrazine in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2) for 24 hours at room temperature), in order to easily identify the peaks of the homodimers and the half-reacted species. The HPLC traces are shown in Figures S131-S133 and all of the expected hydrazones could be identified by mass spectrometry (ESI+). Some hydrazine degradation products were also observed. The results of the library formed from **5b** and three different hydrazines (2 equivalents in total) is shown in Figure S134. The library formation was repeated 3 times and similar HPLC traces were obtained each time. All expected heterodimers were observed, as well as homodimers and half-reacted products (as identified by mass spectrometry). The different hydrazones are expected to have different extinction coefficients and the intensity of the peaks can therefore not be used as an indication of the concentration in the library. To get an estimate of the relative concentrations, preliminary UV-Vis absorbance spectra were obtained for all of the peaks

(Figure S135), which indicated that most compounds have a maximum absorbance around 290 nm. This wavelength was therefore chosen to monitor the HPLC separation process and this indicated that all of the formed dimeric hydrazones are present at roughly equal (statistical) concentrations.

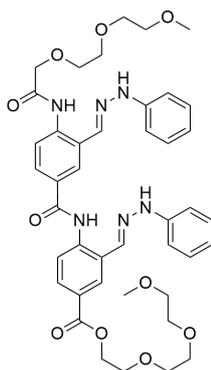


**Figure S125.** Structures of the unreacted species (**5b**) and the half-reacted species that are formed during hydrazone library synthesis. Exact mass (m/z values) are also given, as well as the name of the compounds as referred to in the HPLC traces (in bold). The half-reacted species form 3 pairs of structural isomers.

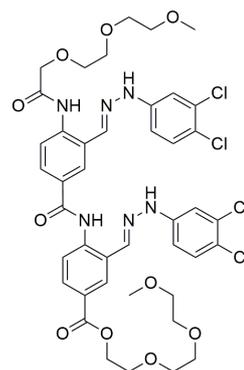
Homodimers



**Bn-Bn**  
m/z: 826.3901 (100.0%),  
827.3935 (47.6%)

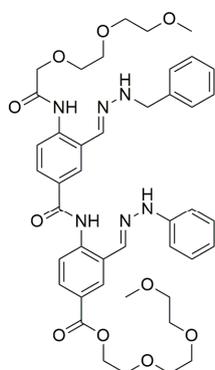


**Ph-Ph**  
m/z: 798.3588 (100.0%),  
799.3622 (45.4%)

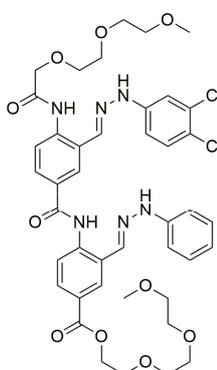


**Cl-Cl**  
m/z: 936.2000 (100.0%),  
934.2030 (78.2%), 938.1971 (47.9%),  
937.2034 (45.4%), 935.2063 (35.5%),  
939.2004 (21.8%)

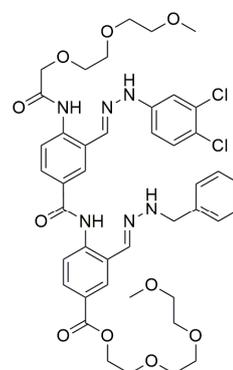
Heterodimers



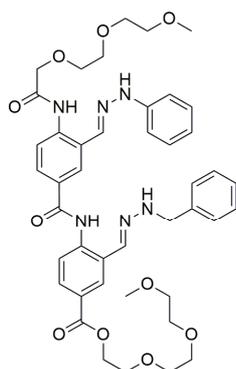
**Bn-Ph**  
m/z: 812.3745 (100.0%),  
813.3778 (46.5%)



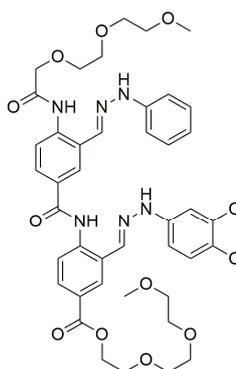
**Cl-Ph**  
m/z: 866.2809 (100.0%),  
868.2779 (63.9%), 867.2843 (45.4%),  
869.2813 (29.0%)



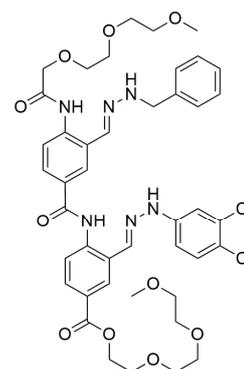
**Cl-Bn**  
m/z: 880.2965 (100.0%),  
882.2936 (63.9%), 881.2999 (46.5%),  
883.2970 (29.7%)



**Ph-Bn**  
m/z: 812.3745 (100.0%),  
813.3778 (46.5%)

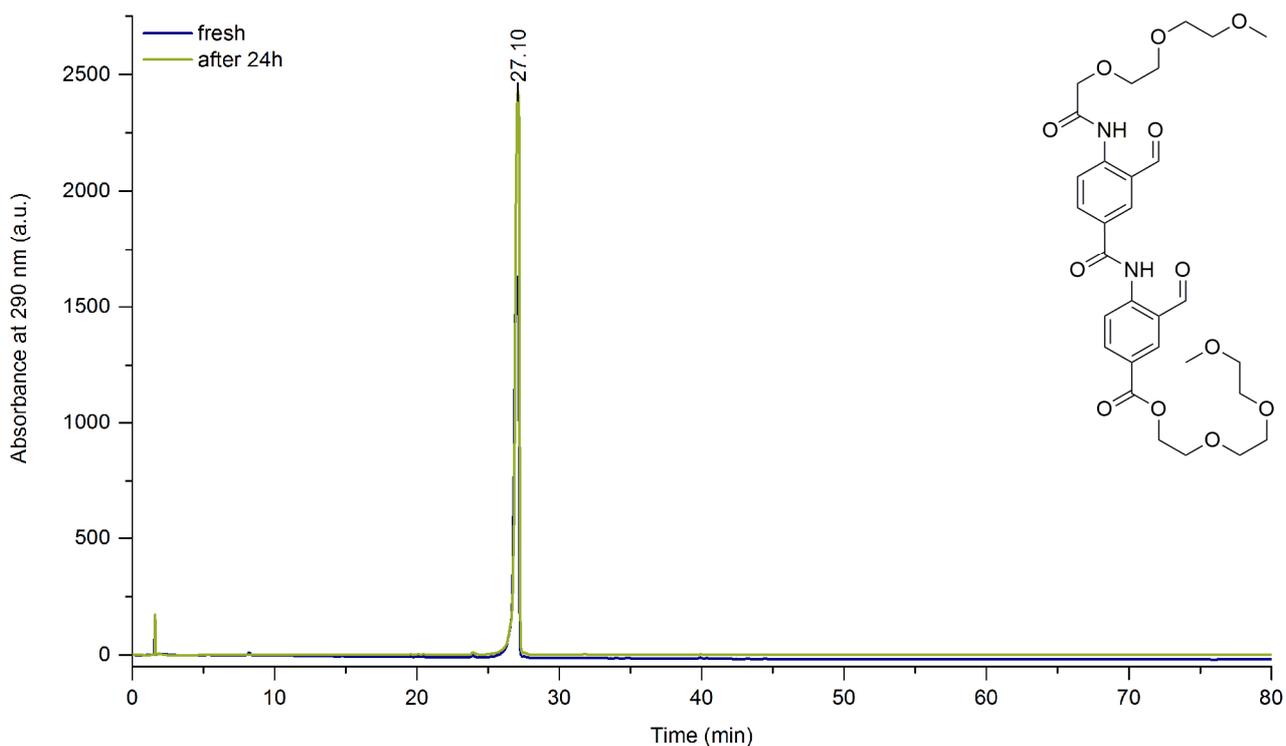


**Ph-Cl**  
m/z: 866.2809 (100.0%),  
868.2779 (63.9%), 867.2843 (45.4%),  
869.2813 (29.0%)

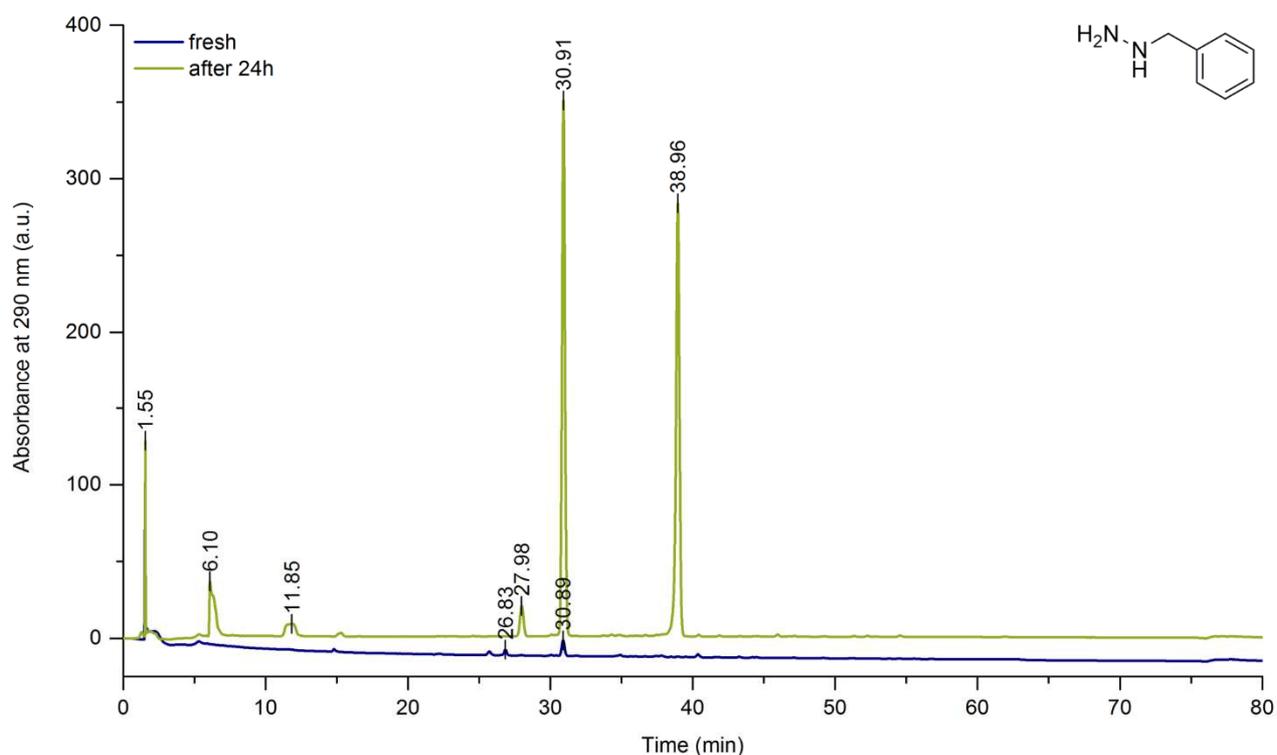


**Bn-Cl**  
m/z: 880.2965 (100.0%),  
882.2936 (63.9%), 881.2999 (46.5%),  
883.2970 (29.7%)

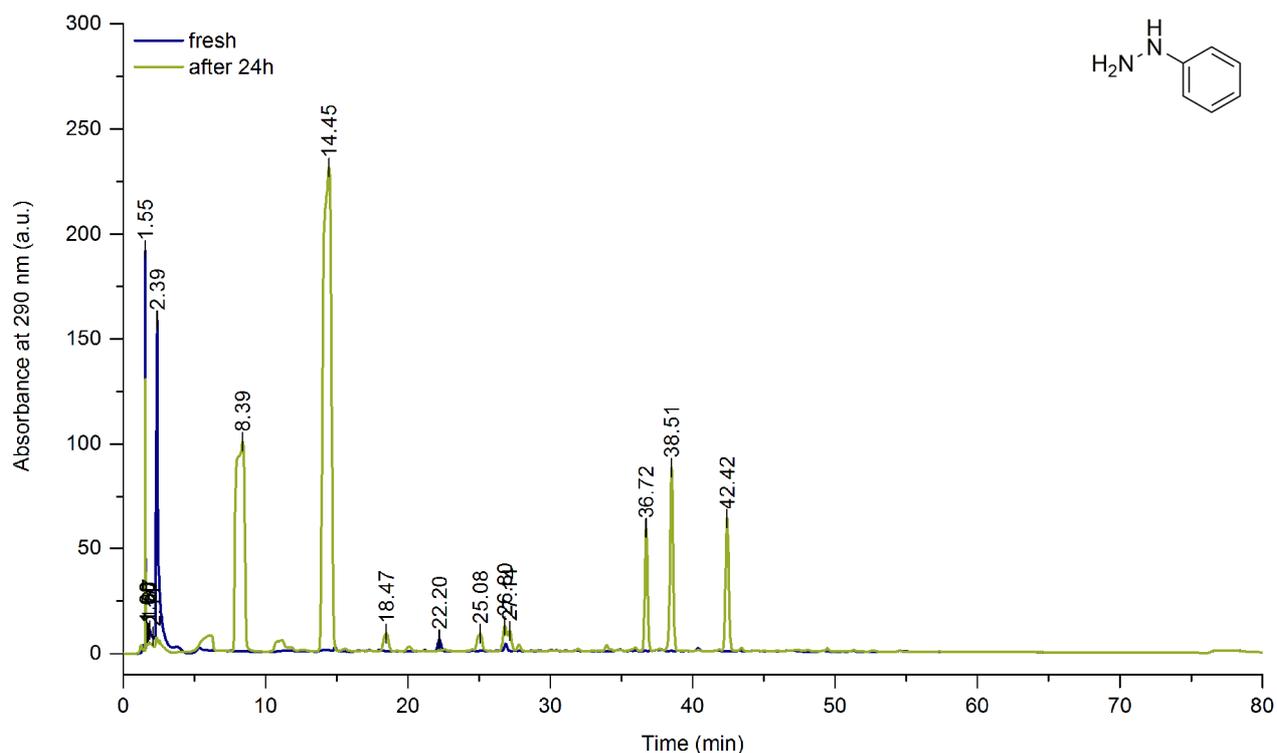
**Figure S126.** Structures of the homodimers and the heterodimers that are formed during hydrazone library synthesis. Exact mass (m/z values) are also given, as well as the name of the compounds as referred to in the HPLC traces (in bold). The heterodimers form 3 pairs of structural isomers.



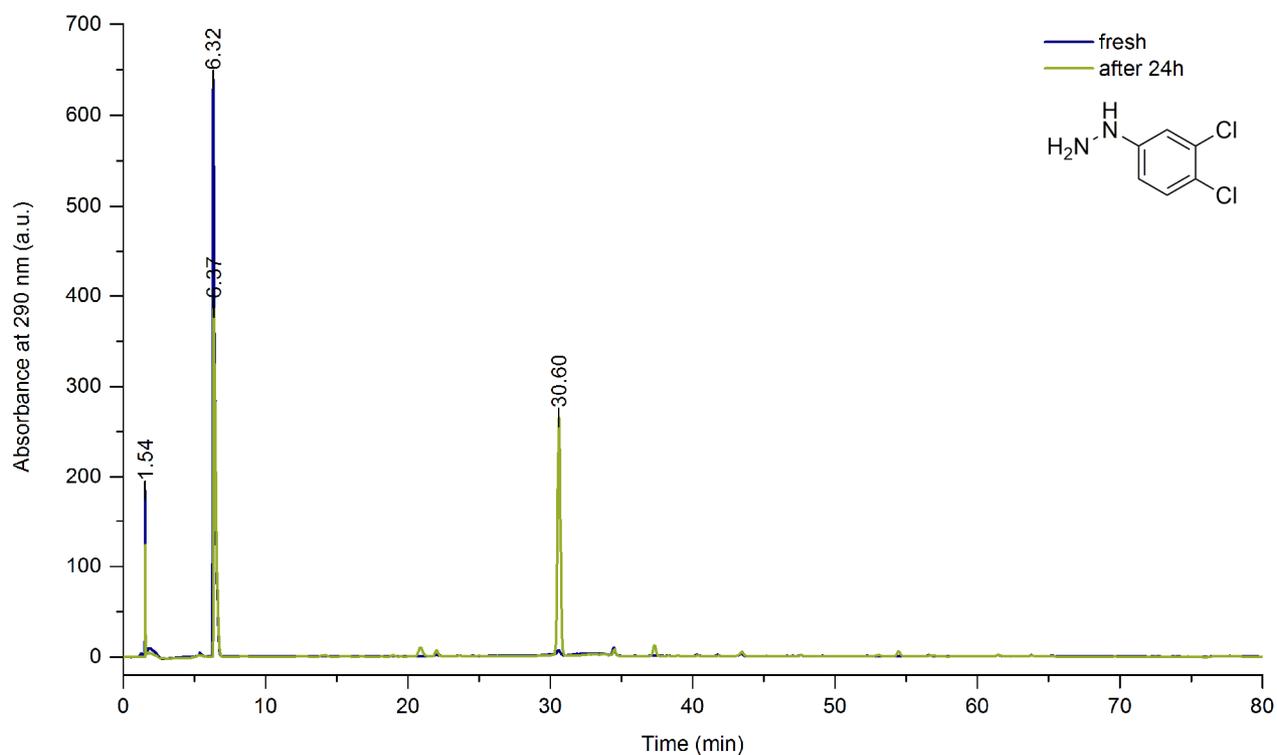
**Figure S127.** HPLC trace of unreacted dimer **5b**, both fresh and after stirring a 0.5 mM solution for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. The aldehyde appears stable in solution



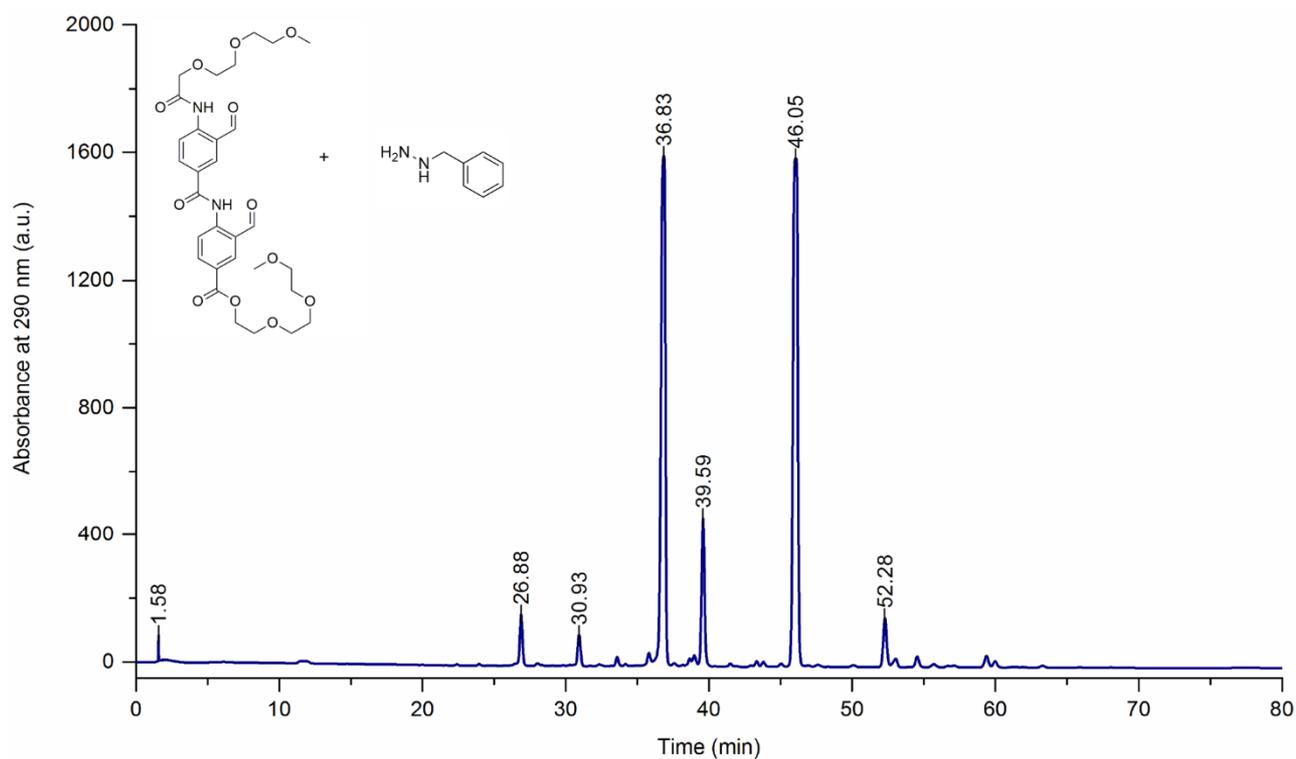
**Figure S128.** HPLC trace of unreacted benzylhydrazine, both fresh and after stirring a 1 mM solution for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. Degradation of the hydrazine is observed in the buffer.



**Figure S129.** HPLC trace of unreacted phenylhydrazine, both fresh and after stirring a 1 mM solution for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. Degradation of the hydrazine is observed in the buffer.

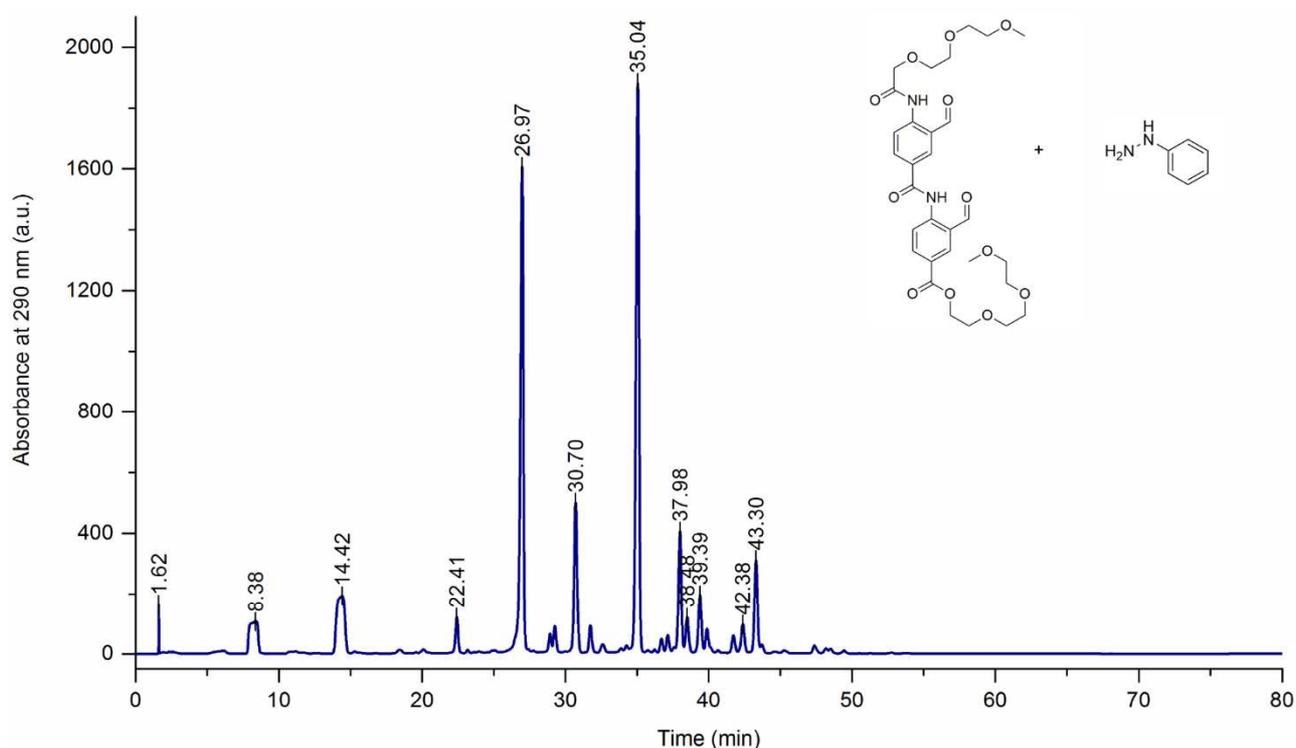


**Figure S130.** HPLC trace of unreacted 3,4-dichlorophenylhydrazine, both fresh and after stirring a 1 mM solution for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. Degradation of the hydrazine is observed in the buffer.



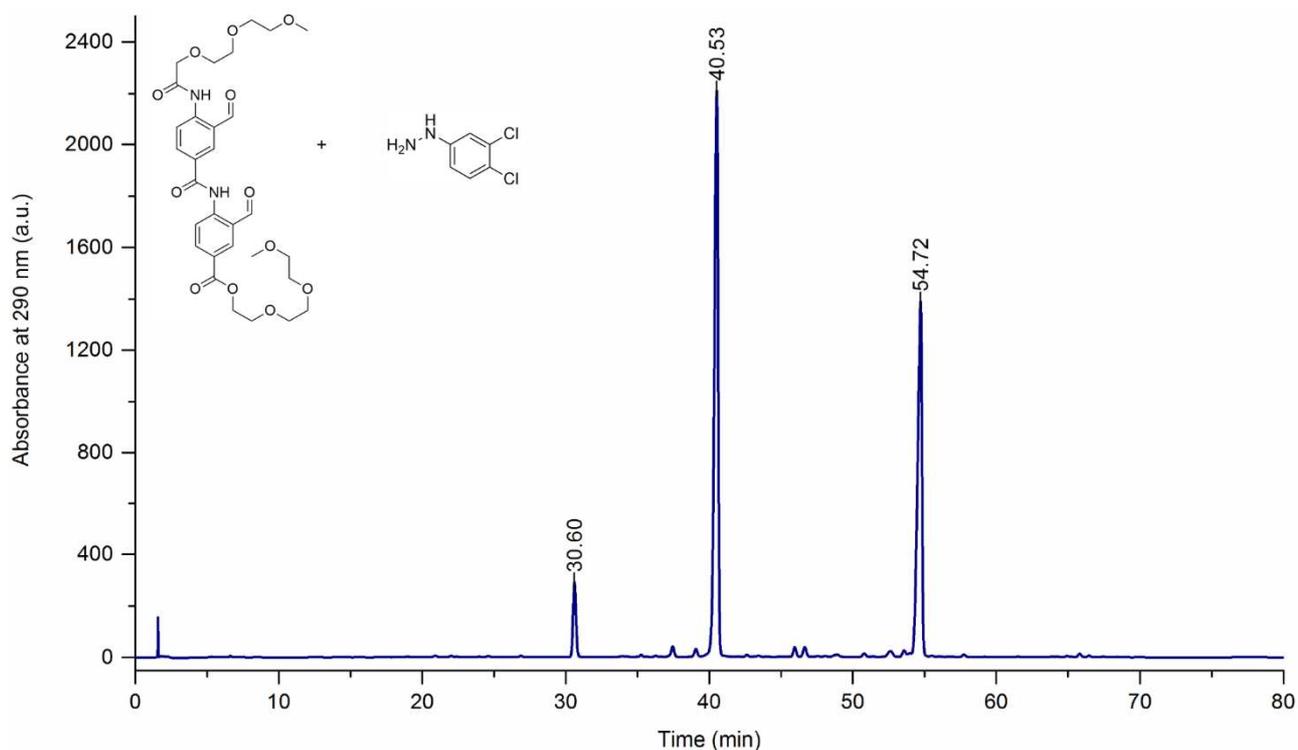
Index	Retention Time (min)	Peak Height (a.u.)	Peak Area (a.u.)	Relative Peak Area (%)	m/z	Compound
1	1.58	103.09	346.90	16.61	/	?
2	26.88	167.64	72.41	3.47	641.6 [M+Na] <sup>+</sup>	<b>Ald-Ald</b>
3	30.93	103.99	45.50	2.18	/	degradation
4	36.83	1607.84	634.57	30.39	745.7 [M+Na] <sup>+</sup>	<b>Bn-Ald or Ald-Bn</b>
5	39.59	472.98	160.69	7.70	745.6 [M+Na] <sup>+</sup>	<b>Bn-Ald or Ald-Bn</b>
6	46.05	1600.53	670.55	32.11	849.8 [M+Na] <sup>+</sup>	<b>Bn-Bn</b>
7	52.28	156.28	134.51	6.44	/	?

**Figure S131.** HPLC trace collected after stirring 0.5 mM **5b** and 1 mM benzylhydrazine for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. The table shows the ESI<sup>+</sup> mass spectrometry results obtained for the different peaks and the identification of the peaks derived from the m/z values. All expected starting materials and hydrazones could be observed, as well as some degradation compounds of the hydrazine.



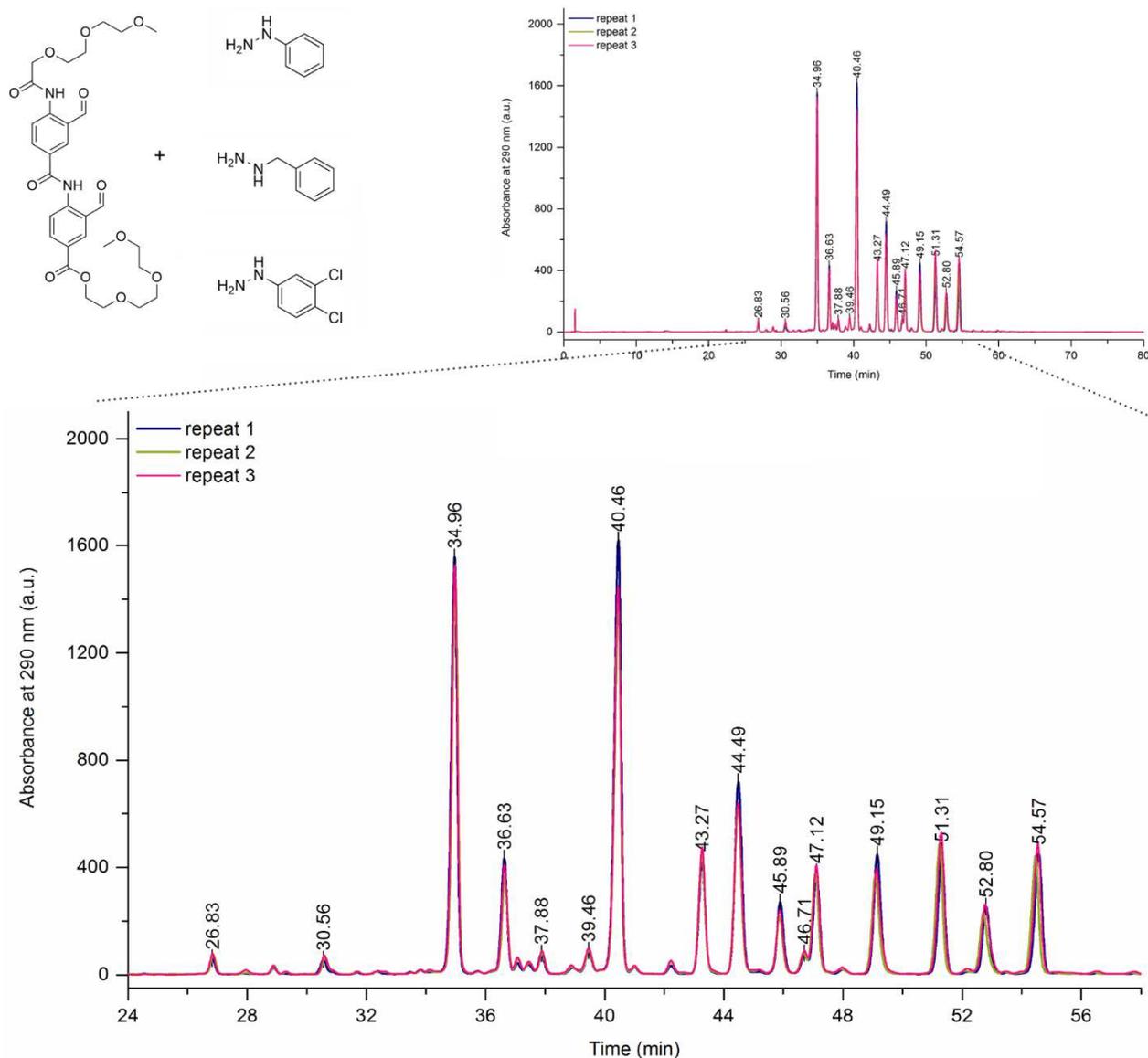
Index	Retention Time (min)	Peak Height (a.u.)	Peak Area (a.u.)	Relative Peak Area (%)	m/z	Compound
1	1.62	165.45	11.01	0.59	/	?
2	8.38	106.82	87.55	4.66	/	degradation
3	14.42	194.02	142.89	7.60	/	degradation
4	22.41	121.04	49.46	2.63	/	degradation
5	26.97	1606.89	474.09	25.22	641.4 [M+Na] <sup>+</sup>	<b>Ald-Ald</b>
6	30.70	501.64	130.40	6.94	725.5, 747.5	?
7	35.04	1882.06	543.26	28.90	731.7 [M+Na] <sup>+</sup>	<b>Ald-Ph or Ph-Ald</b>
8	37.98	406.05	131.54	7.00	731.6 [M+Na] <sup>+</sup>	<b>Ald-Ph or Ph-Ald</b>
9	38.48	121.28	33.31	1.77	/	degradation
10	39.39	196.66	81.24	4.32	785.6, 807.7	?
11	42.38	97.57	45.64	2.43	/	degradation
12	43.30	311.98	143.91	7.65	821.7 [M+Na] <sup>+</sup>	<b>Ph-Ph</b>

**Figure S132.** HPLC trace collected after stirring 0.5 mM **5b** and 1 mM phenylhydrazine for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. The table shows the ESI<sup>+</sup> mass spectrometry results obtained for the different peaks and the identification of the peaks derived from the m/z values. All expected starting materials and hydrazones could be observed, as well as some degradation compounds of the hydrazine.



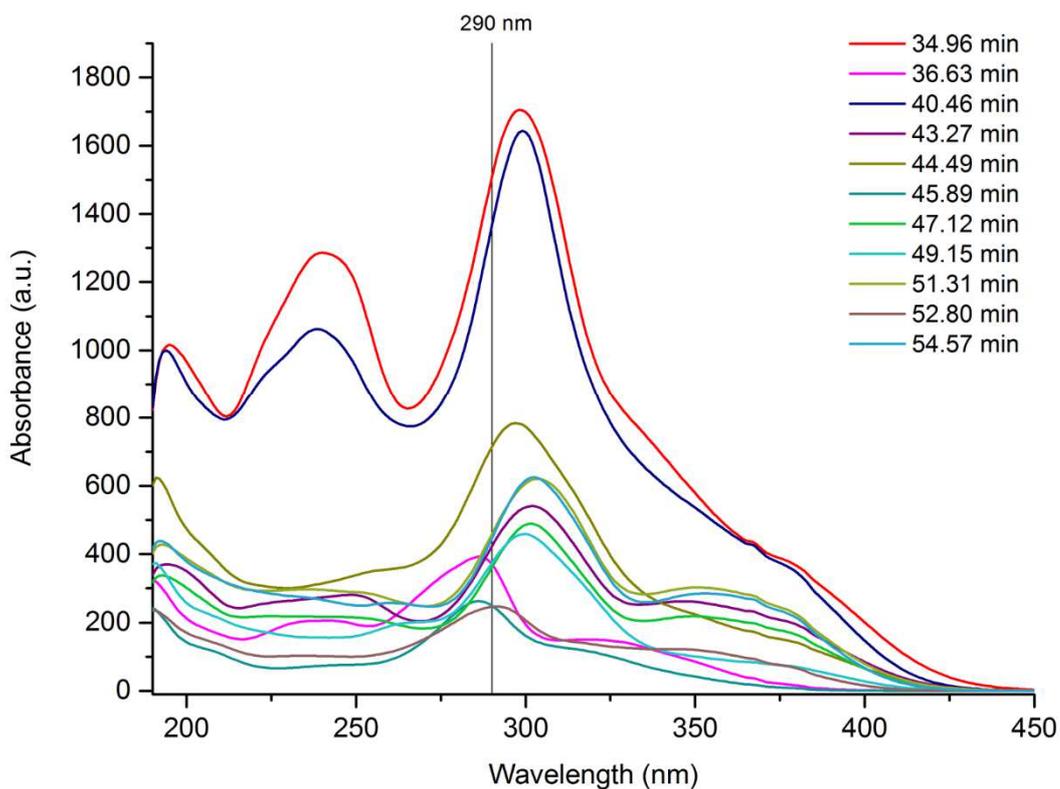
Index	Retention Time (min)	Peak Height (a.u.)	Peak Area (a.u.)	Relative Peak Area (%)	m/z	Compound
1	30.60	292.39	139.56	9.01	/	?
2	40.53	2211.88	797.72	51.52	799.6 [M+Na] <sup>+</sup>	Ald-Cl or Cl-Ald
3	54.72	1391.97	593.09	38.30	959.4 [M+Na] <sup>+</sup>	Cl-Cl

**Figure S133.** HPLC trace collected after stirring 0.5 mM **5b** and 1 mM 3,4-dichlorophenylhydrazine for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. The table shows the ESI<sup>+</sup> mass spectrometry results obtained for the different peaks and the identification of the peaks derived from the m/z values. All expected starting materials and hydrazones could be observed, as well as some degradation compounds of the hydrazine.



Index	Retention Time (min)	Peak Height (a.u.)	Peak Area (a.u.)	Relative Peak Area (%)	m/z	Compound
1	26.83	60.50	24.34	1.16	641.6 [M+Na] <sup>+</sup>	Ald-Ald
2	30.56	55.02	16.44	0.79	/	degradation
3	34.96	1554.53	394.80	18.86	731.6 [M+Na] <sup>+</sup>	Ph-Ald or Ald-Ph
4	36.63	431.66	130.51	6.24	745.6 [M+Na] <sup>+</sup>	Bn-Ald or Ald-Bn
5	37.88	78.23	19.66	0.94	731.6 [M+Na] <sup>+</sup>	Ph-Ald or Ald-Ph
6	39.46	88.67	34.52	1.65	745.6 [M+Na] <sup>+</sup>	Bn-Ald or Ald-Bn
7	40.46	1617.42	450.47	21.52	801.4 [M+Na] <sup>+</sup>	Cl-Ald or Ald-Cl
8	43.27	424.94	120.79	5.77	821.7 [M+Na] <sup>+</sup>	Ph-Ph
9	44.49	716.75	210.25	10.05	835.7 [M+Na] <sup>+</sup>	Ph-Bn or Bn-Ph
10	45.89	269.36	73.12	3.49	849.7 [M+Na] <sup>+</sup>	Bn-Bn
11	46.71	79.85	20.41	0.98	801.5 [M+Na] <sup>+</sup>	Cl-Ald or Ald-Cl
12	47.12	380.23	110.66	5.29	891.4 [M+Na] <sup>+</sup>	Ph-Cl or Cl-Ph
13	49.15	446.42	131.50	6.28	905.4 [M+Na] <sup>+</sup>	Bn-Cl or Cl-Bn
14	51.31	459.41	127.94	6.11	891.4 [M+Na] <sup>+</sup>	Ph-Cl or Cl-Ph
15	52.80	252.11	82.54	3.94	905.5 [M+Na] <sup>+</sup>	Bn-Cl or Cl-Bn
16	54.57	449.82	141.66	6.77	959.5 [M+Na] <sup>+</sup>	Cl-Cl

**Figure S134.** HPLC trace collected after stirring 0.5 mM **5b** and 0.33 mM benzylhydrazine, 0.33 mM phenylhydrazine and 0.33 mM 3,4-dichlorophenylhydrazine for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). Absorbance was measured at 290 nm. The table shows the ESI<sup>+</sup> mass spectrometry results obtained for the different peaks and the identification of the peaks derived from the m/z values. All expected starting materials and hydrazones could be observed, as well as some degradation compounds of the hydrazines.



**Figure S135.** UV-Vis spectra of the major peaks in the HPLC trace collected after stirring 0.5 mM **5b** and 0.33 mM benzylhydrazine, 0.33 mM phenylhydrazine and 0.33 mM 3,4-dichlorophenylhydrazine for 24 hours at room temperature in 75:25 DMSO:buffer (10 mM acetate buffer, pH 4.2). The absorbance of 290 nm was chosen as all of the compounds display a local maximum in this wavelength region. Many peaks have a maximum absorbance near 300 nm, but other peaks (such as the peak at 36.63 min) have a local minimum at 300 nm and therefore 290 nm was considered a better wavelength to compare the various peaks in the HPLC traces.

## 5. References

1. S. Ikeda, M. Shibuya and Y. Iwabuchi, *Chem. Commun.*, 2007, 504-506.
2. APEX2 (version 2014.12-0), *Program for Bruker CCD X-ray Diffractometer Control*, Bruker AXS Inc., Madison, WI, 2014.
3. G. Sheldrick, *Acta Crystallogr., Sect. A*, 2008, **64**, 112-122.
4. L. J. Worrall and M. D. Walkinshaw, *Biochem. Biophys. Res. Commun.*, 2007, **357**, 105-110.
5. M. H. Abraham, R. J. Abraham, W. E. Acree, A. E. Aliev, A. J. Leo and W. L. Whaley, *J. Org. Chem.*, 2014, **79**, 11075-11083.
6. N. H. Andersen, J. W. Neidigh, S. M. Harris, G. M. Lee, Z. Liu and H. Tong, *J. Am. Chem. Soc.*, 1997, **119**, 8547-8561.
7. S. H. Gellman, G. P. Dado, G. B. Liang and B. R. Adams, *J. Am. Chem. Soc.*, 1991, **113**, 1164-1173.