2016

Surprising impact of remote groups on the folding–unfolding and dimer-chain equilibria of bifunctional H-bonding unimers

Rui Liu
State University of New York

Shuang Cheng
Nanjing University

Erin S. Baker
Pacific Northwest National Laboratory

Richard D. Smith
Pacific Northwest National Laboratory

Xiao Cheng Zeng
University of Nebraska-Lincoln, xzeng1@unl.edu

See next page for additional authors

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2016, 52, 3773

Received 9th January 2016,
Accepted 28th January 2016

DOI: 10.1039/c6cc00224b

www.rsc.org/chemcomm

Surprising impact of remote groups on the folding–unfolding and dimer-chain equilibria of bifunctional H-bonding unimers†

Rui Liu,[ab] Shuang Cheng,[c] Erin S. Baker,[d] Richard D. Smith,[d] Xiao Cheng Zeng[e] and Bing Gong*[ab]

Oligoamide 1, consisting of two H-bonding units linked by a trimethylene linker, was previously found to form a very stable, folded dimer. In this work, replacing the side chains and end groups of 1 led to derivatives that show the surprising impact of end groups on the folding and dimer-chain equilibria of the resultant molecules.

The folding and unfolding of biomacromolecules have profound impact on their ability to engage in intermolecular interactions such as molecular recognition, intermolecular association and self-assembly into higher-order structures. For example, the Trp repressor shows that the recognition helix of its helix–turn–helix motif undergoes a conformational stabilization upon binding to cognate DNA.[1] The bZIP domains of leucine zippers[2] turn from random coil to an α-helical conformation upon binding to the target DNA molecules.[2] The water-soluble monomer of α-hemolysin changes conformation and assembles into a membrane-spanning heptamer with a hydrophobic outer surface.[3] The misfolding of amyloid β-peptides (Aβs) from an α-helix to β-sheet conformational transition is suggested for the dysfunctions, fibrillization, and subsequent amyloid deposits in Alzheimer's disease.[4] In contrast to the abundance of biological systems that reflect the conformational change, *e.g.*, the folding or unfolding, of their molecular components on the resultant intermolecular association and supramolecular assembly, synthetic self-assembling systems based on designed molecules have mainly focused on stably folded oligomers, *i.e.*, foldamers,[5] or on adjusting the specificity and strength of intermolecular association.[6,7] Few examples that show altered supramolecular outcomes due to conformational change of molecular components are known.

As part of our long-term interest in developing foldamers and in controlling intermolecular association,[8,9] we previously reported results from a study on H-bond-mediated association of a series of bifunctional molecular building blocks.[10] Compound **1**, consisting of two unsymmetrical 4-H-bonded units[11,12] covalently linked in a head-to-head fashion, represents one such design of this series. If adopting an extended conformation, a molecule of **1** could only partially overlap with another one *via* four intermolecular H-bonds, and would thus lead to linear H-bonded polymers.

In our previous study,[10] examining the assembly of **1** and that of another analogous pair of oligoamide strands having different 4-H-bonded units revealed a surprising outcome. Instead of forming extended polymeric aggregates, compound **1** and its analogs formed H-bonded dimers, which we dubbed "duplex foldamers", in which the component oligoamide strands folded into a conformation that (Fig. 1) led to the formation highly stable 8-H-bonded dimers. The presence of highly stable dimers of **1** and its analogs were shown by the presence of well-resolved and sharp ^{1}H NMR signals, extensive cross-strand NOEs, and also by results from mass spectrometry and vapor-pressure osmometry.

The ^{1}H NMR signals of the methylene protons L1 and L2 of **1** in $CDCl_3$ and DMSO-d_6 provided diagnostic indicators for the folded and open-chain conformations of this molecule and its analogs. The signals of L1 and L2 were found to appear at 2.50 and 2.81 ppm in $CDCl_3$ (5 mM), consistent with the constraining of the trimethylene linker due to the folding of **1**. In DMSO-d_6 that completely interrupts the intermolecular H-bonding of **1**, the signals of L1 and L2 merged into one peak at 2.36 ppm, which indicated an open-chain conformation in which the trimethylene linker becomes more flexible. These observations suggested that the folding and dimerization of **1** were closely coupled.

Thus, instead of undergoing the initially expected polymerization, oligoamide **1** was found to adopt a folded conformation that favors the formation of a highly stable, discrete dimer. Such a folded H-bonded dimer previously uncovered by us integrates the folding and dimerization of the molecular components, which

[a] *Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY 14260, USA. E-mail: bgong@buffalo.edu*
[b] *College of Chemistry, Beijing Normal University, Beijing 100875, China*
[c] *Kuang Yaming Honors School, Nanjing University, Nanjing, Jiangsu 210023, China*
[d] *Earth and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA*
[e] *Department of Chemistry, University of Nebraska–Lincoln, Lincoln, NE 68588, USA*

† Electronic supplementary information (ESI) available: Experimental procedures, 1D and 2D NMR spectra, viscosity measurements, IMS-MS analysis, computational results. See DOI: 10.1039/c6cc00224b

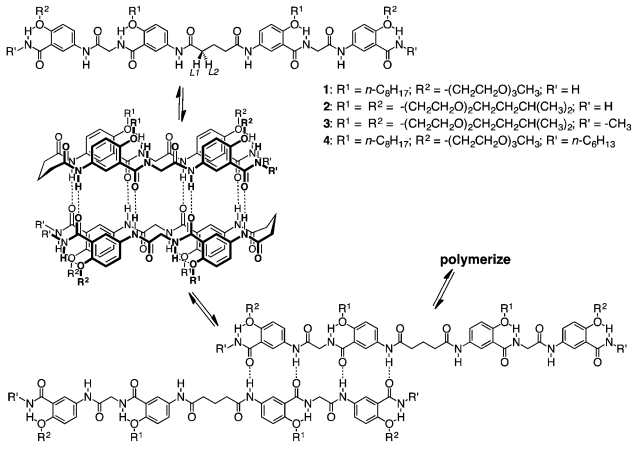


Fig. 1 Oligoamide **1** was previously found to adopt a folded conformation which leads to an 8-H-bonded, folded dimer. In this study, compounds **2**, **3**, and **4**, derived from **1** by modifying side chain R^1 and R^2 and end groups R′, showed greatly enhance solubilities which allowed the concentration-dependent folding–unfolding and the corresponding assembly of these molecules to be probed.

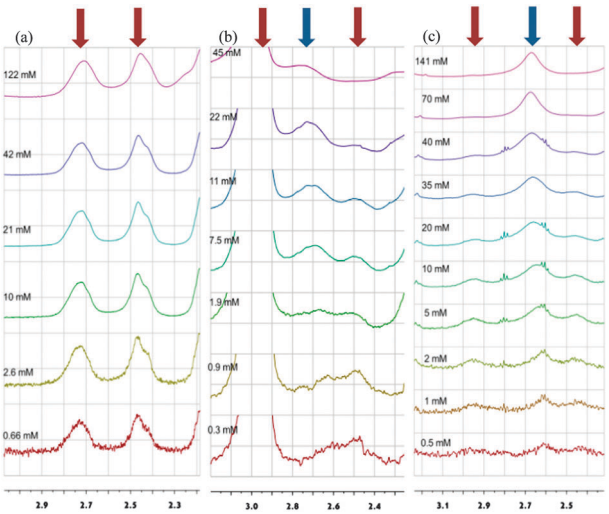


Fig. 2 Stacked partial ^{1}H NMR spectra of **2** (left), **3** (middle) and **4** (right) in $CDCl_3$ as a function of concentration. (Blue arrow: merged signals of L1 and L2; brown arrow: separated signals of L1/L2).

provides a molecular and supramolecular structural motif that may lead to many possibilities such as the development of highly stable H-bonded association units. In addition, the dimer of folded **1** is in fact a supramolecular macrocycle that could undergo ring-opening, leading to H-bond-mediated polymerization under conditions such as high concentration where linear polymeric chains become dominant. Unfortunately, probing the ring-chain equilibrium of **1** has been hampered by the limited solubility of this compound in chloroform and other solvents that promote H-bonding.

In this study, the side chains (*i.e.*, R^1 and R^2 groups) and R′ end groups of **1** were modified, leading to oligoamide strands **2**, **3** and **4** (Fig. 1) that were found to have drastically enhanced solubilities in chloroform and other solvents, which allowed their concentration-dependent behavior to be examined. We report herein the surprising effects of remote structural tuning on the folding and unfolding, and the shift of ring-chain equilibrium of these derivatives of **1**.

Compound **2**, which differs from **1** in its R^1 and R^2 side chains, gave ^{1}H NMR signals that remain sharp and well-resolved over a wide range of concentrations (from 0.03 mM to 122 mM) in $CDCl_3$, suggesting that **2** remains folded and exists as dimers (Fig. S1, ESI†). The dimerization of **2** is further supported by 2D NOESY which revealed numerous cross-strand NOEs (Fig. S9, ESI†). That the two signals of protons L1 and L2 at 2.46 ppm and 2.72 ppm remain unchanged at concentrations up to 122 mM confirms the high stability of the dimer of folded **2** (Fig. 2a). The lack of open-chain conformation and thus higher aggregates suggest that **2**, similar to **1**, must have a high critical concentration below which only the H-bonded dimers consisting of folded molecules exist. Therefore, changing side chains R^1 and R^2 of **1** did not have any detectable effect on the ring-chain equilibrium of the resultant **2** within the concentration range examined.

Replacing one of the two H atoms of each of the terminal primary amide groups of **2** with a methyl group leads to **3**. The ^{1}H NMR spectra of **3** recorded at concentrations from 0.3 mM to 45 mM in $CDCl_3$ reveal a trend that is very different from what observed for **2**. The signals of protons L1 and L2 undergo noticeable change with increasing concentration of **3** (Fig. 2b). Appearing at 2.48 and 2.93 ppm at low (<1 mM) concentrations, the two separate signals of protons L1 and L2 of **3** start to merge into a new peak at 2.72 ppm at 1 mM and higher concentrations, which is accompanied by the appearance of new broad peaks in the region (from 6 to 11 ppm) corresponding to the signals of amide and aromatic protons (Fig. S3, ESI†). Thus, in comparison to the high critical concentration of **2**, that of **3** seems to have been lowered considerably, which allows folded dimeric and open-chain conformations to co-exist in the concentration range of the ^{1}H NMR experiments.

That replacing two terminal amide H atoms of **2** with methyl groups led to the observed conformational change and thus the shift of ring-chain equilibrium for **3** is surprising. Previously, similar tuning on the ring-chain shift of bifunctional H-bonding molecules was achieved by imposing conformational bias on the linker moieties connecting adopted H-bonding units.[13]

To further probe the effect of end (R′) groups on ring-chain equilibrium, compound **4**, which shares the same R^1 and R^2 side chains with **1** but has *n*-hexyl groups attached to its two termini, was prepared. Being very soluble in $CDCl_3$, the concentration-dependent ^{1}H NMR spectra of **4** in $CDCl_3$ reveal that even at 0.5 mM, protons L1 and L2 give the merged signal at 2.67 ppm, and the separate signals at 2.45 and 2.94 ppm (Fig. 2c), which indicates the simultaneous presence of the extended and folded conformations. With increasing concentration, the intensity of the merged signal at 2.67 ppm also increases, which becomes dominant at 10 mM and higher concentrations. Increasing the concentration of **4** is also accompanied by the appearance of new broad peaks in the region corresponding to the signals of amide and aromatic protons (Fig. S3, ESI†).

In the ^{1}H NMR spectra of **3** and **4**, the merging of the otherwise split signals of L1 and L2, along with the appearance of additional

broadened signals at high concentrations, indicates the unfolding and further aggregation of these molecules, *i.e.*, the folded dimers undergo ring-opening, shifting the dimer-chain equilibrium of either compound from dimers to oligomers and polymers as concentration increases.

Compounds **2**, **3**, and **4**, with their different propensities for adopting the folded conformation that promotes dimerization, and open-chain conformation that results in oligomerization or polymerization, were further probed with ion-mobility spectrometry-mass spectrometry (IMS-MS). Species ranging from monomers to decamers were monitored for all three compounds under the condition of the IMS-MS experiments, which revealed different distributions of the measured species (Fig. S4, S5 and Table S1, ESI†). Consistent with the conclusion based on results from ^{1}H NMR, the dimer of compound **2** was found to be the most abundant (>25%) species among other oligomers, while pentamer and trimer were detected as the major oligomers for **3** and **4**, respectively. Thus, compounds **3** and **4**, with their end methyl and hexyl groups, indeed have stronger preferences for forming oligomers higher than dimers.

The concentration-dependent shift of the dimer-chain equilibrium of **4** toward higher aggregates was also demonstrated by diffusion-ordered NMR spectroscopy (DOSY). In $CDCl_3$ at 25 °C, the apparent translational diffusion coefficient (D_4) of **4** decreased considerably as concentrations increased from 25 mM to 141 mM (Table S2 and Fig. S6, ESI†). Ratio D_{TMS}/D_4, which equals to R_4/R_{TMS}, the ratio of the hydrodynamic radii of the aggregates (assuming spherical shapes) of **4** and that of TMS, provides viscosity-independent assessment[14] on the extent of aggregation of **4**. D_{TMS}/D_4 increased from 5.95 to 9.26 as the concentration of **4** changed from 25 mM to 70 mM. Using the solvent signal ($CHCl_3$) as the internal standard revealed the same trend: D_{CHCl_3}/D_4 increased from 6.57 at 0.6 mM to 43.48 at 141 mM. Data from DOSY thus indicate that **4** experiences a transition from mainly discrete dimers at low concentration to polymeric aggregates at high concentration.

Consistent with its concentration-dependent shift of dimer-chain equilibrium, in $CHCl_3$, compound **4** showed high viscosities at high concentrations, and low viscosities at low concentrations. Plotting the specific viscosity (η_{sp}) *vs.* concentration (Fig. 3a) reveals that, at low concentrations, the viscosity-concentration plot has a slope of 0.0072 ± 0.0004 cP mM^{-1} and, above 42 mM, the plot curves upward to a much larger slope of 0.0306 ± 0.0027 cP mM^{-1}. These results indicate that the aggregation of **4** involves two stages, which, in combination with ^{1}H NMR data, can be attributed to the dominance of dimeric species at low concentrations, and the prevalence of polymeric aggregates at high concentrations.

Compound **3** was found to exhibit similar concentration-dependent, two-stage change in viscosities (Fig. S7, ESI†). The viscosity is low at concentrations below 63 mM, above which a much more rapid increase in viscosity is evident as shown by the change of slope in the viscosity-concentration plot.

In contrast to **3** and **4**, compound **2**, which remains dimeric at up to 122 mM based on ^{1}H NMR, displays a linear correlation of viscosity and concentration (Fig. 3b). This further demonstrates that compound **2** remains dimerized and does not engage in concentration-dependent change of its folded conformation. Given that the end hexyl groups of **4** are more effective in shifting the dimer-chain equilibrium than the methyl groups of **3**, while the primary amide groups of **2** exert no influence, the effects of the end alkyl groups seem to be due to steric interaction between the introduced end groups with the trimethylene linkers of these molecules. The dimers of **2**, **3**, and **4** optimized at the level of B3LYP/6-31G(d) are consistent with the role of steric interaction, with the dimer of **2** being the most compact and those of **3** and **4** becoming more twisted (Fig. 4). Out of the eight intermolecular hydrogen bonds of each dimer, the number of those with H···O bond lengths longer than 2 Å changed from 0, 2, and 3, and those with N–H···O bond angles smaller than 170° changed from 0, 5, and 6, for compounds **2**, **3**, and **4**, respectively, indicating that the intermolecular H-bonds are weakened as the end groups become increasingly bulky (Fig. S8, ESI†).

In summary, compound **1**, which was found to adopt a folded conformation and form a very stable H-bonded dimer, was structurally modified to give compounds **2**, **3**, and **4** with drastically increased solubilities in chloroform. Compound **2** behaves similarly with **1** and remains dimeric in a wide range of concentrations, which demonstrates that changing R^1 and R^2 side chains does not alter the dimerization of these molecules to any noticeable extent. In contrast, modifying the end primary amide groups of **1** led to a surprising shift in the dimer-chain equilibrium of the resultant **3** or **4**. Further study based on this discovery should establish a new strategy for controlling the ring-chain equilibria of bi-functional H-bonding molecules.

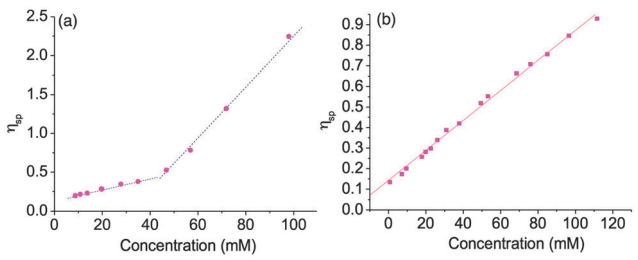


Fig. 3 Specific viscosity of (a) **4**, and (b) **2** *versus* concentration measured in $CHCl_3$ at 298 K.

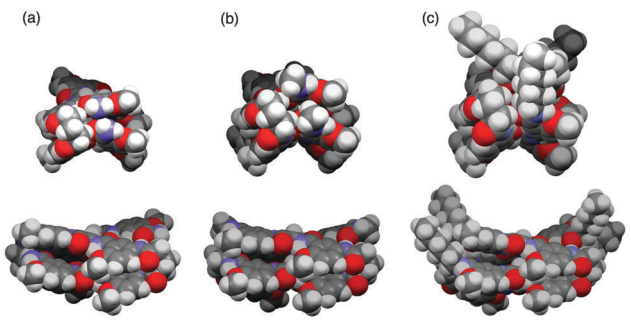


Fig. 4 CPK representations of dimers of (a) **2**, (b) **3**, and (c) **4** optimized at the at the level of B3LYP/6-31G(d). (Top view: along the long axis; bottom view: from the side of the dimer with one molecule placing in front of the other). The R^1 and R^2 side chains are replaced with methyl groups before each dimer was subjected to optimization.

This work was supported by the US National Science Foundation (CHE-1306326 and CBET-1512164), and the National Natural Science Foundation of China (NSFC-91227109). The mass spectrometric measurements were supported by grants from the NIGMS (P41 GM103493) and NIEHS (R01ES022190). This work was performed in the Environmental Molecular Science Laboratory, a DOE/BER national scientific user facility at Pacific Northwest National Laboratory (PNNL). PNNL is operated for the DOE by Battelle under contract no. DE-AC05-76RLO-1830.

Notes and references

1 H. Zhang, D. Zhao, M. Revington, W. Lee, X. Jia, C. Arrowsmith and O. Jardetzky, *J. Mol. Biol.*, 1994, **238**, 592.
2 T. E. Ellenberger, C. J. Brandl, K. Struhl and S. C. Harrison, *Cell*, 1992, **71**, 1223.
3 L. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley and J. E. Gouaux, *Science*, 1996, **274**, 1859.
4 Y. C. Xu, J. H. Shen, X. M. Luo, W. L. Zhu, K. X. Chen, J. P. Ma and H. L. Jiang, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 5403.
5 (*a*) D. Seebach and J. L. Matthews, *J. Chem. Soc., Chem. Commun.*, 1997, 2015; (*b*) S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173; (*c*) D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes and J. S. Moore, *Chem. Rev.*, 2001, **101**, 3893; (*d*) B. Gong, *Chem. – Eur. J.*, 2001, **7**, 4336; (*e*) I. Huc, *Eur. J. Org. Chem.*, 2004, 17; (*f*) K. Kirshenbaum, *ChemBioChem*, 2008, **9**, 345; (*g*) A. Roy, P. Prabhakaran, P. K. Baruah and G. J. Sanjayan, *Chem. Commun.*, 2011, **47**, 11593; (*h*) B. Gong, *Acc. Chem. Res.*, 2012, **45**, 2077; (*i*) R. Spencer, K. H. Chen, G. Manuel and J. S. Nowick, *Eur. J. Org. Chem.*, 2013, 3523.
6 (*a*) T. J. Murray and S. C. Zimmerman, *J. Am. Chem. Soc.*, 1992, **114**, 4010; (*b*) J. Yang, E. K. Fan, S. J. Geib and A. D. Hamilton, *J. Am. Chem. Soc.*, 1993, **115**, 5314; (*c*) J. L. Sessler and R. Z. Wang, *J. Am. Chem. Soc.*, 1996, **118**, 9808; (*d*) R. H. Vreekamp, J. P. M. Duynhoven, M. Hubert, W. Verboom and D. N. Reinhoudt, *Angew. Chem., Int. Ed.*, 1996, **35**, 1215; (*e*) F. H. Beijer, H. Kooijman, A. L. Spek, R. P. Sijbesma and E. W. Meijer, *Angew. Chem., Int. Ed.*, 1998, **37**, 75; (*f*) A. X. Wu, A. Chakraborty, J. C. Fettinger, R. A. Flowers and L. Isaacs, *Angew. Chem., Int. Ed.*, 2002, **41**, 4028; (*g*) H. Fenniri, B. L. Deng and A. E. Ribbe, *J. Am. Chem. Soc.*, 2002, **124**, 11064; (*h*) C. L. Zhan, J. M. Leger and I. Huc, *Angew. Chem., Int. Ed.*, 2006, **45**, 4625; (*i*) W. J. Chu, Y. Yang and C. F. Chen, *Org. Lett.*, 2010, **12**, 3156; (*j*) A. Gooch, A. M. McGhee, M. L. Pellizzaro, C. I. Lindsay and A. J. Wilson, *Org. Lett.*, 2011, **13**, 240; (*k*) B. P. Mudraboyina and J. A. Wisner, *Chem. – Eur. J.*, 2012, **18**, 14157; (*l*) B. A. Blight, C. A. Hunter, D. A. Leigh, H. McNab and P. I. T. Thomson, *Nat. Chem.*, 2011, **3**, 244; (*m*) Y. Li, T. H. Park, J. K. Quansah and S. C. Zimmerman, *J. Am. Chem. Soc.*, 2011, **133**, 17118; (*n*) M. L. Pellizzaro, S. A. Barrett, J. Fisher and A. J. Wilson, *Org. Biomol. Chem.*, 2012, **10**, 4899; (*o*) D. A. Leigh, C. C. Robertson, A. M. Z. Slawin and P. I. T. Thomson, *J. Am. Chem. Soc.*, 2013, **135**, 9939.
7 (*a*) J. S. Nowick and J. S. Chen, *J. Am. Chem. Soc.*, 1992, **114**, 1107; (*b*) Y. Kato, M. M. Conn and J. Rebek, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 1208; (*c*) M. Torneiro and W. C. Still, *J. Am. Chem. Soc.*, 1995, **117**, 5887; (*d*) L. Brunsveld, J. Vekemans, J. Hirschberg, R. P. Sijbesma and E. W. Meijer, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 4977.
8 (*a*) L. Brunsveld, B. J. B. Folmer, E. W. Meijer and R. P. Sijbesma, *Chem. Rev.*, 2001, **101**, 4071–4098; (*b*) A. J. Wilson, *Soft Matter*, 2007, **3**, 409–425; (*c*) S. K. Yang and S. C. Zimmerman, *Isr. J. Chem.*, 2013, **53**, 511–520.
9 (*a*) R. P. Sijbesma, F. H. Beijer, L. Brunsveld, B. J. B. Folmer, J. H. K. K. Hirschberg, R. F. M. Lange, J. K. Lowe and L. E. W. Meijer, *Science*, 1997, **278**, 1601–1604; (*b*) X. W. Yang, F. J. Hua, K. Yamato, E. Ruckenstein, B. Gong, W. Kim and C. Y. Ryu, *Angew. Chem., Int. Ed.*, 2004, **43**, 6471–6474; (*c*) E. M. Todd and S. C. Zimmerman, *J. Am. Chem. Soc.*, 2007, **129**, 14534–14535; (*d*) S. K. Yang, A. V. Ambade and M. Weck, *J. Am. Chem. Soc.*, 2010, **132**, 1637–1645; (*e*) X. Yan, S. Li, J. B. Pollock, T. R. Cook, J. Chen, Y. Zhang, X. Ji, Y. Yu, F. Huang and P. J. Stang, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 15585–15590; (*f*) A. Gooch, N. S. Murphy, N. H. Thomson and A. J. Wilson, *Macromolecules*, 2013, **46**, 9634–9641.
10 X. W. Yang, S. Martinovic, R. D. Smith and B. Gong, *J. Am. Chem. Soc.*, 2003, **125**, 9932.
11 (*a*) B. Gong, *Synlett*, 2001, 582; (*b*) B. Gong, *Polym. Int.*, 2007, **56**, 436; (*c*) B. Gong, *Acc. Chem. Res.*, 2012, **45**, 2077.
12 (*a*) B. Gong, Y. F. Yan, H. Q. Zeng, E. Skrzypczak-Jankunn, Y. W. Kim, J. Zhu and H. Ickes, *J. Am. Chem. Soc.*, 1999, **121**, 5607; (*b*) H. Q. Zeng, R. S. Miller, R. A. Flowers and B. Gong, *J. Am. Chem. Soc.*, 2000, **122**, 2635; (*c*) H. Q. Zeng, H. Ickes, R. A. Flowers and B. Gong, *J. Org. Chem.*, 2001, **66**, 3574; (*d*) H. Q. Zeng, X. W. Yang, R. A. Flowers and B. Gong, *J. Am. Chem. Soc.*, 2002, **124**, 2903; (*e*) M. F. Li, K. Yamato, J. S. Ferguson and B. Gong, *J. Am. Chem. Soc.*, 2006, **128**, 12628–12629; (*f*) M. F. Li, K. Yamato, J. S. Ferguson, K. K. Singarapu, T. Szyperski and B. Gong, *J. Am. Chem. Soc.*, 2008, **130**, 491–500.
13 (*a*) B. J. B. Folmer, R. P. Sijbesma and E. W. Meijer, *J. Am. Chem. Soc.*, 2001, **123**, 2093; (*b*) A. T. ten Cate, H. Kooijman, A. L. Spek, R. P. Sijbesma and E. W. Meijer, *J. Am. Chem. Soc.*, 2004, **126**, 3801.
14 (*a*) E. J. Cabrita and S. Berger, *Magn. Reson. Chem.*, 2001, **39**, S142; (*b*) X. X. Wu, R. Liu, B. Sathyamoorthy, K. Yamato, G. X. Liang, L. Shen, S. F. Ma, D. K. Sukumaran, T. Szyperski, W. H. Fang, L. He, X. B. Chen and B. Gong, *J. Am. Chem. Soc.*, 2015, **137**, 5879.

The Effect of Structural Tuning on the Folding-Unfolding and Dimer-Chain Equilibria of Bifunctionl H-Bonding Unimers

Rui Liu,[ae] Shuang Chen,[b] Erin Baker,[c] Richard D. Smith,[c] Xiao Cheng Zeng,[d] and Bing Gong*[ae]

[a] *Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, NY 14260, USA*

[b] *Kuang Yaming Honors School, Nanjing University, Nanjing, Jiangsu 210023, China.*

[c] *Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA.*

[d] *Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588, USA*

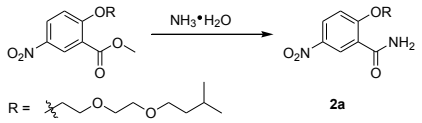
[e] *College of Chemistry, Beijing Normal University, Beijing 100875, China.*

Electronic Supplementary Information

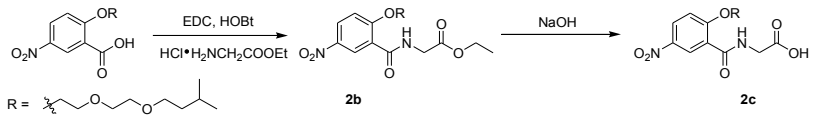
1. Materials and Methods

All reagents were purchased from commercial sources (Fisher Scientific, Acros, Alfa Aesar and Aldrich) and were used as received unless otherwise noted. Silica gel column chromatography was carried out with silica gel 60 (mesh 230-400) and products were detected as single spots by thin-layer chromatography (precoated 0.25 mm silica plates from Sorbent). All ^{1}H NMR and ^{13}C NMR data were recorded on Varian Inova 500 (or 400) Spectrometers (500 MHz or 400 MHz) and Varian Mercury 300 Spectrometer (300 MHz). NMR chemical shifts are reported in ppm relative to internal standard TMS, and coupling constant, *J*, is reported in Hertz (Hz). For the ^{1}H NMR experiments, $CDCl_3$ (99.8% D), DMF-d_7(99.5%) and DMSO-d_6 (99.8% D) were purchased from Carmbridge Isotope Laboratory and used without further purification. Low-resolution electrospray ionization (LRESI) mass spectra were obtained on a Bruker Esquire 3000 plus mass spectrometer (Bruker-Franzen Analytik GmbH, Bremen, Germany) equipped with an ESI interface and an ion trap analyzer.

2. Syntheses

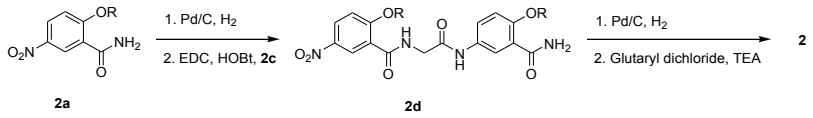


Compound **2a**: Ammonium hydroxide (18 mL) was added to a methanol solution (9 mL) of methyl 2-(2-(2-(isopentyloxy)ethoxy)ethoxy)-5-nitrobenzoate [1](1.62 g, 4.56 mmol). The reaction mixture was allowed to react at 40˚C overnight. Most of the solvent is removed in a reduced pressure. Water and dichloromethane were added. The aqueous layer was extracted with dichloromethane three times. Organic layer was combined and solvent was removed in vaccum, White powder was obtained as the product (1.24g, 80%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.61 (d, *J* = 3.0 Hz, 1H), 8.34 (dd, *J* = 9.1, 3.0 Hz, 1H), 7.89 (s, 1H), 7.71 (s, 1H), 7.40 (d, *J* = 9.2 Hz, 1H), 4.43 – 4.37 (m, 2H), 3.87 – 3.81 (m, 2H), 3.59 (m, 2H), 3.47 (m, 2H), 3.36 (t, *J* = 6.8 Hz, 2H), 1.57 (m, 1H), 1.32 (m, 2H), 0.90 – 0.71 (m, 6H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 164.55, 161.78, 141.21, 128.30, 126.93, 123.96, 114.88, 70.21, 69.95, 69.53, 69.16, 68.68, 38.46, 25.00, 22.91. ESI MS: calculated 363.4, found 363.3 $(M + Na)^+$.



Compound **2b**: **2b** was prepared following a reported method.[2] ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.05 (d, *J* = 2.8 Hz, 1H), 8.49 (s, 1H), 8.29 (dd, *J* = 9.1, 2.8 Hz, 1H), 7.07 (d, *J* = 9.1 Hz, 1H), 4.39 (m, 2H), 4.29 – 4.15 (m, 4H), 4.07 – 3.91 (m, 2H), 3.75 – 3.61 (m, 2H), 3.59 – 3.48 (m, 2H), 3.43 (t, *J* = 6.9 Hz, 2H), 1.61 (m, 1H), 1.42 (m, 2H), 1.29 (t, *J* = 7.1 Hz, 3H), 0.85 (m, 6H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.81, 163.01, 161.28, 141.92, 128.45, 128.10, 122.19, 112.96, 70.75, 70.02, 69.92, 69.32, 68.69, 61.39, 42.08, 38.29, 25.05, 22.58, 14.17. ESI MS: calculated 449.5, found 449.3 $(M + Na)^+$.

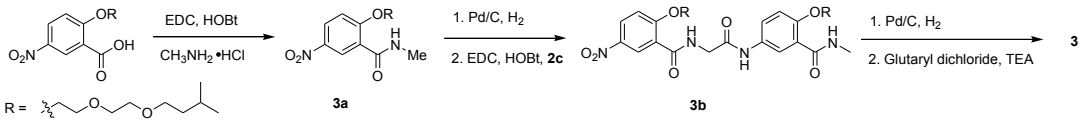
Compound **2c**: **2c** was prepared following a reported method.[2] ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.09 (d, *J* = 2.9 Hz, 1H), 8.61 (t, *J* = 4.5 Hz, 1H), 8.35 (dd, *J* = 9.1, 2.9 Hz, 1H), 7.09 (d, *J* = 9.1 Hz, 1H), 4.49 – 4.38 (m, 2H), 4.32 (m, 2H), 4.03 – 3.95 (m, 2H), 3.79 – 3.69 (m, 2H), 3.69 – 3.62 (m, 2H), 3.55 (t, *J* = 7.1 Hz, 2H), 1.77 – 1.57 (m, 1H), 1.52 (m, 2H), 1.01 – 0.81 (m, 6H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.05, 163.51, 161.30, 141.91, 128.34, 121.81, 112.99, 70.71, 70.12, 70.03, 69.36, 68.71, 42.23, 38.01, 25.03, 22.56. ESI MS: calculated 399.4, found 399.1$(M + H)^+$.



Compound **2d**: A solution of **2a** (0.60 g, 1.76 mmol) in dichloromethane /methanol was shaken in the presence of Pd/C (10%) under a hydrogen atmosphere for 2 hours, and then catalyst was filtered. The filtrate was evaporated in vacuum, yielding the corresponding amine. To a solution **2c** (0.70 g, 1.76 mmol), EDC (0.38 g, 2.47 mmol) and HOBt (0.26 g, 1.94 mmol) in CH_2Cl_2 (20 mL), the CH_2Cl_2 solution (6 mL) of freshly prepared amine was added. The reaction was allowed to proceed for 4 hours at room temperature. The mixture was then washed with diluted HCl and solvent was removed in vacuum. Purification was accomplished by chromatography on silica gel using chloroform/methanol to afford **2d** (0.64 g, 78%) as a light yellow solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.13 (s, 1H), 8.71 – 8.64 (m, 2H), 8.36 (dd, *J* = 9.1, 2.9 Hz, 1H), 8.05 (d, *J* = 2.8 Hz, 1H), 7.79 (dd, *J* = 8.7, 2.8 Hz, 1H), 7.74 (s, 1H), 7.53 (s, 1H), 7.44 (d, *J* = 9.1 Hz, 1H), 7.12 (d, *J* = 8.9 Hz, 1H), 4.49 – 4.41 (m, 2H), 4.24 – 4.18 (m, 2H), 4.16 (m, 2H), 3.93 – 3.87 (m, 2H), 3.80 – 3.74 (m, 2H), 3.63 – 3.58 (m, 2H), 3.58 – 3.55 (m, 2H), 3.50 – 3.45 (m, 2H), 3.46 – 3.41 (m, 2H), 3.38 (t, *J* = 6.8 Hz, 2H), 3.32 – 3.26 (m, 2H), 1.64 – 1.55 (m, 1H), 1.55 – 1.46 (m, 1H), 1.34 (q, *J* = 6.9 Hz, 2H), 1.27 (q, *J* = 6.7 Hz, 2H), 0.82 (d, *J* = 6.6 Hz, 6H), 0.76 (d, *J* = 6.6 Hz, 6H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 167.16, 165.96, 163.15, 161.90, 153.02, 141.28, 132.74, 128.53, 126.99, 123.93, 122.96, 122.91, 122.60, 115.03, 114.60, 70.34, 70.17, 70.12, 69.93, 69.89, 69.16, 69.09, 69.00, 68.80, 68.72, 48.70, 48.68, 43.88, 38.48, 38.43, 25.01, 24.97, 22.93, 22.87. ESI MS: calculated 713.8, found 713.5 $(M + Na)^+$.

Compound **2**: A solution of **2d** (0.28 g, 0.41 mmol) in dichloromethane /methanol was shaken in the presence of Pd/C (10%) under a hydrogen atmosphere for 2 hours, and catalyst was then filtered. The filtrate was evaporated in vacuum, yielding the corresponding amine. The amine and triethylamine (0.05 g, 0.49 mmol) were dissolved in dry dichloromethane (10 mL), followed by a solution of glutaryl dichloride (34.8 mg, 0.20 mmol) in dichloromethane (5 mL). The resulting mixture was allowed to react overnight at room temperature. The crude was washed with diluted HCl and then solvent was removed in vacuum. Purification was accomplished by chromatography on silica gel using chloroform/methanol to afford **2** (0.21 g, 71%) as a pale yellow solid. ^{1}H NMR (300 MHz, DMF-d_7) δ 10.15 (s, 2H), 10.02 (s, 2H), 8.88 (t, *J* = 5.0 Hz, 2H), 8.30 (dd, *J* = 9.3, 2.4 Hz, 4H), 7.99 – 7.91 (m, 8H), 7.44 (s, 2H), 7.22 (d, *J* = 8.7 Hz, 2H), 7.18 (d, *J* = 9.0 Hz, 2H), 4.32 (m, 12H), 4.00 – 3.93 (m, 4H), 3.93 – 3.85 (m, 4H), 3.76 – 3.58 (m, 8H), 3.58 – 3.51 (m, 8H), 3.45 (t, *J* = 6.7 Hz, 4H), 3.39 (t, *J* = 6.8 Hz, 4H), 2.47 (t, *J* = 7.2 Hz,

4H), 2.10 – 1.96 (m, 2H), 1.75 – 1.49 (m, 4H), 1.37 (m, 8H), 0.86 (d, *J* = 6.7 Hz, 12H), 0.81 (d, *J* = 6.6 Hz, 12H). ^{13}C NMR (75 MHz, DMF-d_7) δ 170.86, 167.59, 165.75, 164.60, 153.19, 153.01, 133.72, 133.05, 123.87, 123.75, 122.89, 122.79, 122.32, 122.07, 114.23, 114.19, 70.38, 70.23, 69.97, 69.32, 69.17, 69.14, 69.07, 69.02, 68.83, 43.75, 38.49, 38.45, 35.84, 24.93, 22.20, 21.46. ESI MS: calculated 1418.7, found 1418.5 $(M + H)^+$.

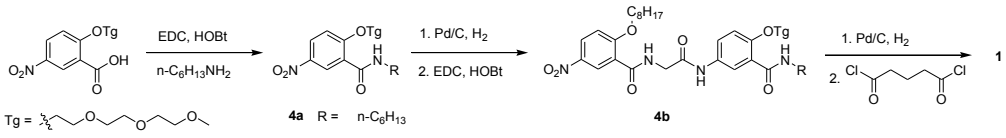


Compound **3a**: The product was prepared following the procedure reported previously.[2] ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.56 (d, *J* = 3.0 Hz, 1H), 8.30 (dd, *J* = 9.0, 2.5 Hz, 1H), 8.29 – 8.09 (m, 1H), 7.38 (d, *J* = 9.1 Hz, 1H), 4.47 – 4.35 (m, 2H), 3.92 – 3.80 (m, 2H), 3.69 – 3.58 (m, 2H), 3.55 – 3.44 (m, 2H), 3.36 (t, *J* = 6.7 Hz, 2H), 2.83 (d, *J* = 4.6 Hz, 3H), 1.55 (m, 1H), 1.31 (m, 2H), 0.84 – 0.77 (m, 6H). ^{13}C NMR (75 MHz, DMSO-d_6) δ 164.29, 162.20, 141.89, 128.68, 127.26, 124.67, 115.49, 71.04, 70.68, 70.28, 69.82, 69.33, 39.12, 27.45, 25.63, 23.50. ESI MS: calculated 355.5, found 355.3 $(M + H)^+$.

Compound **3b**: The product was prepared following the procedure reported previously.[2] ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.16 (s, 1H), 8.68 (s, 2H), 8.37 (d, *J* = 9.1 Hz, 1H), 8.23 (d, *J* = 3.4 Hz, 1H), 8.02 (s, 1H), 7.78 (d, *J* = 8.6 Hz, 1H), 7.45 (d, *J* = 9.1 Hz, 1H), 7.14 (d, *J* = 9.1 Hz, 1H), 4.61 – 4.35 (m, 2H), 4.35 – 4.09 (m, 4H), 3.90 (s, 2H), 3.79 (s, 2H), 3.66 – 3.54 (m, 4H), 3.54 – 3.31 (m, 8H), 2.81 (s, 3H), 1.70 – 1.38 (m, 2H), 1.38 – 1.15 (m, 4H), 0.96 – 0.67 (m, 12H). ^{13}C NMR (75 MHz, DMSO-d_6) δ 167.19, 165.06, 163.17, 161.91, 152.69, 141.28, 132.92, 128.55, 126.98, 123.59, 123.10, 122.92, 122.19, 115.04, 114.92, 70.34, 70.12, 70.01, 69.88, 69.15, 69.08, 68.96, 68.71, 43.87, 38.47, 38.42, 26.65, 24.98, 22.91, 22.87. ESI MS: calculated 705.8, found 705.4 $(M+H)^+$.

Compound **3**: The product was prepared following the procedure reported previously.[2] The crude product was purified using flash column chromatography to yield a pale powder (0.11g, 62%). ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.12 (s, 2H), 9.93 (s, 1H), 8.70 (s, 2H), 8.23 (m, 2H), 8.13 – 7.93 (m, 3H), 7.80 (m, 3H), 7.68 (s, 1H), 7.24 (s, 3H), 7.12 (d, *J* = 8.5 Hz, 3H), 4.37 – 4.04 (m, 12H), 3.94 – 3.81 (m, 4H), 3.81 – 3.70 (m, 4H), 3.64 – 3.54 (m, 12H), 3.47 (m, 8H), 3.41 – 3.31 (m, 6H), 2.81 (m, 6H), 2.75 – 2.62 (m, 4H), 2.33 (m, 4H), 2.06 – 1.90 (m, 2H), 1.89 – 1.74 (m, 2H), 1.57 (m, 4H), 1.38 – 1.22 (m, 8H), 0.87 – 0.68 (m, 24H). ^{13}C NMR (75 MHz, DMSO-d_6) δ 173.32, 170.74, 167.47, 165.06, 164.37, 156.55, 152.64, 133.60, 133.02, 131.94, 129.47, 123.58,

123.05, 122.18, 114.87, 114.42, 70.34, 70.29, 70.00, 69.88, 69.31, 69.13, 69.04, 68.96, 51.68, 43.83, 38.43, 35.56, 33.08, 32.85, 26.63, 24.98, 22.88, 20.86, 16.99. ESI MS: calculated 1446.7, found 1446.9 $(M + H)^+$.



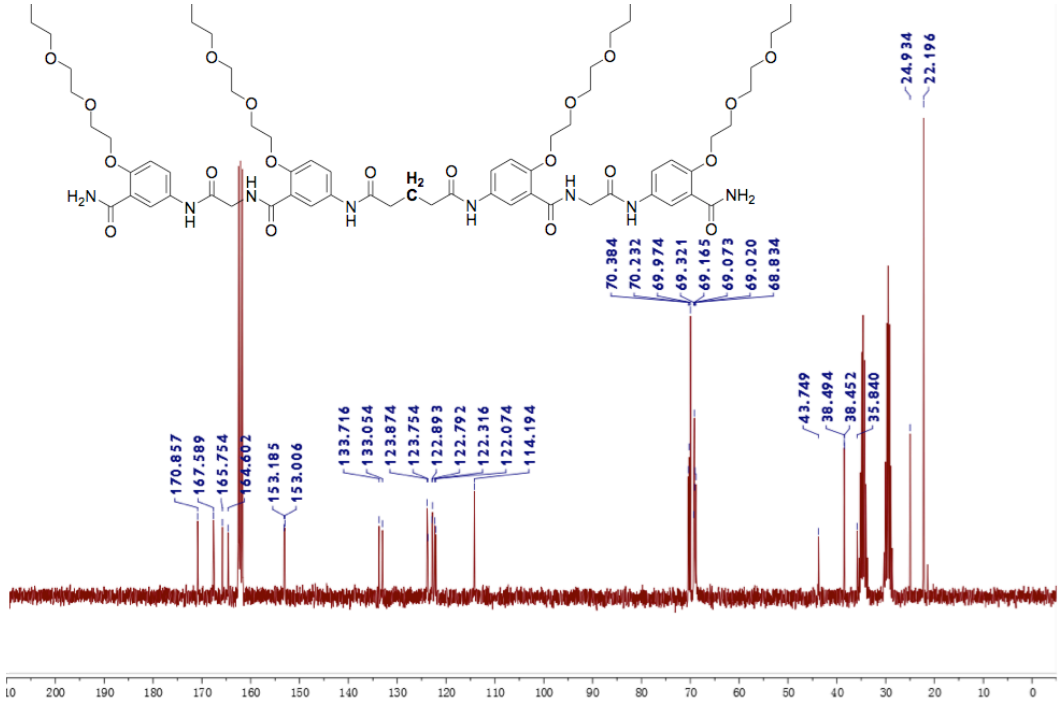
Compound **4a**: To a solution of 2-(2-(2-(2-methoxyethoxy)ethoxy)ethoxy)-5-nitrobenzoic acid [2] (1.0 g, 3.04 mmol), EDC (0.57 g, 3.65 mmol) and HOBt (0.45 g, 3.34 mmol) in CH_2Cl_2 (20 mL) was added the CH_2Cl_2 solution (3 mL) of hexylamine (0.32 g, 3.19 mmol). The reaction was allowed to proceed for 6 hours at room temperature. The mixture was then washed with diluted HCl and solvent was removed in vacuum. Purification was accomplished by chromatography on silica gel using hexane/acetone to afford **4a** (1.09 g, 87%) as a light yellow solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.02 (s, 1H), 8.25 (d, *J* = 8.8 Hz, 1H), 7.92 (s, 1H), 7.03 (d, *J* = 9.0 Hz, 1H), 4.42 – 4.27 (m, 2H), 3.94 (m, 2H), 3.80 – 3.55 (m, 6H), 3.52 – 3.36 (m, 4H), 3.37 – 3.25 (m, 3H), 1.71 – 1.49 (m, 2H), 1.45 – 1.09 (m, 6H), 0.97 – 0.73 (m, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.82, 160.85, 141.96, 128.33, 127.60, 123.24, 112.69, 71.85, 70.65, 70.60, 68.82, 58.99, 40.15, 31.53, 29.47, 26.75, 22.59, 14.03. ESI MS: calculated 413.5, found 413.4 $(M + Na)^+$.

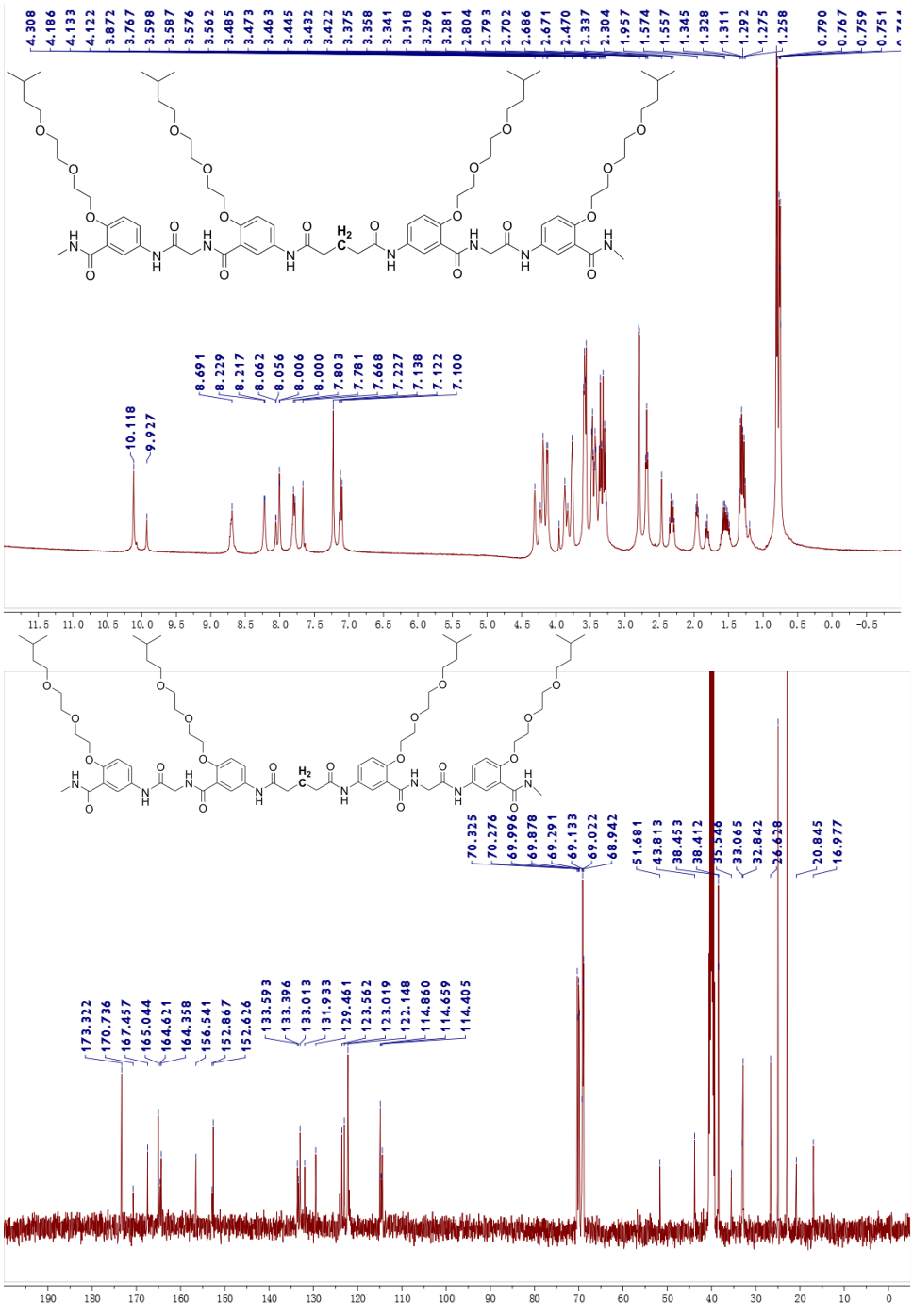
Compound **4b**: A solution of **4a** (1.09 g, 2.84 mmol) in dichloromethane /methanol was shaken in the presence of Pd/C (10%) under a hydrogen atmosphere for 2 hours, and then catalyst was filtered. The filtrate was evaporated in vacuum, yielding the corresponding amine. To a solution of 2-(5-nitro-2-(octyloxy)benzamido)acetic acid [2](1.0 g, 2.84 mmol), EDC (0.48 g, 3.09 mmol) and HOBt (0.42 g, 3.09 mmol) in CH_2Cl_2 (45 mL) was added the CH_2Cl_2 solution (5 mL) of freshly prepared amine. The reaction was allowed to proceed for 4 hours at room temperature. The mixture was then washed with diluted HCl and solvent was removed in vacuum. Purification was accomplished by chromatography on silica gel using chloroform/methanol to afford **4b** (1.44 g, 71%) as a light yellow solid. ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.16 (s, 1H), 8.65 (d, *J* = 3.1 Hz, 2H), 8.33 (dd, *J* = 9.2, 3.0 Hz, 1H), 8.21 (t, *J* = 5.6 Hz, 1H), 7.99 (d, *J* = 2.8 Hz, 1H), 7.80 (dd, *J* = 8.9, 2.8 Hz, 1H), 7.39 (d, *J* = 9.3 Hz, 1H), 7.10 (d, *J* = 9.0 Hz, 1H), 4.28 (t, *J* = 6.4 Hz, 2H), 4.23 – 4.13 (m, 4H), 3.85 – 3.71 (m, 2H), 3.64 – 3.57 (m, 2H), 3.56 – 3.44 (m, 4H), 3.42 – 3.35 (m, 2H), 3.31 – 3.22 (m, 2H), 3.20 (s, 3H), 1.96 – 1.74 (m, 2H), 1.57 – 1.12 (m, 18H), 0.95 – 0.69 (m, 6H). ^{13}C NMR (75 MHz, DMSO-d_6) δ 167.07, 164.35, 163.12, 161.99, 152.62, 140.95, 132.79, 128.48, 126.86, 123.47, 123.09, 122.86, 122.22, 114.48, 114.33, 71.69, 70.79, 70.27,

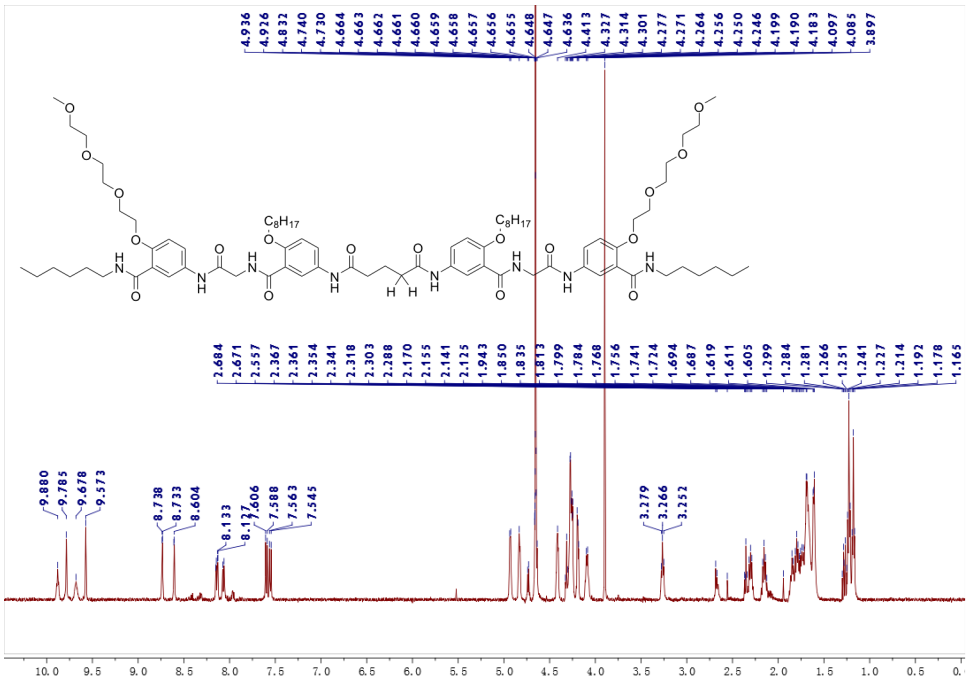
70.09, 69.09, 68.74, 58.48, 43.91, 39.55, 31.69, 31.49, 29.53, 29.20, 29.13, 28.69, 26.63, 26.03, 22.53, 14.36. ESI MS: calculated 739.9, found 739.6 $(M + Na)^+$.

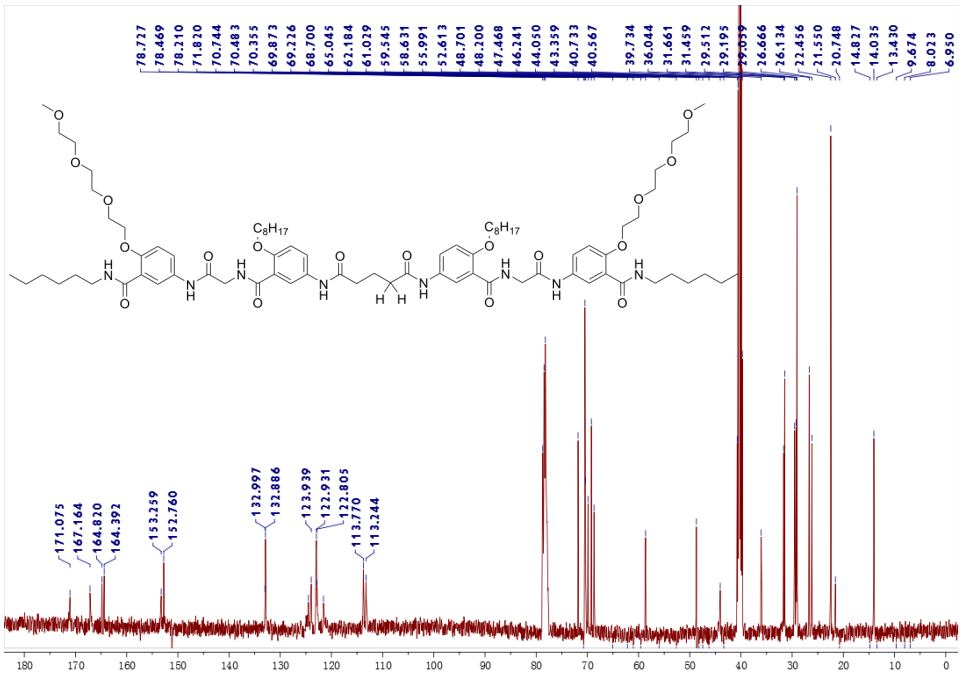
Compound **1**: A solution of **4b** (0.12 g, 0.17 mmol) in dichloromethane /methanol was shaken in the presence of Pd/C (10%) under a hydrogen atmosphere for 2 hours, and catalyst was then filtered. The filtrate was evaporated in vacuum, yielding the corresponding amine. The amine and triethylamine (0.18 g, 1.8 mmol) were dissolved in dry dichloromethane (20 mL), followed by a solution of glutaryl dichloride (14.8 mg, 0.08 mmol) in dichloromethane (5 mL). The resulting mixture was allowed to react overnight at room temperature. The crude was washed with diluted HCl and then solvent was removed in vacuum. Purification was accomplished by chromatography on silica gel using chloroform/methanol to afford **1** (0.17 g, 66%) as a pale yellow solid. ^{1}H NMR (500 MHz, CF_3COOH/D_2O) δ 9.88 (s, 1H), 9.78 (s, 1H), 9.68 (s, 1H), 9.57 (s, 1H), 8.74 (d, *J* = 2.2 Hz, 1H), 8.61 (d, *J* = 2.4 Hz, 1H), 8.14 (dd, *J* = 9.0, 2.9 Hz, 1H), 8.07 (dd, *J* = 9.0, 2.2 Hz, 1H), 7.60 (d, *J* = 9.2 Hz, 1H), 7.55 (d, *J* = 9.1 Hz, 1H), 4.96 – 4.90 (m, 4H), 4.86 – 4.80 (m, 4H), 4.45 – 4.38 (m, 4H), 4.29 – 4.23 (m, 8H), 4.22 – 4.16 (m, 4H), 4.13 – 4.06 (m, 4H), 3.90 (s, 6H), 3.26 (t, *J* = 6.8 Hz, 4H), 2.73 – 2.61 (m, 2H), 2.35 – 2.26 (m, 2H), 2.21 – 2.10 (m, 2H), 1.90 – 1.53 (m, 36H), 1.28 – 1.13 (m, 12H). ^{13}C NMR (126 MHz, 30%DMSO-d_6/70%$CDCl_3$) δ 171.08, 167.16, 164.82, 164.39, 153.26, 152.76, 151.17, 132.89, 124.48, 123.94, 123.00, 122.93, 122.80, 121.55, 113.77, 113.24, 71.82, 70.48, 69.87, 69.23, 68.70, 58.63, 48.70, 36.04, 31.66, 31.46, 29.51, 29.19, 29.06, 26.66, 26.13, 22.46, 14.03. MS-MALDI: calculated 1491.9, found 1491.5 $(M + Na)^+$.

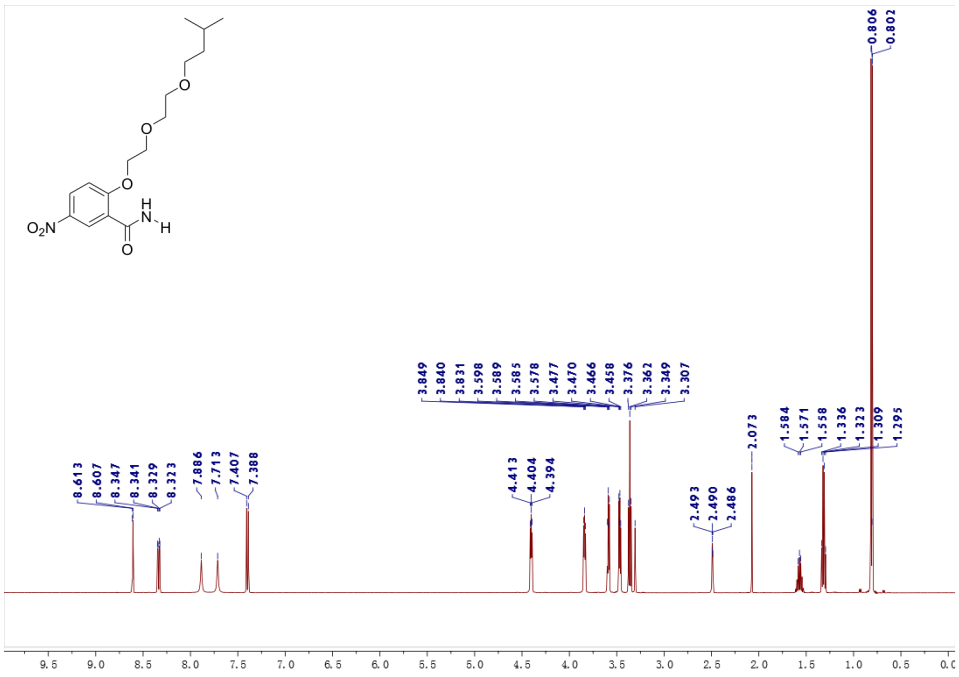
3. ^{1}H and ^{13}C NMR spectra

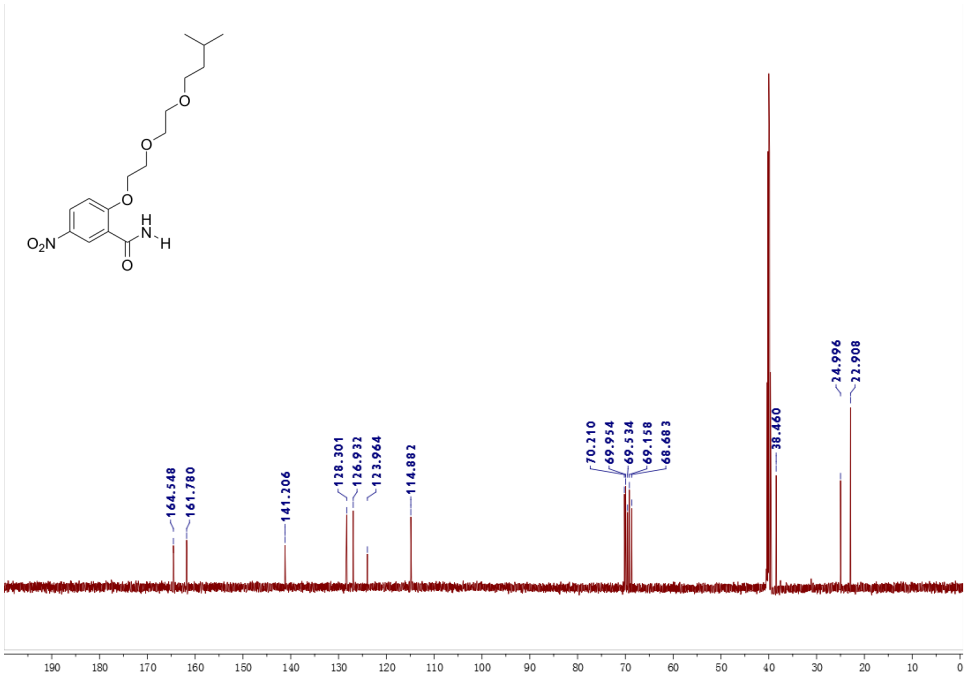


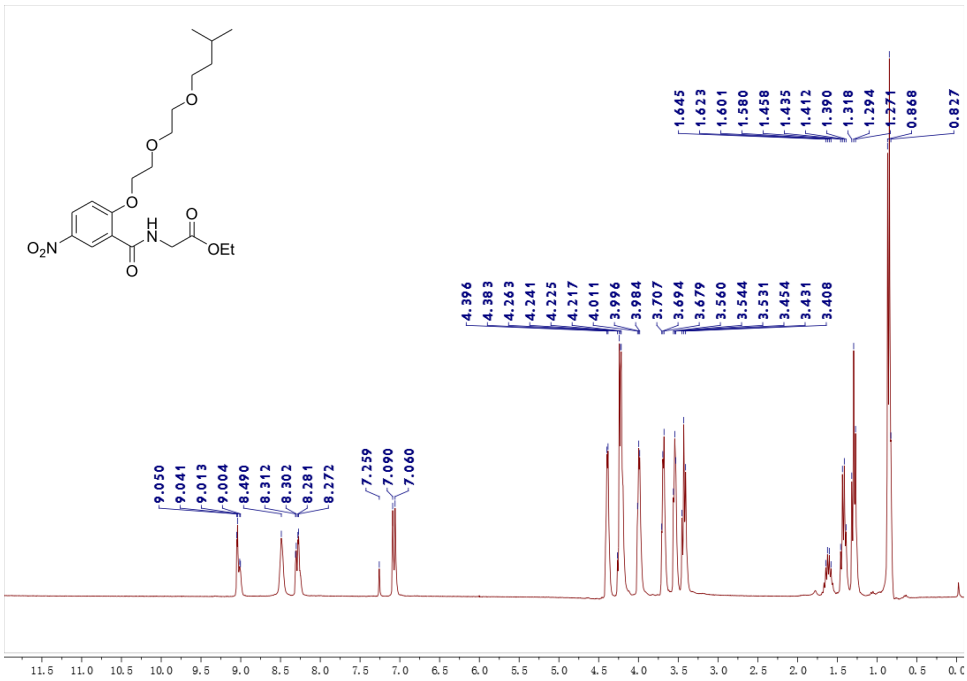


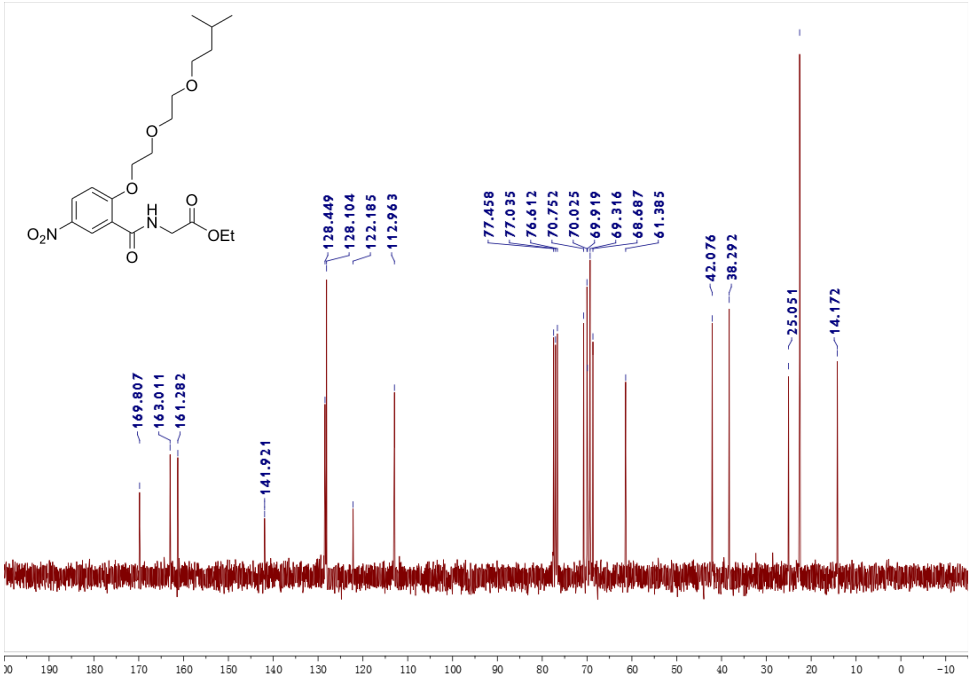


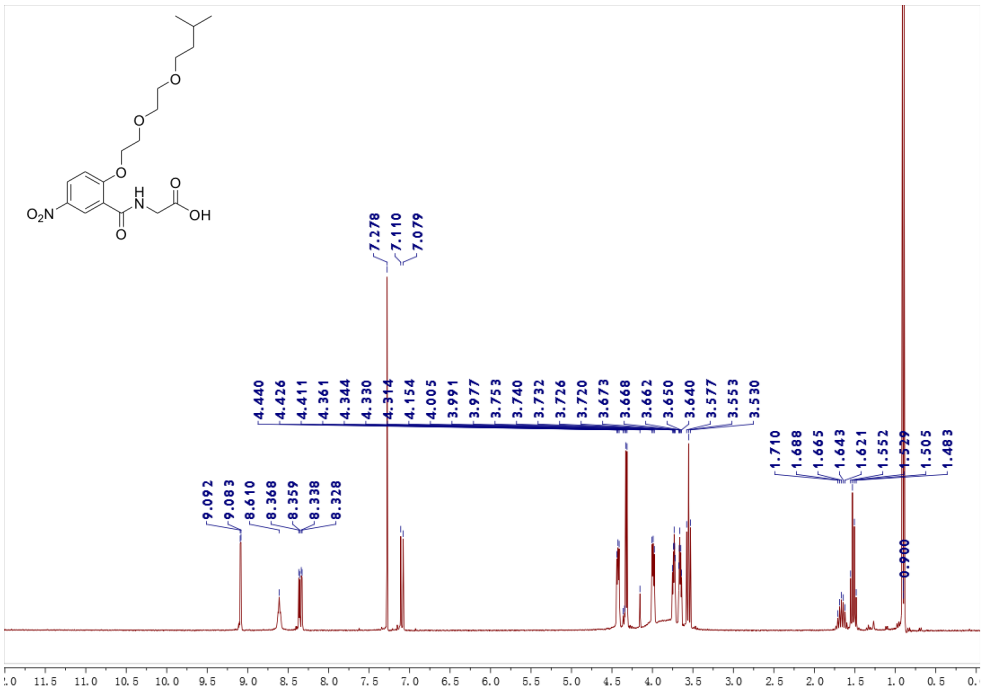


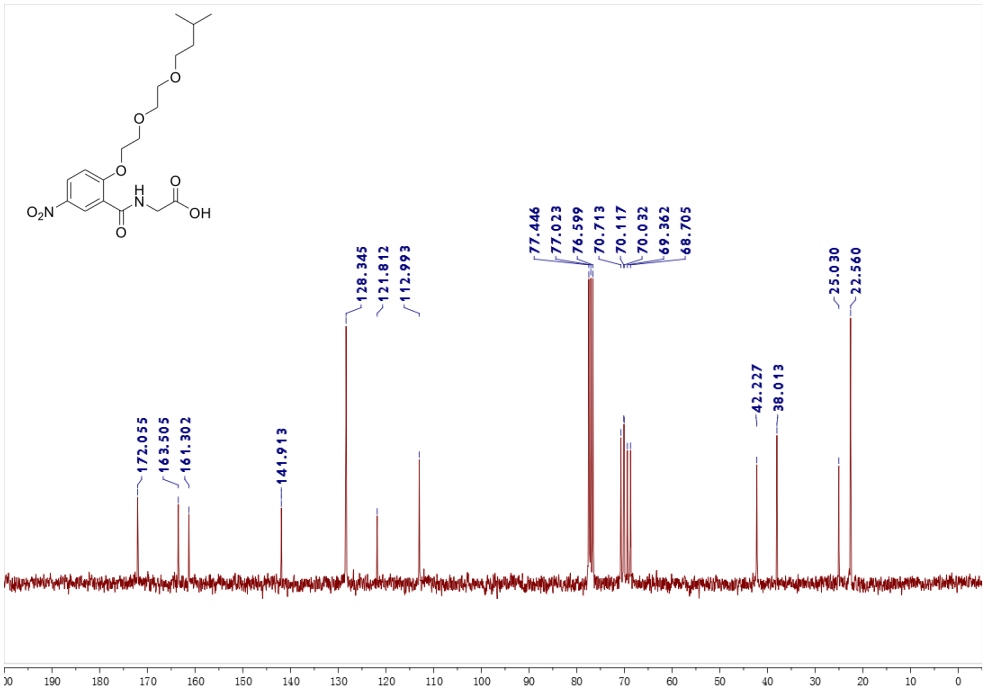

0.806
0.802
3.849
3.840
3.831
3.598
3.589
3.585
3.578
3.477
3.470
3.466
3.458
3.376
3.362
3.349
3.307
2.073
1.584
1.571
1.558
1.336
1.323
1.309
1.295
8.613
8.607
8.347
8.341
8.329
8.323
7.886
7.713
7.407
7.388
4.413
4.404
4.394
2.493
2.490
2.486
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5

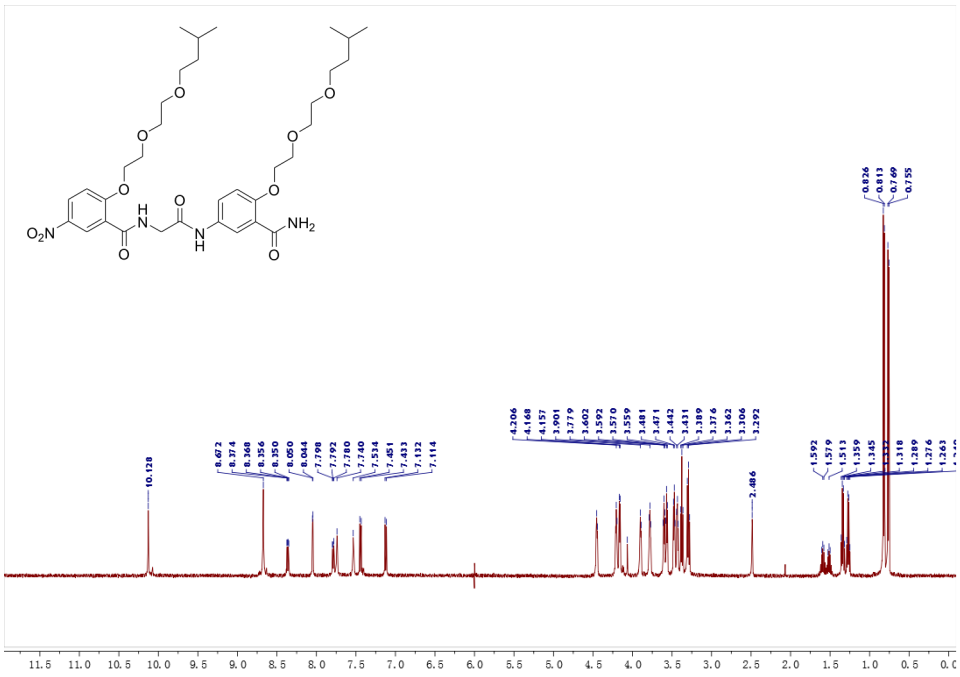

164.548
161.780
141.206
128.301
126.932
123.964
114.882
70.210
69.954
69.534
69.158
68.683
38.460
24.996
22.908
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0

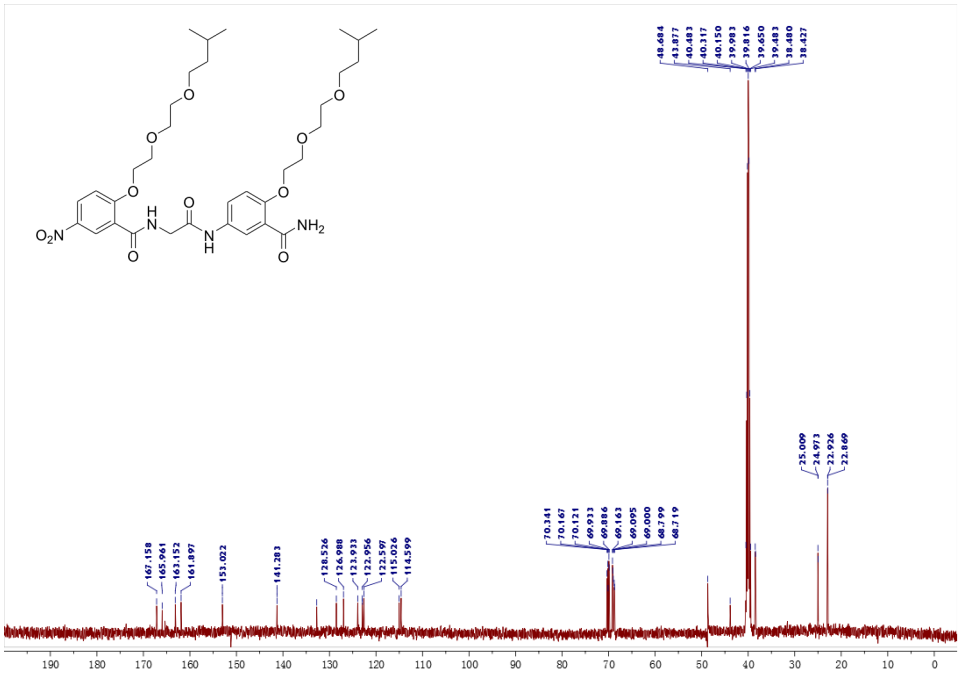

1.645
1.623
1.601
1.580
1.458
1.435
1.412
1.390
1.318
1.294
1.271
0.868
0.827
4.396
4.383
4.263
4.241
4.225
4.217
4.011
3.996
3.984
3.707
3.694
3.679
3.560
3.544
3.531
3.454
3.431
3.408
9.050
9.041
9.013
9.004
8.490
8.312
8.302
8.281
8.272
7.259
7.090
7.060
O2N
O
H
N
O
OEt

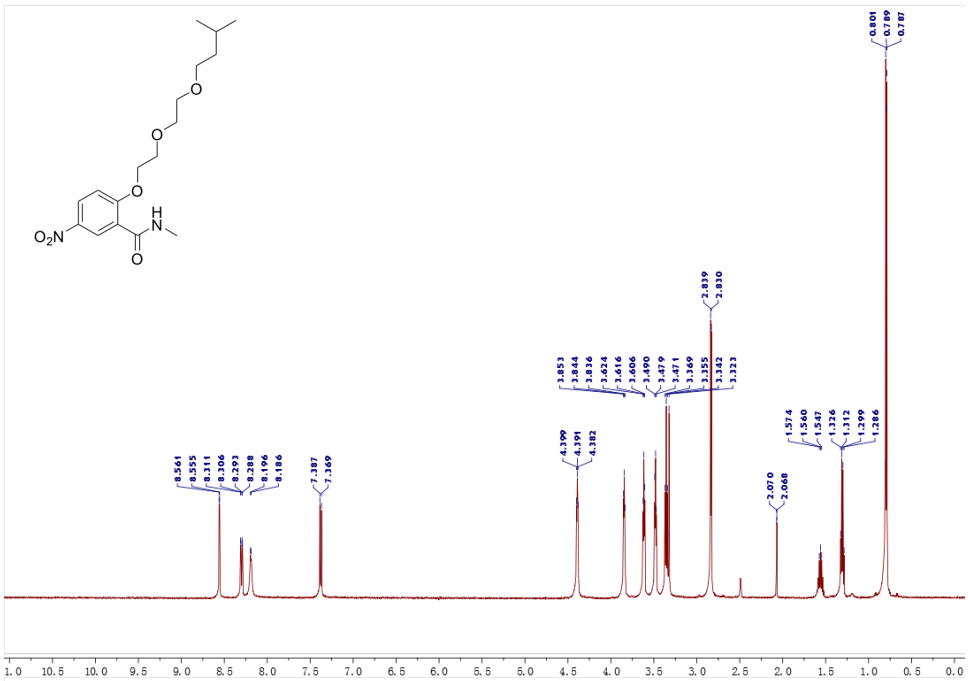

128.449
128.104
122.185
112.963
77.458
77.035
76.612
70.752
70.025
69.919
69.316
68.687
61.385
42.076
38.292
25.051
14.172
169.807
163.011
161.282
141.921
O2N
O
H
N
O
OEt

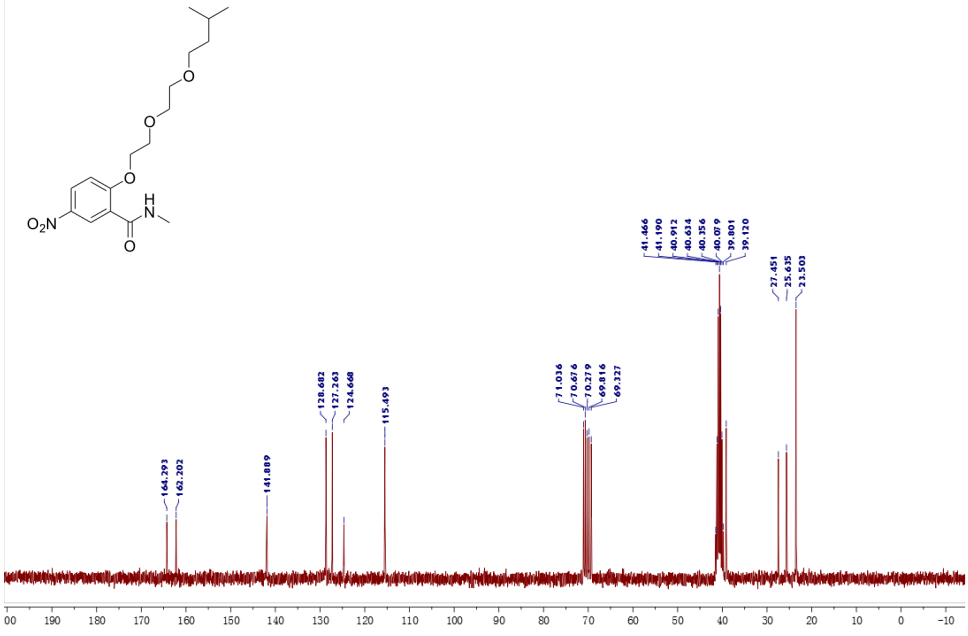

7.278
7.110
7.079
4.440
4.426
4.411
4.361
4.344
4.330
4.314
4.154
4.005
3.991
3.977
3.753
3.740
3.732
3.726
3.720
3.673
3.668
3.662
3.650
3.640
3.577
3.553
3.530
1.710
1.688
1.665
1.643
1.621
1.552
1.529
1.505
1.483
9.092
9.083
8.610
8.368
8.359
8.338
8.328
0.900

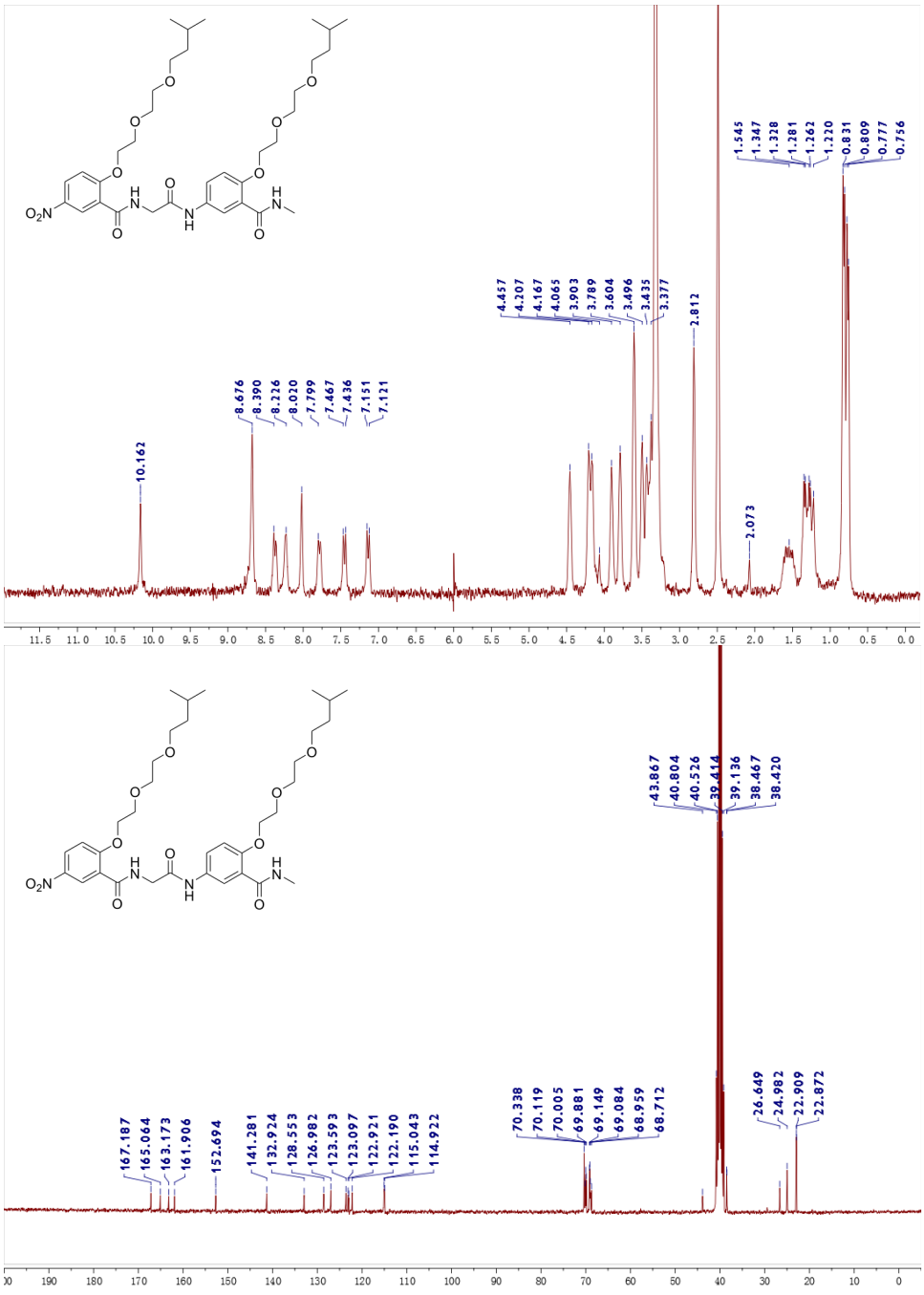

77.446
77.023
76.599
70.713
70.117
70.032
69.362
68.705
128.345
121.812
112.993
172.055
163.505
161.302
141.913
42.227
38.013
25.030
22.560

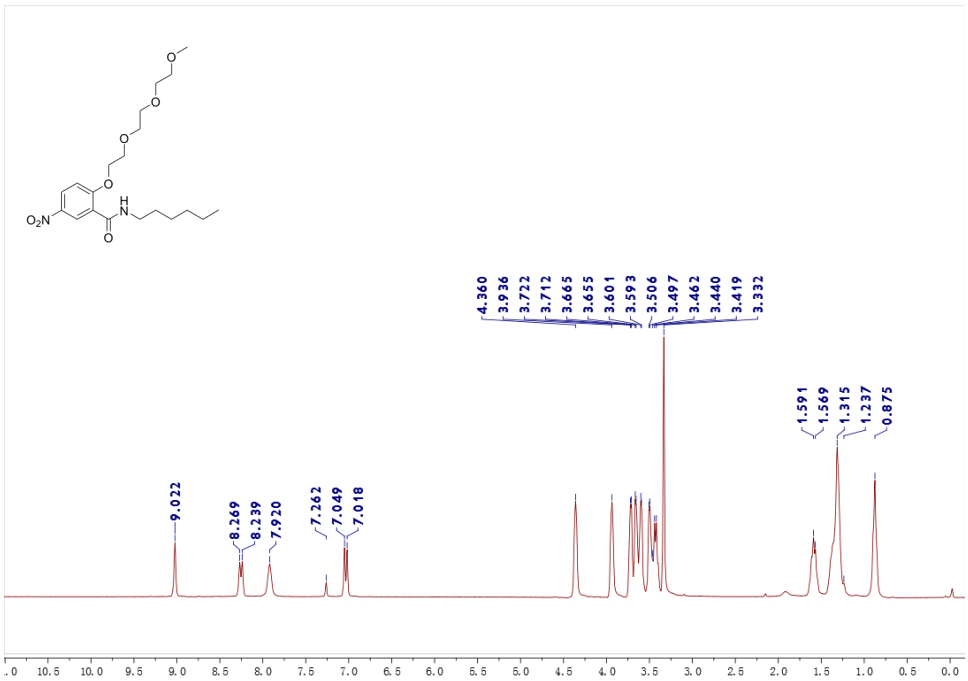


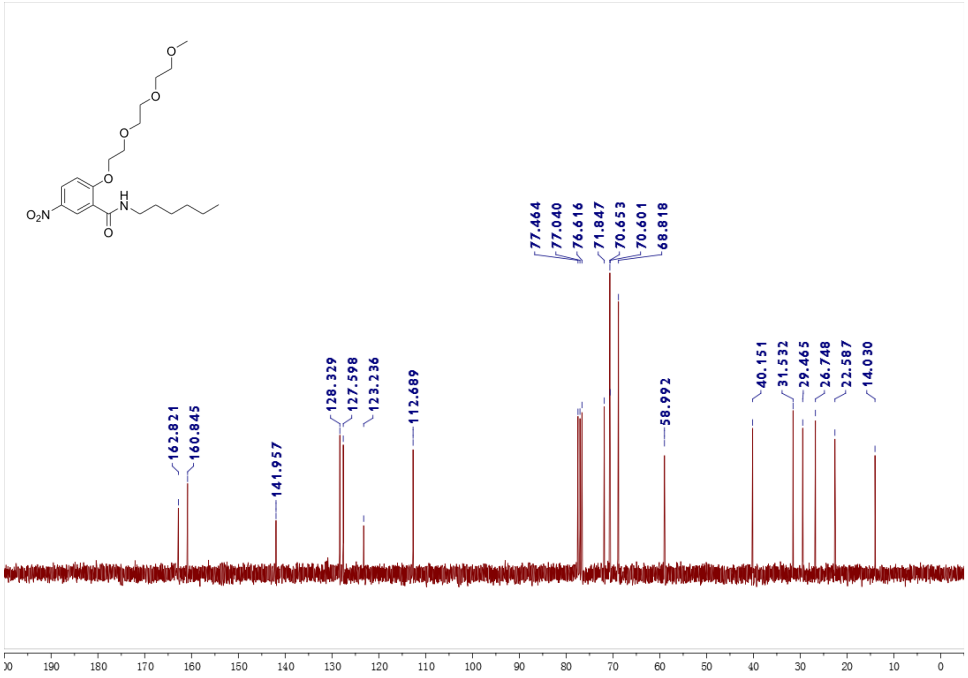


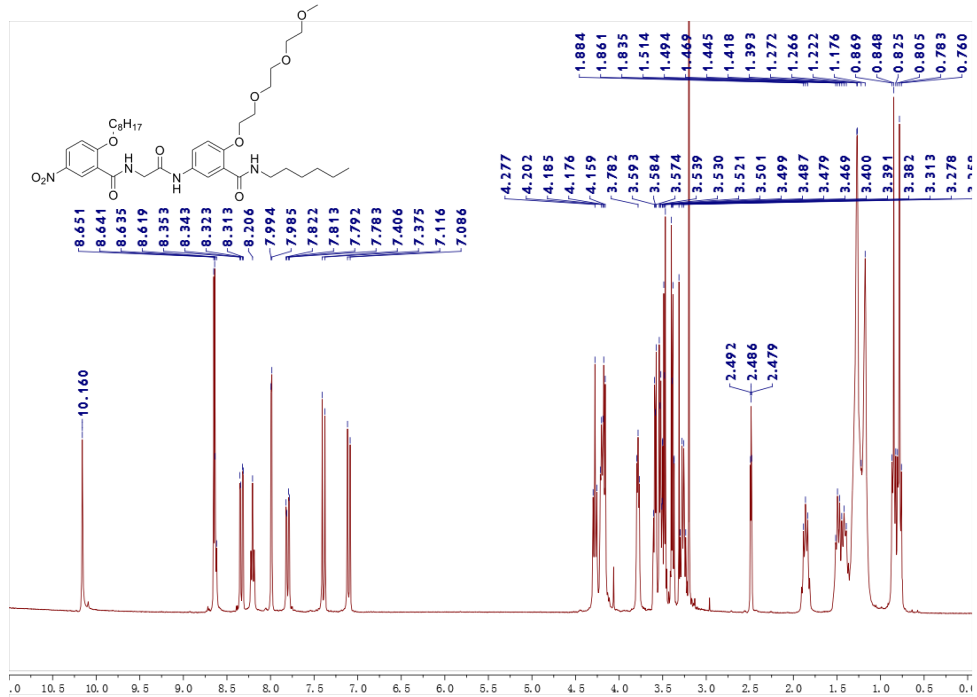

O
O
O
O
H
N
O2N
O
8.561
8.555
8.311
8.306
8.293
8.288
8.196
8.186
7.387
7.369
4.399
4.391
4.382
3.853
3.844
3.836
3.624
3.616
3.606
3.490
3.479
3.471
3.369
3.355
3.342
3.323
2.839
2.830
2.070
2.068
1.574
1.560
1.547
1.326
1.312
1.299
1.286
0.801
0.789
0.787
1.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

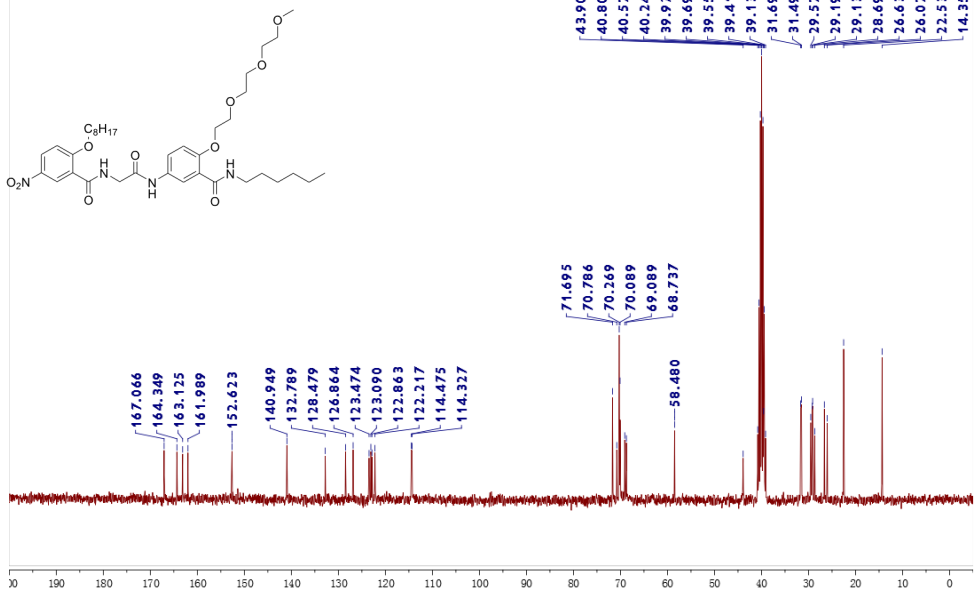

O
O
O
O
H
N
O2N
O
164.293
162.202
141.889
128.682
127.263
124.668
115.493
71.036
70.676
70.279
69.816
69.327
41.466
41.190
40.912
40.634
40.356
40.079
39.801
39.120
27.451
25.635
23.503
00 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10


10.162
8.676
8.390
8.226
8.020
7.799
7.467
7.436
7.151
7.121
4.457
4.207
4.167
4.065
3.903
3.789
3.604
3.496
3.435
3.377
2.812
2.073
1.545
1.347
1.328
1.281
1.262
1.220
0.831
0.809
0.777
0.756
43.867
40.804
40.526
39.414
39.136
38.467
38.420
167.187
165.064
163.173
161.906
152.694
141.281
132.924
128.553
126.982
123.593
123.097
122.921
122.190
115.043
114.922
70.338
70.119
70.005
69.881
69.149
69.084
68.959
68.712
26.649
24.982
22.909
22.872


4.360
3.936
3.722
3.712
3.665
3.655
3.601
3.593
3.506
3.497
3.462
3.440
3.419
3.332
1.591
1.569
1.315
1.237
0.875
9.022
8.269
8.239
7.920
7.262
7.049
7.018


77.464
77.040
76.616
71.847
70.653
70.601
68.818
162.821
160.845
141.957
128.329
127.598
123.236
112.689
58.992
40.151
31.532
29.465
26.748
22.587
14.030


C8H17
O2N
1.884
1.861
1.835
1.514
1.494
1.469
1.445
1.418
1.393
1.272
1.266
1.222
1.176
0.869
0.848
0.825
0.805
0.783
0.760
4.277
4.202
4.185
4.176
4.159
3.782
3.593
3.584
3.574
3.539
3.530
3.521
3.501
3.499
3.487
3.479
3.469
3.400
3.391
3.382
3.313
3.278
8.651
8.641
8.635
8.619
8.353
8.343
8.323
8.313
8.206
7.994
7.985
7.822
7.813
7.792
7.783
7.406
7.375
7.116
7.086
10.160
2.492
2.486
2.479


C8H17
O2N
43.908
40.807
40.527
40.249
39.971
39.693
39.551
39.416
39.139
31.690
31.490
29.526
29.197
29.130
28.691
26.630
26.025
22.533
14.356
71.695
70.786
70.269
70.089
69.089
68.737
58.480
167.066
164.349
163.125
161.989
152.623
140.949
132.789
128.479
126.864
123.474
123.090
122.863
122.217
114.475
114.327

4. ^{1}H NMR Spectra Recorded at Different Concentrations

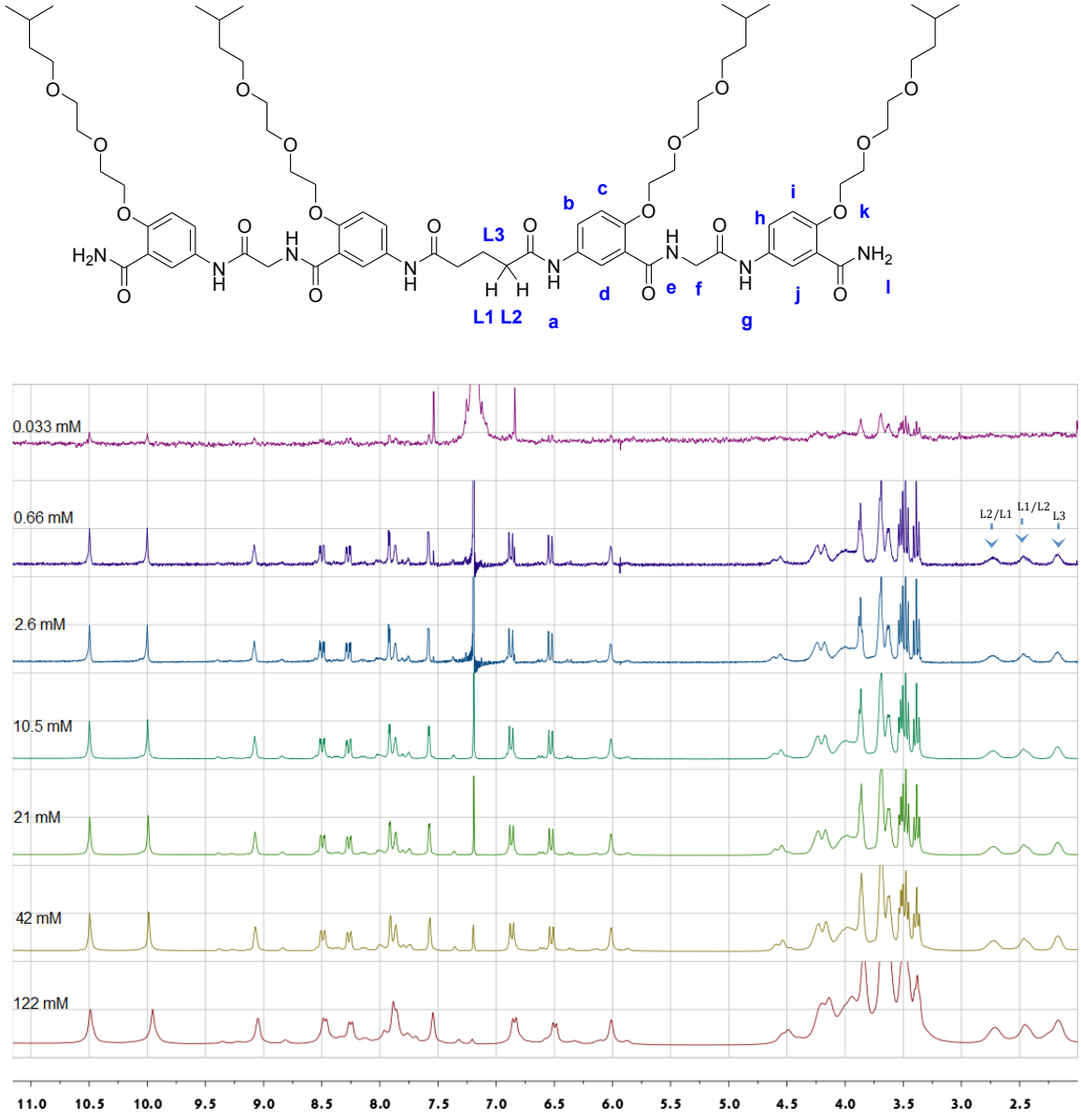


Figure S1. Concentration-dependent ^{1}H NMR experiment of **2** (400 MHz, $CDCl_3$, 298K). All the ^{1}H NMR signals remain sharp and well dispersed. The two signals of protons *L*1 and L2 remain unchanged with increasing concentrations.

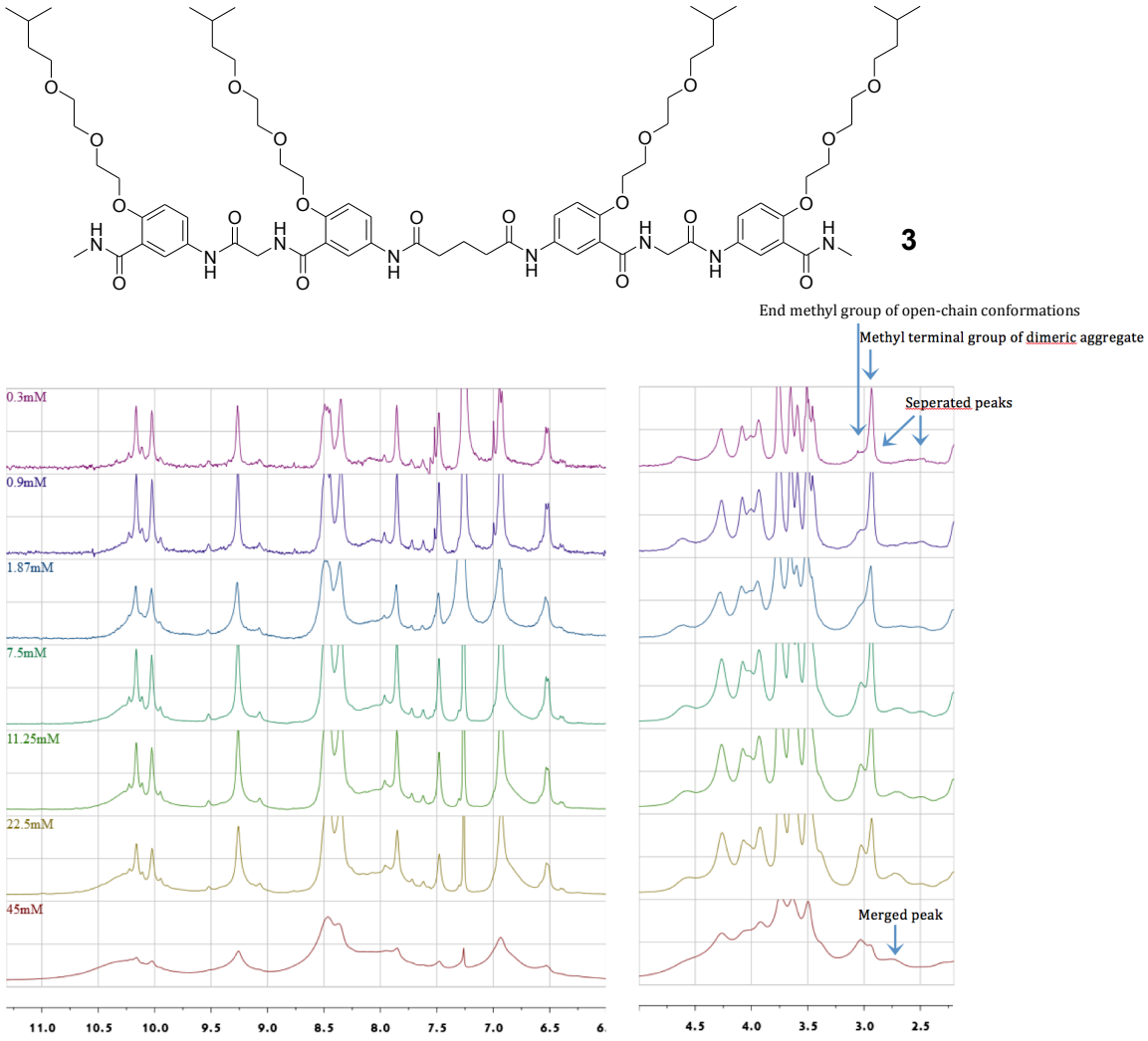


Figure S2. Concentration-dependent ^{1}H NMR experiment of **3** in $CDCl_3$ (400 MHz, 298K). One of the two signals of *L*1 and *L*2 overlaps with that of the end methyl groups, making it difficult to discern the merged peak (indicated) at ~2.7 ppm. Another observation is that the end methyl groups of the folded dimer give one signal, while an extra peak at 3.1 ppm becomes more and more prominent with increasing concentration, indicating the oligomers/polymers formation.

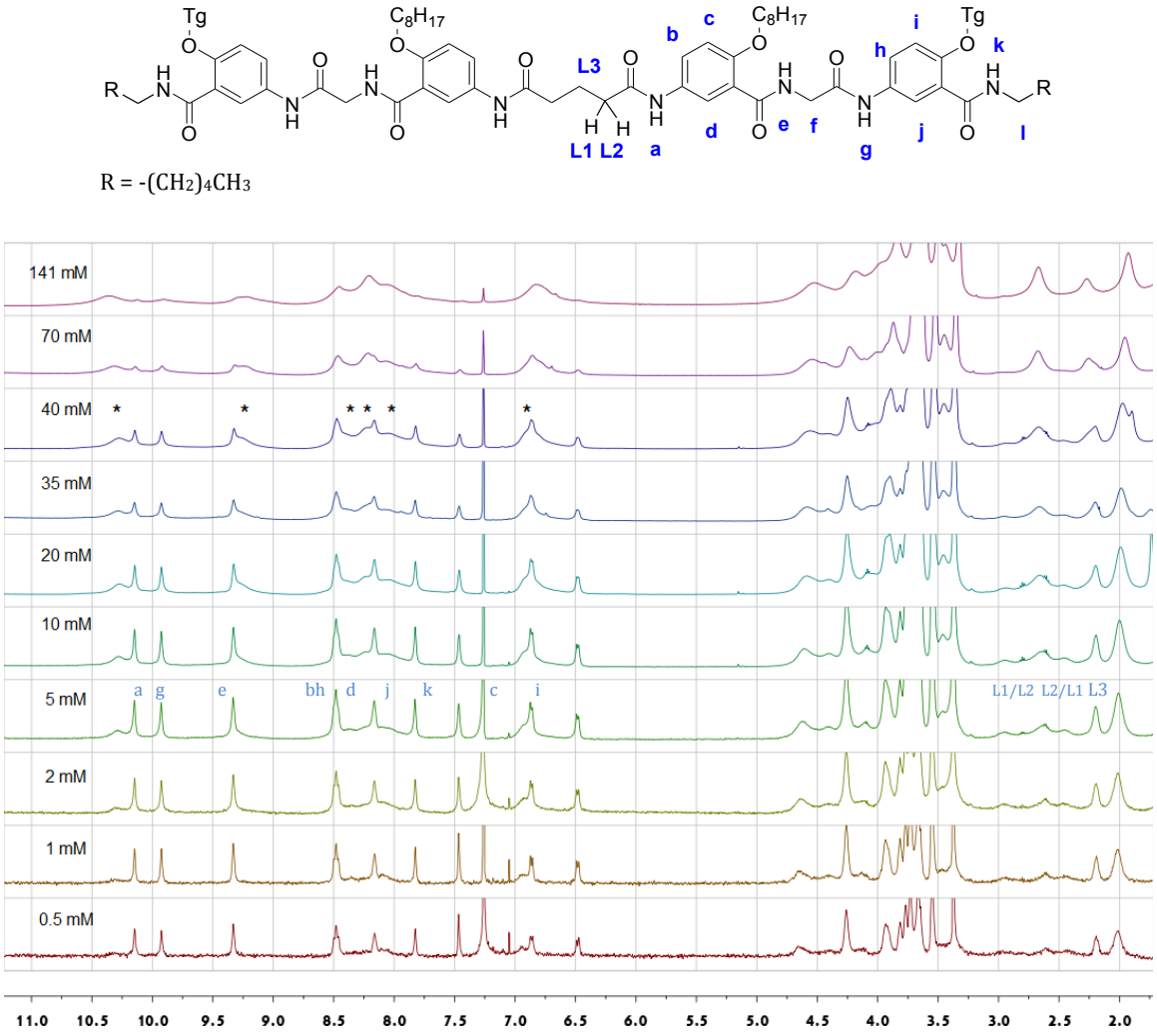


Figure S3. Concentration-dependent NMR spectra of compound **4** (500 MHz, $CDCl_3$, 298 K). Backbone signals of aggregates higher than dimer are marked with "*". Those aggregates are clearly observed at the concentration as low as 1 mM. The line-broadening of oligomers is probably due to fast exchange among assembled structures. For dimer, protons *L*1 and *L*2 show two peaks at 2.9 and 2.5 ppm respectively. The signals of *L*1 and *L*2 merge into a new peak at 2.7 ppm. The intensity of this new peak increases with increasing concentration.

5. IMS-MS Instrumental Analysis

Analysis of the samples of **2**, **3** and **4** was performed using an in-house PNNL built IMS-MS instrument which has previously been described.[3] Briefly, this instrumental platform couples a 1-m IMS separation with an Agilent 6224 TOF MS upgraded to a 1.5 meter flight tube providing resolution of ~25,000. Sample solutions were directly injected into the instrument using a chemically etched fused-silica emitter (20 µm I.D./150 µm O.D., at a potential of 2.6 kV)[4] and transported through a heated capillary inlet (0.43 mm I.D. x 64 mm at 120°C).[5] Once through the heated capillary, the ions were transmitted into the IMS drift cell via ion funnels. Following IMS separation, the ions were refocused using a rear ion funnel and transmitted into the TOF MS, which was set to collect data from 50-14300 *m/z* for each sample. The signal from the TOF detector was routed to 8-bit Analog-to-Digital converter (ADC) (AP240, Agilent Technologies, Switzerland) and processed using a custom control-software written in C#.[6]

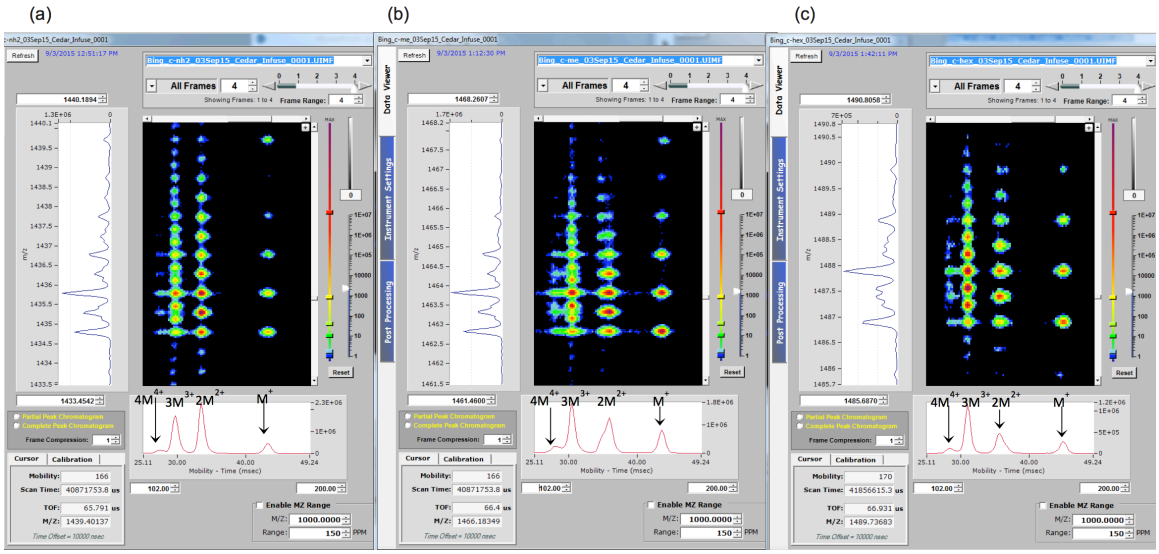
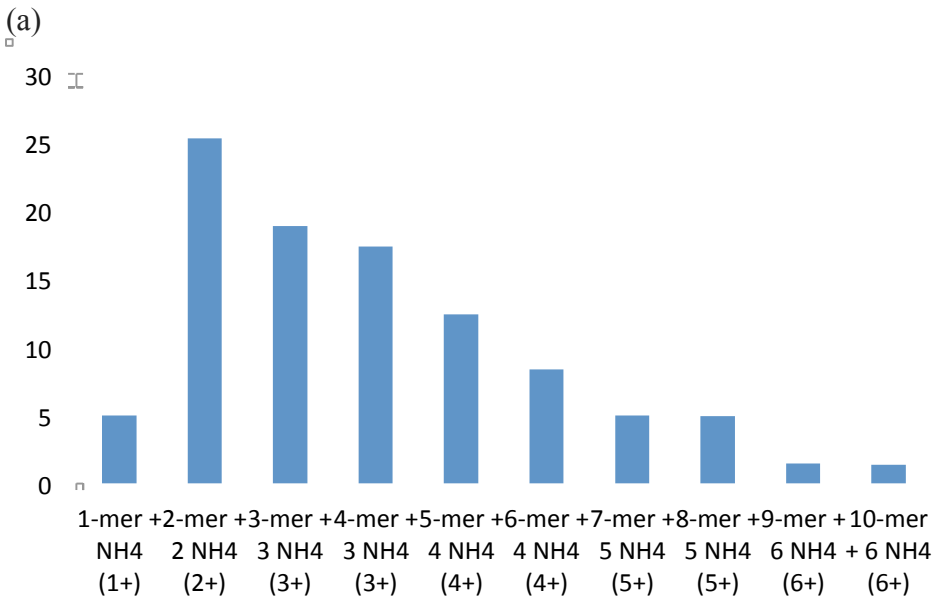


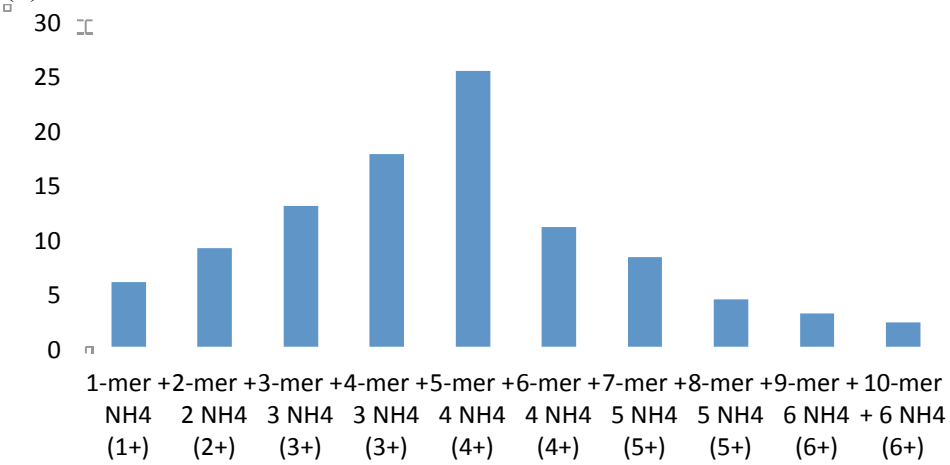
Figure S4. Partial IMS-MS spectra of (a) **2**, (b) **3**, and (c) **4** (NH_4^+ adducts).

Table S1. The Ten Most Abundant Complexes with NH_4^+ Adducts

Compound **2**	Pk Area	%	Compound **3**	Pk Area	%	Compound **4**	Pk Area	%
1-mer + NH_4 (1+)	450244	4.97	1-mer + NH_4 (1+)	826770	5.97	1-mer + NH_4 (1+)	281433	5.81
2-mer + 2 NH_4 (2+)	2293357	25.29	2-mer + 2 NH_4 (2+)	1258813	9.09	2-mer + 2 NH_4 (2+)	468164	9.66
3-mer + 3 NH_4 (3+)	1709768	18.86	3-mer + 3 NH_4 (3+)	1791863	12.94	3-mer + 3 NH_4 (3+)	1147577	23.68
4-mer + 3 NH_4 (3+)	1574243	17.36	4-mer + 3 NH_4 (3+)	2453069	17.71	4-mer + 3 NH_4 (3+)	870718	17.96
5-mer + 4 NH_4 (4+)	1125114	12.41	5-mer + 4 NH_4 (4+)	3512411	25.36	5-mer + 4 NH_4 (4+)	831246	17.15
6-mer + 4 NH_4 (4+)	758415	8.36	6-mer + 4 NH_4 (4+)	1524561	11.01	6-mer + 4 NH_4 (4+)	548115	11.31
7-mer + 5 NH_4 (5+)	451474	4.98	7-mer + 5 NH_4 (5+)	1144799	8.27	7-mer + 5 NH_4 (5+)	276366	5.70
8-mer + 5 NH_4 (5+)	446606	4.93	8-mer + 5 NH_4 (5+)	604358	4.36	8-mer + 5 NH_4 (5+)	204080	4.21
9-mer + 6 NH_4 (6+)	131772	1.45	9-mer + 6 NH_4 (6+)	422315	3.05	9-mer + 6 NH_4 (6+)	117315	2.42
10-mer + 6 NH_4 (6+)	125851	1.39	10-mer + 6 NH_4 (6+)	311961	2.25	10-mer + 6 NH_4 (6+)	101972	2.10
Total	9066844	100	Total	13850920	100	Total	4846986	100



(b)



(c)

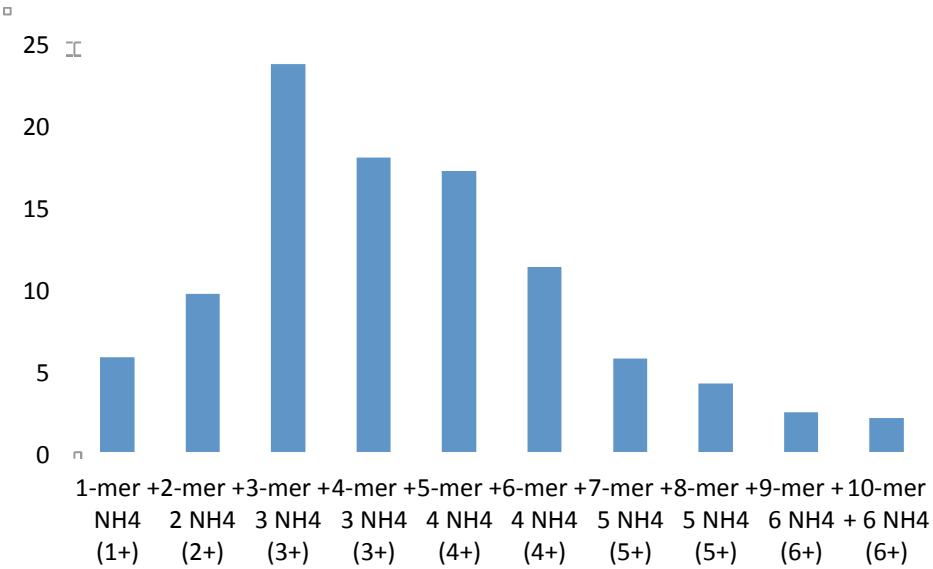
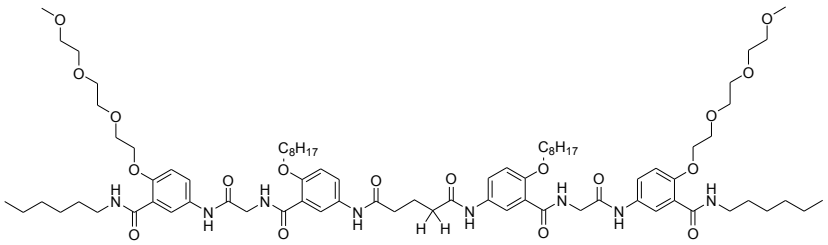


Figure S5. The relative abundance of the oligomeric aggregates of (a) **2**, (b) **3** and (c) **4** with ammonium adducts. The numbers on the vertical axes are shown in percentage.

6. Diffusion-Ordered Spectroscopy (DOSY)

Diffusion-ordered spectroscopy (DOSY) experiments were performed on a Varian Inova 500 MHz spectrometer under regulated temperature (298 K), with a 5 mm probe. The pulse sequence employed was a bipolar pulse pair simulated echo (BPPSTE). Additional parameters: gradient strength array has 15 increments from 3% to 66% of the maximum gradient strength in a linear ramp, diffusion gradient length is set to 2 ms, and diffusion delay is 100 ms.



4

Table S2. DOSY results of **4** in $CDCl_3$ at 298K

Concentration of **4** (hexyl terminals)	D_4 ($\times 10^{-10}$ m^2/s)	D_{TMS} ($\times 10^{-10}$ m^2/s)	D_{TMS}/D_4	D_{CHCl3} ($\times 10^{-10}$ m^2/s)	D_{CHCl3}/D_4
141 mM	0.24	N/A [a]	------	10.4	43.48
70 mM	1.11	10.3	9.26	15. 5	13.89
35 mM	2.20	17.3	8.00	19.7	9.00
25 mM	2.29	17.1	7.46	20.6	9.00
12 mM	2.70	18.5	6.85	20.7	7.69
0.5 mM	3.28	19.5	5.95	21.6	6.57

[a]Value can not be measured accurately.

D_4: averaged diffusion coefficient of **4**.

D_{TMS}: diffusion coefficient of TMS.

D_{CHCl3}: diffusion coefficient of $CHCl_3$.

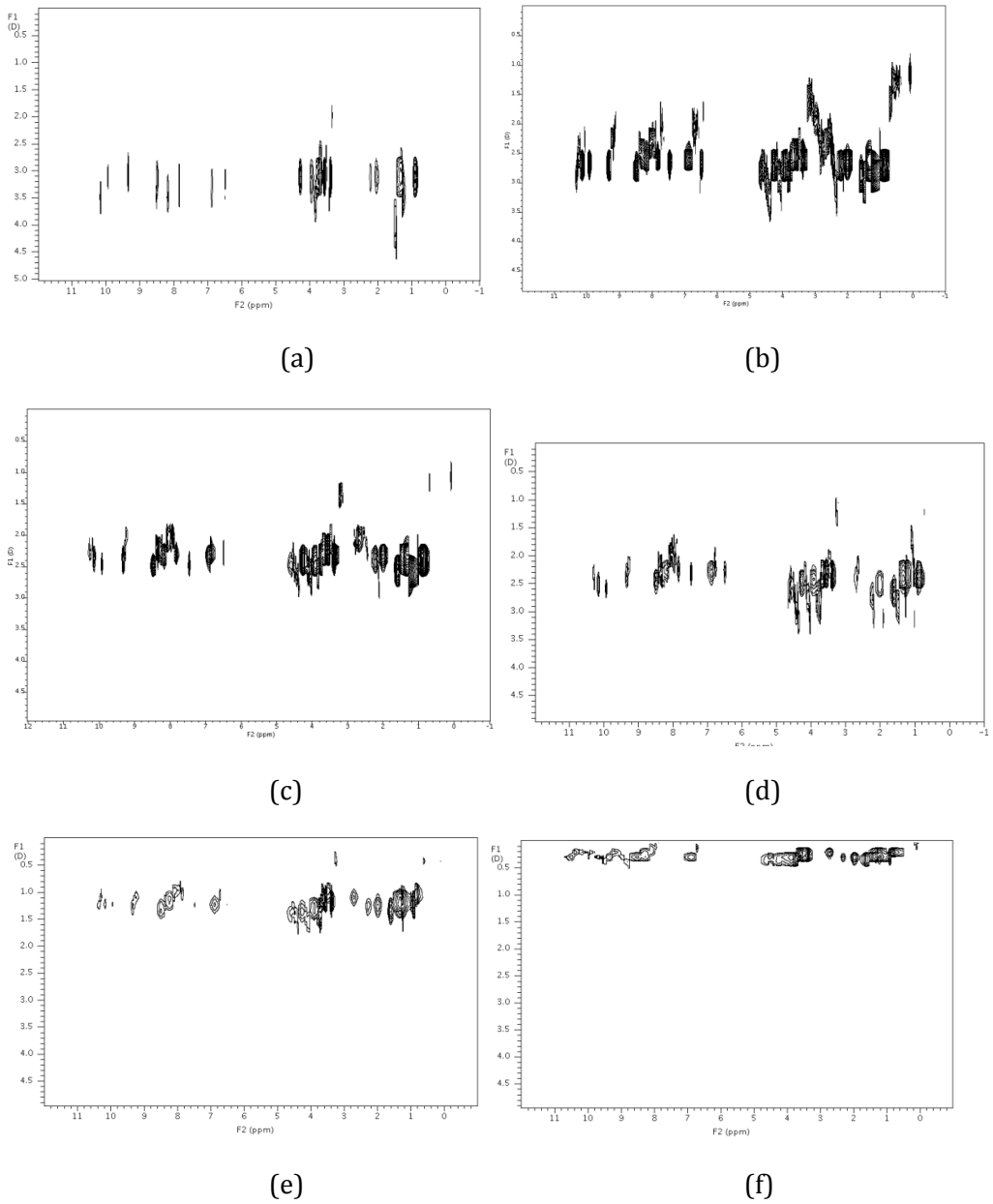


Figure S6. Partial DOSY spectra of **4** in $CDCl_3$ (a) 0.5 mM, (b) 12 mM, (c) 25 mM, (d) 35 mM, (e) 70 mM, and (f) 141 mM.

7. Viscosity Measurements

Viscosity measurements were carried out at ambient temperature (298K) on a Brookfield DV2+Pro Viscometer with a thermostat attached. The sample solutions in $CHCl_3$ were filtered through a 0.45 μm filter to remove dust and debris. The resulting solution was left to stand for 1 hour before measurements.

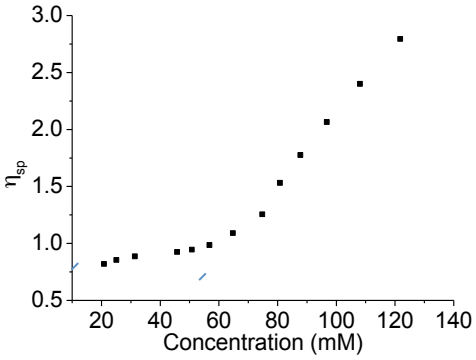
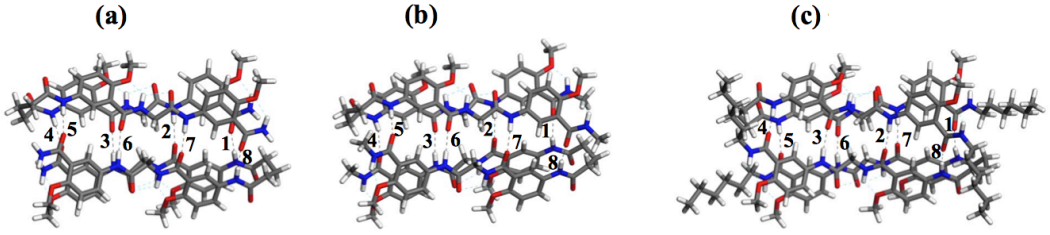


Figure S7. Specific viscosity of **3** versus concentration in $CHCl_3$. The results suggest that there are two-stage aggregations. Combined with NMR data, dimeric assembly dominates at low concentrations, while larger aggregates are favored at high concentrations. The slopes are 0.0063 ± 0.0004 cP/mM at low concentrations and 0.0506 ± 0.0020 cP/mM at high concentrations, respectively.

8. Computational Study

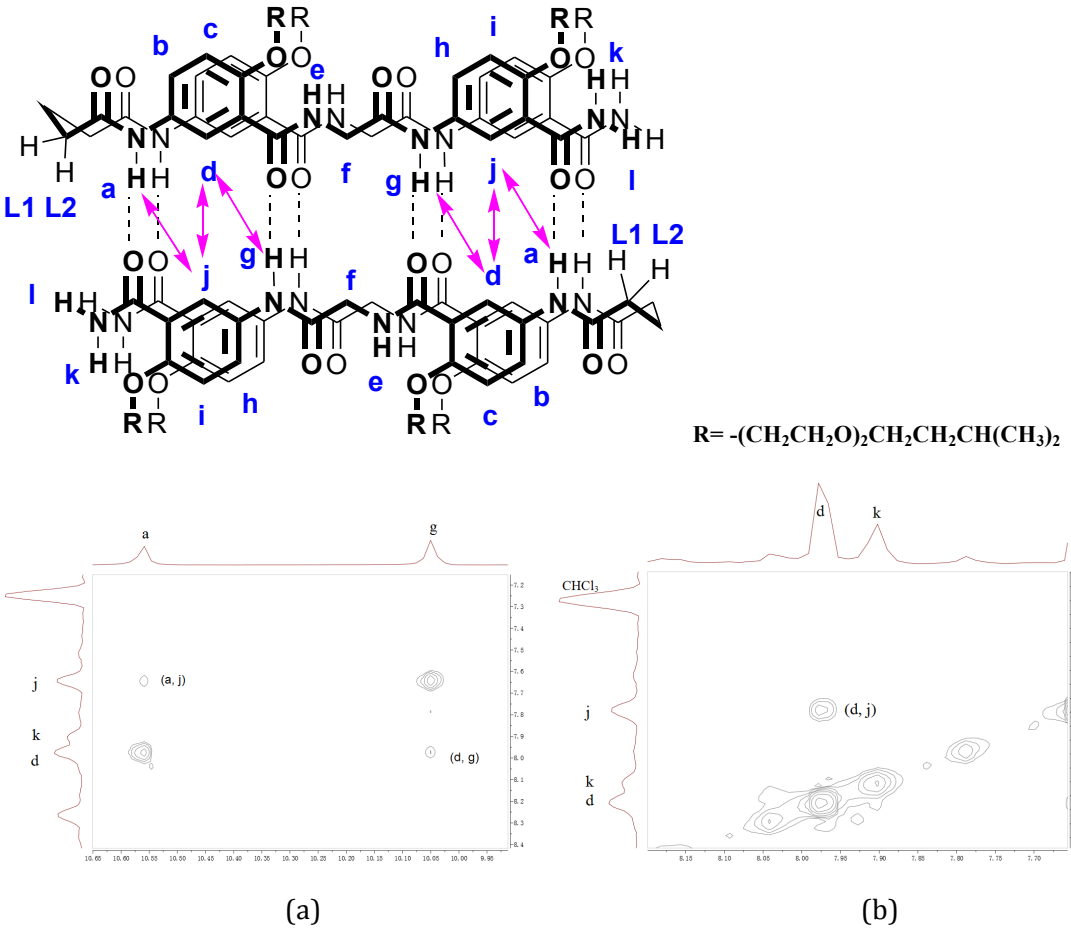
B3LYP/6-31G(d) optimizations were performed on the dimers of compounds 2, 3, and 4 (with all R1 and R2 groups being replaced by methyl groups) in Gaussian 09 software package. The three dimers with eight intermolecular hydrogen bonds are obtained after optimizations both shown in Figures S9 and S10.



	(a)		(b)		(c)	
	Length (Å)	Angle (°)	Length (Å)	Angle (°)	Length (Å)	Angle (°)
HB1	1.986	171.7	2.025	161.9	2.071	169.3
HB2	1.900	174.1	1.912	169.4	1.913	160.2
HB3	1.902	173.6	1.914	160.8	1.904	173.0
HB4	1.985	171.4	2.103	161.3	1.996	166.0
HB5	1.961	175.0	1.963	172.3	2.003	169.6
HB6	1.908	175.7	1.924	167.7	1.872	171.7
HB7	1.909	175.8	1.886	175.7	1.965	167.6
HB8	1.961	174.6	1.978	173.1	2.002	163.1

Figure S8. Optimized structures of the dimers of compounds (a) **2**, (b) **3**, and (c) **4**. The bond lengths and bond angles of the eight intermolecular hydrogen bonds (HB) of each dimers are listed.

9. 2D-NOESY Spectra



(a) (b)

Figure S9. Partial NOESY spectra of **2** in $CDCl_3$ (500MHz, 298K, 3 mM).

10. Reference

1. R. Liu, A. L. Connor, F. Y. Al-mkhaizim and B. Gong, *New J. Chem.* 2015, **39**, 3217.
2. X. W. Yang, S. Martinovic, R. D. Smith and B. Gong, *J. Am. Chem. Soc.* 2003, **125**, 9932.
3. E. S. Baker, B. H. Clowers, F. Li, K. Tang, A. V. Tolmachev and D. C. Prior, *J. Am. Soc. Mass Spectrom*. 2007, **18**, 1176-87.
4. R. T. Kelly, J. S. Page, Q. Luo, R. J. Moore, D. J. Orton, K. Tang and R. D. Smith, Anal. Chem., 2006, **78**, 7796.
5. T. Kim, A. V. Tolmachev, R. Harkewicz, D. C. Prior, G. Anderson and H. R. Udseth, *Anal. Chem*. 2000, **72**, 2247.
6. Y. M. Ibrahim, E. S. Baker, W. F. 3rd Danielson, R. V. Norheim, D. C. Prior, G. A. Anderson M. E. Belov and R. D. Smith, *Int. J. Mass Spectrom*. 2015, **377**, 655.