Supporting information for:

Universal Relationship between Molecular Structure and Crystal Structure in Peptoid Polymers and Prevalence of the *Cis* Backbone Conformation

Douglas R. Greer[1,2], Michael A. Stolberg[1], Joyjit Kundu[3†], Ryan K. Spencer[4], Tod Pascal[3], David Prendergast[3], Nitash P. Balsara*[1,2], Ronald N. Zuckermann*[3]

[1]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 United States.

[2]College of Chemistry, University of California, Berkeley, Berkeley, CA 94720 United States.

[3]Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 United States.

[4]Department of Chemistry and Department of Chemical Engineering & Materials Science, University of California, Irvine, Irvine, California.

[†]Present address: Department of Chemistry, Duke University, Durham, North Carolina 27708 United States.

*nbalsara@berkeley.edu; rnzuckermann@lbl.gov

Peptoid Synthesis

The mPEG amine submonomers for Nde ($mPEG_2$-NH_2) were purchased from PurePEG (98% purity), and for Nte ($mPEG_3$-NH_2) were purchased from Peptide Solutions, Inc. (98% purity). Linear alkyl amines were purchased from TCI (>98% purity). All submonomers were used without further purification. Automated solid-phase submonomer synthesis of the polypeptoids was performed on a Symphony X peptide synthesizer at a scale of 200 mg Rink amide resin (0.64 mmol/g) following published procedures.[1] Displacement reactions were performed at amine concentrations of 1 M in N,N'-dimethylformamide (DMF) for 30 min at room temperature. Bromoacetylation reactions were performed with bromoacetic acid and N,N'-diisopropylcarbodiimide (both at 0.8 M in DMF) for 20 min at room temperature.

Peptoids were cleaved from resin by treating with an acid cleavage cocktail of dichloromethane/trifluoroacetic acid/water (50/45/5, v/v) for 10 min at room temperature. After resin filtration and washing, the cocktail was evaporated, and the peptoids were lyophilized from acetonitrile/water (1:1, v/v).

Acetylation was performed on the crude, cleaved peptoid (~200 mg). The peptoid was dissolved in 2 mL DMF/tetrahydrofuran (THF) (1:5, v/v), followed by the addition of 100 µL acetic anhydride and 100 µL pyridine, and allowed to stir at room temperature for 20 minutes. The volatiles were then removed by evaporation, and the peptoid lyophilized from acetonitrile/water (1:1, v/v).

Purification was performed on a reverse phase HPLC Waters prep system using a XSelect HSS cyano column (5 µm, 18 x 150 mm). A linear, binary elution gradient was used (where solvent A was 10% isopropyl alcohol (IPA) in water, and solvent B was 10% IPA in ACN), from 50 – 95% B in 20 minutes at a flow rate of 20 mL/min). Note that TFA was not included in the buffer system, since the peptoid *N*-acetyl group exhibits some acid lability.[2] 50 mg of peptoid dissolved in 3 mL of ACN/water (1:1, v/v), was used per injection. The fractions were collected and analyzed by MALDI using a superDHB matrix. The collected fractions containing pure product were combined, evaporated, and lyophilized from ACN/water (1:1, v/v) to afford a fluffy white powder. The product was then analyzed by reverse phase HPLC equipped with an analytical cyano column and MicroTOF electrospray mass spectrometer. 50 mg of final product was obtained in most cases.

Nanosheet Formation

To form nanosheets, each peptoid we synthesized was dissolved in 50/50 v/v tetrahydrofuran (THF)/water at a concentration of 2 mg/mL. The THF was evaporated under a vacuum of 500 torr at room temperature, over a period of 24 hours, leaving behind the aqueous peptoid solution at a concentration of about 6 mg/mL. The solutions turned turbid, consistent with the formation of nanosheets, which was verified by atomic force microscopy. WAXS measurements were performed as described above on solutions encapsulated in Kapton windows. The nanosheet samples were not annealed. Dry nanosheets were obtained by putting drops of aqueous nanosheet solution on Kapton and drying the sample in a fume hood.

X-ray Scattering

Wide angle X-ray scattering (WAXS) measurements were performed at ALS beamline 7.3.3.3 Small angle X-ray scattering (SAXS) measurements were performed at ALS beamline 7.3.3 and SSRL beamline 1-5. Prior to measurement, lyophilized peptoids were placed between Kapton windows separated by a rubber gasket spacer, annealed at 125 °C for half an hour, and cooled slowly to remove thermal history. The diblock peptoids synthesized for this study are known to be isotropic above 100 °C and known to be thermally equilibrated with slow cooling, as measured by DSC.[4] The X-ray beam path length through the sample is not standardized, which means that absolute peak intensities between samples should not be quantitatively compared.

Molecular Dynamics

Based on the experimentally observed X-ray scattering, we prepared a number of low energy, ordered bulk phase peptoid configurations as initial structures for our molecular dynamics simulations. The simulations were performed with the simulation package NAMD.[5] To explore the underlying conformations of the peptoids within the bulk-phase, we used the polymer Ac-Ndc_9-Nte_9 as our model. We considered a bulk phase consisting of all-*cis* and all-*trans* conformations of this polymer, expecting one of these conformations to be dominant due to free energy differences. We also considered that the polymer molecules have a polar N terminal block and a hydrophobic C terminal block, meaning each molecule has a backbone directionality. With this directionality in mind, we considered that each molecule's neighbor in the *a* and *c* direction could either be assembled parallel or antiparallel to it. The structure is mainly stabilized by the strong interaction in the close packing direction (*a* direction). Thus, we quickly found that bulk phases in which adjacent molecules are arranged antiparallel in the *a* direction are disfavored due to higher internal energy configurations. We completed simulations of bulk phases with parallel-parallel assemblies (designated PP), and with parallel assembly along the *a* direction but antiparallel along the *c* direction (designated PA). We note that the physical Ac-Ndc_9-Nte_9 material will likely contain contributions from both PP and PA assemblies.

For each simulation, a periodic box consisting of 288 polymers (16 along *a*, 3 along *b*, and 6 along *c*-axis) was considered. Each system is equilibrated for at least 100 ns of molecular dynamics simulations in the isothermal-isobaric (NPT) ensemble at 298 K and 1 atm pressure where all the three orthogonal box dimensions were allowed to fluctuate independently. The constant pressure is maintained using the Nosé-Hoover Langevin piston method[6,7] with an oscillation period of 100 fs and decay time of 50 fs. All the bonds involving hydrogen atoms were kept rigid (using SHAKE[8]). Particle-mesh Ewald summations[9] were used to evaluate long-range electrostatic interactions with a real-space cut-off of 12 Å. Van der Waals interactions were calculated up to a distance of 12 Å , with a smoothing function applied from 10 Å to 12 Å to ensure vanishing energies and forces at the cutoff. All the systems were simulated using a leap-frog integration algorithm with a 2-fs timestep. The relaxation for each model is checked by ensuring that the final configuration is independent of the initial preparations. The data was recorded in the steady state (energy and box dimensions fluctuate around the respective mean values). The equilibrated

configuration (in the NPT ensemble) was again evolved in the NVT (constant number of particles, volume, and temperature) ensemble to estimate the internal energy. The absolute value of Gibbs free energies were determined using the Two Phase Thermodynamics (2PT) method as discussed in the main text. Here we estimate the absolute entropy and quantum (zero point energy and heat capacity) corrections to the enthalpy from the vibrational density of states function (vDOS), determined from a Fourier transform for the atomic velocity autocorrelation function, *i.e.,* the power spectrum. When determining the vDOS, we ran an additional 4 ps canonical (constant particle, constant temperature or NVT) simulation, saving snapshots of the system (atomic velocities and coordinates) every 1 fs. In the 2PT formalism, each discrete frequency in the vDOS is modeled as a quantum harmonic oscillator. The value of internal energy over time is shown in Figure S1. The *cis* backbone conformations have much lower energy than the *trans*-backbone conformation. The two models (PP and PA) of the *cis* conformation are approximately energetically degenerate.

Using the equilibrated configurations obtained from MD simulations, we calculated the histograms of *a*, *b*, and *c* spacings for Ac-Ndc$_9$-Nte$_9$. The spacings calculated from MD simulations of cis-PP and cis-PA are in good agreement with spacings found with X-ray scattering measurements, as seen in Figure S1. The *a* distance was calculated by measuring the distance between each nitrogen atom and the nearest nitrogen atom on an adjacent molecule in the *a* direction. The *b* dimension was calculated by measuring the distance between each carbon atom at the N terminus and the carbon atom on an adjacent molecule in the *b* direction. The *c* dimension was calculated by measuring the distance between each carbon atom on the backbone and the nearest carbon atom on an adjacent molecule in the *c* direction. The *c* dimension distribution for the *trans* conformation model has a broader distribution indicative of disorder and slight bending of the backbones within the structure.

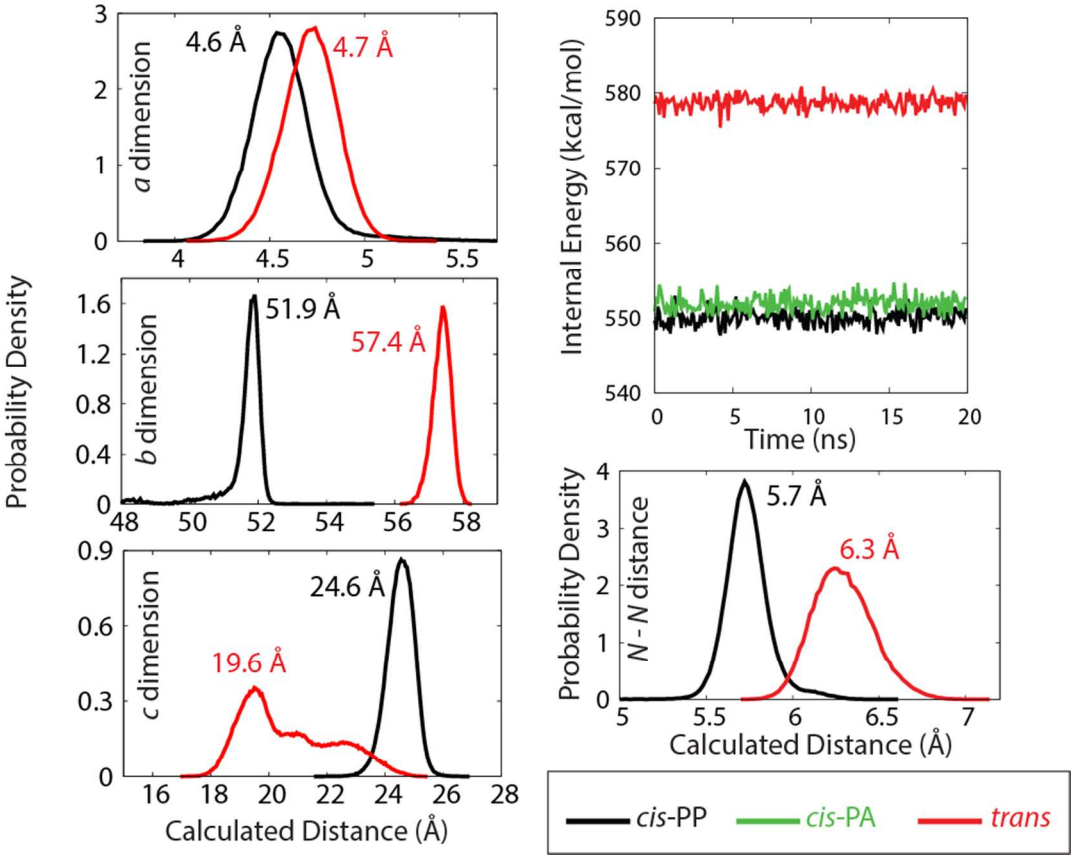


Figure S1. Simulation data from a molecular model of compound **1**.

(Left Panel) The probability densities of *a*, *b* and *c* dimensions averaged over 100 well separated time points after relaxation. Probability maxima are labeled.

(Right Panel) (Top) The internal energy (without the Zero Point energy) is plotted versus time for after equilibration. (Bottom) Probability distribution of the distance between intramolecular backbone N atoms on the same side backbone (see Figure 2, main text)

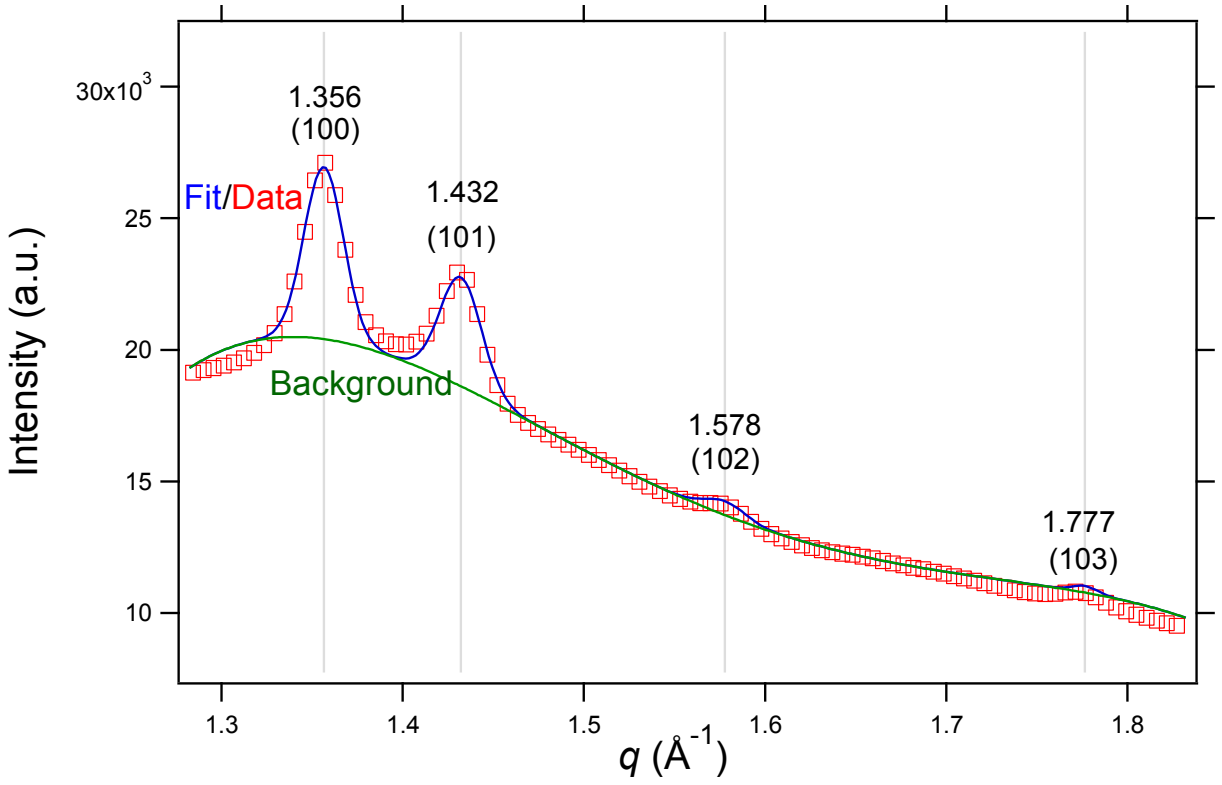


Figure S2. WAXS detail for compound **4**. The high-*q* WAXS data and fit for Ac-Ndc$_9$-Nte$_9$, with the peaks indexed and their *q* positions labeled. The fit is plotted in blue on top of the data, and the background fit (log poly5) is shown in green. The angle between *a* and *c* is 96.1 degrees.

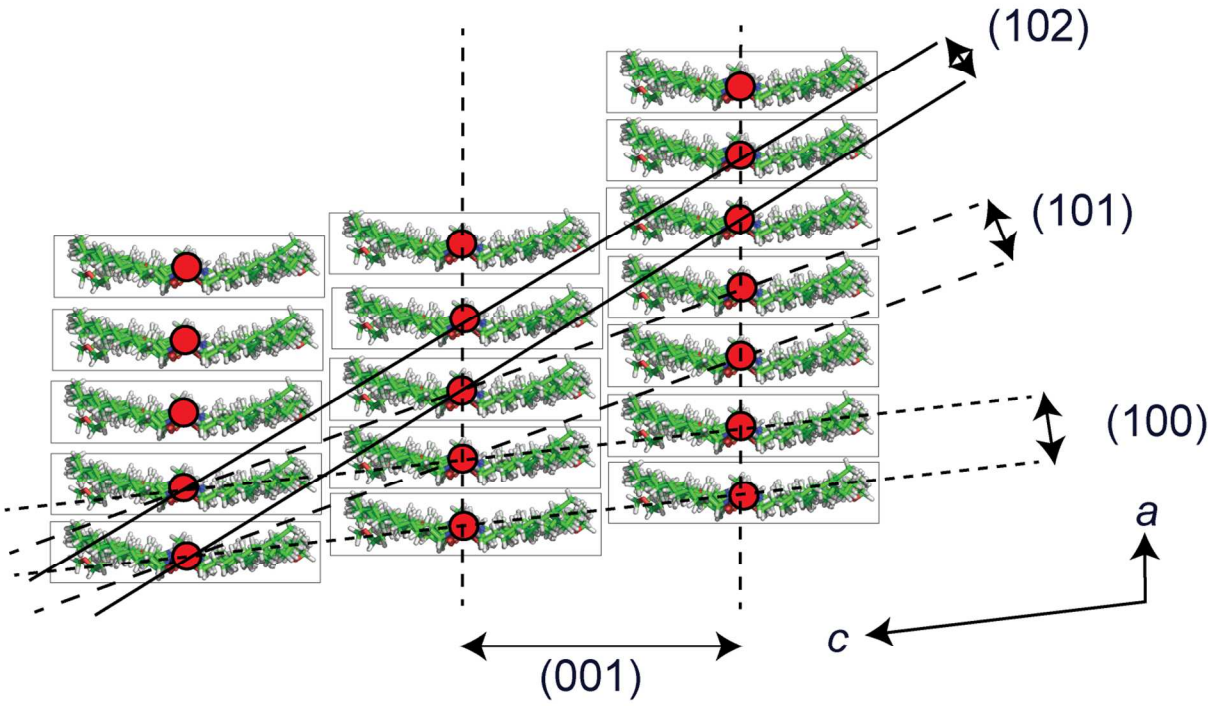
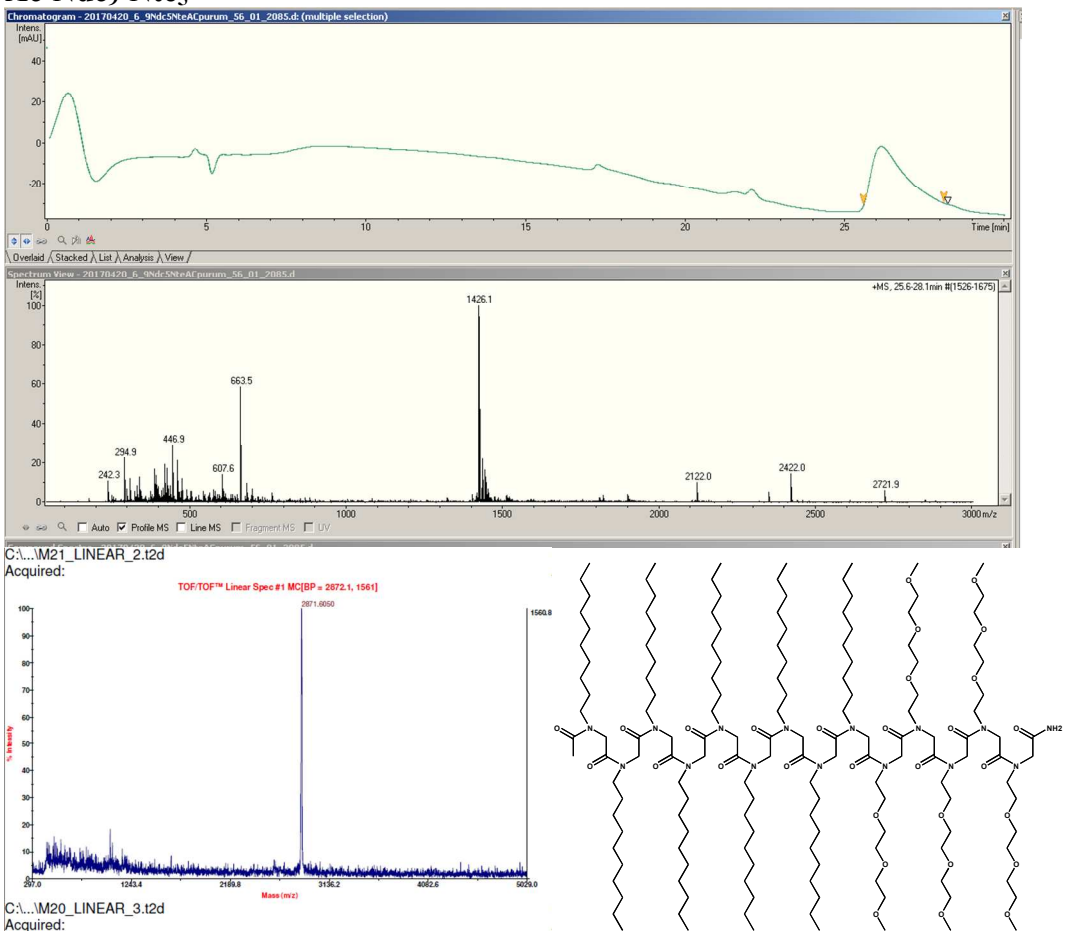


Figure S3. A heuristic for the crystal planes seen in the WAXS data. A simplified real-space lattice in the *a* and *c* directions is made from compound **1** molecules viewed along the *b* direction (along the backbone). The origin of the non-orthogonality between *a* and *c* can be seen as resulting from the half-offset shift in the *a* direction when progressing in the direction along the sidechains. This shift also explains why the non-orthogonal angle decreases with increasing sidechain length. The higher order reflections in compound **4**, which has shorter sidechains, are shifted to higher *q* compared to compound **1** since its higher order planes will have a larger angle away from the (100) plane. The origin of the (100), (101), and (102) reflections are shown as planes, and the plane corresponding to the (103) reflection can be extrapolated from the pattern. When compared to the actual *cis* backbone simulation, this heuristic has exaggerated *a* spacing for clarity.

Peptoid Characterization Data

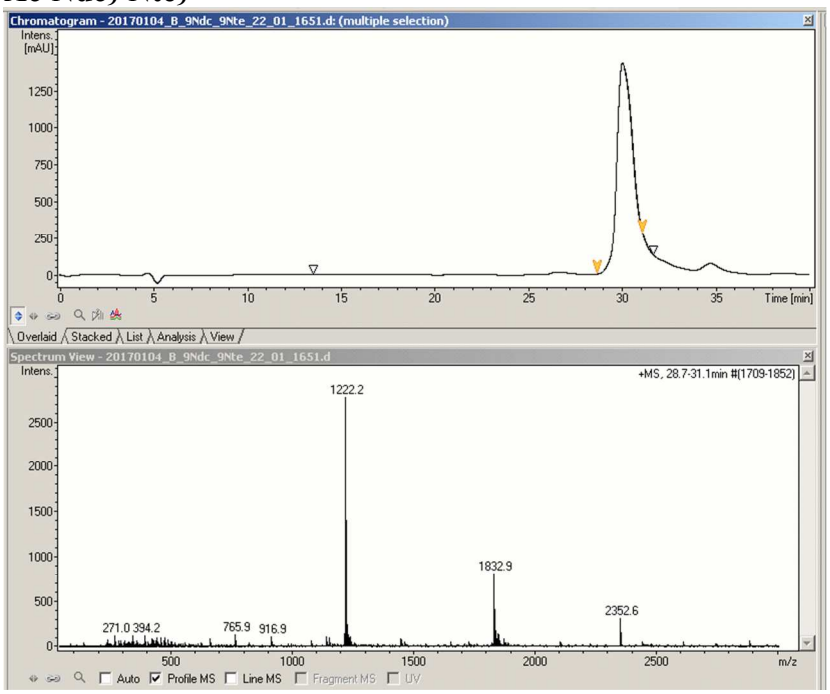
Ac-Ndc$_9$-Nte$_5$

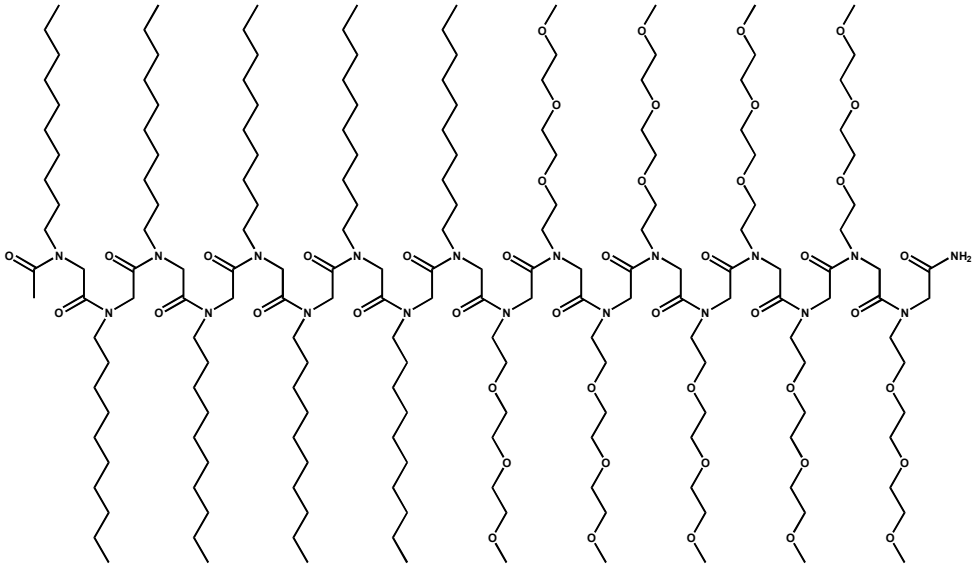


M/Z = 2850.22

Observed Mass in MicroTOF: 1426.1 (+2 H^+); 951.1 (+3 H^+);
The peak at 2352.6 is a contaminant in the MicroTOF, and is present in all runs for all users. The Observed Mass in MALDI is 2871.6, which is within instrument sensitivity (1 part per thousand m/z) of + ^{23}Na

Ac-Ndc$_9$-Nte$_9$

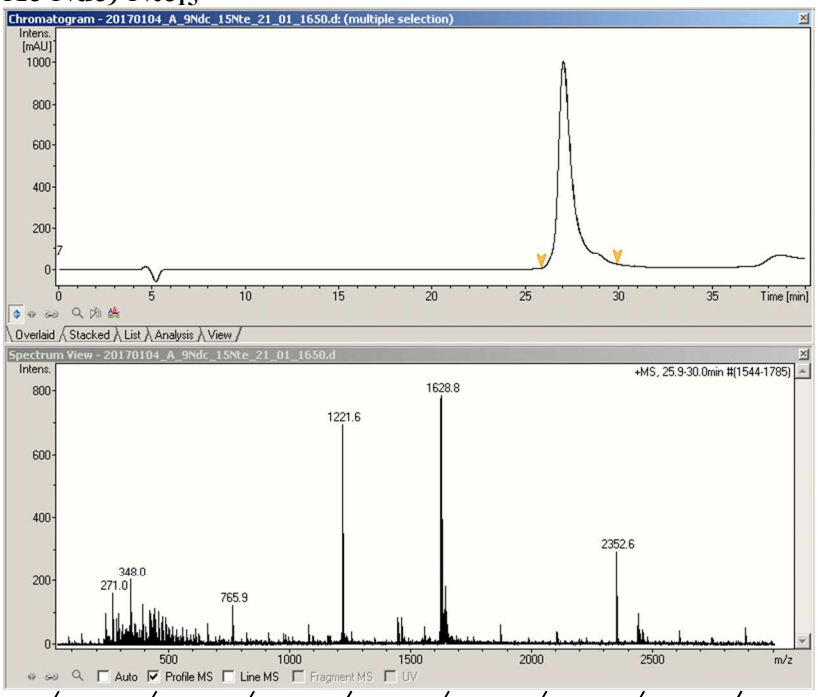


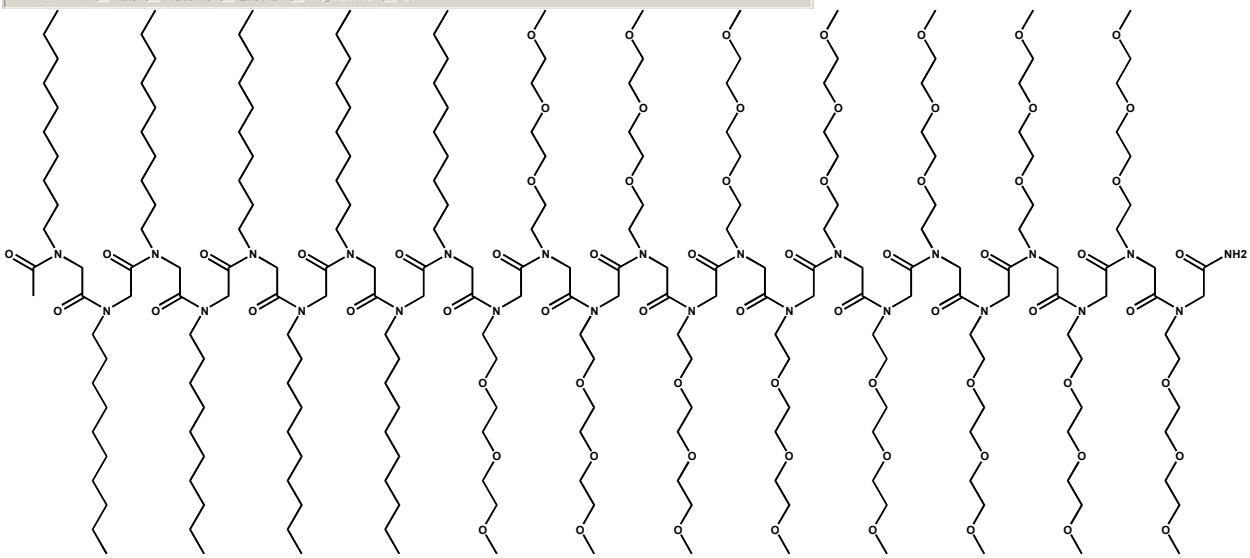


M/Z = 3663.69

Observed Mass: 1832.9 (+2 H^+); 1222.2 (+3 H^+)

Ac-Ndc$_9$-Nte$_{15}$

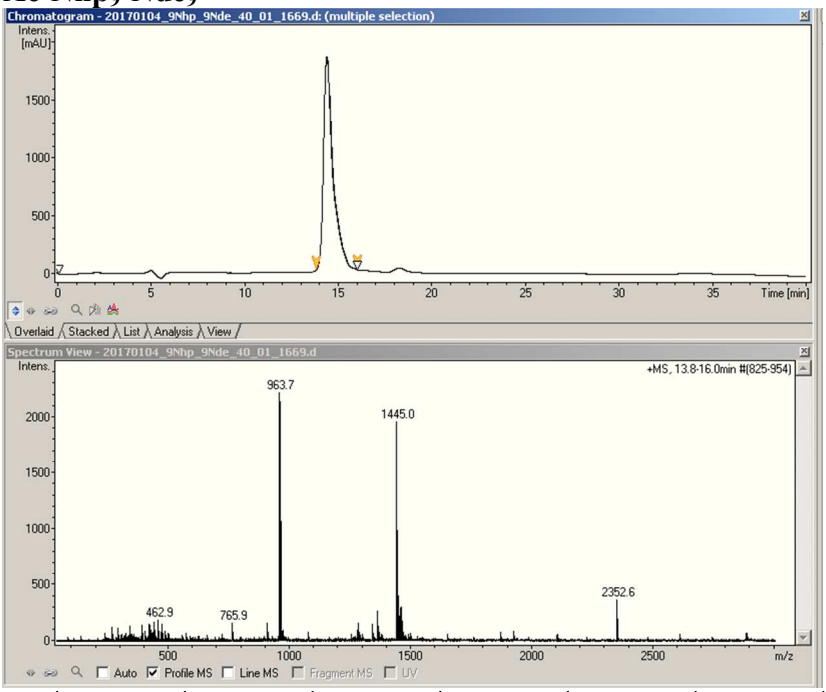


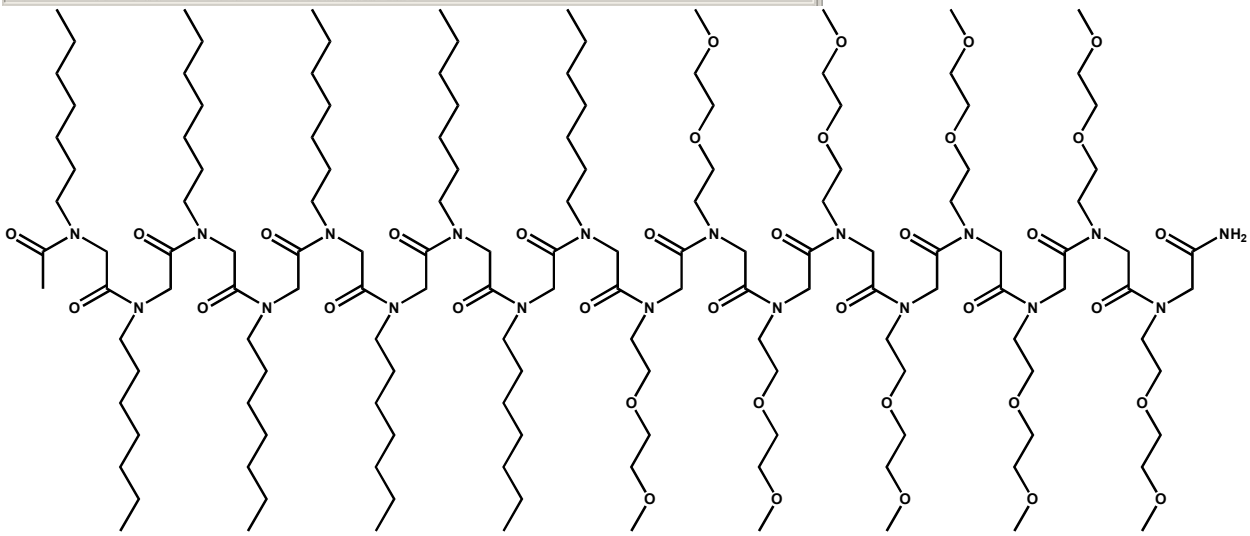


M/Z = 4882.38

Observed Mass: 1221.6 (+4 H^+); 1628.8 (+3 H^+); 2352.6 (+2 H^+)

Ac-Nhp$_9$-Nde$_9$





M/Z = 2888.03

Observed Mass: 1445.0 (+2 H^+); 963.7 (+3 H^+);

References

(1) Tran, H.; Gael, S. L.; Connolly, M. D.; Zuckermann, R. N. *J. Vis. Exp.* **2011**, No. 57.

(2) Kim, S.; Biswas, G.; Park, S.; Kim, A.; Park, H.; Park, E.; Kim, J.; Kwon, Y.-U. *Org. Biomol. Chem.* **2014**, *12*, 5222.

(3) Hexemer, A.; Bras, W.; Glossinger, J.; Schaible, E.; Gann, E.; Kirian, R.; MacDowell, A.; Church, M.; Rude, B.; Padmore, H. *J. Phys. Conf. Ser.* **2010**, *247*, 012007.

(4) Sun, J.; Jiang, X.; Lund, R.; Downing, K. H.; Balsara, N. P.; Zuckermann, R. N. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 3954.

(5) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.

(6) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.

(7) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.

(8) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(9) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.