Supplementary Material for:

Synthesis, Structural Investigation and Computational Modelling of Water-Binding Aquafoldamers

Huaiqing Zhao,[a] Wei Qiang Ong,[a] Xiao Fang,[a] Feng Zhou,[b] Meng Ni Hii,[a] Sam Fong Yau Li,[a] Haibin Su,[b] and Huaqiang Zeng*,[a]

[a] *Department of Chemistry, National University of Singapor, 3 Science Drive 3, Singapore. Fax: (+) 65-6779-1691; Tel: (+) 65-6516-2683.*

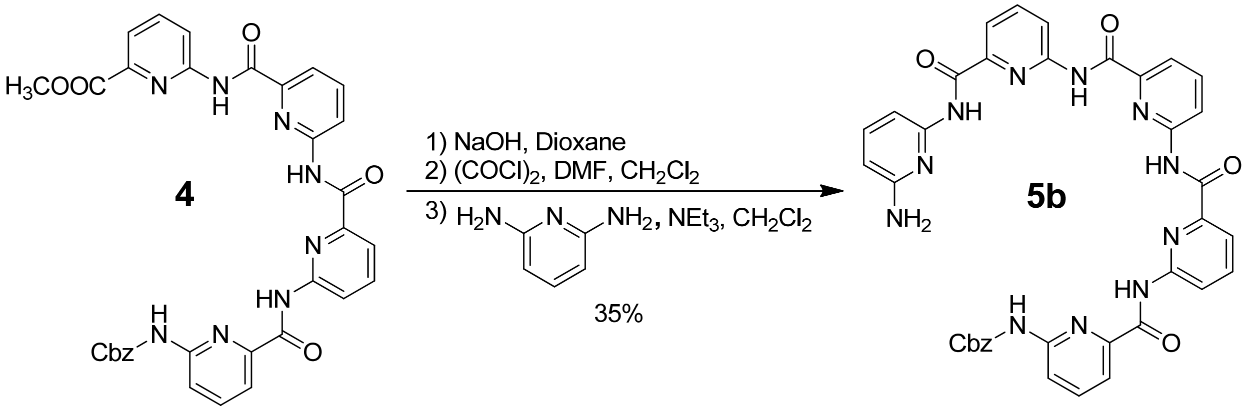
[b] *Division of Materials Science, 50 Nanyang Avenue, Nanyang Technological University, Singapore 639798.*

Corresponding Author: chmzh@nus.edu.sg

1. General Remarks

All the reagents were obtained from commercial suppliers and used as received unless otherwise stated. Aqueous solutions were prepared from distilled water. The organic solutions from all liquid extractions were dried over anhydrous sodium sulphate for a minimum of 15 minutes before filtration. Reactions were monitored by thin-layer chromatography (TLC) on silica gel pre-coated glass plates (0.25 mm thickness, 60F-254, E., Merck). Chemical yields refer to pure isolated substances. Mass spectra were obtained using instrumentation which includes Finnigan MAT95XL-T and Micromass VG7035. Single crystal X-ray diffraction study was recorded on Bruker-AXS Smart Apex CCD single crystal diffractometer. ^{1}H NMR, ^{1}H variable temperature (VT) NMR and ^{13}C NMR spectra were recorded on a Bruker AMX500 (500 MHz) spectrometer. The solvent signal of $CDCl_3$ in ^{1}H NMR was referenced at $\delta = 7.26$ ppm. Coupling constants (*J* values) are reported in Hertz (Hz). ^{1}H NMR data are recorded in the order: chemical shift value, multiplicity (s: singlet; d: doublet; t: triplet; q: quartet; m: multiplet; br: broad), coupling constant and number of protons that gave rise to the signal, where applicable. ^{13}C NMR spectra are proton-decoupled and the solvent peaks of $CDCl_3$ was referenced at $\delta = 77.0$ ppm. 2D NOESY were recorded either on a Bruker AMX500 (500 MHz) or Bruker DRX500 (500 MHz) spectrometer. $CDCl_3$ was purchased from Cambridge Isotope Laboratories, Inc. and used without further purification unless otherwise stated. The normal $CDCl_3$ used in the ^{1}H VT NMR refers to the $CDCl_3$ that was used directly from the bottle. Dry $CDCl_3$ used in the ^{1}H VT NMR was prepared by standing over 4A molecular sieves before use. “Wetted” $CDCl_3$ used in the ^{1}H VT NMR was prepared by adding 1 drop of water into 1 ml of $CDCl_3$ and allowing it to equilibrate for 2 days before use.

2. Scheme S1. Synthetic Route that Affords Pentamer 5b

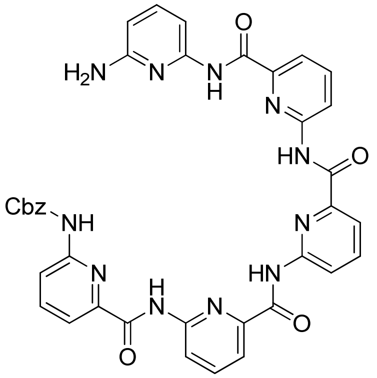


3. Synthetic Procedures & Characterization

The synthesis of the oligomers **1–5a** had been reported in the following journal articles:

(1) Ong, W. Q.; Zhao, H.; Du, Z.; Yeh, J. Z. Y.; Ren, C.; Tan, L. Z. W.; Zhang, K.; Zeng, H. *Chem. Commun.* **2011**, *47*, 6416.

(2) Ong, W. Q.; Zhao, H.; Fang, X.; Woen, S.; Zhou, F.; Yap, W.; Su, H.; Li, S. F. Y.; Zeng, H. *Org. Lett.* **2011**, *13*, 3194.



Pentamer **5b**:

Solid NaOH (0.16 g, 4.0 mmol) was dissolved in minimal amount of deionized water and was then added into the round bottom flask containing tetramer **4** (1.29 g, 2.0 mmol) in dioxane (20 ml). The mixture was stirred at room temperature overnight and the solvent was then removed *in vacuo*. Water (20 ml), MeOH (20 ml) and solid $KHSO_4$ (0.54 g, 4.0 mmol) was then added. The suspension was then filtered, washed and the residue obtained was dried to give the acid intermediate. The acid intermediate (0.63 g, 1.0 mmol) was dissolved in dry dichloromethane (10ml) in a rbf. DMF (0.1 ml) was added, followed by the dropwise addition of oxalyl chloride (0.3 ml, 2.0 mmol) into the rbf. The reaction mixture was stirred for 2 hrs and the solvent and excess oxalyl chloride were removed *in vacuo* and dry dichloromethane (20 ml) was added to the acid chloride in nitrogen atmosphere. Minimal amount of dry dichloromethane was added to a mixture containing 2,6-diaminopyridine (0.22 g, 2.0 mmol) and triethylamine (0.28 ml, 2.0 mmol) and the mixture was injected into the acid chloride. The mixture was allowed to stir overnight at room temperature and after the reaction, the product was washed with water (2 x 10ml), 1M HCl (2 x 10ml) and 1M NaOH (2 x 10ml). The solvent was removed *in vacuo* to give the crude product and the flash column chromatography was used to afford the pure product as a white solid (Yield: 0.25 g, 35 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 10.48 (s, 2H), 10.36 (s, 1H), 10.04 (s, 1H), 8.73–8.77 (m, 2H), 8.65 (d, 1H, *J* = 8.2 Hz), 8.13–8.16 (m, 2H), 7.94–8.07 (m, 5H), 7.90 (t, 1H, *J* = 7.6 Hz), 7.78 (t, 1H, *J* = 7.9 Hz), 7.73 (d, 1H, *J* = 8.2 Hz), 7.57 (s, 1H), 7.36 (t, 1H, *J* = 8.1 Hz), 7.06–7.17 (m, 5H), 5.88 (d, 1H, *J* = 7.5 Hz), 5.03 (s, 2H), 3.88 (s, 2H). ^{13}C NMR (125 MHz, DMSO-d_6, 50 °C): δ 162.76, 162.51, 162.31, 161.67, 157.12, 152.41, 150.66, 150.26, 150.22, 149.90, 149.57, 148.10, 147.66, 147.62, 147.04, 140.23, 140.11, 140.03, 139.97, 139.54, 135.63, 128.37, 128.10, 127.86, 119.06, 118.94, 118.55, 117.87, 117.69, 117.60, 116.91, 115.94, 104.21, 103.32, 67.36. HRMS-ESI: calculated for $[M+Na]^+$ ($C_{37}H_{29}N_{11}O_6{}^{23}Na$): *m/z* 746.2194, found: *m/z* 746.2188.

4. *Ab Initio* Molecular Modeling

All the calculations were carried out by utilizing the either the Gaussian 03[1] or Gaussian 09[2] program package. The geometry optimizations were performed at the density functional theory (DFT) level, and the Becke's three parameter hybrid functional with the Lee-Yang-Parr correlation functional (B3LYP)[3] method was employed to do the calculations. Unless otherwise stated, the 6-311+G(2d, p)[4] basic from the Gaussian basis set library has been used in all the structural optimization and single point energy calculations. The harmonic vibrational frequencies and zero-point energy corrections were calculated at the same level of theory.

5. Theoretical Computations by dreiding force field[5]

In this paper, the dreiding force field [1] is used to optimize the geometry and calculate the energy of all the structural motifs taken from the respective crystal structures as tabulated in Table 1. The convergence tolerance is $2X10^{-5}$ kcal/mol for the energy, 0.001 kcal/mol/A for the force, 0.001 GPa for the stress and 10^{-5} A for the displacement. The Ewald method is used for calculating the electrostatic and the van der Waals terms. The accuracy is 10^{-5} kcal/mol. The repulsive cutoff is 6 Å for the van der Waals term. For the hydrogen bond term, the summation method is atom based and the truncation method is cubic spline with cutoff distance of 4.5 Å. For the isolated helical pentamer or dimer made up of two pentamers, these molecules are put in a three dimensional cubic box with every lattice length to be 100 Å to avoid the interaction of the molecules with their mirror counterparts. During the optimization the box is fixed and only the molecules are optimized for the isolated case. For the 1D or 2D cases, where the molecules repeat periodically in one or two directions, the box vectors are also optimized together with the molecules.

1. Frisch, M. J.; et al. *Gaussian 03*; Gaussian , Inc.: Wallingford CT, 2004.
2. Frisch, M. J.; et al. *Gaussian 09*; Gaussian , Inc.: Wallingford CT, 2009.
3. Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648.
4. Petersson, G. A.; Al-Laham, M. A. *J. Chem. Phys.* **1991***,* 94, 6081; Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. *J. Chem. Phys.* **1988***,* 89, 2193.
5. Mayo, S. L., Olafson, B. D. & Goddard III, W. A. "Dreiding - A generic force-field for molecular simulations," *J. Phys. Chem.*, **94**, 8897-909 (1990).

6. X-Ray Crystal Data

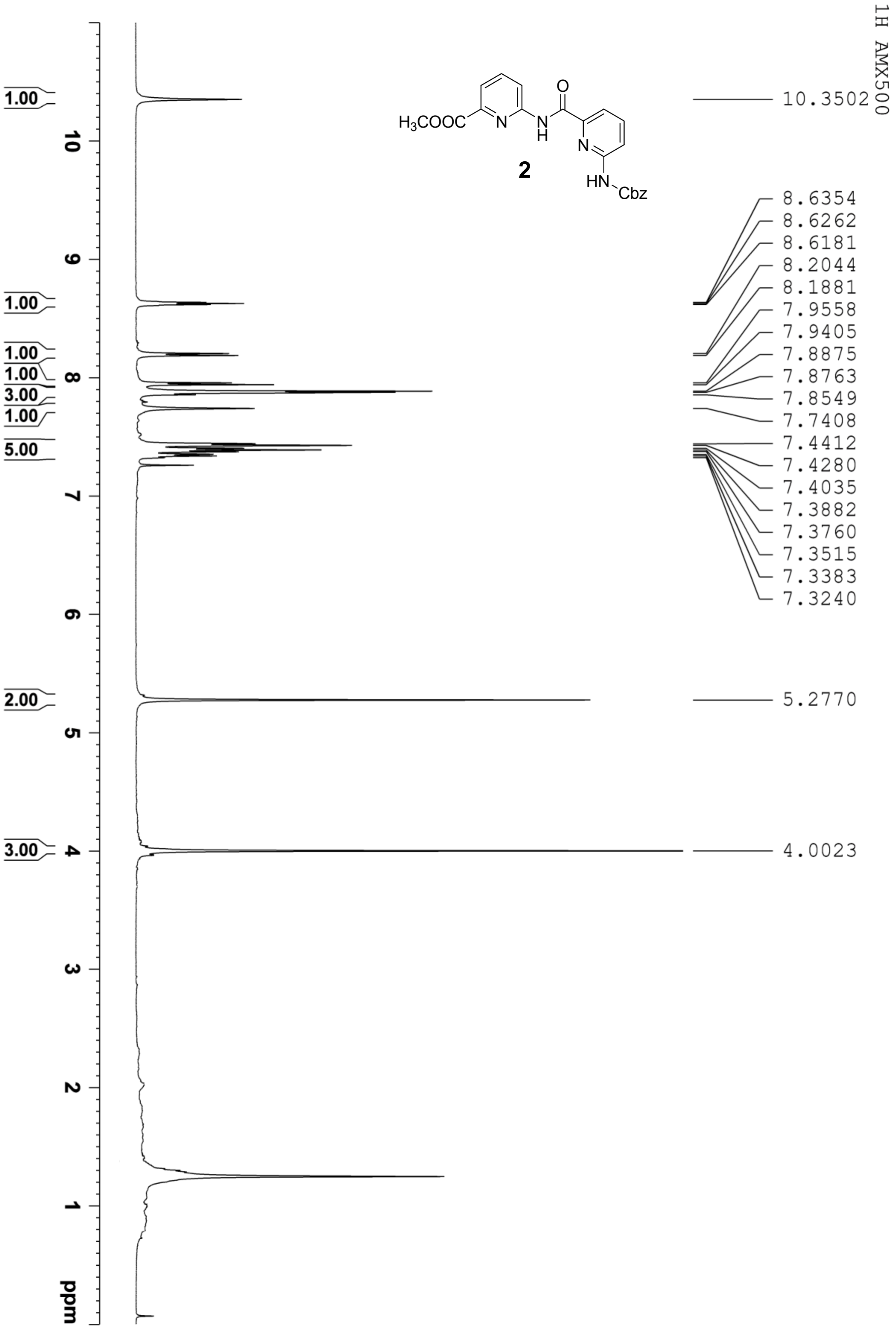
(a) Pentamer 5b

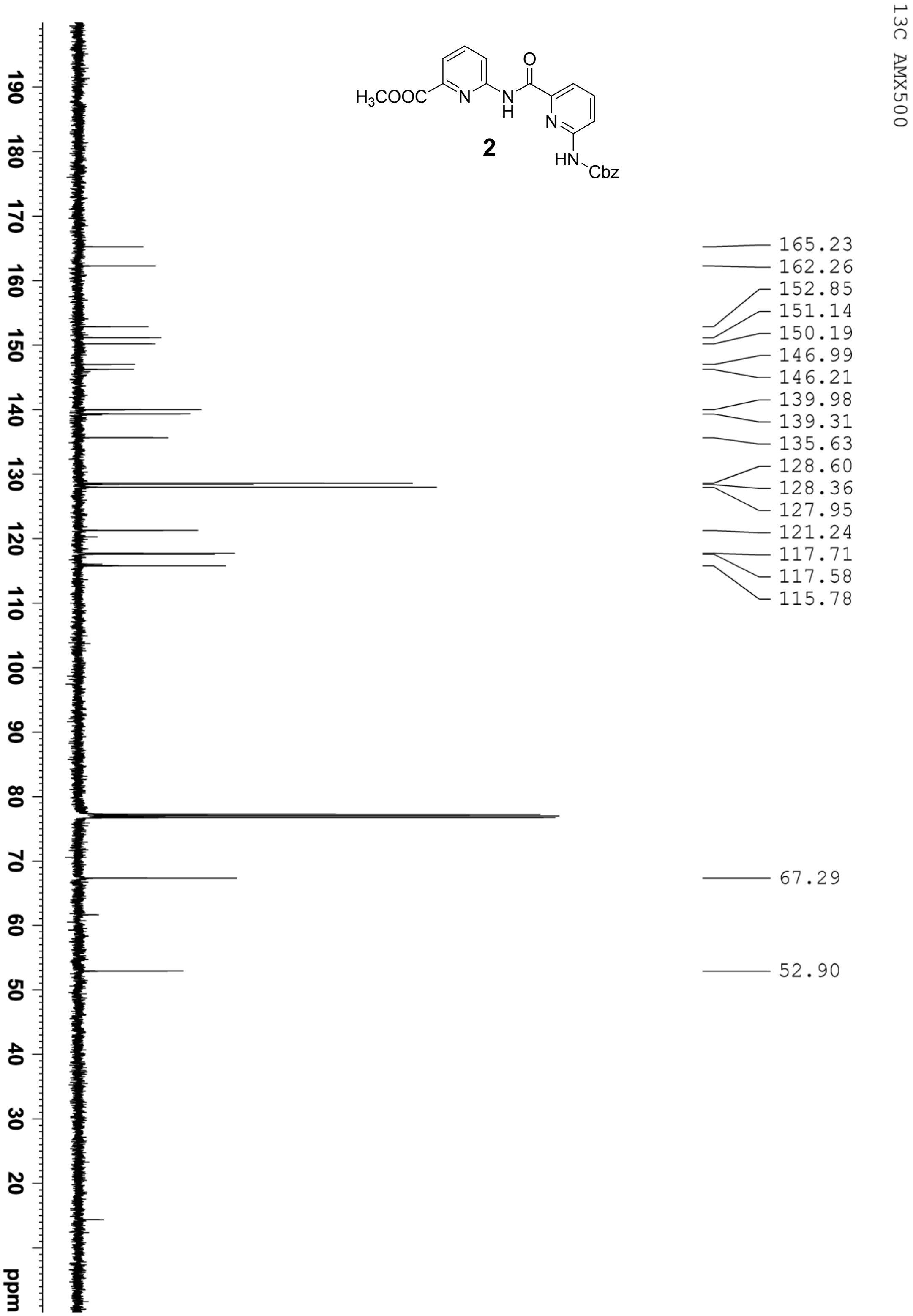
CCDC	834710	
Empirical formula	$C_{37} H_{33} N_{11} O_8$	
Formula weight	759.74	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Triclinic	
Space group	P-1	
Unit cell dimensions	a = 9.2092(10) Å	α= 90.884(3)°.
	b = 13.6899(14) Å	β= 99.028(3)°.
	c = 14.6484(16) Å	γ = 103.338(3)°.
Volume	1772.1(3) $Å^3$	
Z	2	
Density (calculated)	1.424 Mg/m^3	
Absorption coefficient	0.104 mm^{-1}	
F(000)	792	
Crystal size	0.40 x 0.30 x 0.24 mm^3	
Theta range for data collection	1.53 to 27.50°.	
Index ranges	-11<=h<=11, -17<=k<=16, -18<=l<=15	
Reflections collected	12748	
Independent reflections	8020 [R(int) = 0.0301]	
Completeness to theta = 27.50°	98.8 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.9755 and 0.9596	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	8020 / 4 / 549	
Goodness-of-fit on F^2	0.961	
Final R indices [I>2sigma(I)]	R1 = 0.0530, wR2 = 0.1169	
R indices (all data)	R1 = 0.0885, wR2 = 0.1292	
Largest diff. peak and hole	0.378 and -0.238 e.$Å^{-3}$	

(b) Other oligomers have been reported with following CCDC numbers:

Dimer 2: CCDC# 817347; **Trimer 3b:** CCDC# 821681; **Tetramer 4:** CCDC# 817349; **Pentamer 5a:** CCDC# 817350

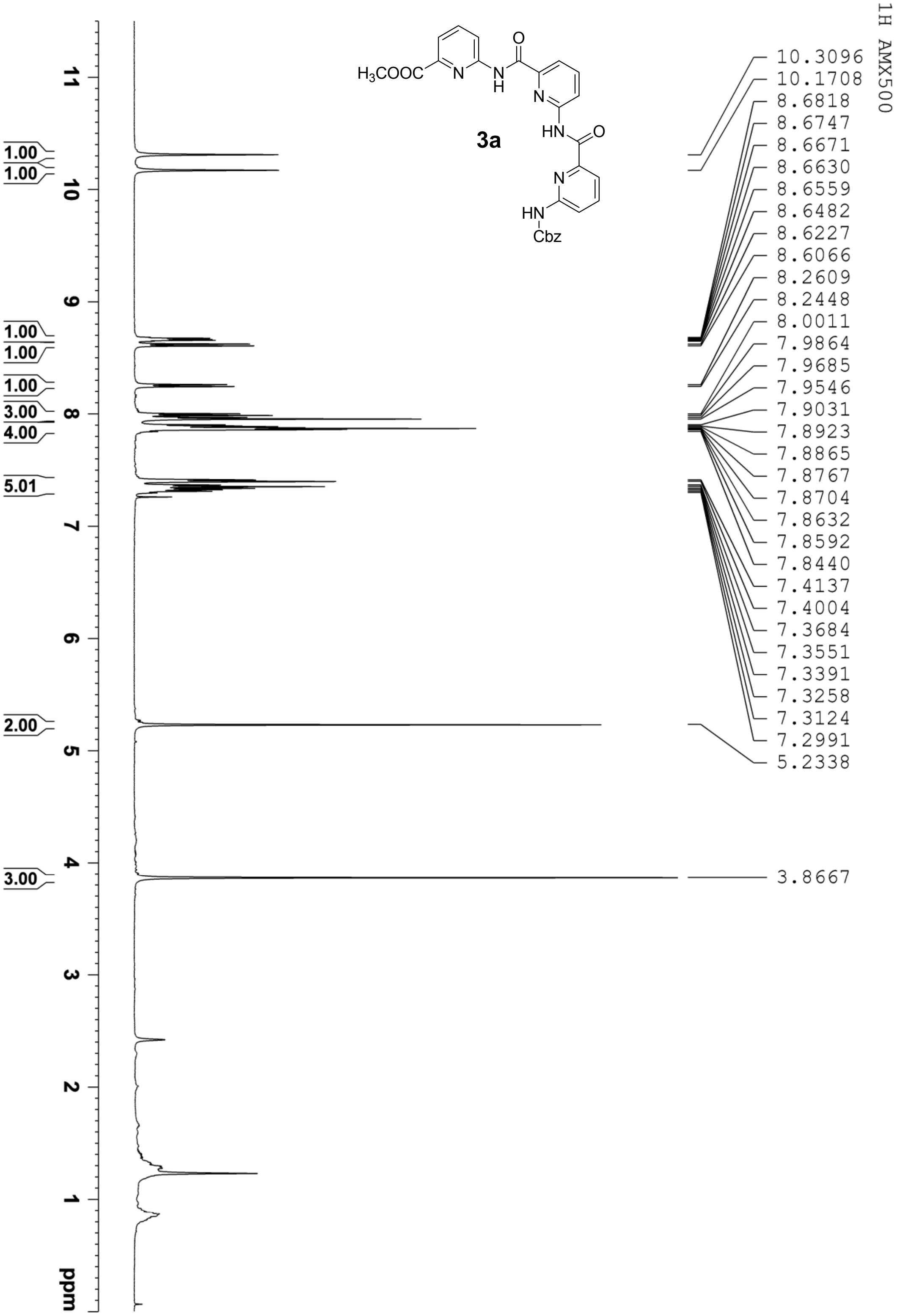
7. ^{1}H and ^{13}C NMR spectra of Oligomers 2–5b

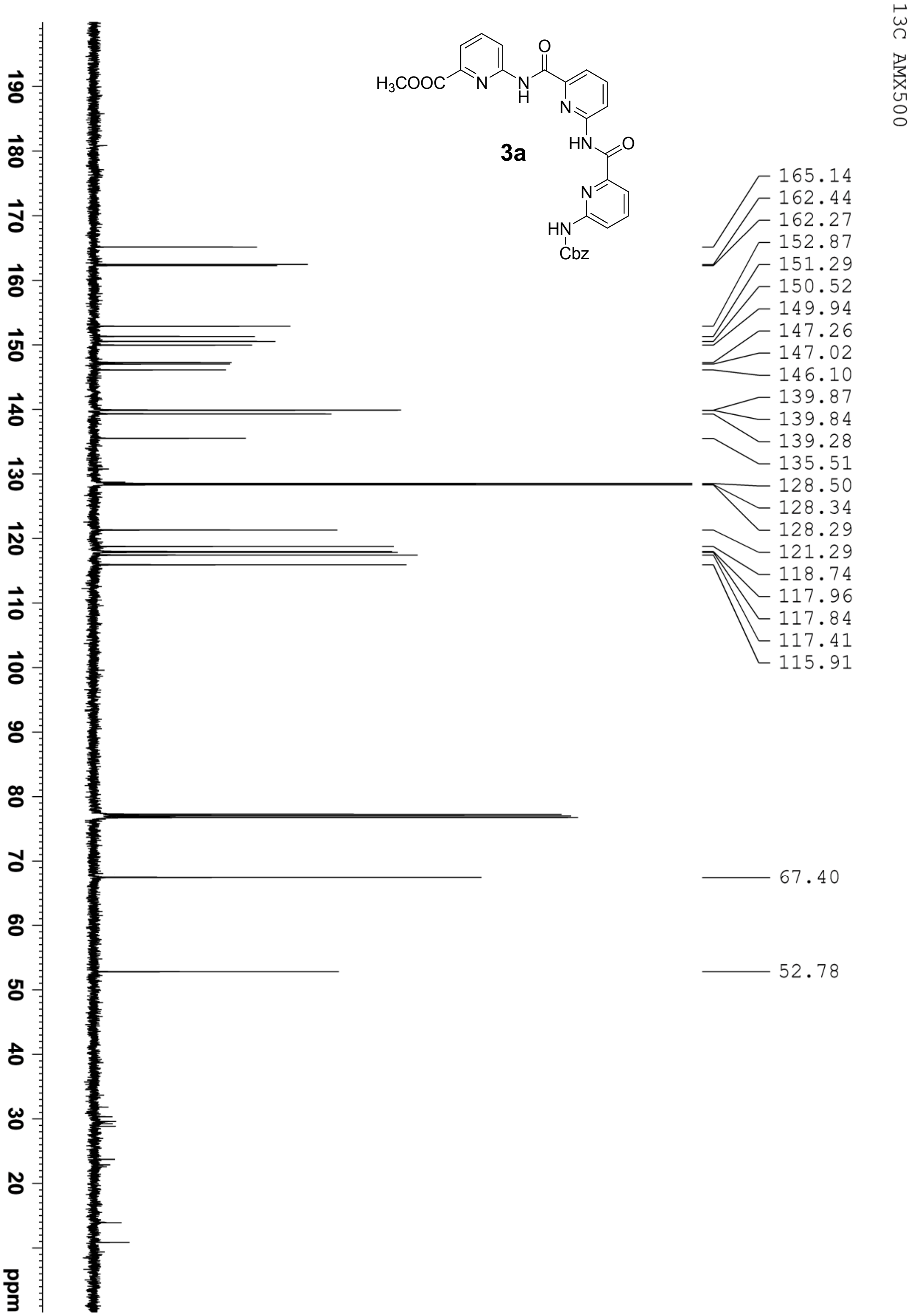


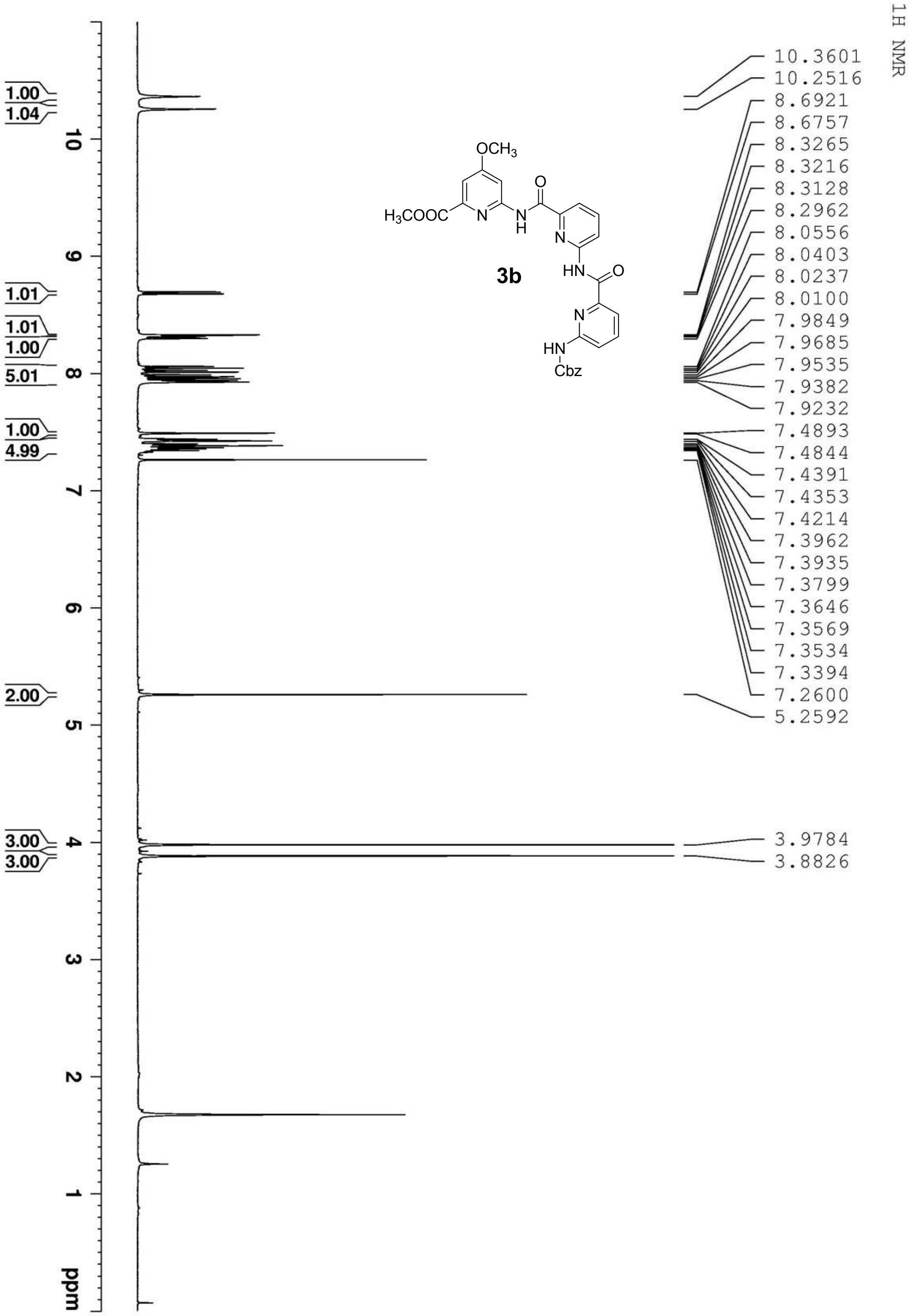


2

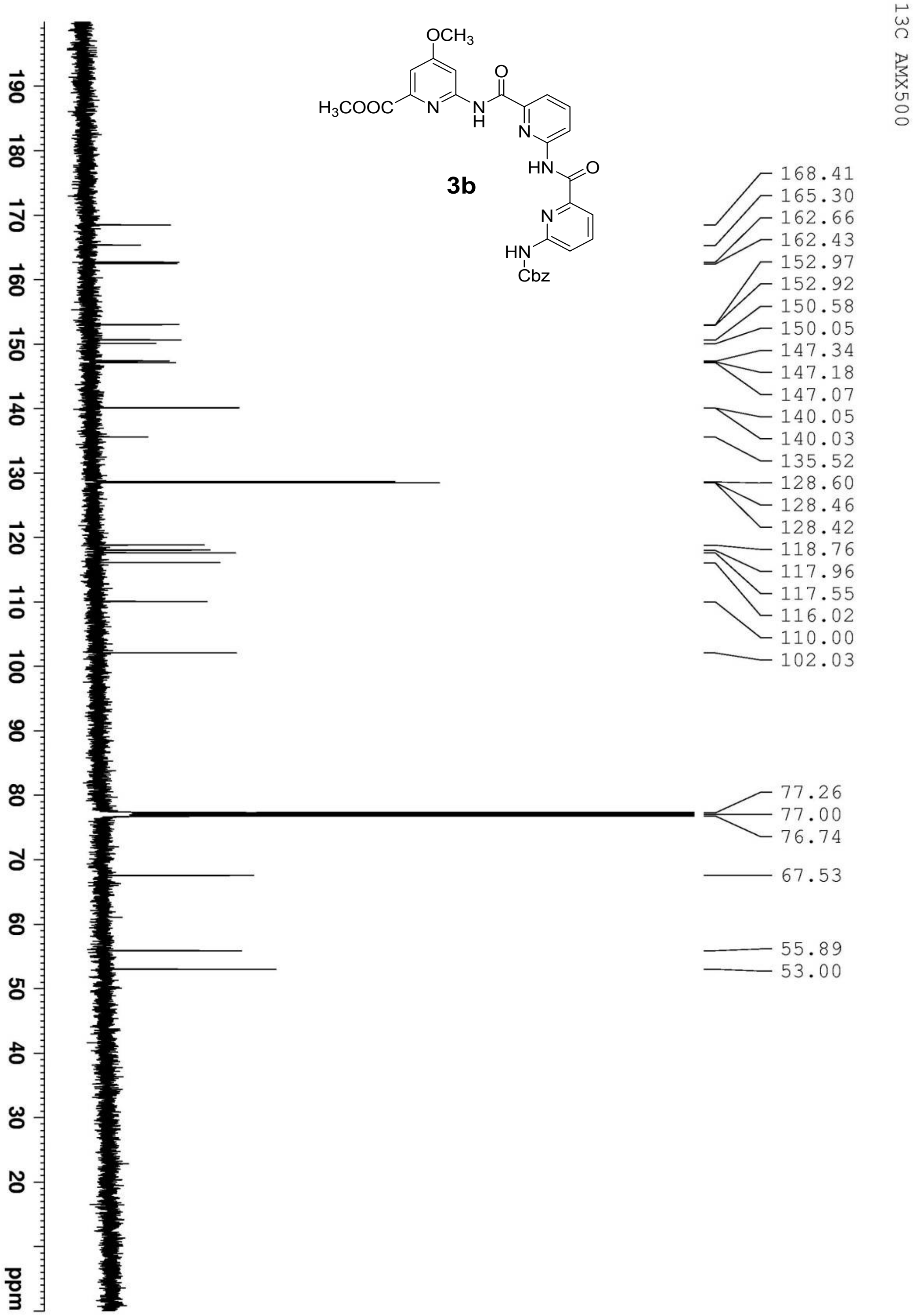
13C AMX500
H3COOC
N
N
H
O
N
HN
Cbz
165.23
162.26
152.85
151.14
150.19
146.99
146.21
139.98
139.31
135.63
128.60
128.36
127.95
121.24
117.71
117.58
115.78
67.29
52.90
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm


1H AMX500
H_3COOC
N
H
O
HN
Cbz
3a
10.3096
10.1708
8.6818
8.6747
8.6671
8.6630
8.6559
8.6482
8.6227
8.6066
8.2609
8.2448
8.0011
7.9864
7.9685
7.9546
7.9031
7.8923
7.8865
7.8767
7.8704
7.8632
7.8592
7.8440
7.4137
7.4004
7.3684
7.3551
7.3391
7.3258
7.3124
7.2991
5.2338
3.8667
1.00
1.00
1.00
1.00
1.00
3.00
4.00
5.01
2.00
3.00
11
10
9
8
7
6
5
4
3
2
1
ppm

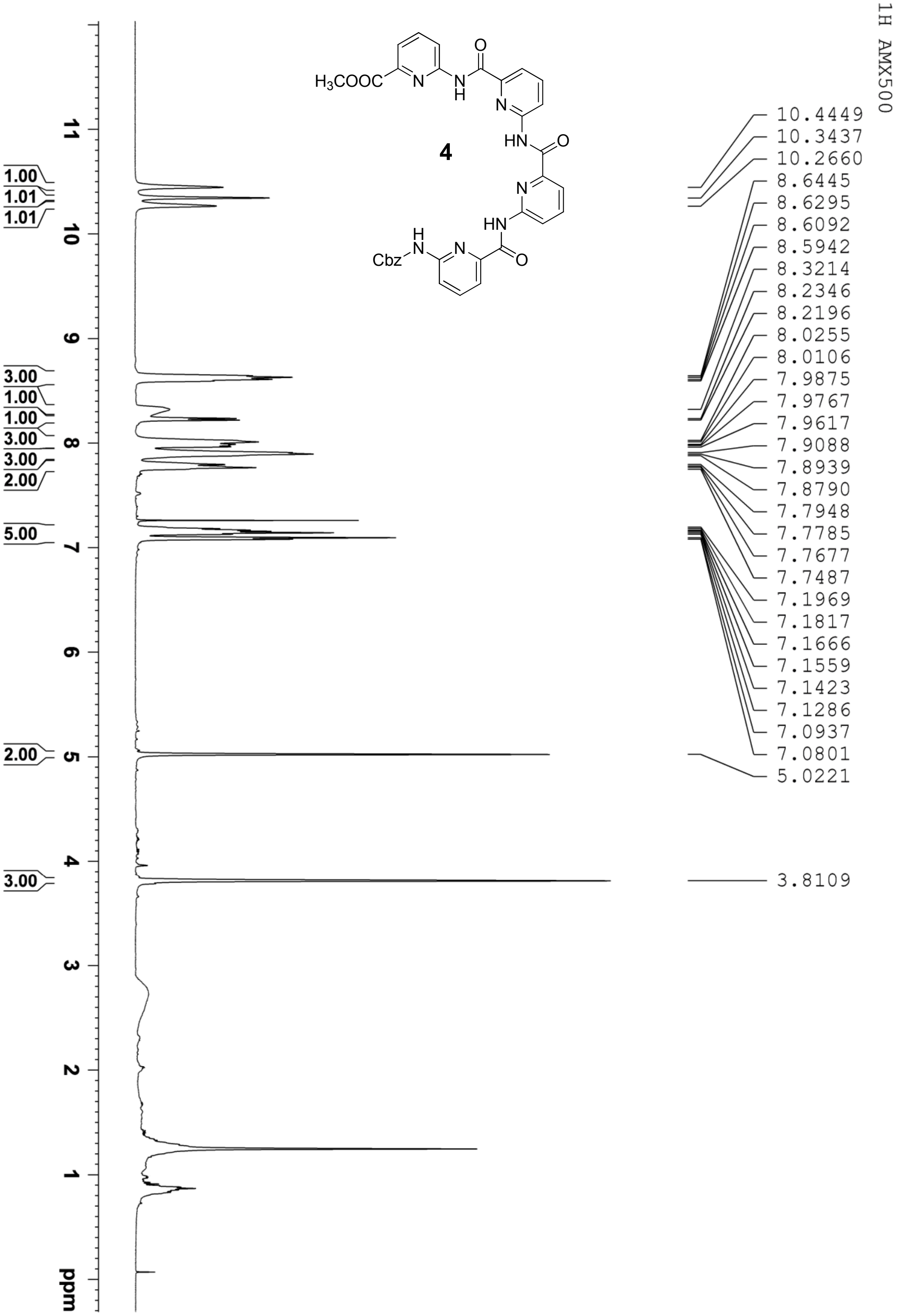

13C AMX500
H_3COOC
N
N
H
O
N
HN
O
N
HN
Cbz
3a
165.14
162.44
162.27
152.87
151.29
150.52
149.94
147.26
147.02
146.10
139.87
139.84
139.28
135.51
128.50
128.34
128.29
121.29
118.74
117.96
117.84
117.41
115.91
67.40
52.78
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm

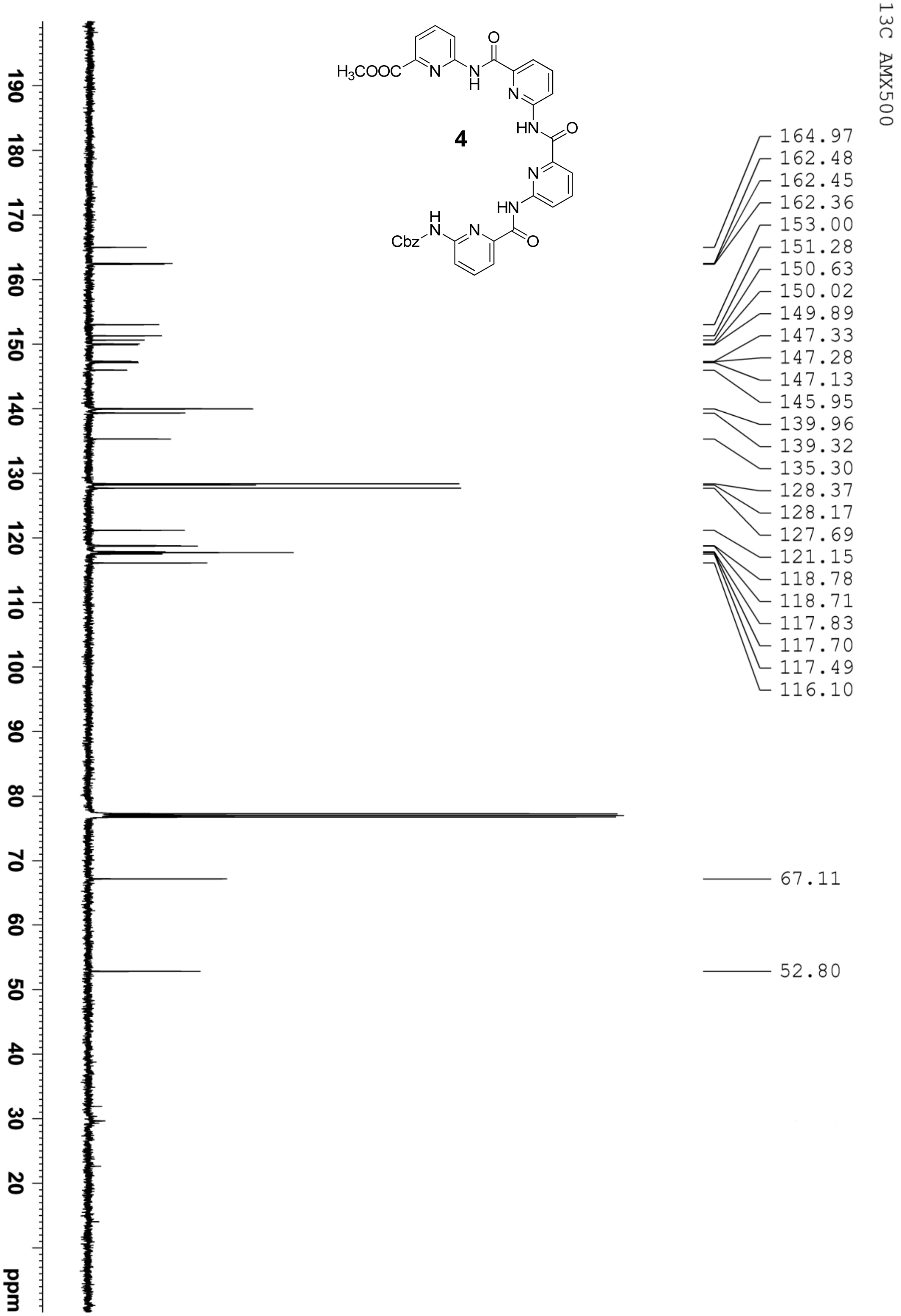

1H NMR
10.3601
10.2516
8.6921
8.6757
8.3265
8.3216
8.3128
8.2962
8.0556
8.0403
8.0237
8.0100
7.9849
7.9685
7.9535
7.9382
7.9232
7.4893
7.4844
7.4391
7.4353
7.4214
7.3962
7.3935
7.3799
7.3646
7.3569
7.3534
7.3394
7.2600
5.2592
3.9784
3.8826
1.00
1.04
1.01
1.01
1.00
5.01
1.00
4.99
2.00
3.00
3.00
10
9
8
7
6
5
4
3
2
1
ppm
OCH3
H3COOC
N
N
H
O
N
HN
O
N
HN
Cbz

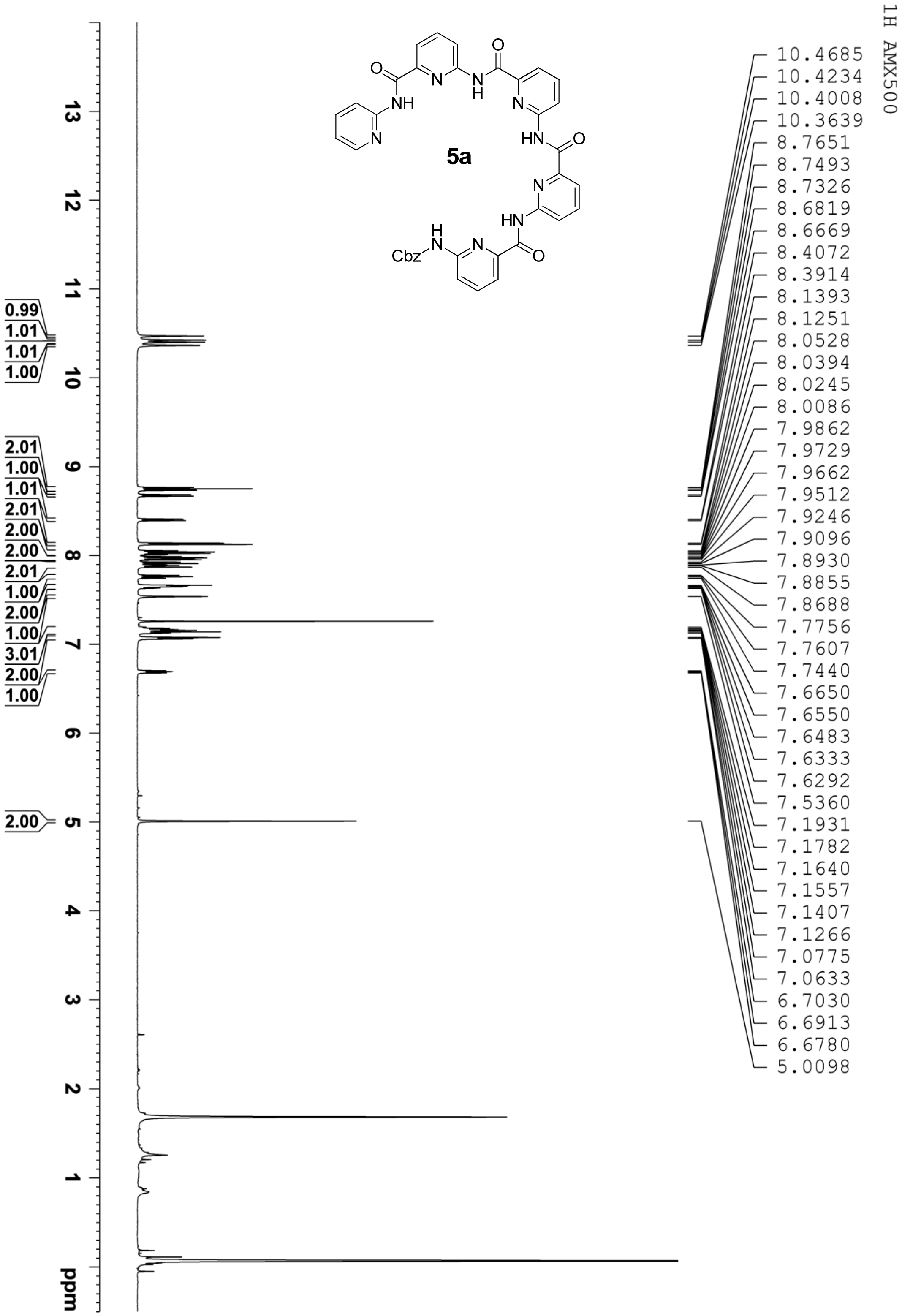
3b

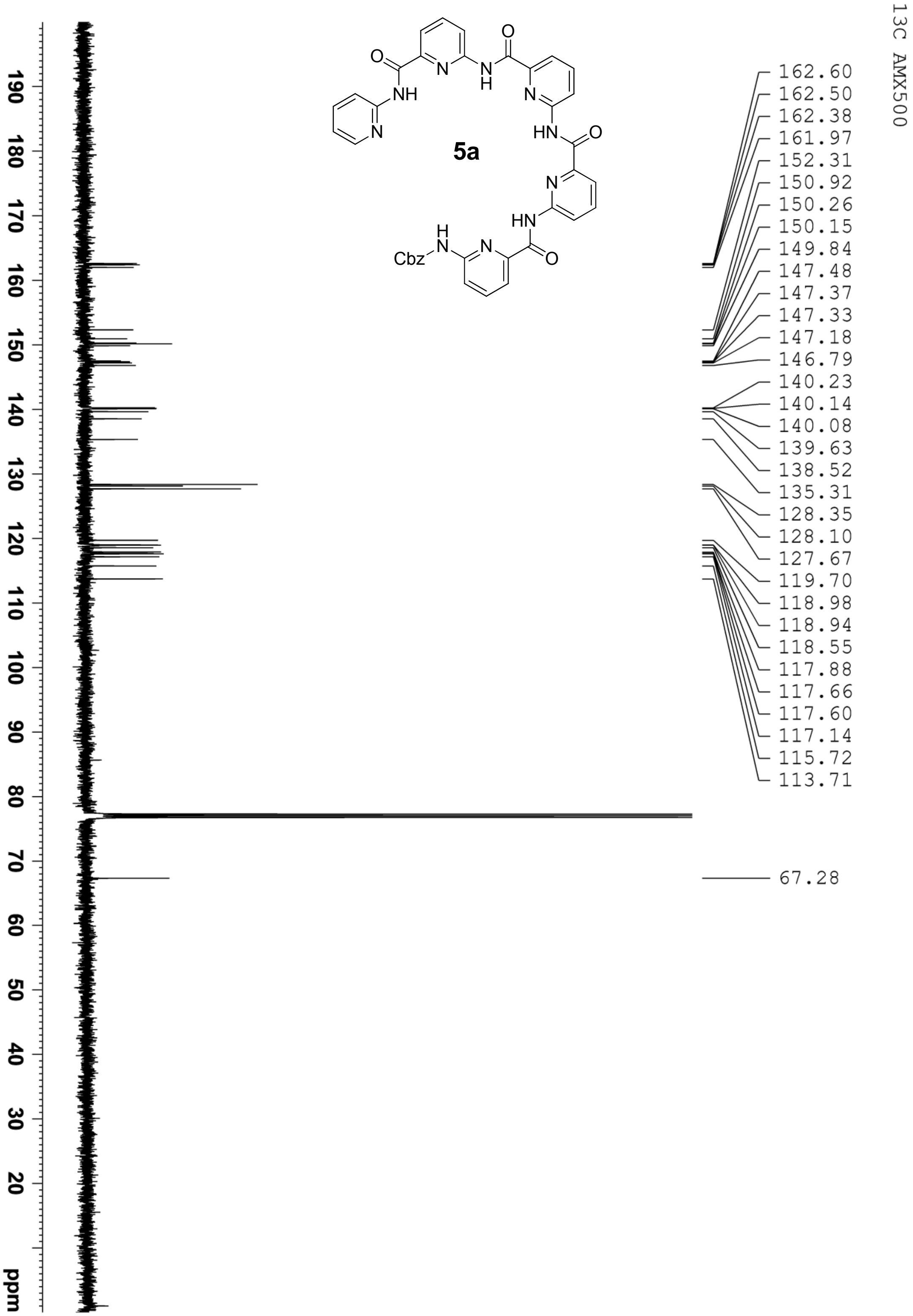

3b

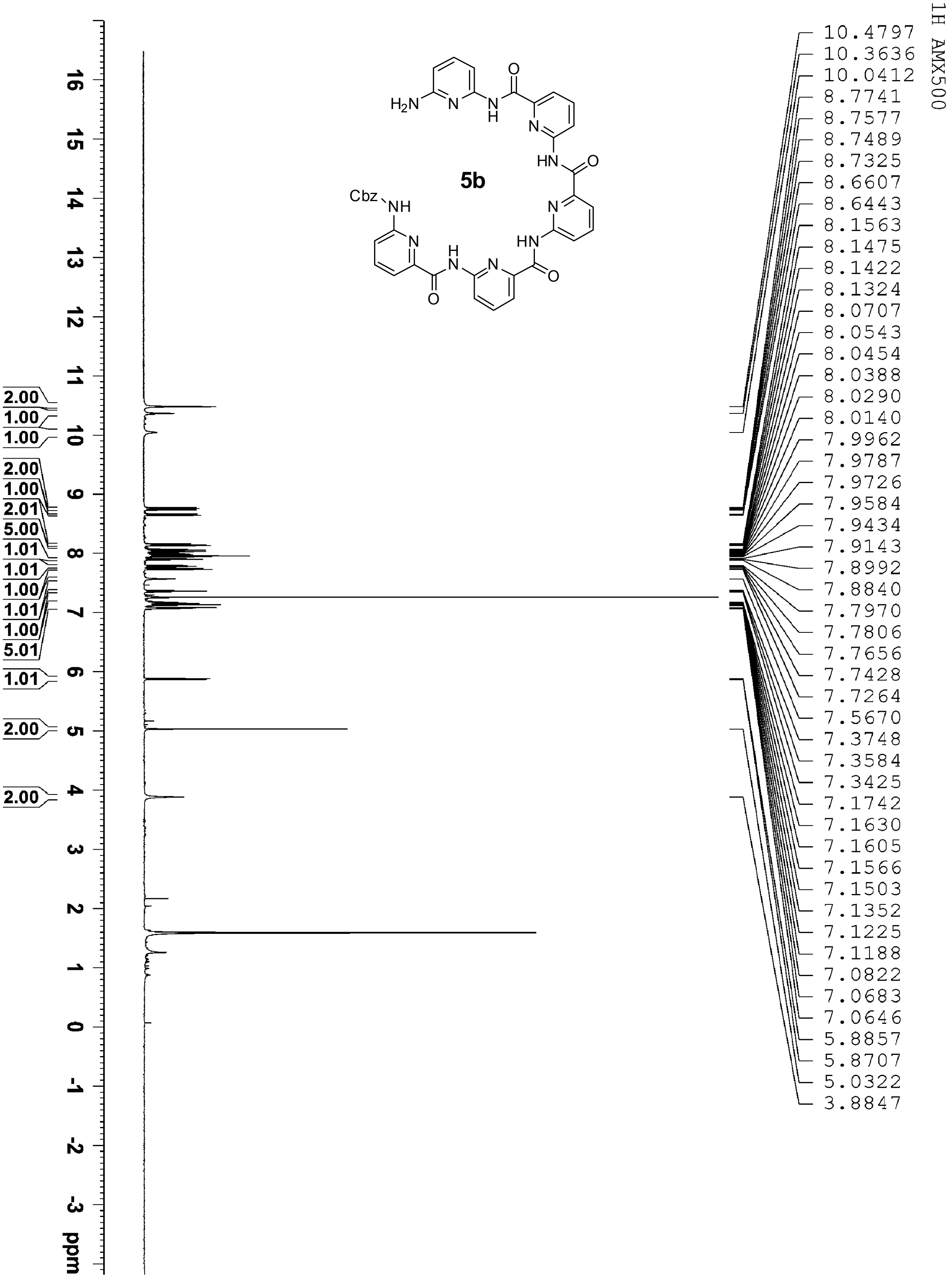
OCH3
O
H3COOC
N
N
H
N
HN
O
N
HN
Cbz
13C AMX500
168.41
165.30
162.66
162.43
152.97
152.92
150.58
150.05
147.34
147.18
147.07
140.05
140.03
135.52
128.60
128.46
128.42
118.76
117.96
117.55
116.02
110.00
102.03
77.26
77.00
76.74
67.53
55.89
53.00
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm

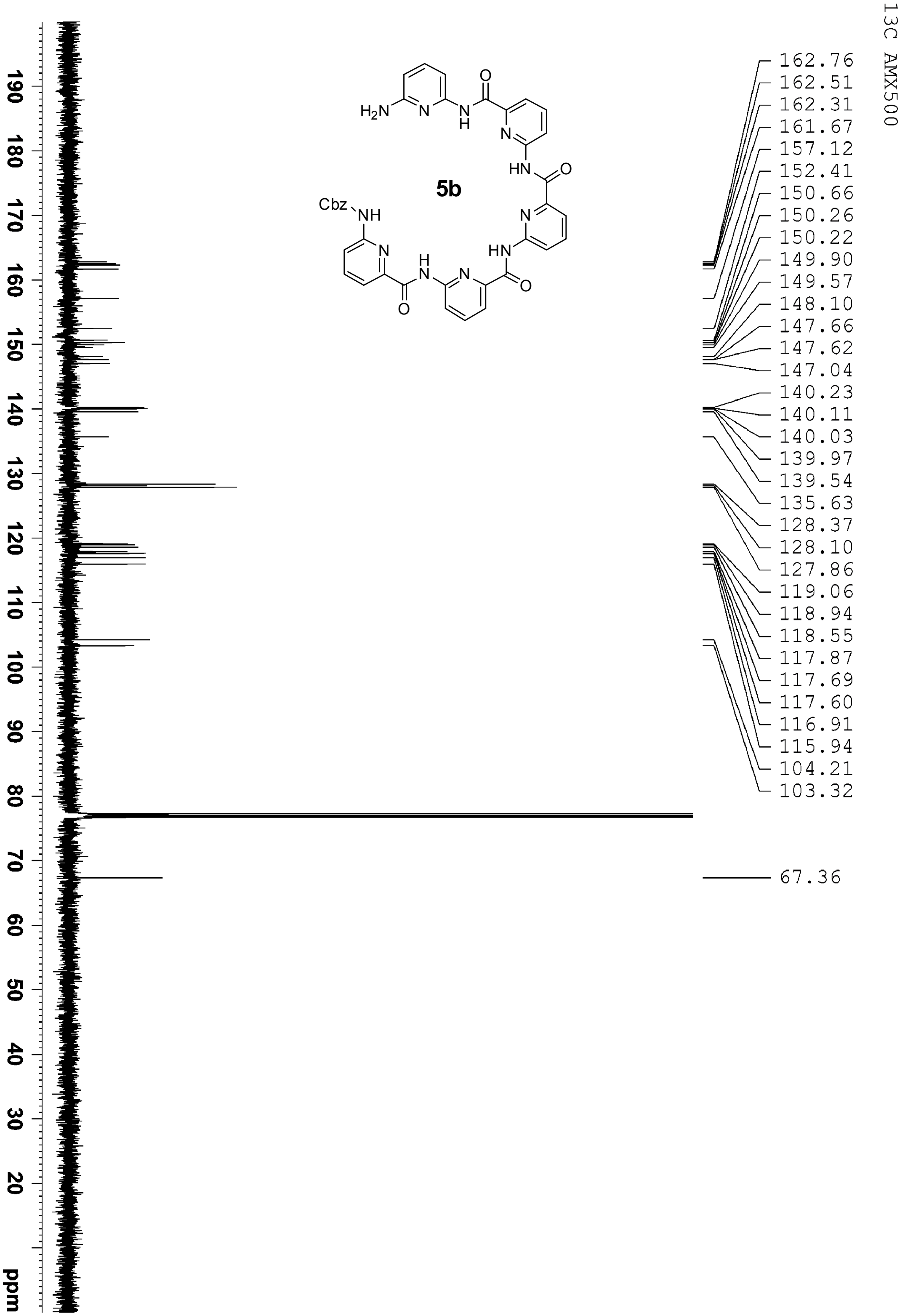

1H AMX500
H3COOC
N
N
H
O
N
HN
O
4
N
HN
O
H
N
Cbz
N
10.4449
10.3437
10.2660
8.6445
8.6295
8.6092
8.5942
8.3214
8.2346
8.2196
8.0255
8.0106
7.9875
7.9767
7.9617
7.9088
7.8939
7.8790
7.7948
7.7785
7.7677
7.7487
7.1969
7.1817
7.1666
7.1559
7.1423
7.1286
7.0937
7.0801
5.0221
3.8109
1.00
1.01
1.01
3.00
1.00
1.00
3.00
3.00
2.00
5.00
2.00
3.00
11
10
9
8
7
6
5
4
3
2
1
ppm


13C AMX500
H_3COOC
N
H
O
HN
4
Cbz
164.97
162.48
162.45
162.36
153.00
151.28
150.63
150.02
149.89
147.33
147.28
147.13
145.95
139.96
139.32
135.30
128.37
128.17
127.69
121.15
118.78
118.71
117.83
117.70
117.49
116.10
67.11
52.80
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm


1H AMX500
5a
Cbz
10.4685
10.4234
10.4008
10.3639
8.7651
8.7493
8.7326
8.6819
8.6669
8.4072
8.3914
8.1393
8.1251
8.0528
8.0394
8.0245
8.0086
7.9862
7.9729
7.9662
7.9512
7.9246
7.9096
7.8930
7.8855
7.8688
7.7756
7.7607
7.7440
7.6650
7.6550
7.6483
7.6333
7.6292
7.5360
7.1931
7.1782
7.1640
7.1557
7.1407
7.1266
7.0775
7.0633
6.7030
6.6913
6.6780
5.0098
0.99
1.01
1.01
1.00
2.01
1.00
1.01
2.01
2.00
2.00
2.01
1.00
2.00
1.00
3.01
2.00
1.00
2.00
13
12
11
10
9
8
7
6
5
4
3
2
1
ppm


13C AMX500
5a
Cbz
162.60
162.50
162.38
161.97
152.31
150.92
150.26
150.15
149.84
147.48
147.37
147.33
147.18
146.79
140.23
140.14
140.08
139.63
138.52
135.31
128.35
128.10
127.67
119.70
118.98
118.94
118.55
117.88
117.66
117.60
117.14
115.72
113.71
67.28
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm


1H AMX500
10.4797
10.3636
10.0412
8.7741
8.7577
8.7489
8.7325
8.6607
8.6443
8.1563
8.1475
8.1422
8.1324
8.0707
8.0543
8.0454
8.0388
8.0290
8.0140
7.9962
7.9787
7.9726
7.9584
7.9434
7.9143
7.8992
7.8840
7.7970
7.7806
7.7656
7.7428
7.7264
7.5670
7.3748
7.3584
7.3425
7.1742
7.1630
7.1605
7.1566
7.1503
7.1352
7.1225
7.1188
7.0822
7.0683
7.0646
5.8857
5.8707
5.0322
3.8847
2.00
1.00
1.00
2.00
1.00
2.01
5.00
1.01
1.01
1.00
1.01
1.00
5.01
1.01
2.00
2.00
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
-3
ppm
5b
H_2N
N
H
O
HN
Cbz
NH


5b
13C AMX500
162.76
162.51
162.31
161.67
157.12
152.41
150.66
150.26
150.22
149.90
149.57
148.10
147.66
147.62
147.04
140.23
140.11
140.03
139.97
139.54
135.63
128.37
128.10
127.86
119.06
118.94
118.55
117.87
117.69
117.60
116.91
115.94
104.21
103.32
67.36
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm
H_2N
Cbz
NH
HN
N
O