Characterization of aggregated morphologies derived from *mono*- and *bis*-arylbenzamides – potential alpha helix mimetics

Oleg V. Kulikov,* Yulia V. Sevryugina,[a] Arshad Mehmood,[a] Ishu Saraogi[b]

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA; [a]Department of Chemistry, Texas Christian University, Fort Worth, TX, 76129, USA; [b]Department of Chemistry, Indian Institute of Science Education and Research Bhopal, Bhopal 462066, Madhya Pradesh, India;
**Corresponding author. Tel.:+1* (215) 470-3581*; E-mail: oleg.kulikov.chem@gmail.com, okulikov@mit.edu*

Supplementary Data, vol. 1:

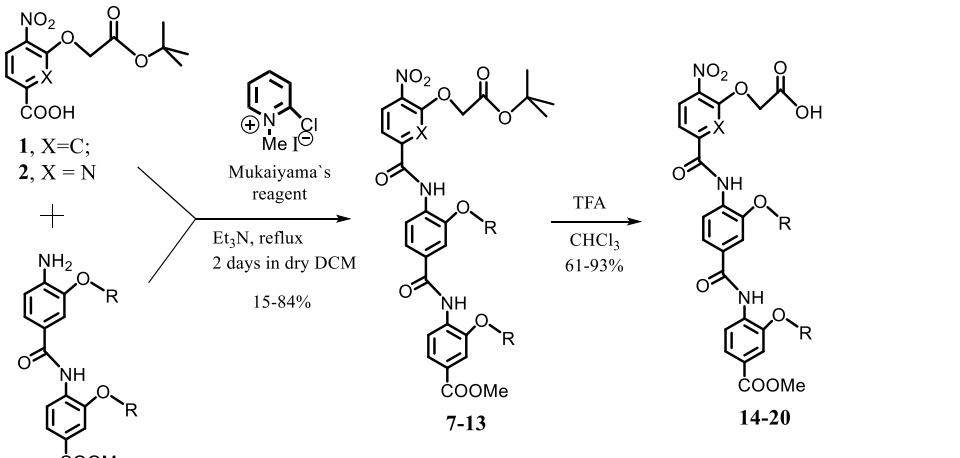
General Methods: All starting materials used were obtained from Sigma-Aldrich, Fluka or TCI America and were used without further purification unless otherwise noted. Thin layer chromatography (TLC) was performed on Sigma-Aldrich TLC Plates (silica gel on aluminum, 200 mm layer thickness, 2-25 mm particle size, 60 Å pore size). The synthesis of monomers **1**, **2** and NH_2-dimers **3-6** was reported earlier.[4,14] Column chromatography was performed using silica gel (230-400 mesh) from Solvent technologies. ^{1}H nuclear magnetic resonance spectra were recorded at 400 MHz and 500 MHz on either Bruker 400 Ultra Shield or DPX-500 spectrometers at room temperature; ^{13}C NMR spectra were recorded on the same instruments at 100 and 125 MHz, correspondingly. ^{1}H and ^{13}C chemical shifts are reported in parts per million relative to the corresponding residual solvent peak. High-resolution Fourier transform mass spectra (FT-ICR MS) were obtained from the W. M. Keck Biotechnology Resource Laboratory at Yale University, MALDI-TOF MS spectra were acquired by using Voyager-DE PRO instrument. Secondary electron images (SEI) were collected with the SEM system on the JEOL JXA-8530F (field-emission-gun, FEG) electron microprobe in the Yale University Department of Geology and Geophysics. The samples of **7**, **9** and **11** were mounted on double-stick carbon tape, then were coated with conductive carbon or gold to minimize charge buildup and heating of the sample from the electron beam, a requisite for both imaging and compositional analysis under high vacuum. All images were collected at an accelerating voltage of 15 kV and a beam current between 5 nA and 500 pA.

X-ray crystal structure of **6** was obtained on Rigaku Mercury2 CCD area detector with filtered Mo-Ka radiation. The structure was solved by direct methods and expanded using Fourier techniques. The non-hydrogen atoms were refined anisotropically. Hydrogen atoms were refined using the riding model. All calculations were performed using the CrystalStructure[S1] crystallographic software package except for refinement, which was performed using SHELXL-97.[S2] X-ray crystal structures of **21** and **22** were obtained on Bruker PHOTON 100 CMOS instrument equipped with the "fine-focus sealed tube" using Mo-Ka radiation. The frames for compounds **21** and **22** were integrated with either SHELXL-2014/6 or SHELXL-2014/7 software packages. X-ray crystal structure of 5-amino-6-*tert*-butoxycarbonylmethoxypyridine-2-carboxylic acid methyl ester was acquired on Nonius Kappa CCD diffractometer with graphite monochromated Mo-Kα radiation. The data frames were processed and scaled using the DENZO software package.[S3] The structure was solved by direct methods and expanded using Fourier techniques.[S4] Crystallographic data for **6**, **21**(**ys091**), **22** (**ys092**), and 5-amino-6-*tert*-butoxycarbonylmethoxypyridine-2-carboxylic acid methyl ester (**S1**) have been deposited with the Cambridge Crystallographic Data Centre. Compound **6**: CCDC 852853, $C_{17}H_{18}N_2O_5$, M = 330.34, monoclinic, a = 4.6091(5), b = 14.1634(17), c = 24.427(3) Å, β = 97.974(3)°, U = 1579.2(3) Å^3, μ = 1.033 mm^{-1}, T = -50±1°C, space group P2$_1$/c (#14), Z = 4, GOF = 1.051, final R-indices (R_1 = 0.0688, ωR_2=0.1614), 3557 reflections collected, crystal size 0.35 x 0.15 x 0.15 mm^3; compound **21**: CCDC 1435026, $2(C_{67}H_{78}N_4O_{13}) \cdot 2(C_2H_6O_1) \cdot H_2O_1$, M = 1202.40, triclinic, a = 10.0873(13), b = 16.344(2), c = 41.016(5) Å, α = 80.339(4)°, β = 83.426(4)°, γ = 72.997(4)°, U = 6359.4(15) Å^3, μ = 0.088 mm^{-1}, T = 100(2) K, space group P-1, Z = 4, GOF = 1.073, final R-indices (R_1 = 0.1055, ωR_2=0.2567), 15277 reflections collected, crystal size

0.218 x 0.297 x 0.322 mm^3; compound **22** (**ys092**), DMF: CCDC 1435025, $C_{71}H_{86}N_4O_{13}$, M = 1203.43, monoclinic, a = 8.4076(3), b = 37.3169(17), c = 21.4037(10) Å, β = 95.689(1)°, U = 6682.2(5) $Å^3$, μ = 0.082 mm^{-1}, T = 100(2) K, space group $P\ 2_1/n$, Z = 4, GOF = 1.061, final R-indices (R_1 = 0.0514, ωR_2=0.1257), 7450 reflections collected, crystal size 0.156 x 0.247 x 0.324 mm^3; **5-amino-6-*tert*-butoxycarbonyl-methoxypyridine-2-carboxylic acid methyl ester** (**S1**): CCDC 1438119, $C_{13}H_{18}N_2O_5$, M = 282.29, triclinic, a = 7.7375(15), b = 8.0667(16), c = 12.449(3) Å, β = 88.96(3)°, U = 701.9(2) $Å^3$, μ = 0.103 mm^{-1}, T = 173(2) K, space group P-1, Z = 2, GOF = 1.032, crystal size 0.20 x 0.10 x 0.10 mm^3.

Synthesis and characterization of *mono*-arylamides 14-20

The preparation of the trimer molecules, **14-20,** having the only one CH_2COOH-functionality protected by *t*-But group was conducted by using Mukaiyama`s reagent (2-chloro-1-methylpyridinium iodide) induced amide coupling of known precursors **1, 2** with NH_2-dimers **3-6** and then TFA-catalyzed deprotection of the formed trimers **7-13** (Scheme 1, Figures S69-S82).



Scheme S1. Synthesis of arylamide trimers

General protocol for an amide coupling in the presence of Mukaiyama reagent (2-chloro-1-methyl-pyridinium iodide):

The aryl carboxylic acid monomer (6.0 mmol, **1** or **2**), Mukaiyama reagent (6.0 mmol) and Et_3N (12 mmol) were dissolved in an anhydrous dichloromethane (100 mL). The solution was refluxed for 15 min. under N_2. Then a solution of the corresponding arylamine (5.0 mmol, **3-6**) in an anhydrous dichloromethane (20 mL) was added to the reaction mixture and resulting solution was refluxed for 2 days. The solvent was evaporated under vacuum and the residue was subjected to the column of silica gel using the mixture of hexanes/ dichloromethane (2:3) as an eluent to give amide coupling products in the moderate to good yields.

Compound 7: yield 84%; ^{1}H-NMR (500 Hz, $CDCl_3$): δ 1.38 (d, J = 6.2 Hz, 6H, $Me_{isoprop}$), 1.40 (d, J = 5.7 Hz, 6H, $Me_{isoprop}$), 1.42 (s, 9H, *t*-Butyl), 3.84 (s, 3H, COOMe), 4.71-4.73 (m, 3H, $OC\underline{H}_2CO$+$OC\underline{H}Me_2$), 4.75-4.78

(m, 1H, OC<u>H</u>Me$_2$), 7.33-8.57 (m, 9H, ArH), 8.70 (s, 1H, NHCO), 8.79 (s, 1H, NHCO). ^{13}C NMR (125 MHz, $CDCl_3$): δ 21.2, 27.0, 51.1, 65.6, 70.8, 71.0, 82.4, 110.8, 112.1, 113.8, 117.3, 117.6, 117.7, 118.2, 122.3, 124.1, 125.2, 129.6, 130.2, 131.9, 138.6, 141.0, 144.7, 145.7, 150.7, 161.7, 163.3, 165.2, 165.8. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{34}H_{39}N_3O_{11}$ (m/z = 666.2657). Observed m/z = 666.2663. MALDI-TOF, M = $C_{34}H_{39}N_3O_{11}$ ($[M+H]^+$, m/z = 666.5). Rf ~ 0.6 (DCM).

Compound 8: yield 37%; ^{1}H-NMR (500 Hz, $CDCl_3$): δ 1.03 (d, *J* = 6.6 Hz, 6H, Me$_{isobut}$), 1.05 (d, *J* = 6.7 Hz, 6H, Me$_{isobut}$), 1.42 (s, 9H, *t*-Butyl), 2.11-2.20 (m, 2H, OCH$_2$C<u>H</u>Me$_2$), 3.84 (s, 3H, COOMe), 3.86 (d, *J* = 7.0 Hz, 2H, OC<u>H$_2$</u>CHMe$_2$), 3.90 (d, *J* = 6.6 Hz, 2H, OC<u>H$_2$</u>CHMe$_2$), 4.70 (s, 2H, OC<u>H$_2$</u>CO), 7.35-8.56 (m, 9H, ArH), 8.70 (s, 1H, NHCO), 8.77 (s, 1H, NHCO). ^{13}C NMR (125 MHz, $CDCl_3$): δ 19.3, 19.4, 28.0, 28.2, 28.3, 52.1, 66.5, 75.1, 75.2, 83.4, 110.4, 111.5, 114.7, 118.1, 118.3, 118.9, 119.0, 123.4, 125.1, 126.2, 130.6, 132.1, 139.5, 142.0, 147.0, 147.8, 151.7, 162.7, 164.2, 166.1, 166.7. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{36}H_{43}N_3O_{11}$ (m/z = 694.2970). Observed m/z = 694.2969. MALDI-TOF, M = $C_{36}H_{43}N_3O_{11}$ ($[M + 2H]^{2+}$, m/z = 695.3). Rf ~ 0.7 (DCM).

Compound 9: yield 33%; ^{1}H-NMR (500 Hz, $CDCl_3$): δ 0.91-0.94 (m, 12H, Me), 1.42 (s, 9H, *t*-Butyl), 1.45-1.52 (m, 8H, CHC<u>H$_2$</u>Me), 1.70-1.77 (m, 2H, OCH$_2$C<u>H</u>Et$_2$), 3.85 (s, 3H, COOMe), 4.01 (d, *J* = 4.5 Hz, 2H, OC<u>H$_2$</u>CHEt$_2$), 4.04 (d, *J* = 6.8 Hz, 2H, OC<u>H$_2$</u>CHEt$_2$), 4.70 (s, 2H, OC<u>H$_2$</u>CO), 7.34-8.57 (m, 9H, ArH), 8.71 (s, 1H, NHCO), 8.77 (s, 1H, NHCO). ^{13}C NMR (125 MHz, $CDCl_3$): δ 10.2, 10.3, 22.8, 27.0, 39.9, 40.0, 51.1, 65.5, 69.7, 70.1, 82.4, 109.5, 110.4, 113.7, 117.0, 117.4, 117.8, 122.4, 124.1, 125.2, 129.6, 129.7, 131.2, 138.6, 141.0, 146.1, 147.0, 150.7, 161.7, 163.3, 165.1, 165.8. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{40}H_{51}N_3O_{11}$ (m/z = 750.3596). Observed m/z = 750.3596. MALDI-TOF, M = $C_{40}H_{51}N_3O_{11}$ ($[M]^+$, m/z = 749.5). Rf ~ 0.7 (DCM).

Compound 10: yield 44%; ^{1}H-NMR (400 Hz, $CDCl_3$): δ 1.34 (s, 9H, *t*-Butyl), 1.38 (d, *J* = 6.0 Hz, 6H, Me$_{isoprop}$), 1.45 (d, *J* = 6.0 Hz, 6H, Me$_{isoprop}$), 3.86 (s, 3H, COOMe), 4.68-4.74 (m, 1H, OC<u>H</u>Me$_2$), 4.77-4.83 (m, 1H, OC<u>H</u>Me$_2$), 5.96 (s, 2H, OC<u>H$_2$</u>CO), 7.33-8.67 (m, 8H, ArH), 8.81 (s, 1H, NHCO), 10.15 (s, 1H, NHCO). ^{13}C NMR (100 MHz, $CDCl_3$): δ 22.2, 22.3, 28.0, 52.2, 64.0, 71.7, 72.0, 83.3, 112.0, 113.1, 116.9, 118.7, 119.5, 123.4, 125.1, 130.9, 133.1, 136.1, 137.3, 145.9, 147.2, 150.3, 154.4, 159.7, 164.5, 165.9, 166.9. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{33}H_{38}N_4O_{11}$ (m/z = 667.2610). Observed m/z = 667.2616. MALDI-TOF, M = $C_{33}H_{38}N_4O_{11}$ ($[M+H]^+$, m/z = 667.3). Rf ~ 0.6 (DCM).

Compound 11: yield 42%; ^{1}H-NMR (500 Hz, $CDCl_3$): δ 1.05 (d, *J* = 7.4 Hz, 12H, Me$_{isobut}$), 1.33 (s, 9H, *t*-Butyl), 2.11-2.18 (m, 1H, OCH$_2$C<u>H</u>Me$_2$), 2.25-2.31 (m, 1H, OCH$_2$C<u>H</u>Me$_2$), 3.85 (s, 3H, COOMe), 3.87 (d, *J* = 6.9 Hz, 2H, OC<u>H$_2$</u>CHMe$_2$), 3.93 (d, *J* = 7.0 Hz, 2H, OC<u>H$_2$</u>CHMe$_2$), 5.00 (s, 2H, OC<u>H$_2$</u>CO), 7.34-8.64 (m, 8H, ArH), 8.80 (s, 1H, NHCO), 10.11 (s, 1H, NHCO). ^{13}C NMR (125 MHz, $CDCl_3$): δ 19.4, 27.9, 28.3, 52.1, 64.0, 75.1, 75.5, 83.1, 110.6, 111.5, 116.8, 118.4, 118.8, 119.4, 123.4, 125.2, 130.1, 130.9, 132.1, 136.1, 137.2, 147.0, 148.4, 150.1, 154.3, 159.8, 164.3, 165.9, 166.8. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{35}H_{42}N_4O_{11}$ (m/z = 695.2923). Observed m/z = 695.2905. MALDI-TOF, M = $C_{35}H_{42}N_4O_{11}$ ($[M+H]^+$, m/z = 695.4). Rf ~ 0.7 (DCM).

Compound 12: yield 15%; ^{1}H-NMR (500 Hz, $CDCl_3$): δ 0.89-0.94 (m, 12H, Me), 1.31 (s, 9H, *t*-Butyl), 1.46-1.53 (m, 8H, CHC<u>H$_2$</u>Me), 1.69-1.75 (m, 1H, OCH$_2$C<u>H</u>Et$_2$), 1.87-1.95 (m, 1H, OCH$_2$C<u>H</u>Et$_2$), 3.85 (s, 3H, COOMe), 4.01 (d, *J* = 5.6 Hz, 2H, OC<u>H$_2$</u>CHEt$_2$), 4.06 (d, *J* = 6.4 Hz, 2H, OC<u>H$_2$</u>CHEt$_2$), 4.99 (s, 2H, OC<u>H$_2$</u>CO), 7.34-8.61 (m, 8H, ArH), 8.78 (s, 1H, NHCO), 10.01 (s, 1H, NHCO). ^{13}C NMR (125 MHz, $CDCl_3$): δ 9.7, 10.1, 10.3, 22.0, 22.8, 26.9, 28.7, 39.1, 40.0, 51.1, 63.1, 69.7, 70.4, 82.2, 109.8, 110.5, 115.9, 117.4, 117.6, 118.5, 122.4, 124.2, 129.2, 129.9, 131.2, 135.2, 136.2, 146.1, 147.6, 149.2, 153.4, 158.8, 163.3, 164.9, 165.8. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{39}H_{50}N_4O_{11}$ (m/z = 751.3549). Observed m/z = 751.3518. Rf ~ 0.7 (DCM).

Compound 13: yield 20%; ^{1}H-NMR (500 Hz, $CDCl_3$): δ 1.31 (s, 9H, *t*-Butyl), 3.85 (s, 3H, COOMe), 3.95 (s, 3H, OMe), 4.04 (s, 3H, OMe), 5.03 (s, 2H, OC$\underline{H_2}$CO), 7.39-8.61 (m, 8H, ArH), 8.68 (s, 1H, NHCO), 10.24 (s, 1H, NHCO). ^{13}C NMR (125 MHz, $CDCl_3$): δ 27.9, 52.1, 56.2, 56.3, 64.1, 83.1, 109.9, 110.7, 116.5, 118.6, 118.7, 119.1, 123.5, 125.2, 130.2, 130.8, 132.0, 135.9, 137.2, 147.6, 148.8, 150.0, 154.2, 159.5, 164.6, 166.1, 166.7. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{29}H_{30}N_4O_{11}$ (m/z = 611.1984). Observed m/z = 611.1996. MALDI-TOF, M = $C_{29}H_{30}N_4O_{11}$ ($[M]^+$, m/z = 610.7). Rf ~ 0.6 (DCM).

General protocol for the removal of protective *tert*-butyl group:

The corresponding *tert*-butyl-protected trimer (0.1 mmol, **7-13**) was dissolved in chloroform (1 mL), then 0.5 mL of trifluoroacetic acid (TFA) was added and the reaction mixture was stirred for 3-5 hr. Solvent was evaporated under reduced pressure to the dryness and the residue was sonicated with a small volume of acetonitrile. Suspension formed was filtered off to give a yellow solid of mono-acid trimer (**14-20**) in the nearly quantitative yield.

Compound 14: yield 76%; ^{1}H-NMR (500 Hz, DMSO-d6): δ 1.40 (d, *J* = 5.8 Hz, 12H, $Me_{isoprop}$), 3.90 (s, 3H, COOMe), 4.77-4.84 (m, 2H, OC$\underline{H}Me_2$), 5.09 (s, 2H, OC$\underline{H_2}$CO), 7.67-8.27 (m, 9H, ArH), 9.44 (s, 1H, NHCO), 9.77 (s, 1H, NHCO), 13.35 (br s, 1H, COOH). ^{13}C NMR (125 MHz, DMSO-d6): δ 21.6, 21.7, 52.1, 65.5, 71.4, 112.8, 113.7, 114.2, 119.7, 120.2, 121.5, 122.0, 123.6, 125.1, 125.5, 130.8, 131.4, 132.7, 138.9, 141.4, 147.7, 149.1, 150.1, 163.4, 164.3, 165.8, 169.1. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{30}H_{31}N_3O_{11}$ (m/z = 610.2031). Observed m/z = 610.2029. MALDI-TOF, M = $C_{30}H_{31}N_3O_{11}$ ($[M]^+$, m/z = 609.7), Rf ~ 0.1 (DCM).

Compound 15: yield 71%; ^{1}H-NMR (500 Hz, DMSO-d6): δ 0.99 (d, *J* = 7.1 Hz, 6H, Me_{isobut}), 1.03 (d, *J* = 6.8 Hz, 6H, Me_{isobut}), 2.06-2.17 (m, 2H, $OCH_2C\underline{H}Me_2$), 3.86 (s, 3H, COOMe), 3.90-3.93 (m, 4H, OC$\underline{H_2}$CHMe$_2$), 5.02 (s, 2H, OC$\underline{H_2}$CO), 7.66-8.15 (m, 9H, ArH), 9.48 (s, 1H, NHCO), 9.84 (s, 1H, NHCO), 13.29 (br s, 1H, COOH). ^{13}C NMR (125 MHz, DMSO-d6): δ 19.0, 27.7, 27.8, 52.1, 65.5, 74.5, 74.6, 111.1, 112.0, 114.1, 119.8, 120.1, 121.9, 122.0, 124.3, 125.1, 125.9, 129.8, 131.8, 138.9, 141.4, 149.5, 150.1, 150.9, 163.3, 164.2, 165.8, 169.0. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{32}H_{35}N_3O_{11}$ (m/z = 638.2344). Observed m/z = 638.2333. MALDI-TOF, M = $C_{32}H_{35}N_3O_{11}$ ($[M]^+$, m/z = 637.2). Rf ~ 0.1 (DCM).

Compound 16: yield 79%; ^{1}H-NMR (500 Hz, DMSO-d6): δ 0.92-0.96 (m, 12H, Me), 1.42-1.59 (m, 8H, CHC$\underline{H_2}$Me), 1.73-1.80 (m, 2H, $OCH_2C\underline{H}Et_2$), 3.92 (s, 3H, COOMe), 4.08-4.10 (m, 4H, OC$\underline{H_2}$CHEt$_2$), 5.07 (s, 2H, OC$\underline{H_2}$CO), 7.71-8.19 (m, 9H, ArH), 9.51 (s, 1H, NHCO), 9.87 (s, 1H, NHCO), 13.34 (br s, 1H, COOH). ^{13}C NMR (125 MHz, DMSO-d6): δ 10.9, 11.0, 22.9, 52.1, 65.5, 70.3, 70.4, 111.1, 112.0, 114.1, 119.8, 120.0, 122.0, 122.1, 124.4, 125.0, 125.9, 129.8, 131.8, 131.9, 138.9, 141.5, 149.7, 150.1, 151.2, 163.3, 164.2, 165.8, 169.0. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{36}H_{43}N_3O_{11}$ (m/z = 694.2970). Observed m/z = 694.2964. MALDI-TOF, M = $C_{36}H_{43}N_3O_{11}$ ($[M]^+$, m/z = 693.7). Rf ~ 0.1 (DCM).

Compound 17: yield 80%; ^{1}H-NMR (500 Hz, DMSO-d6): δ 1.41 (d, *J* = 5.7 Hz, 6H, $Me_{isoprop}$), 1.51 (d, *J* = 6.5 Hz, 6H, $Me_{isoprop}$), 3.90 (s, 3H, COOMe), 4.76-4.82 (m, 1H, OC$\underline{H}Me_2$), 4.95-5.01 (m, 1H, OC$\underline{H}Me_2$), 5.32 (s, 2H, OC$\underline{H_2}$CO), 7.66-8.79 (m, 8H, ArH), 9.39 (s, 1H, NHCO), 10.19 (s, 1H, NHCO), 13.41 (br s, 1H, COOH). ^{13}C NMR (125 MHz, DMSO-d6): δ 21.5, 21.6, 52.1, 63.3, 71.3, 71.4, 111.5, 113.7, 116.7, 118.8, 120.0, 121.4, 122.1, 125.5, 130.1, 130.2, 132.8, 136.1, 137.8, 146.3, 147.7, 149.1, 153.5, 159.4, 164.1, 165.8, 168.7. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{29}H_{30}N_4O_{11}$ (m/z = 611.1984). Observed m/z = 611.1981. Rf ~ 0.1 (DCM).

Compound 18: yield 61%; ^{1}H-NMR (500 Hz, DMSO-d6): δ 1.09-1.12 (m, 12H, Me_{isobut}), 2.15-2.24 (m, 1H, $OCH_2C\underline{H}Me_2$), 2.30-2.38 (m, 1H, $OCH_2C\underline{H}Me_2$), 3.92 (s, 3H, COOMe), 3.98 (d, *J* = 7.8 Hz, 2H, OC$\underline{H_2}$CHMe$_2$), 4.08 (d, *J* = 7.8 Hz, 2H, OC$\underline{H_2}$CHMe$_2$), 5.29 (s, 2H, OC$\underline{H_2}$CO), 7.63-8.80 (m, 8H, ArH), 9.51 (s,

1H, NHCO), 10.19 (s, 1H, NHCO), 13.39 (br s, 1H, COOH). ^{13}C NMR (125 MHz, DMSO-d6): δ 19.0, 19.1, 27.3, 27.8, 52.1, 63.3, 74.5, 74.8, 110.5, 112.0, 116.7, 119.1, 119.8, 120.3, 121.9, 122.0, 125.8,129.5, 130.3, 131.8, 136.2, 137.7, 147.8, 149.1, 149.4, 153.5, 159.6, 164.0, 165.8, 168.7. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{31}H_{34}N_4O_{11}$ (m/z = 639.2297). Observed m/z = 639.2305. Rf ~ 0.1 (DCM).

Compound 19: yield 96%; ^{1}H-NMR (500 Hz, DMSO-d6): δ 0.94-0.99 (m, 12H, Me), 1.46-1.61 (m, 8H, CHC$\underline{H}_2$Me), 1.76-1.82 (m, 1H, OCH$_2$C$\underline{H}$Et$_2$), 1.97-2.04 (m, 1H, OCH$_2$C$\underline{H}$Et$_2$), 3.92 (s, 3H, COOMe), 4.10 (d, *J* = 5.4 Hz, 2H, OC$\underline{H}_2$CHEt$_2$), 4.18 (d, *J* = 6.0 Hz, 2H, OC$\underline{H}_2$CHEt$_2$), 5.29 (s, 2H, OC$\underline{H}_2$CO), 7.67-8.79 (m, 8H, ArH), 9.48 (s, 1H, NHCO), 10.12 (s, 1H, NHCO), 13.37 (br s, 1H, COOH). ^{13}C NMR (125 MHz, DMSO-d6): δ 10.5, 11.0, 22.5, 22.9, 52.1, 63.3, 70.4, 70.9, 110.5, 112.0, 116.8, 119.4, 120.2, 122.0, 125.9, 129.5, 130.5, 131.8, 136.2, 137.7, 148.1, 149.2, 149.7, 153.5, 159.7, 164.1, 165.8, 168.7. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{35}H_{42}N_4O_{11}$ (m/z = 695.2923). Observed m/z = 695.2927. Rf ~ 0.1 (DCM).

Compound 20: yield 93%; ^{1}H-NMR (500 Hz, DMSO-d6): δ 3.82 (s, 3H, COOMe), 3.91 (s, 3H, OMe), 4.00 (s, 3H, OMe), 5.20 (s, 2H, OC$\underline{H}_2$CO), 7.60-8.70 (m, 8H, ArH), 9.50 (s, 1H, NHCO), 10.20 (s, 1H, NHCO), 13.33 (br s, 1H, COOH). ^{13}C NMR (125 MHz, DMSO-d6): δ 52.6, 56.6, 56.7, 64.0, 109.5, 110.5, 111.7, 116.9, 118.6, 120.9, 122.3, 123.0, 126.5, 130.0, 130.6, 132.1, 136.5, 138.3, 148.8, 149.6, 150.9, 153.9, 159.9, 164.9, 166.3, 169.5. FT-ICR MS (MeCN): Calculated for $[M+H]^+$, M = $C_{25}H_{22}N_4O_{11}$ (m/z = 555.1358). Observed m/z = 555.1360. Rf ~ 0.1 (DCM).

Bis-arylamides **21-28** have been prepared as described earlier.[9a]

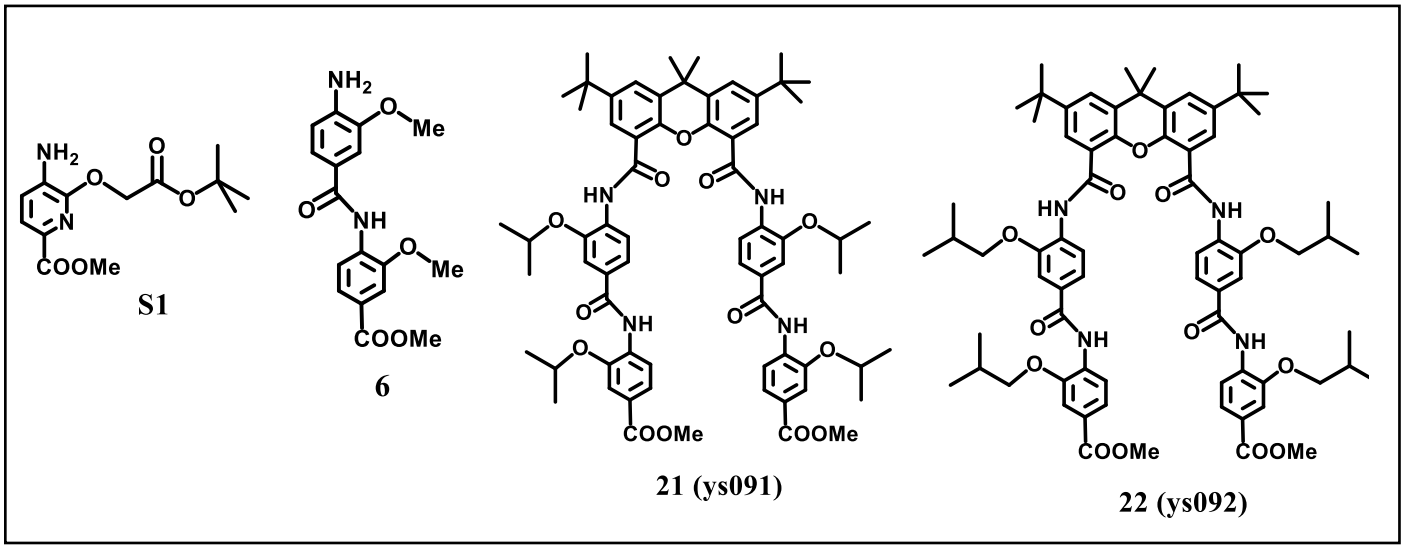
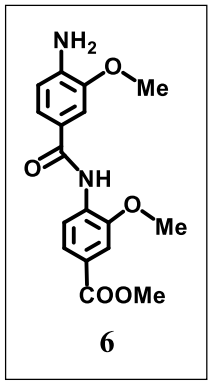


Chart 1. Compounds examined by single crystal X-ray analysis



6

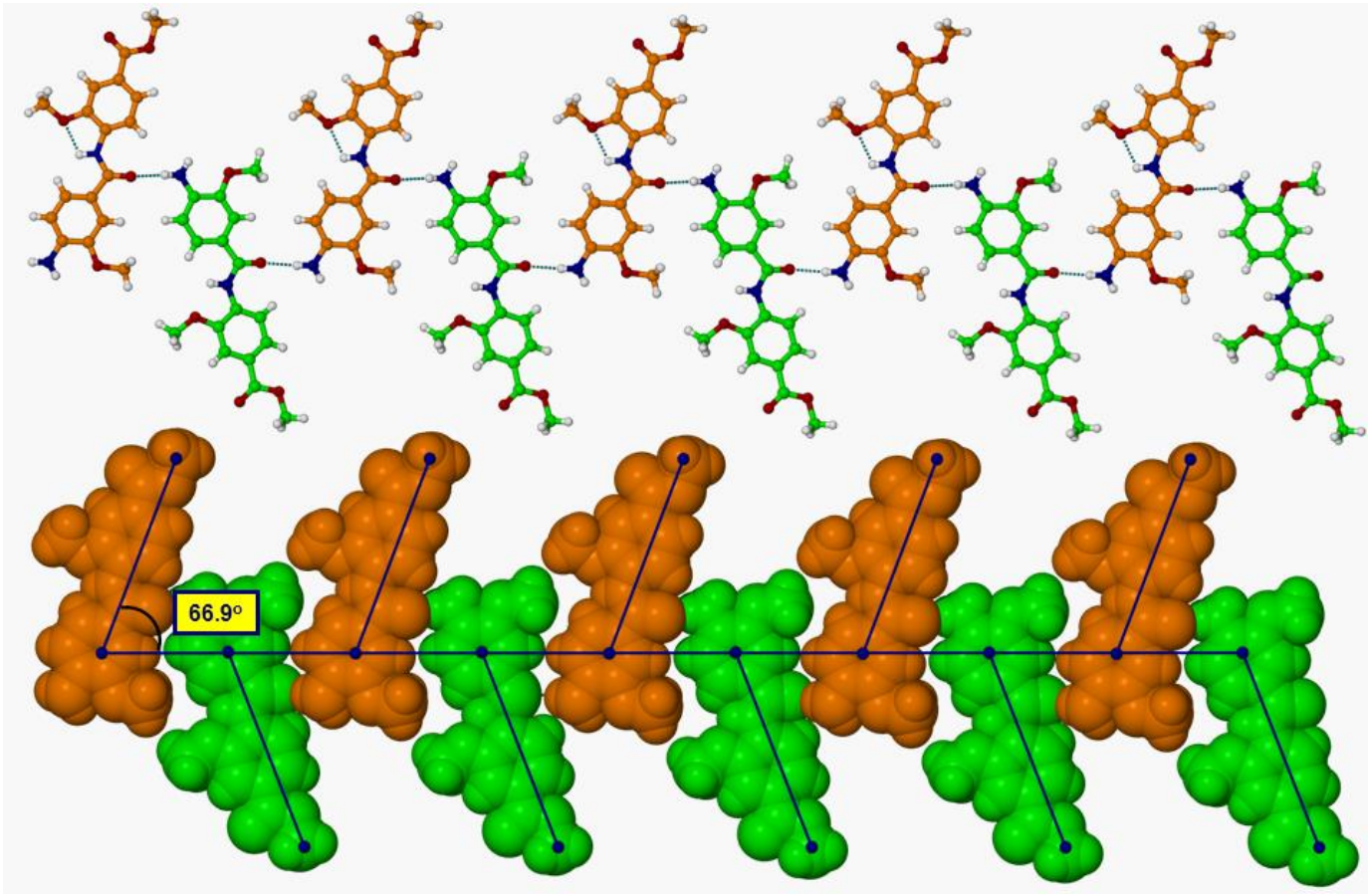
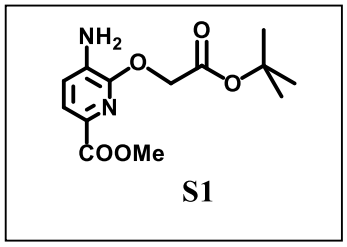
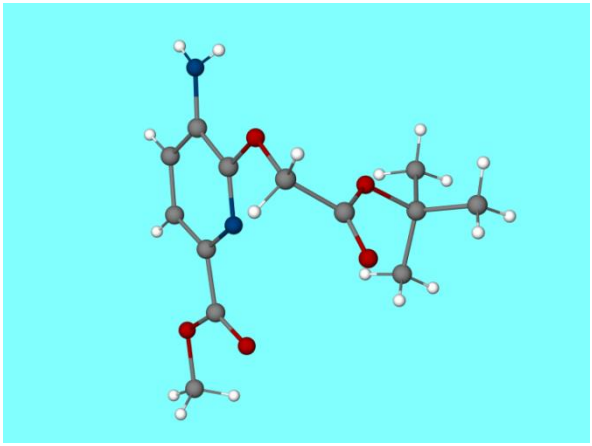


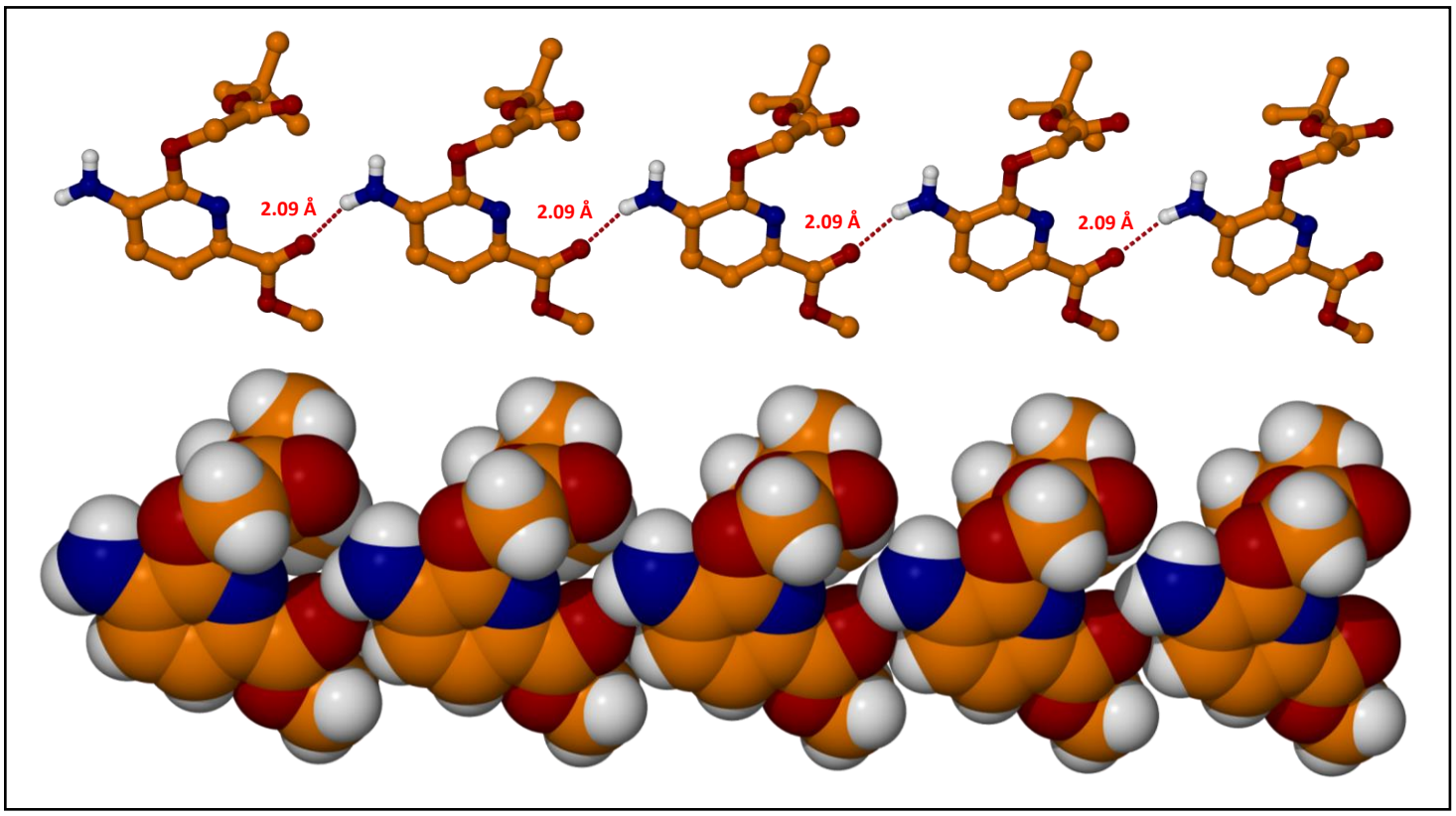
Figure S1. Crystal packing motif of **6**: formation of hydrogen-bonded network (ball-and-stick and CPK representation)

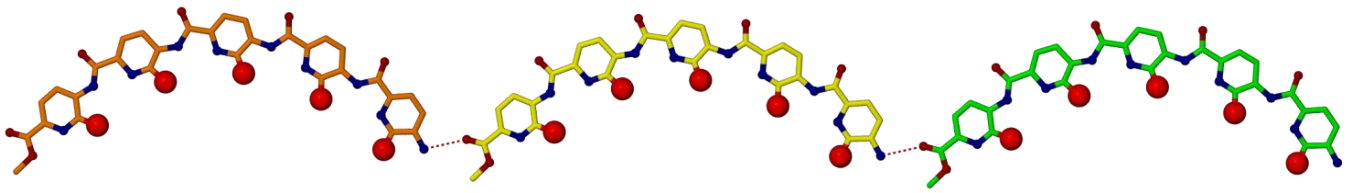


S1



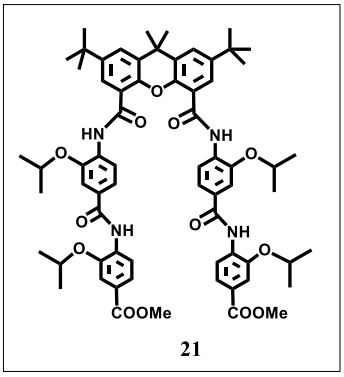
The simplest *mono*-NH_2-arylamide (1 monomer unit) displaying intermolecular H-bonding



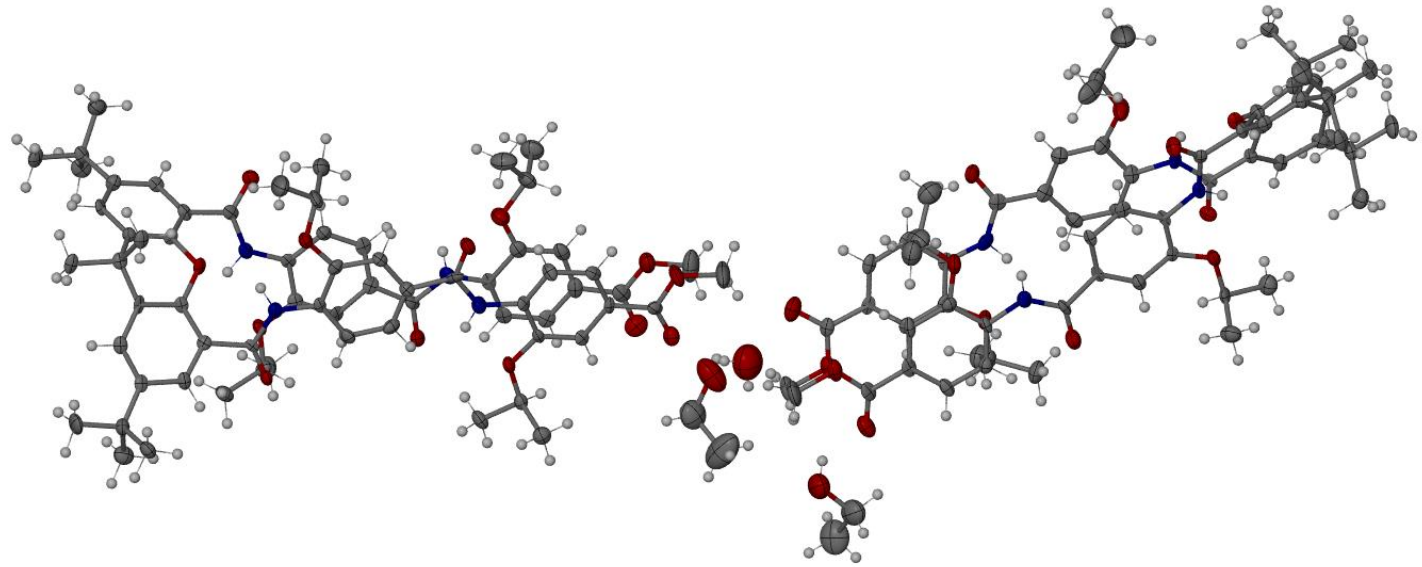


The most complex *mono*-NH_2-arylamide (comprised of 5 monomer units) reported earlier[4] that showed intermolecular H-bonding

Figure S2. NH_2-monomer (**S1**) *vs*. NH_2-pentamer: 5-amino-6-*tert*-butoxycarbonylmethoxypyridine-2-carboxylic acid methyl ester molecular structure and hydrogen-bonded pattern (ball-and-stick and CPK representation); red balls represent disordered side chains in the structure of known NH_2-pentamer[4]



21



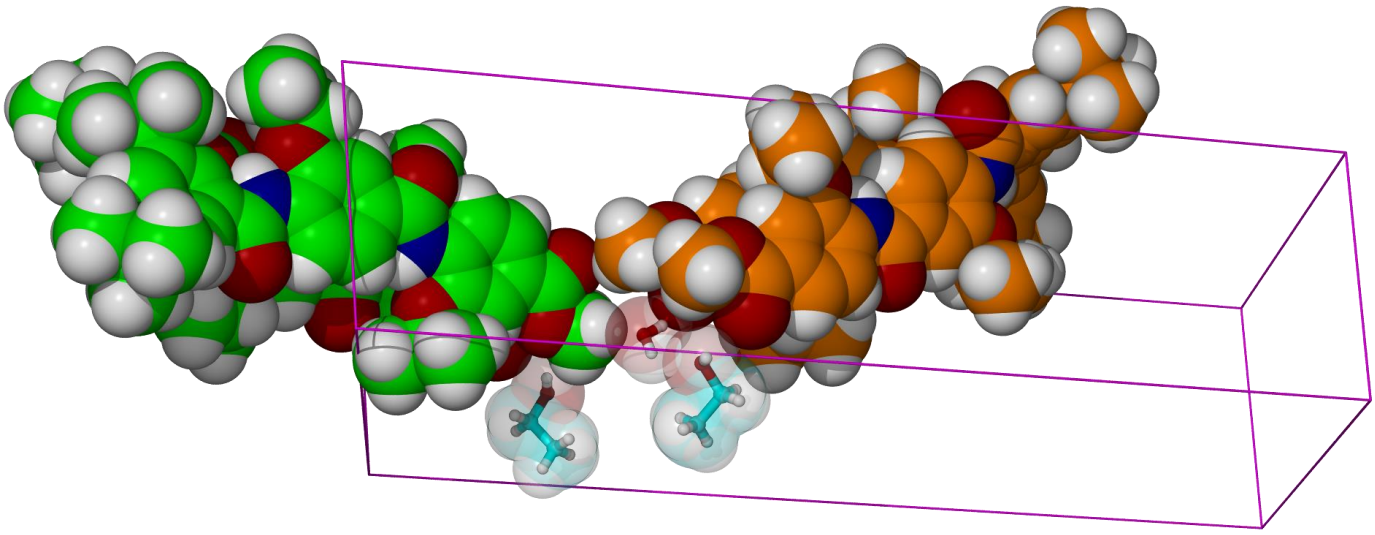


Figure S3. Asymmetric unit of *i*-Pr-*bis*-dimer **21** showcasing two crystallographically independent molecules and a solvent matrix associated with them (*i.e.* two EtOH and one water depicted as semitransparent spheres)

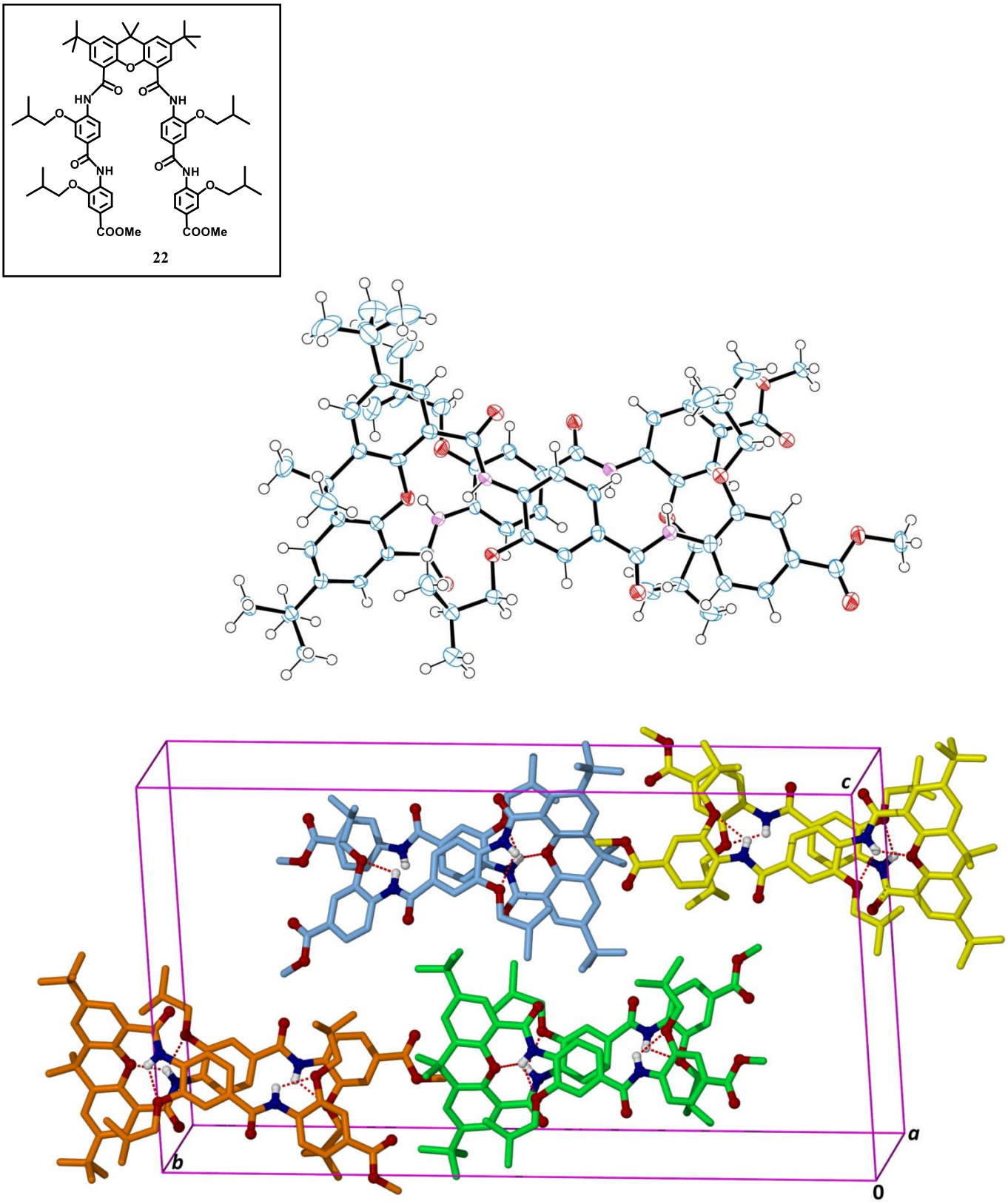


Figure S4. Molecular structure and unit cell of compound **22**

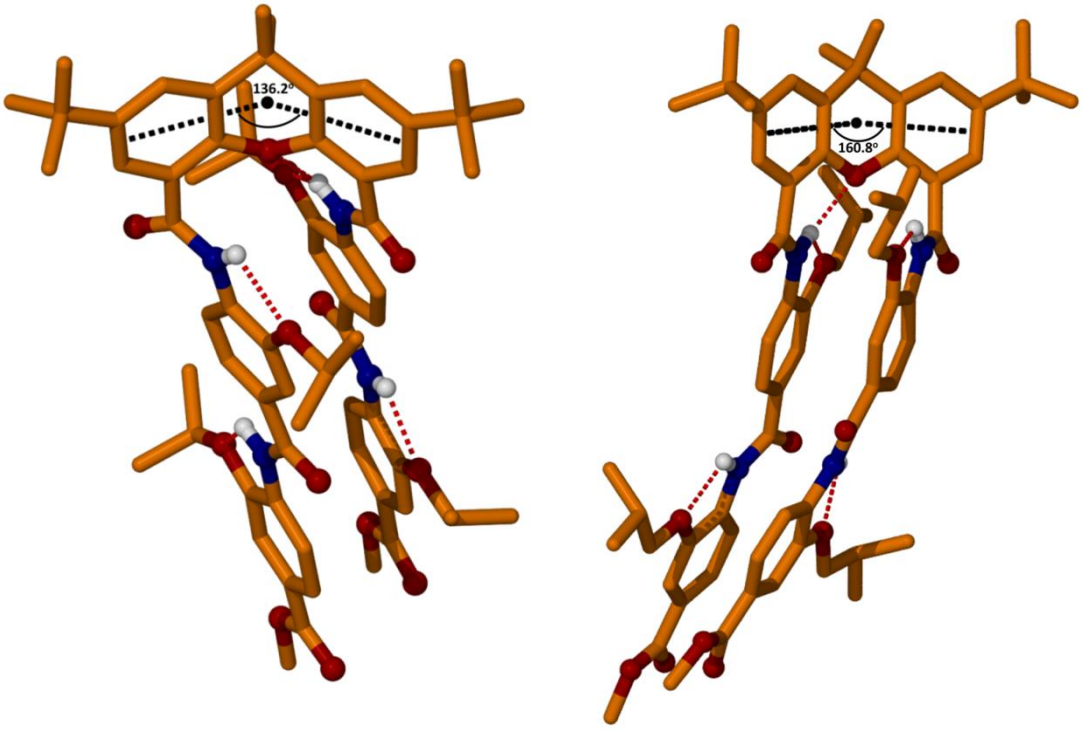


Figure S5. Molecular structure of *bis*-dimers **21** (left – **ys091**) and **22** (right – **ys092**); non-polar hydrogen atoms omitted for clarity

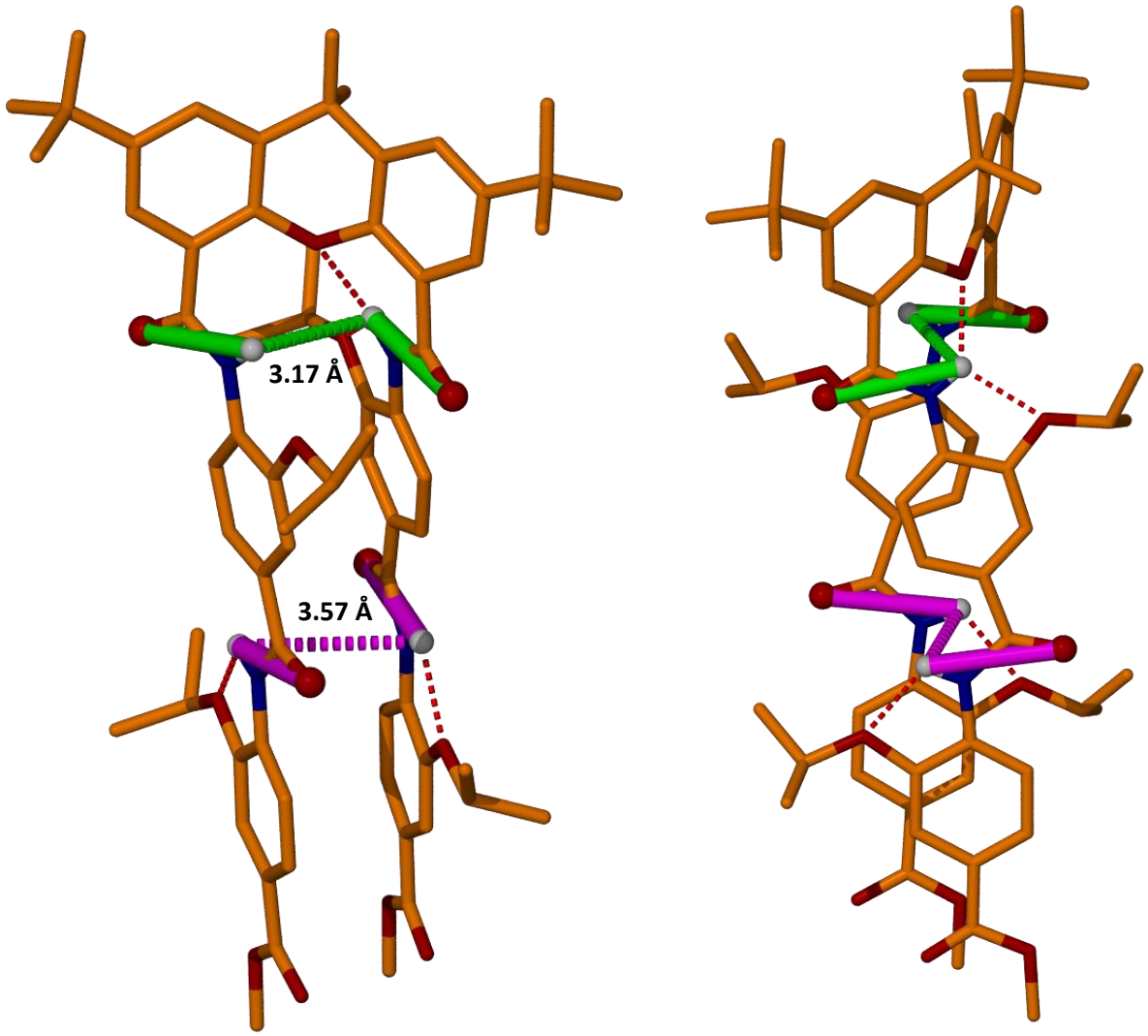


Figure S5-2. Intramolecular amide bond dipole vectors orientation in **21** (structure **ys091**, one of two crystallographically independent molecules in unit cell), front and side views; torsion angle between amide bond vectors (green) appended to xanthene moiety = **158.6°**; torsion angle between amide bond vectors connected to isobutyloxybenzoic acid methyl ester segments (violet) = **163.4°**; red dashed lines represent intramolecular H-bonding, violet and green dashed lines connect amide bond vectors (*i.e.*, solid violet and green lines, correspondingly) forming anti-parallel dipole patterns (for clarity only amide hydrogen atoms are shown).

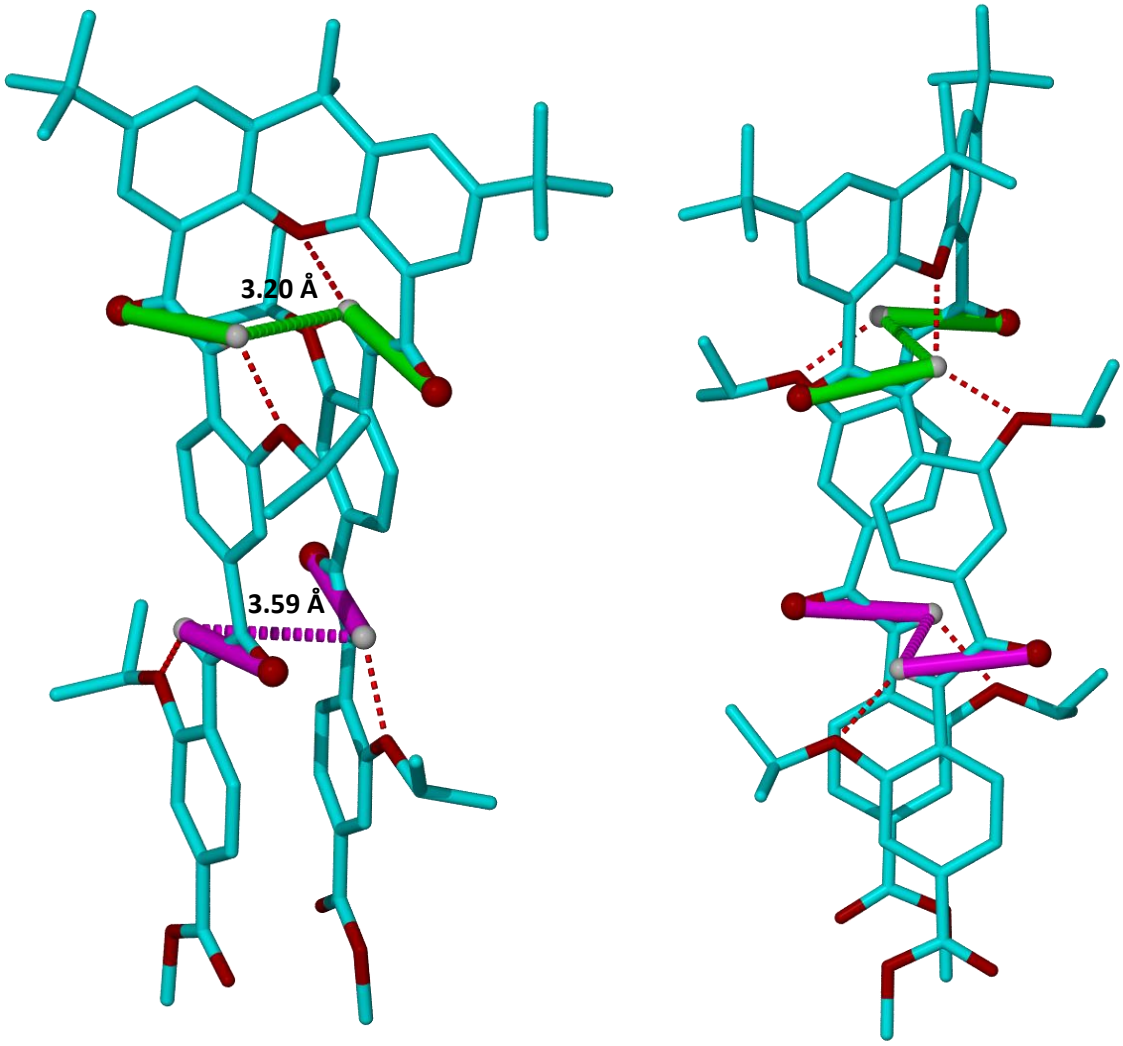


Figure S5-3. Intramolecular amide bond dipole vectors orientation in **21** (structure **ys091**, another crystallographically independent molecule in unit cell), front and side views; torsion angle between amide bond vectors (green) appended to xanthene moiety = **159.6°**; torsion angle between amide bond vectors connected to isobutyloxybenzoic acid methyl ester segments (violet) = **164.6°**; red dashed lines represent intramolecular H-bonding, violet and green dashed lines connect amide bond vectors (*i.e.*, solid violet and green lines, correspondingly) forming anti-parallel dipole patterns (for clarity only amide hydrogen atoms are shown).

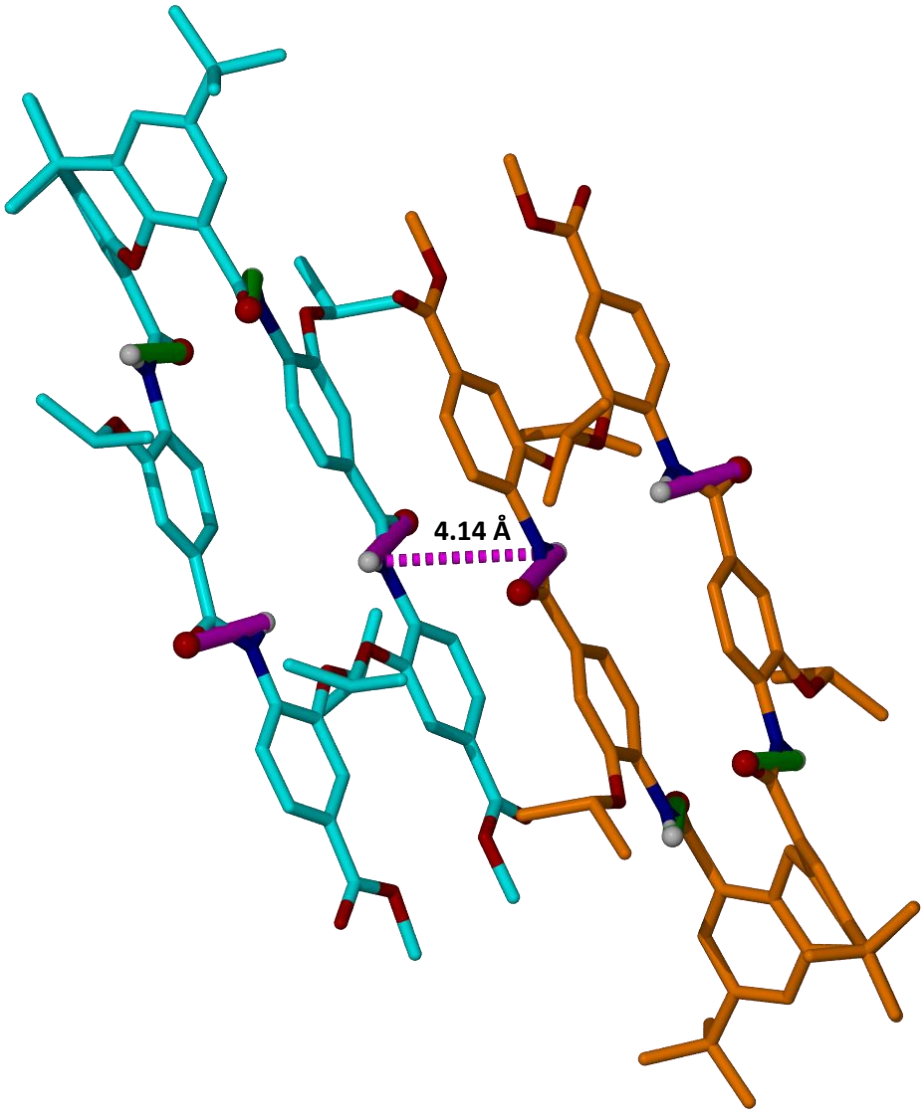


Figure S5-4. Intermolecular amide bond dipole vectors orientation in **21** (structure **ys091**, one of two crystallographically unique molecules in the crystal lattice), torsion angle between "central" amide bond vectors (solid violet) = **180.0°**; violet dashed lines represent connection of amide bond vectors belonging to the adjacent molecules (for clarity only amide hydrogen atoms are shown).

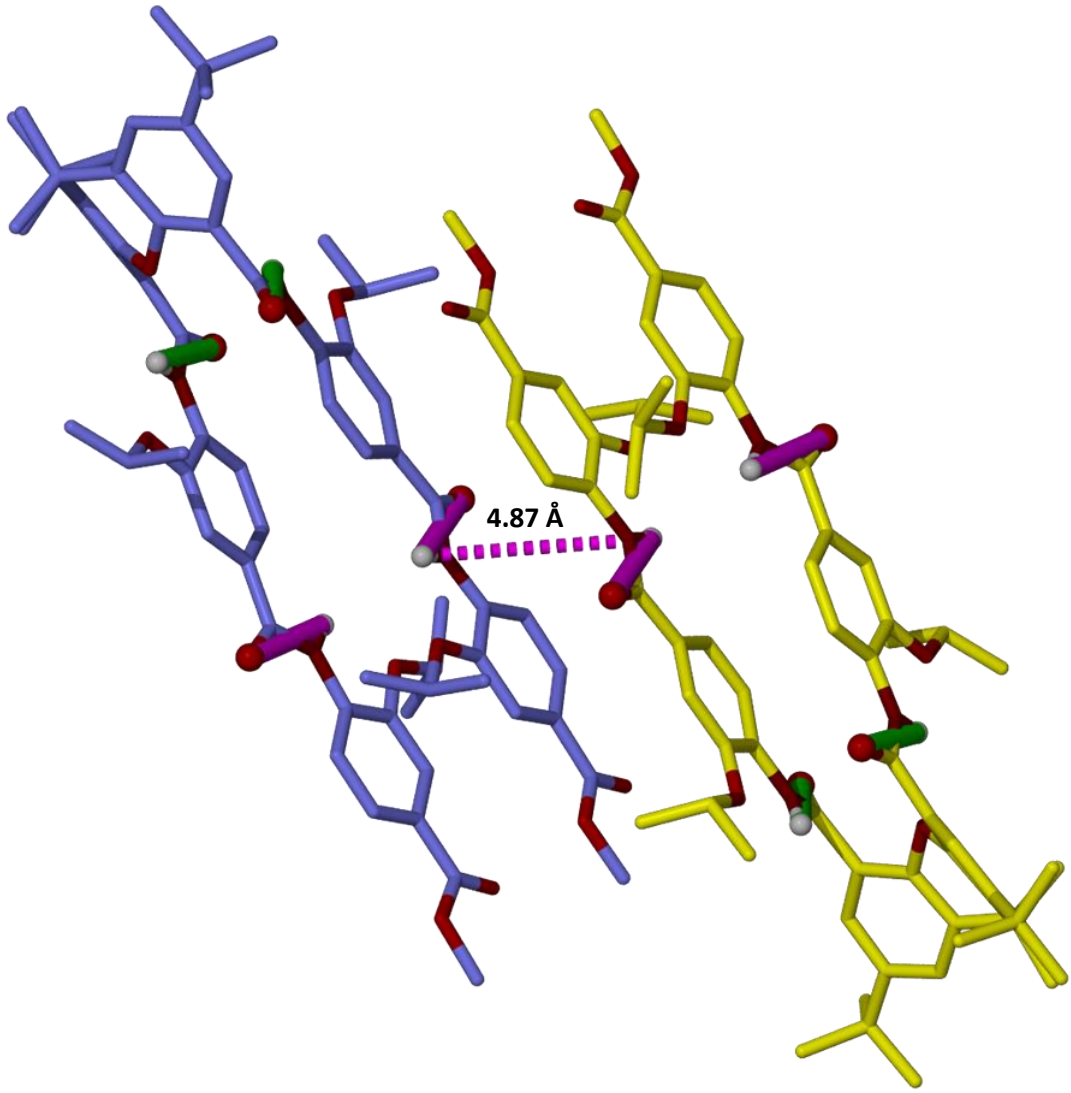


Figure S5-5. Intermolecular amide bond dipole vectors orientation in **21** (structure **ys091**, another crystallographically unique molecule in a crystal lattice), torsion angle between "central" amide bond vectors (solid violet) = **180.0°**; violet dashed lines represent connection of amide bond vectors belonging to the adjacent molecules (for clarity only amide hydrogen atoms are shown).

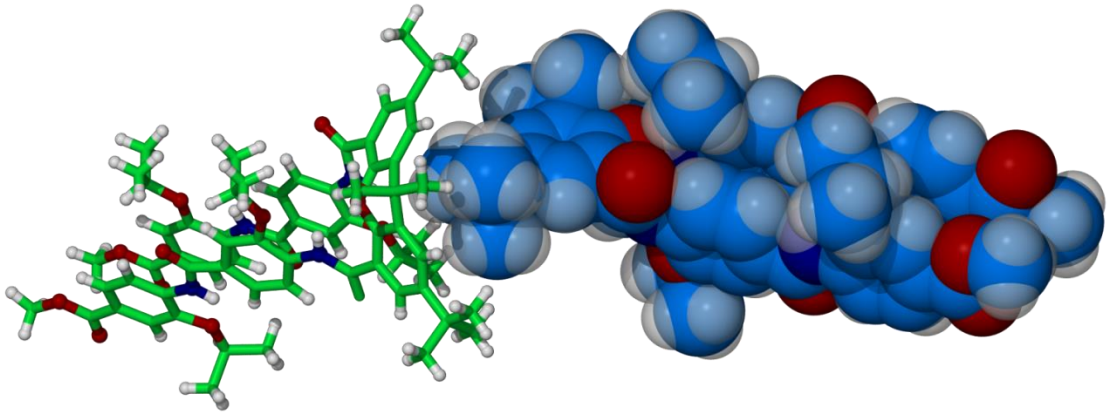


Figure S6. Partial packing diagram of *bis*-dimer **21**: hydrophobic interactions of the xanthene fragments belonging to adjacent molecules. One methyl and two *tert*-butyl groups of xanthene fragment (showed in green) form a "pocket" for *tert*-butyl group of the neighboring *bis*-dimer molecule (depicted in blue, CPK format)

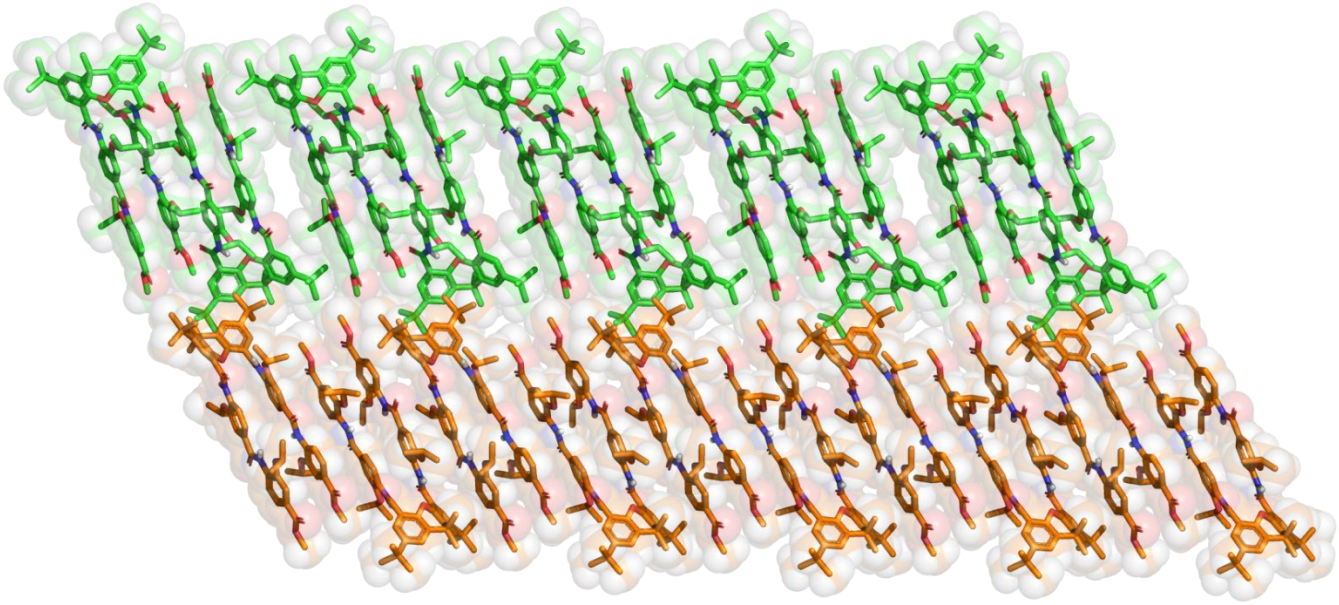


Figure S7. Partial crystal packing diagram of *bis*-dimer **21**, solvents omitted for clarity

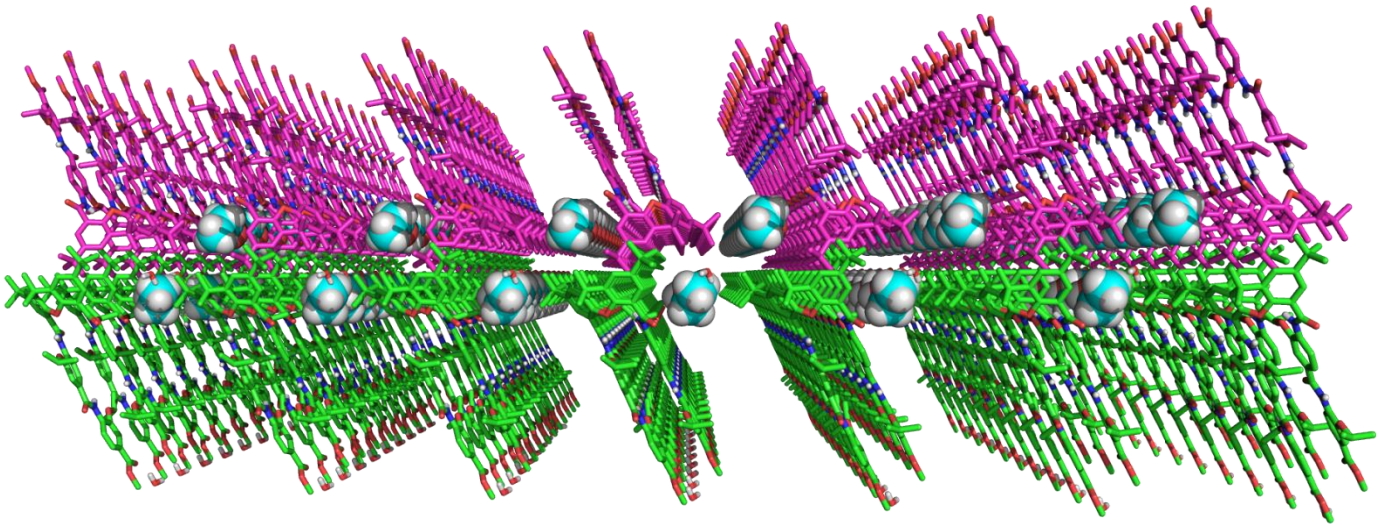


Figure S8. Partial crystal packing diagram of *bis*-dimer **21**: a bilayer arrangement, EtOH molecules (cyan) shown in CPK format

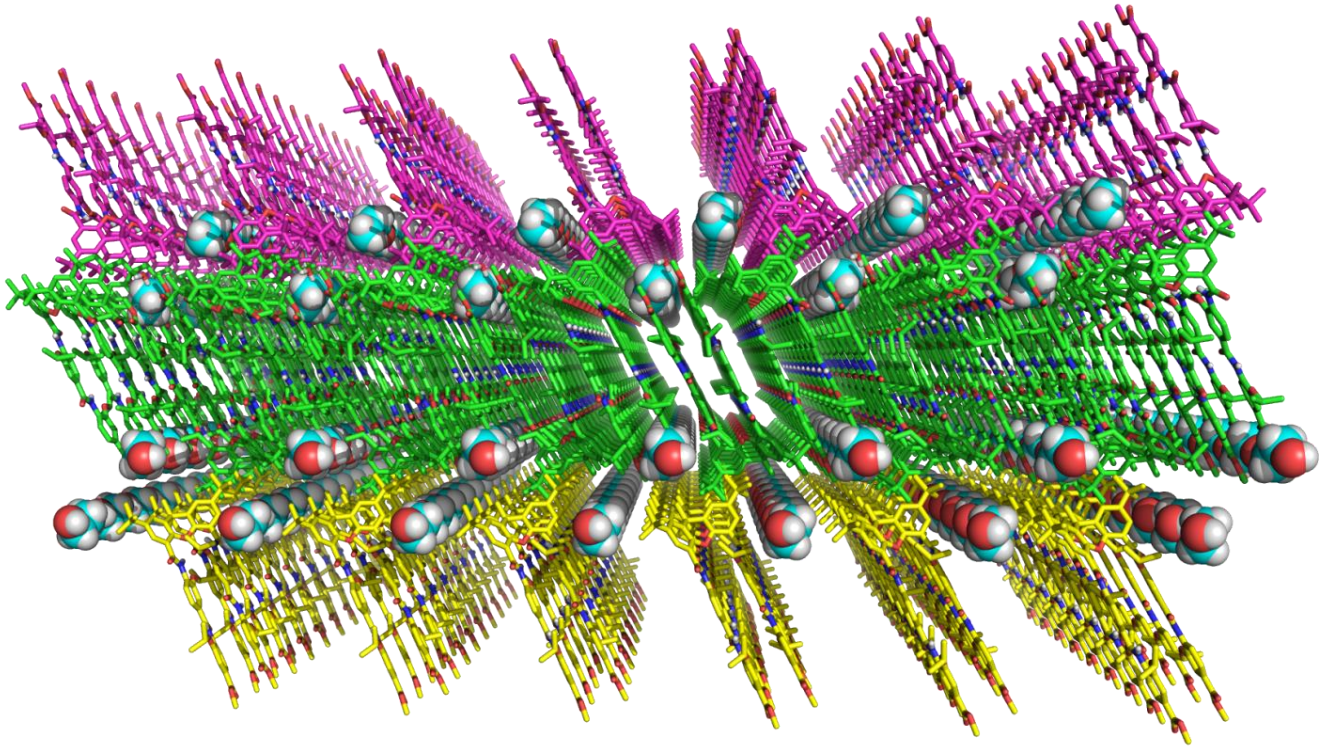


Figure S9. Partial crystal packing diagram of *bis*-dimer **21**: a multilayer arrangement, EtOH molecules (cyan) shown in CPK format

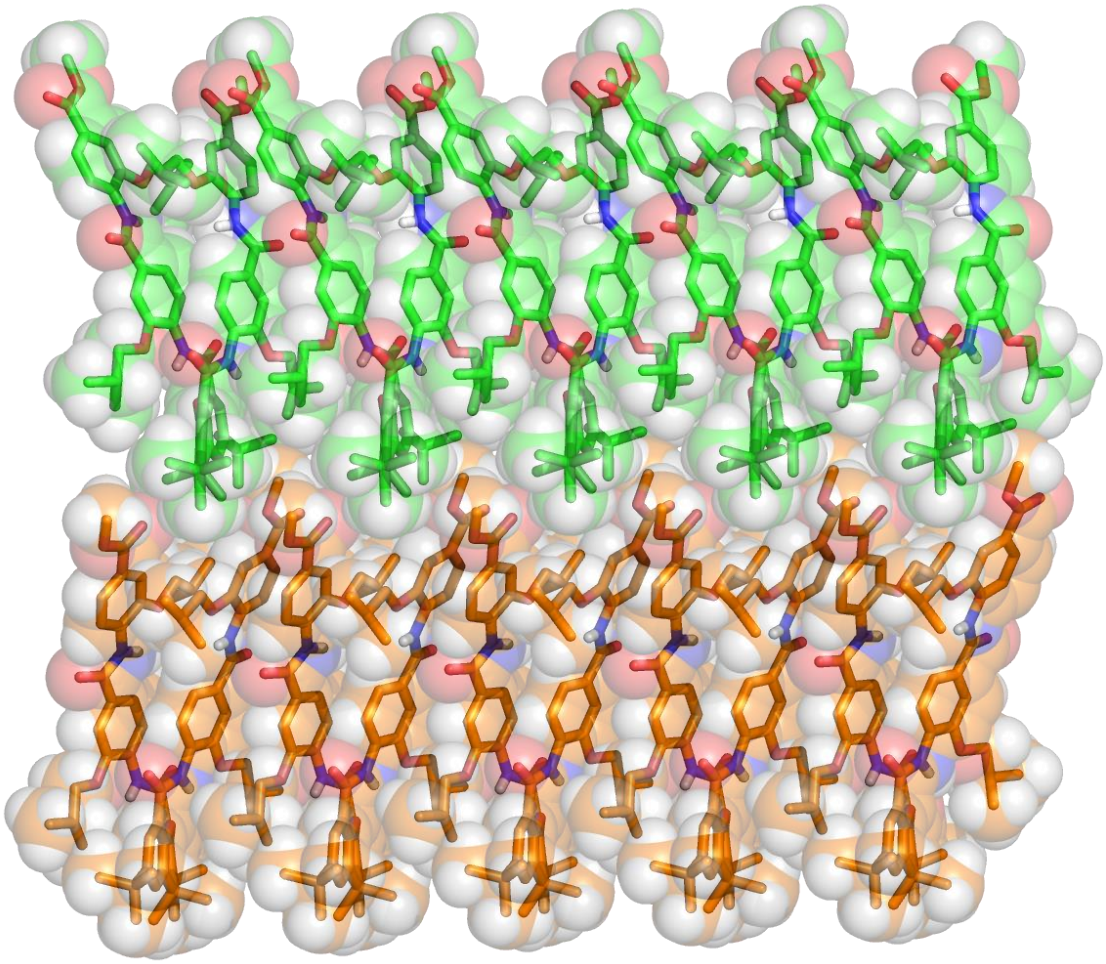


Figure S10. Partial crystal packing diagram of *bis*-dimer **22**

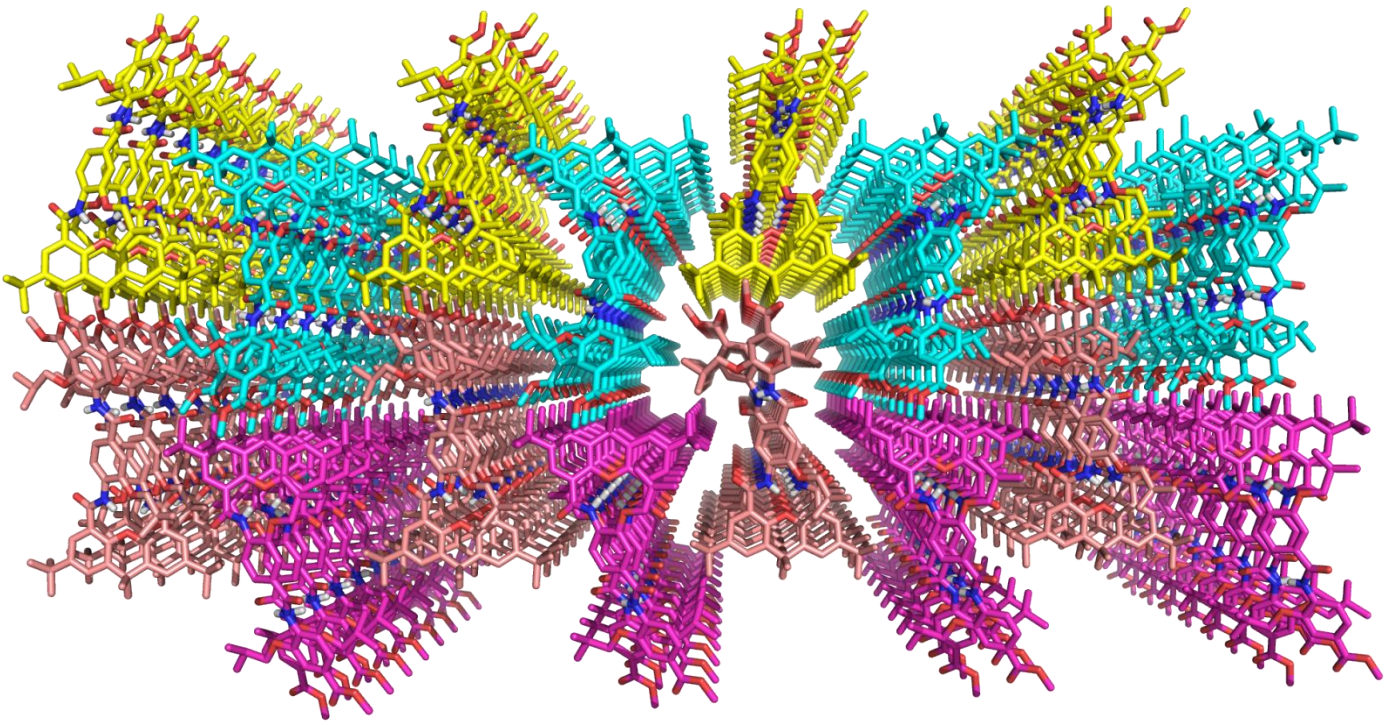


Figure S11. Partial crystal packing diagram of *bis*-dimer **22**: a multilayer arrangement

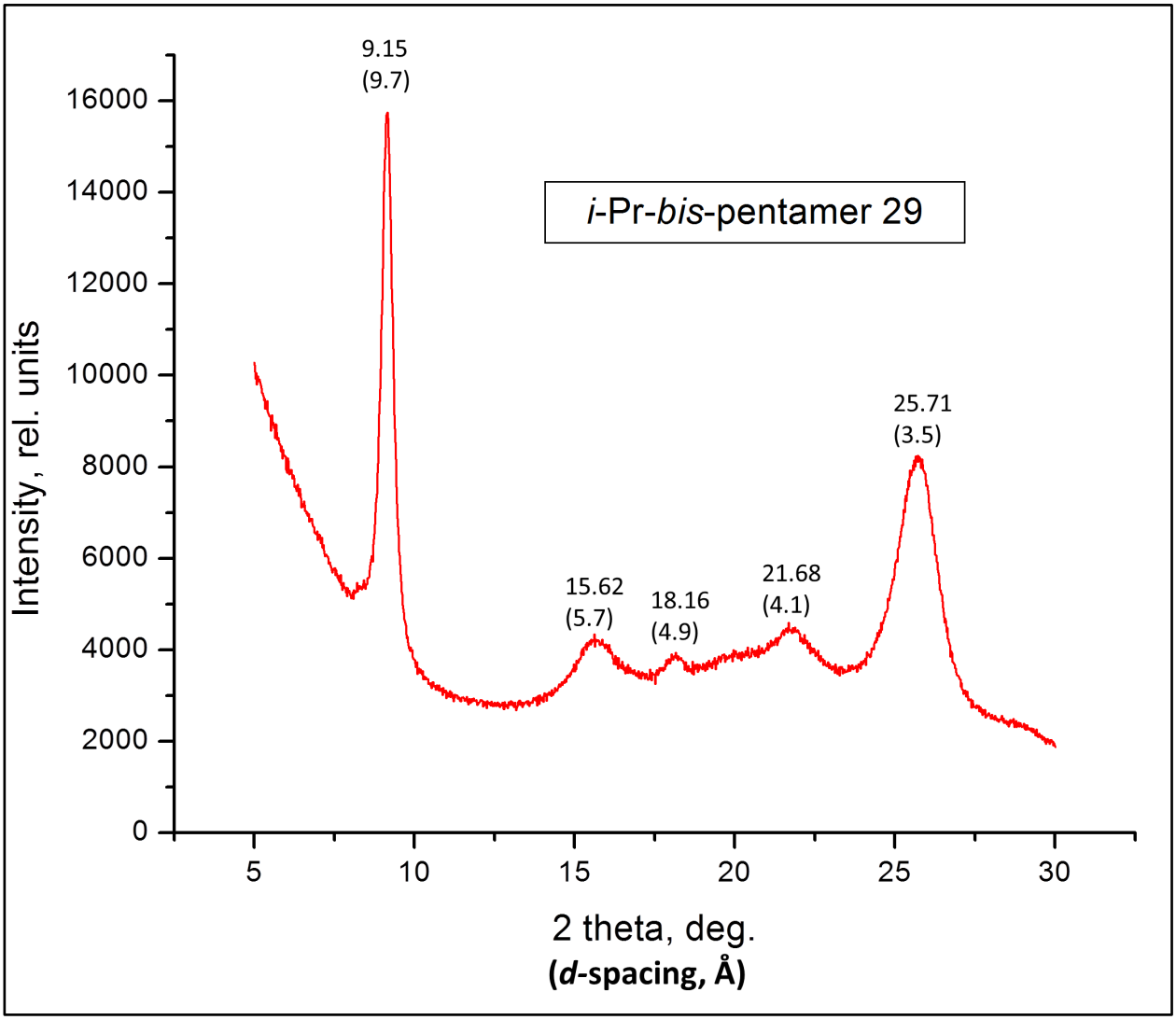


Figure S12. Powder XRD pattern of *bis*-pentamer **29**

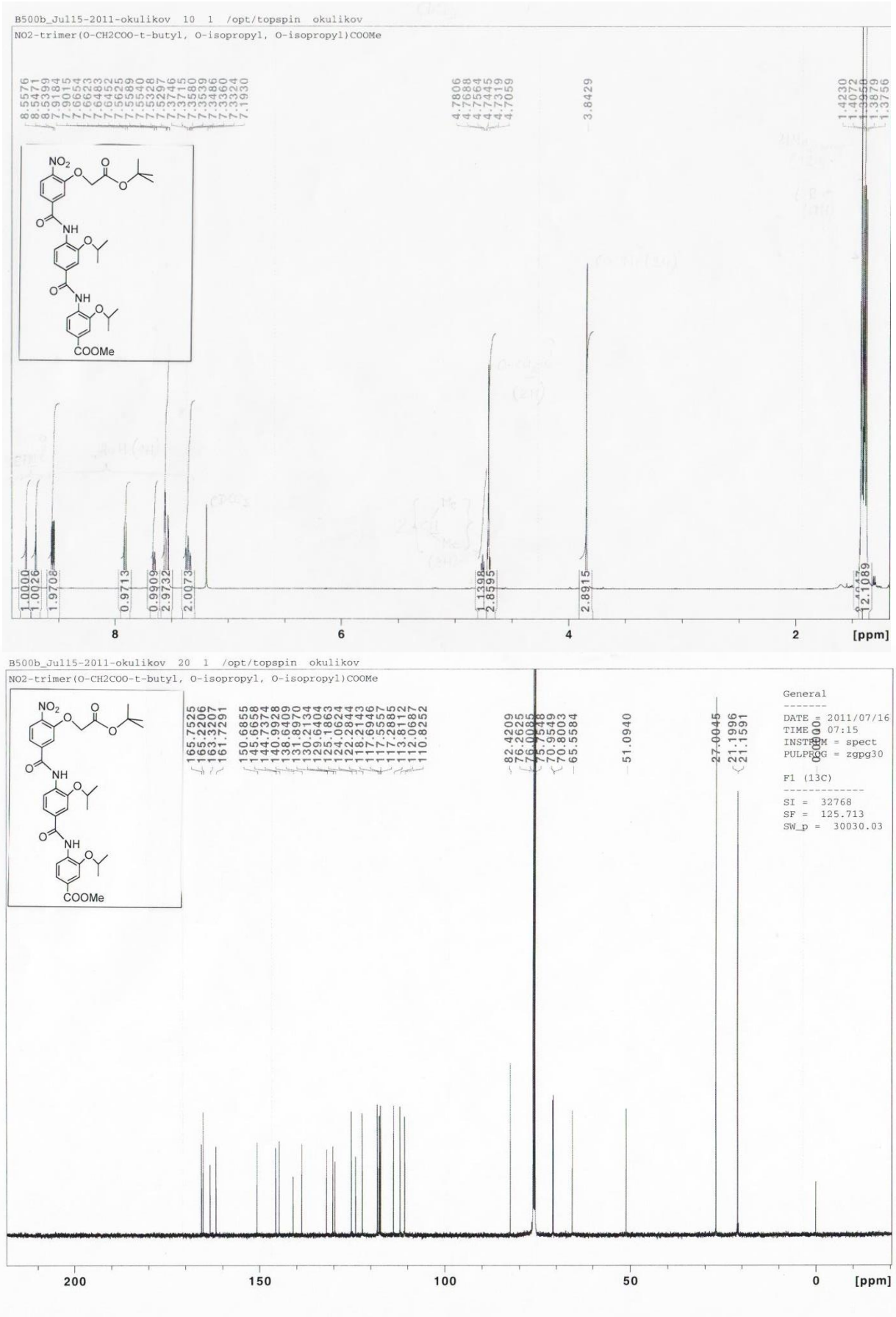


Figure S13. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **7** in $CDCl_3$

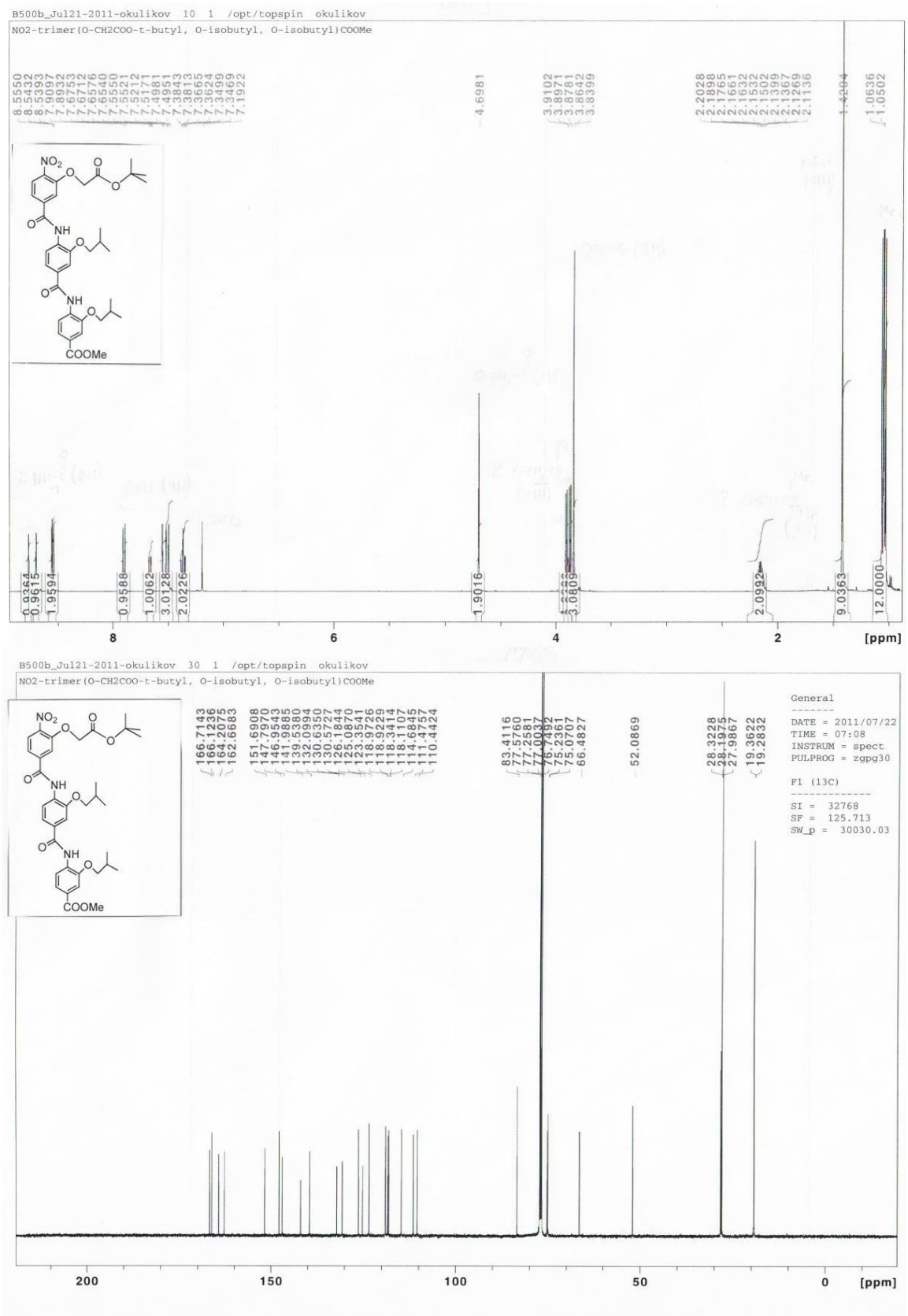


Figure S14. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **8** in $CDCl_3$

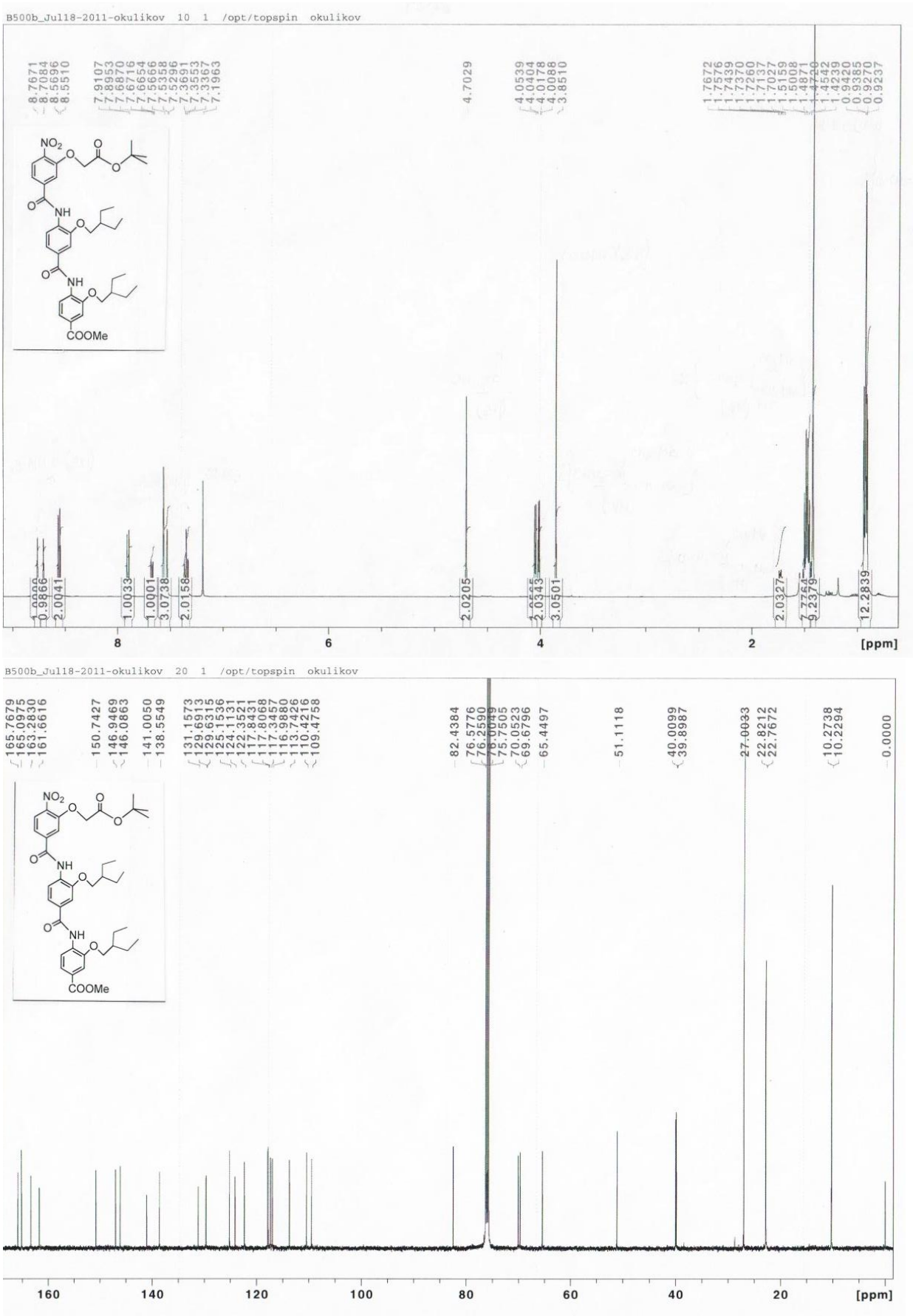


Figure S15. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **9** in $CDCl_3$

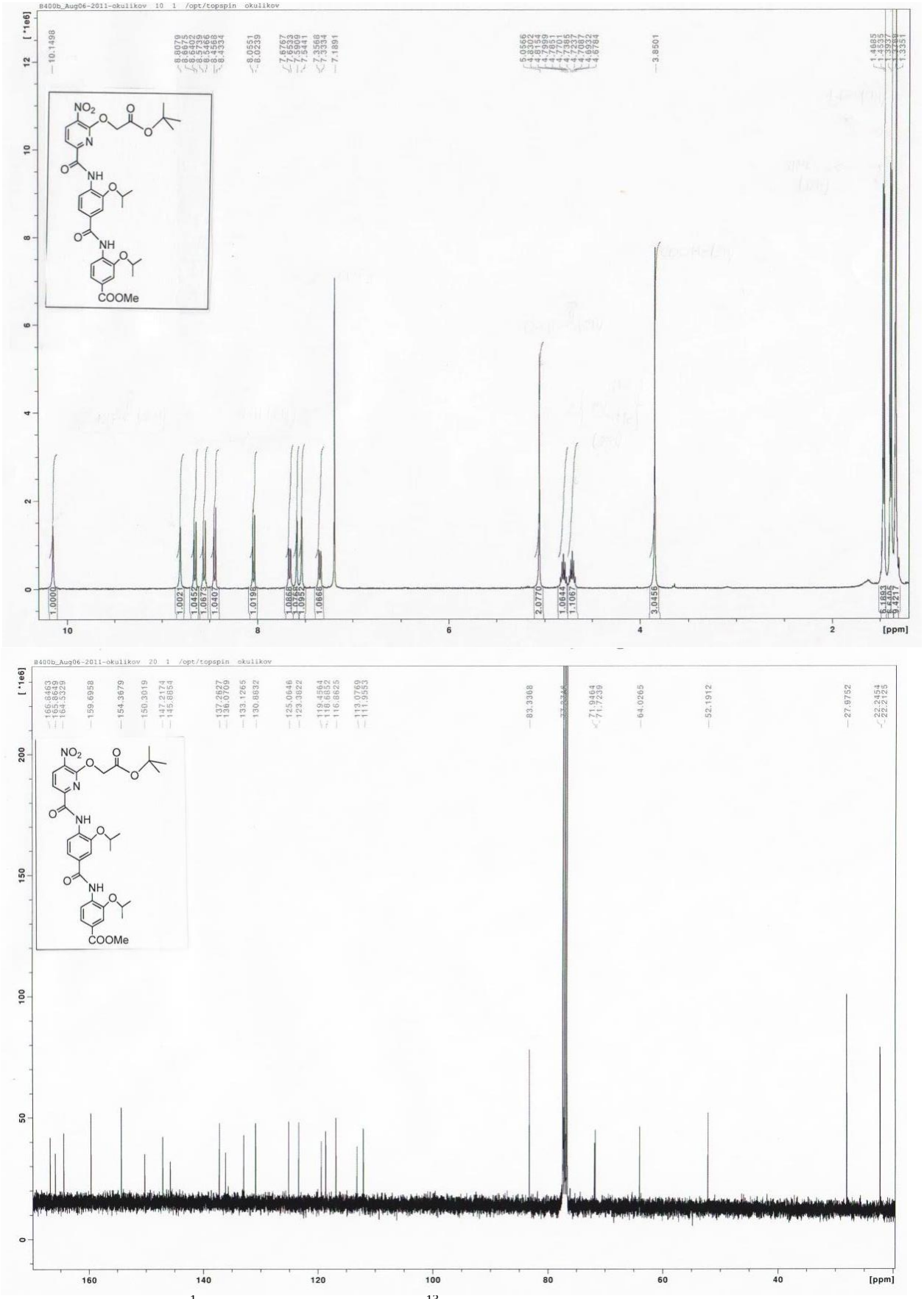


Figure S16. ^{1}H NMR (top, 400 MHz) and ^{13}C NMR (bottom, 100 MHz) spectra of **10** in $CDCl_3$

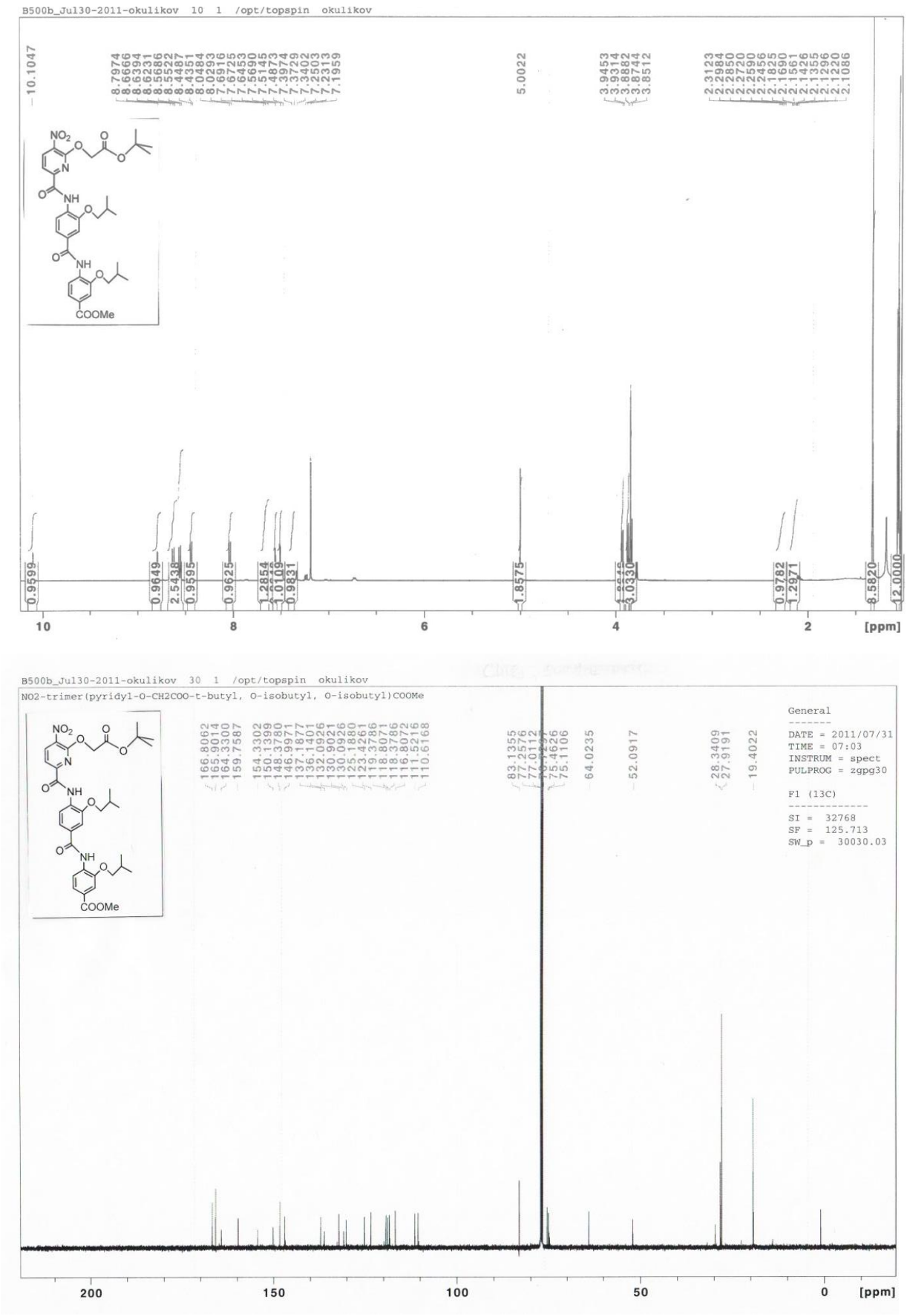


Figure S17. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **11** in $CDCl_3$

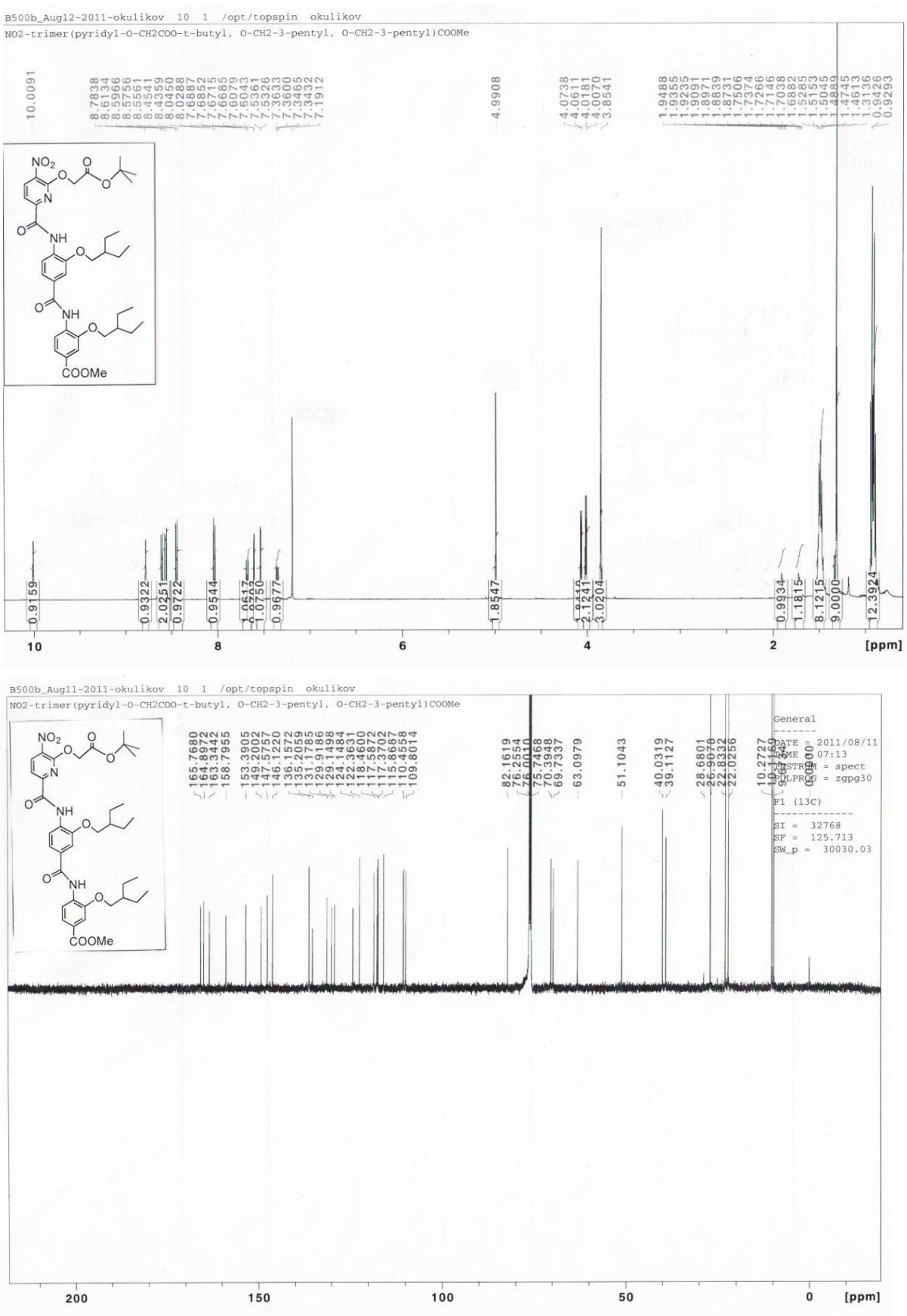


Figure S18. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **12** in $CDCl_3$

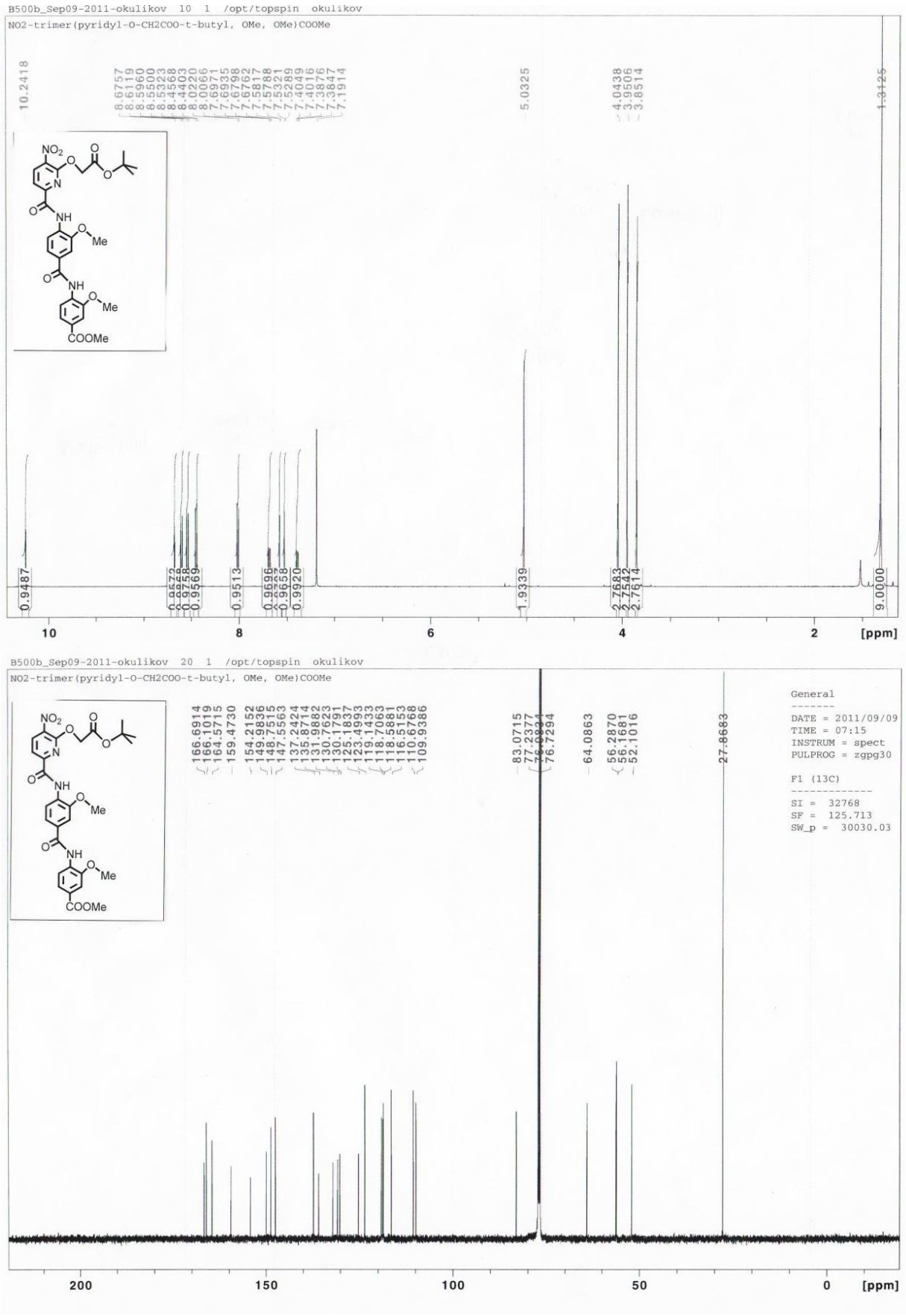


Figure S19. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **13** in $CDCl_3$

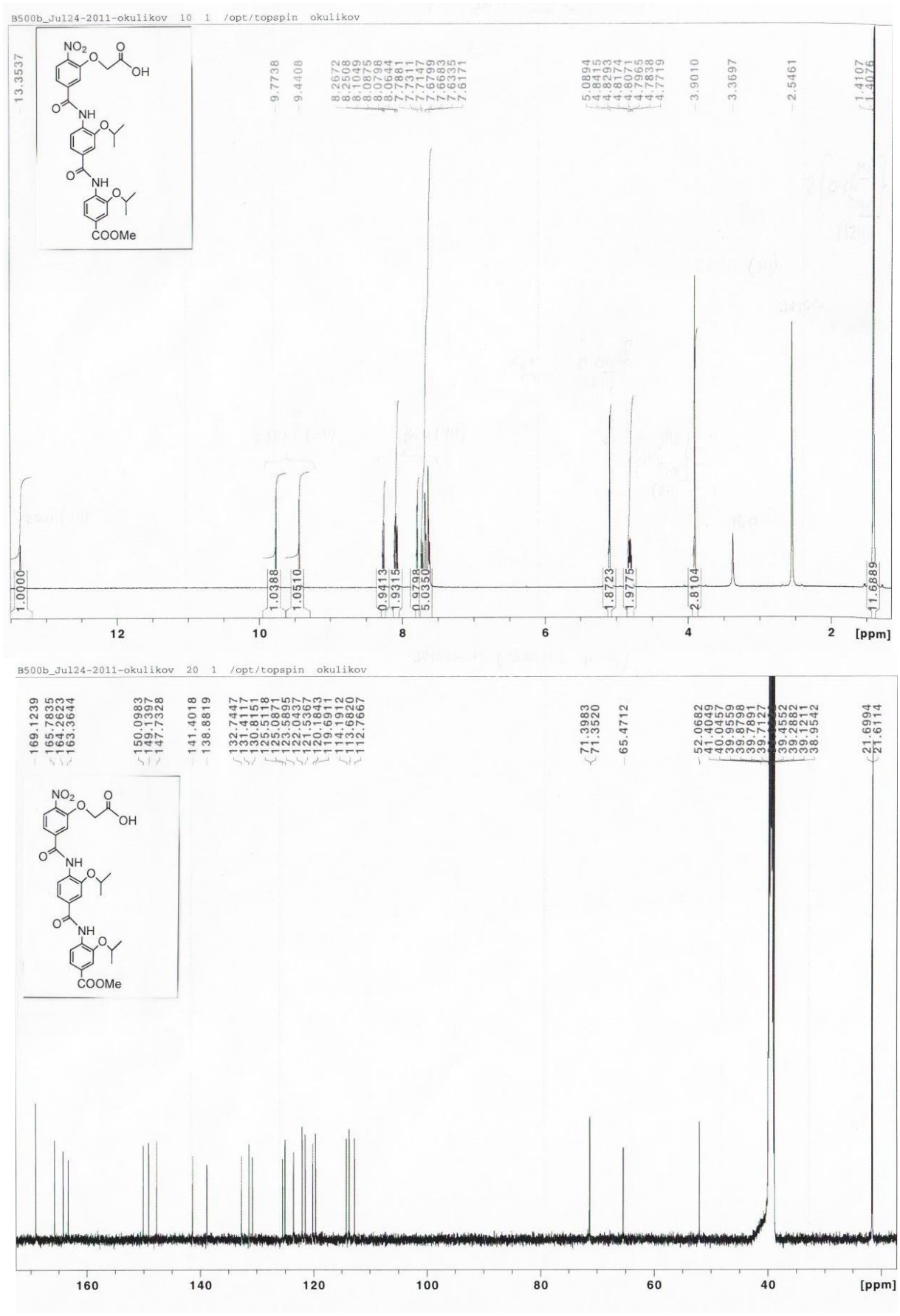


Figure S20. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **14** in DMSO-d_6

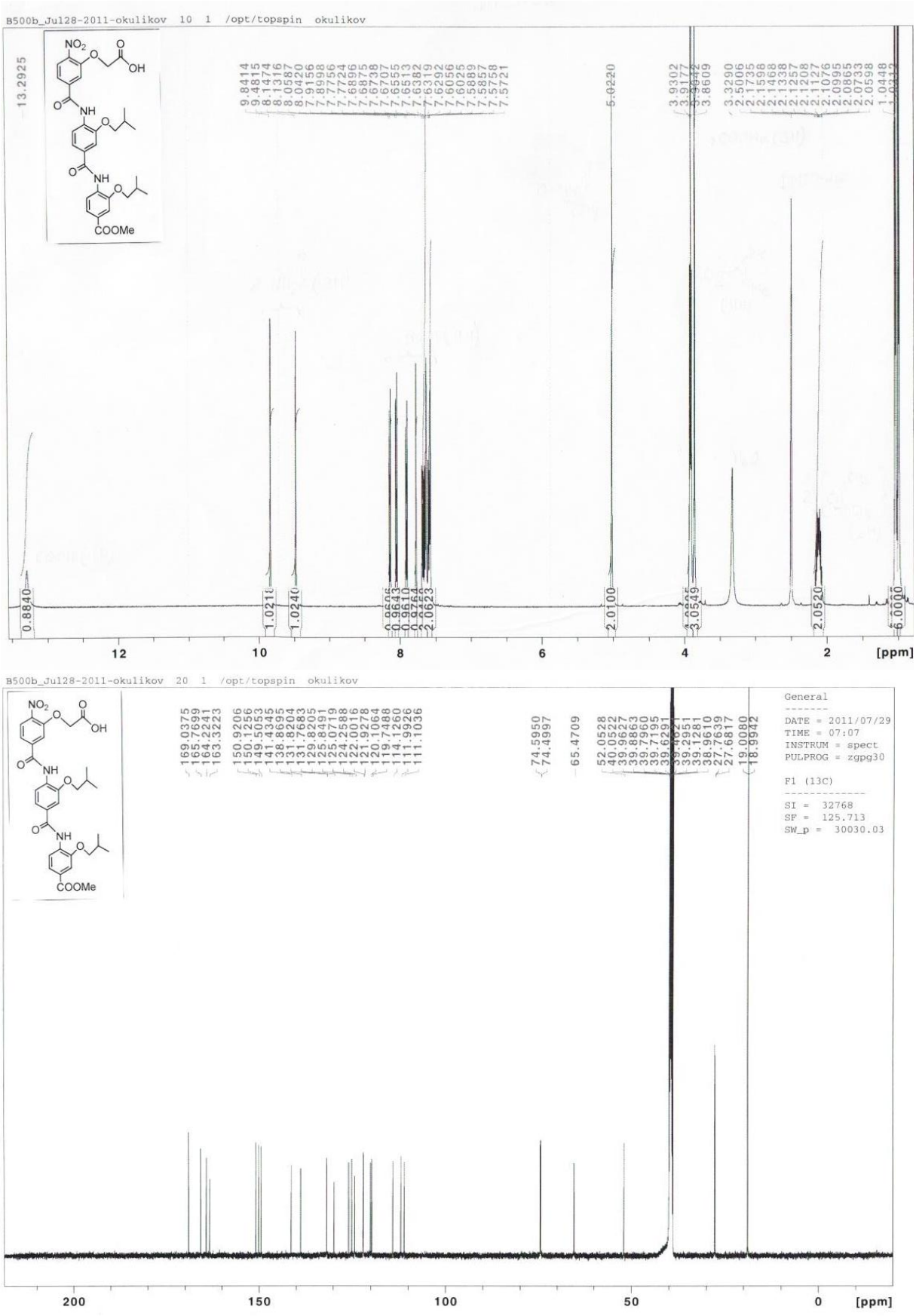


Figure S21. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **15** in DMSO-d_6

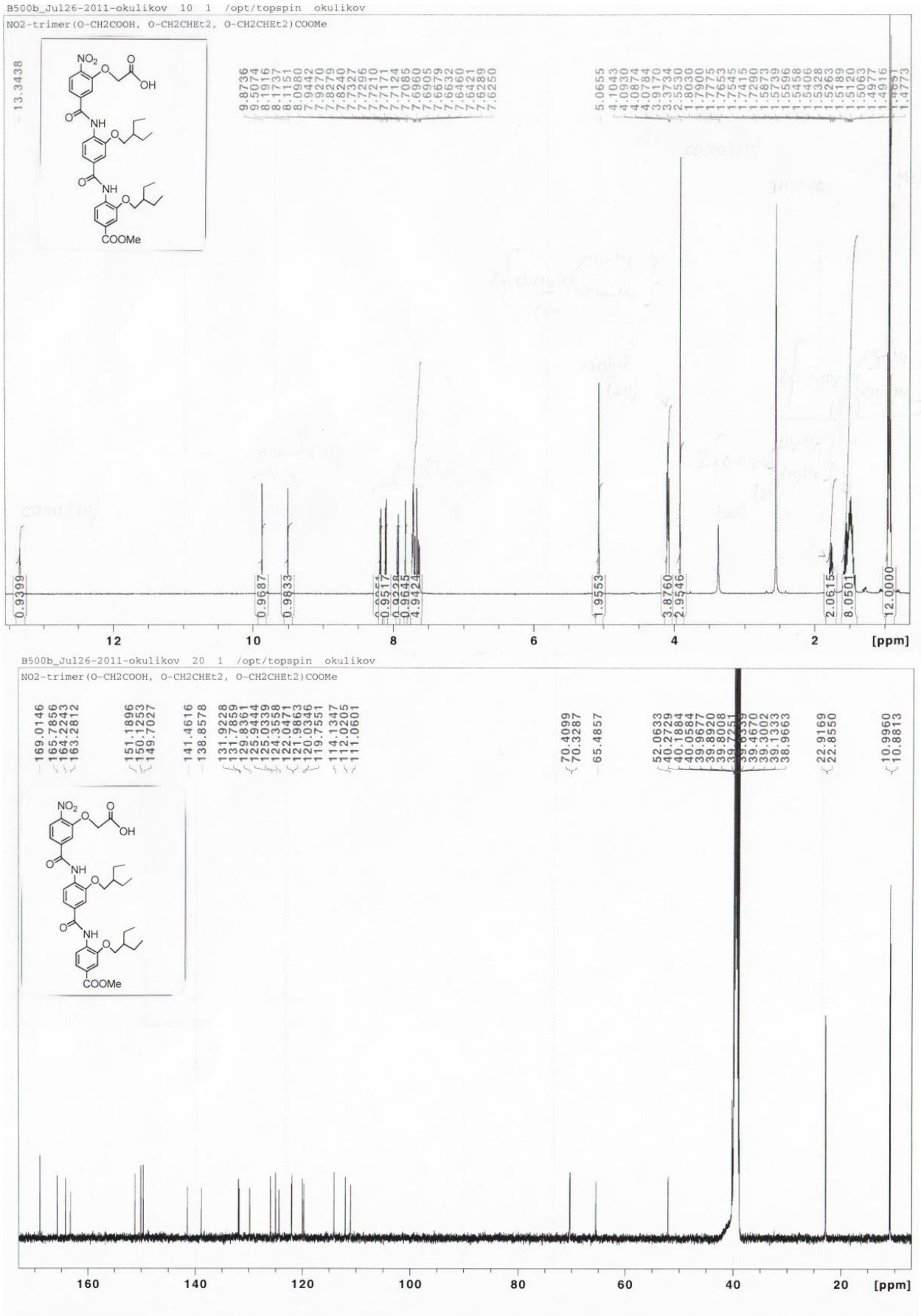


Figure S22. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **16** in DMSO-d_6

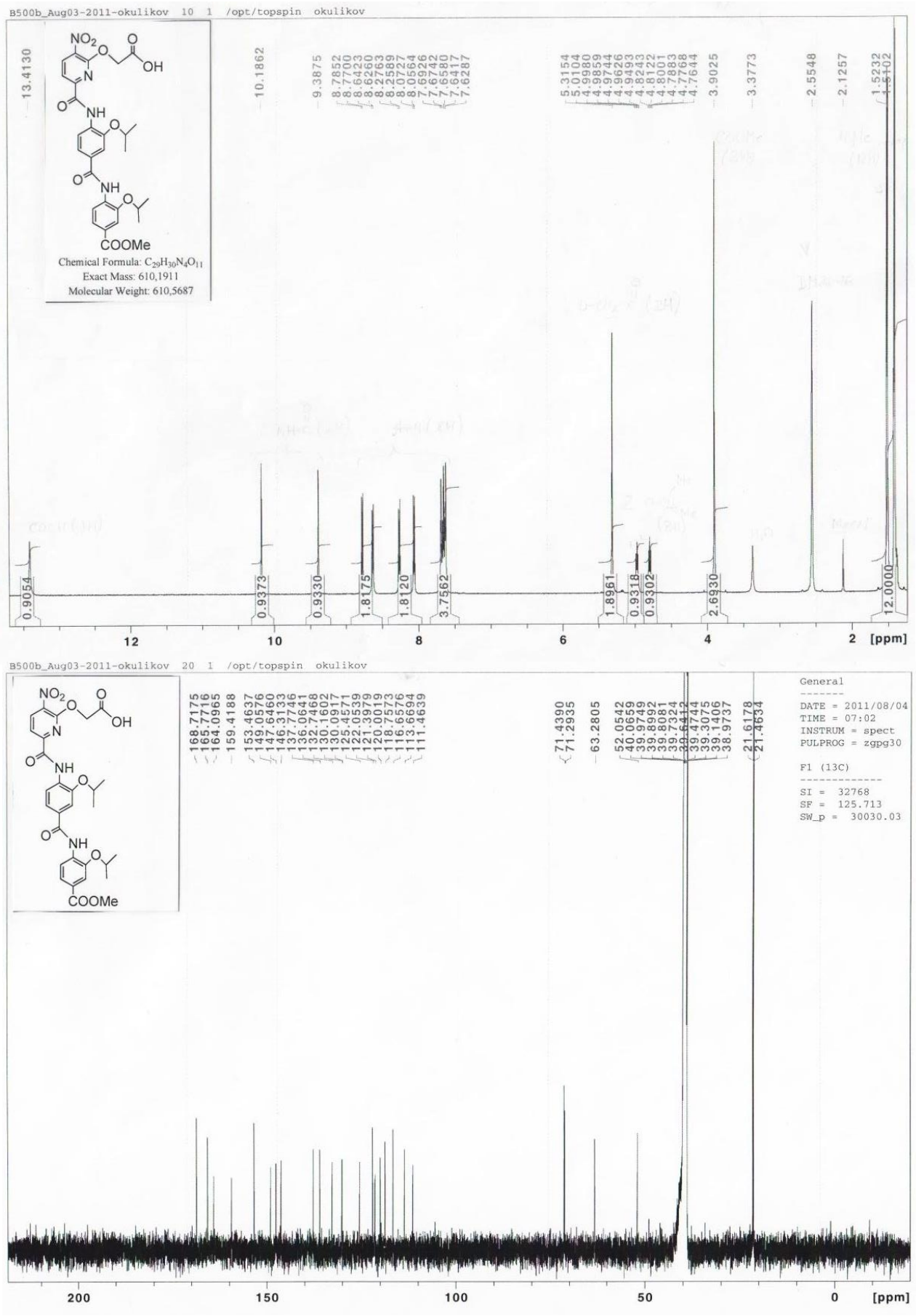


Figure S23. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **17** in DMSO-d_6

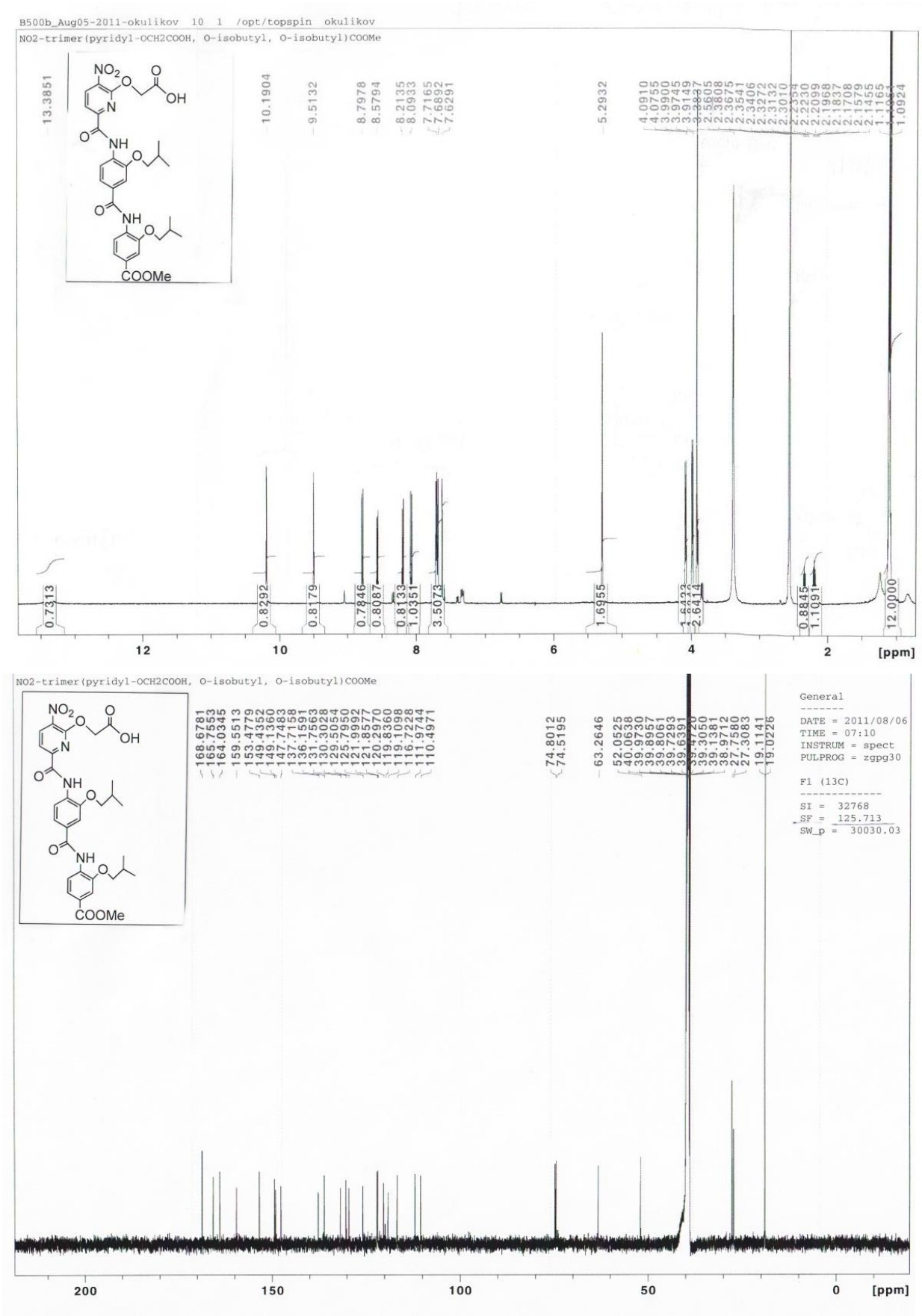


Figure S24. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **18** in DMSO-d_6

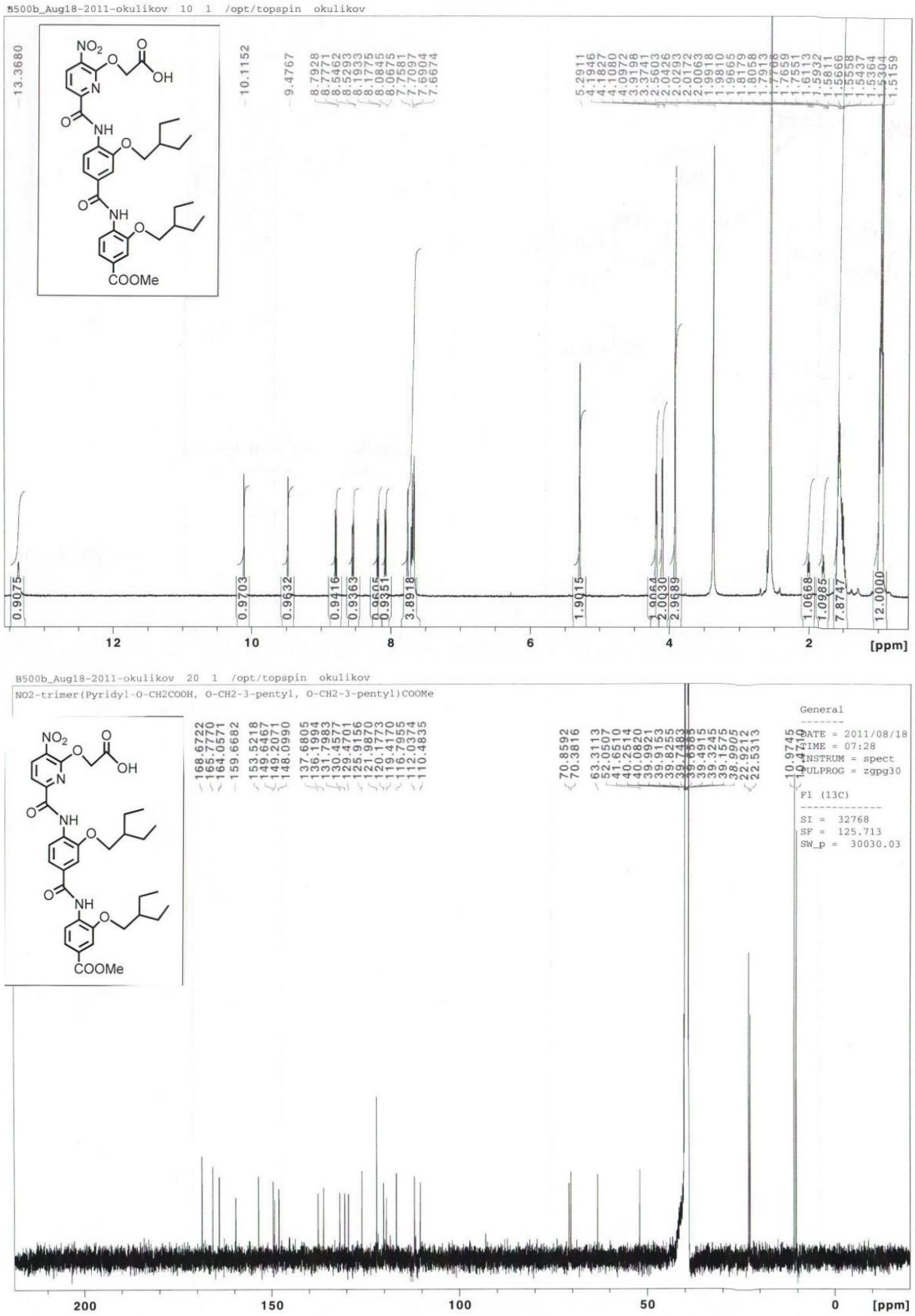


Figure S25. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **19** in DMSO-d_6

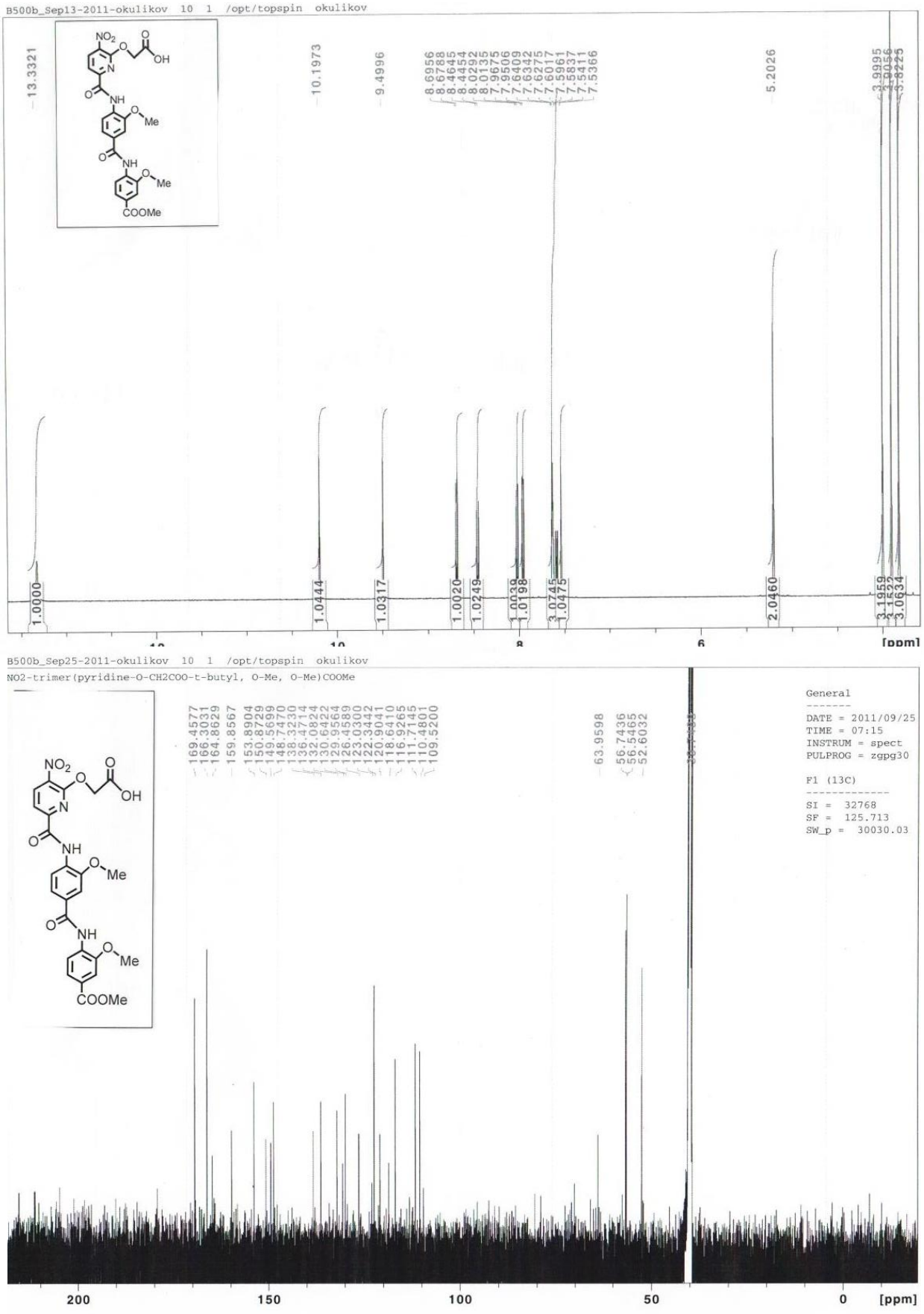


Figure S26. ^{1}H NMR (top, 500 MHz) and ^{13}C NMR (bottom, 125 MHz) spectra of **20** in DMSO-d_6

References:

(S1) CrystalStructure 3.8: Crystal Structure Analysis Package, Rigaku and Rigaku Americas (2000-2007). 9009 New Trails Dr. The Woodlands TX 77381 USA.

(S2) SHELX97: Sheldrick, G.M. (1997).

(S3) Z. Otwinowski and W. Minor, "Processing of X-Ray Diffraction Data Collected in Oscillation Mode," Methods in Enzymology, vol. 276: Macromolecular Crystallography, part A, 307-326, 1997, C.W. Carter, Jr. & R.M. Sweet, Eds., Academic Press.

(S4) *Acta Cryst.* **A46** (1990) 467-473.