

Synthesis and Structural Studies of Homo-Oligomers of Geminally Disubstituted $\beta^{2,2}$ -Amino Acids with Carbohydrate Side Chain

Gangavaram V. M. Sharma,*^a Post Sai Reddy,^{a,c} Deepak Chatterjee,^{b,c}

Ajit C. Kunwar*^b

- a. Organic Chemistry Division III, Indian Institute of Chemical
Technology (CSIR), Hyderabad 500 607, India
- b. Centre for Nuclear Magnetic Resonance, Indian Institute of
Chemical Technology (CSIR), Hyderabad 500 607, India
- c. These authors contributed equally to this work

Email: esmvee@iict.res.in; kunwar@iict.res.in

Contents

Conformational analysis	S2-3
Solvent titration plots	S4
NMR spectra	S5-S43
Molecular Dynamics	S44-S47
CD Spectra	S48

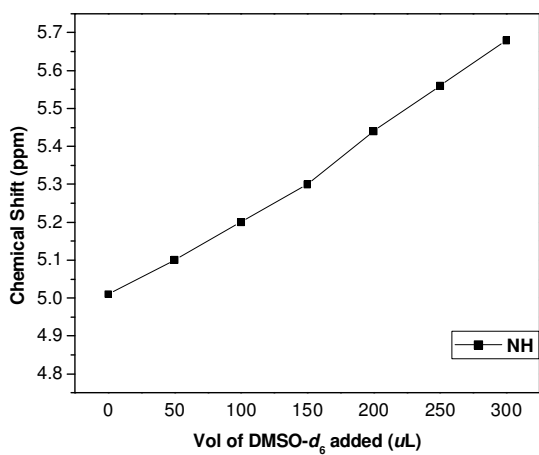
Conformational analysis:

Solvent titration studies, when upto 33% (v/v) of DMSO- d_6 was added to the $CDCl_3$ solution, showed that for the monomer (**1**), the change in the amide proton chemical shift ($\Delta\delta_{NH}$) is 0.67ppm. This does not provide clear cut picture about its involvement in H-bonding. From the studies on the dimer (**2**), on the other hand, a clearer picture emerges. The value of $\delta_{NH(2)} > 7.07$ ppm supports its likely participation in H-bonding. Solvent titration studies (Figure S1), where both the amide protons showed $\Delta\delta_{NH} < 0.20$ ppm, emphatically support the participation of both theamide protons in H-bonding. The couplings, $^3J_{NH-C\beta H}$, are quite distinctive, with one large value (> 8.0 Hz) and another rather small (< 2 Hz), thus providing strong evidence for an anti-periplanar arrangement of NH and $C\beta H$, corresponding to a value of $\sim \pm 120^\circ$ for the torsion angle ϕ ($C(O)-N-C\beta-C\alpha$). Because of the absence of protons at the $C\beta$, the information on torsion angle θ ($N-C\beta-C\alpha-CO$) had to be obtained through the nOe correlations only. Intra-residue nOe cross peaks, NH/ $C3H$ and $C3H/C\beta H_{(pro-R)}$ for both the residues, provided adequate support for $\theta \sim -60^\circ$. Additionally, NH(1)/ $C3H(1)$ and NH(2)/ $C1H(1)$ enabled us to deduce that

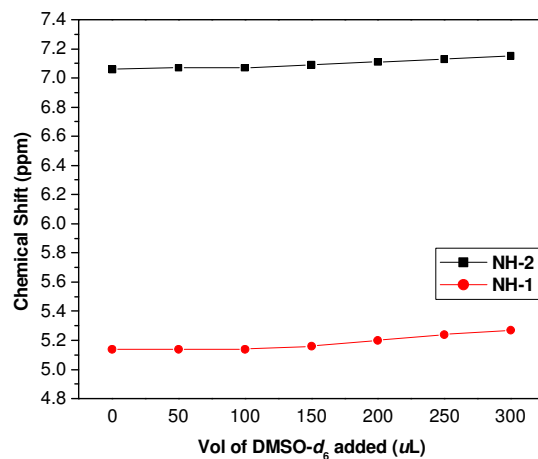
ψ ($C\beta$ - $C\alpha$ -CO-N) is constrained to $\sim \pm 180^\circ$. These observations indicating presence of a rigid backbone in **2**, were the basis to extend the studies to the larger oligomers **3** and **4**.

Resonance Assignments

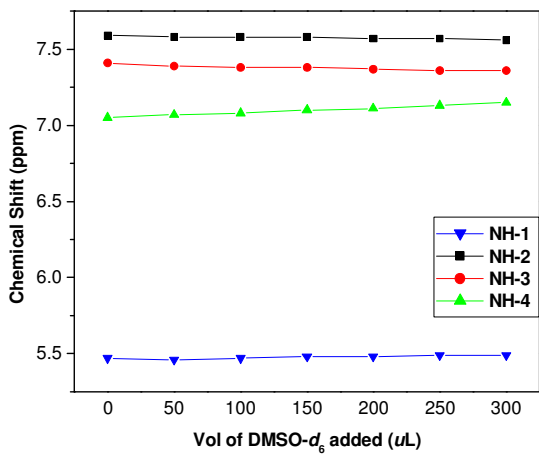
The resonance assignment in oligomers with one type of residue is difficult due to the likely overlap. However in the present scenario, we were able to achieve the task without much difficulty. The assignments were carried out with the help of TOCSY and HSQC/HMBC experiments. It is straight forward to assign Boc-NH, as it resonates at higher field $\sim 5 - 6$ ppm. Using Boc-NH to begin with, TOCSY experiment was used to assign protons at $C\beta$ (we have stereo-specific assignments for both these protons and the rational has been included in the text already). HSQC experiment in turn assigns the $C\beta$, which shows a cross peak in the HMBC experiment with $C3H$ in the sugar side chain. $C3H$ shows a weak correlation with the $C2H$ in the TOCSY spectrum (due to very small, not measurable $^3J_{C2H-C3H}$). The $C1H$ proton was further assigned from the TOCSY correlation with the $C2H$ proton. This thus completes the assignments within a residue, the first residue in this case. The HSQC experiment provides complete ^{13}C assignments (excluding the carbonyl carbon). Intra residue carbonyl carbon was assigned from their HMBC correlation with $C\beta$ Hs. Sequential resonance assignments were achieved utilizing the correlation between $C(O)$ of the preceding residue with HN of the following residue. Thus having assigned the amide proton of the second residue, we can follow the assignment right through the peptide chain.



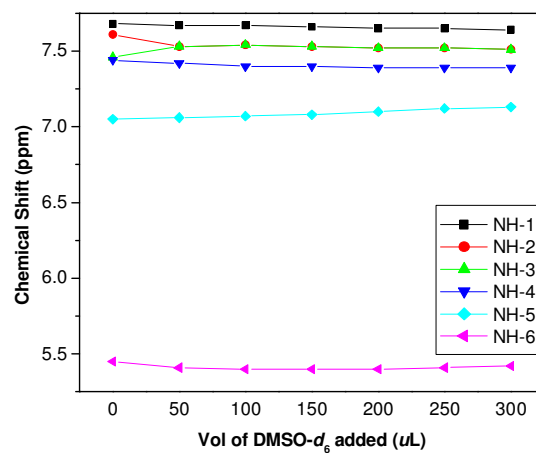
1



Peptide 2



Peptide 3



Peptide 4

Figure S1. Solvent Titration Plots for **1-4** (X-axis shows the amount of DMSO added to 600 μ L of CDCl_3 solution).

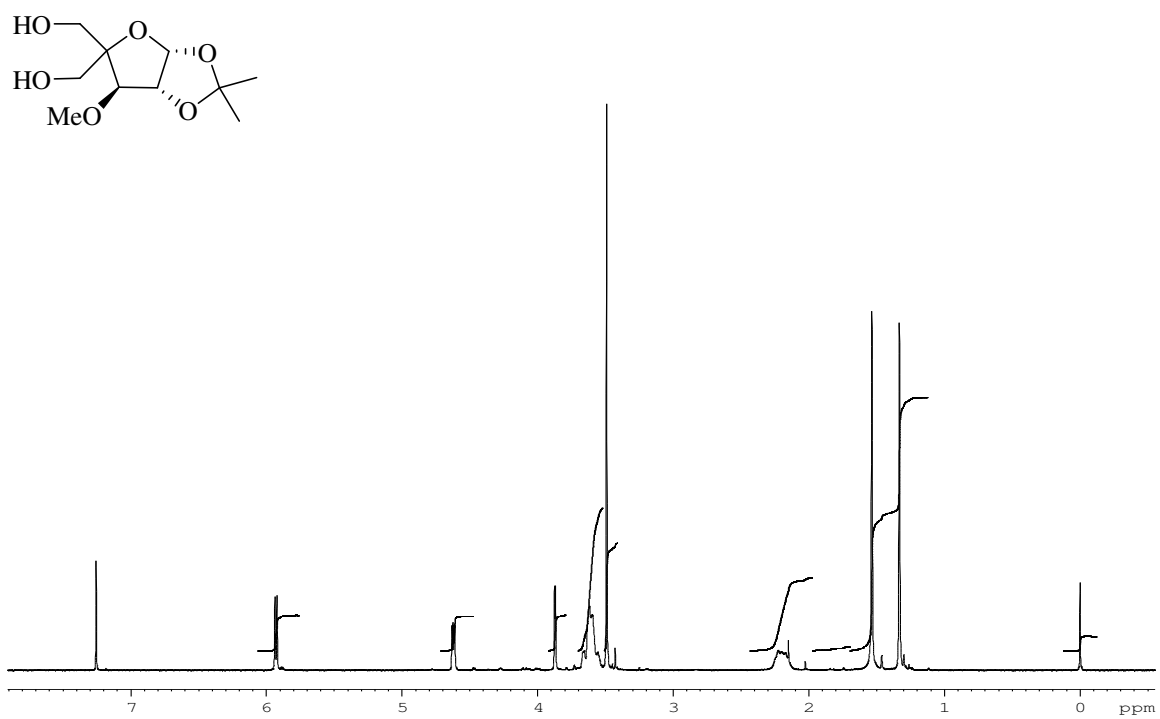


Figure S2. ^1H NMR Spectrum of **6** (300 MHz, CDCl_3 , 295K).

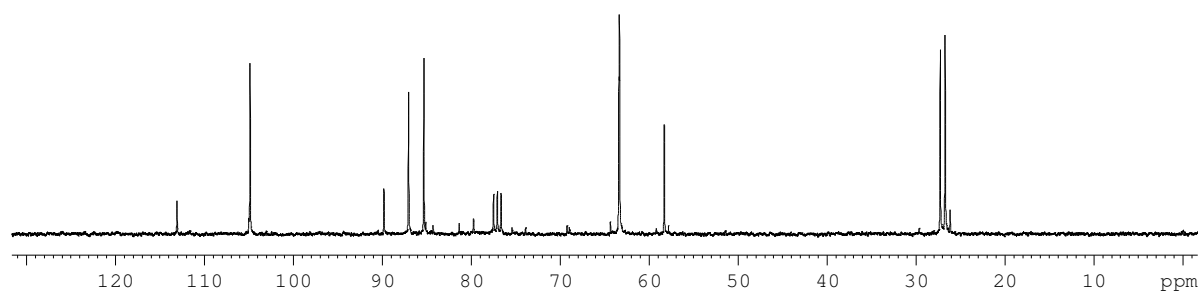
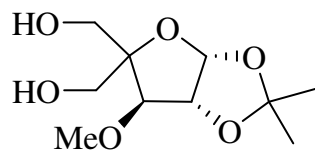


Figure S3. ^{13}C NMR Spectrum of **6** (75 MHz, CDCl_3 , 295K).

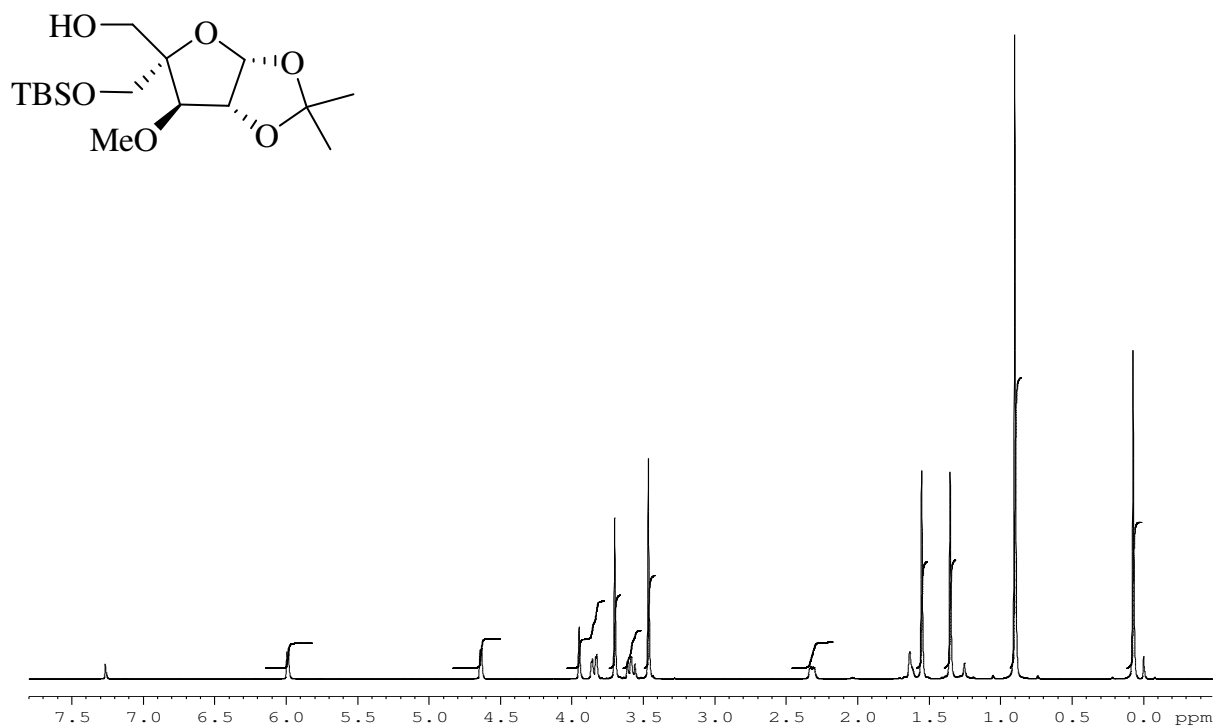


Figure S4. ^1H NMR Spectrum of **7** (400 MHz, CDCl_3 , 295K).

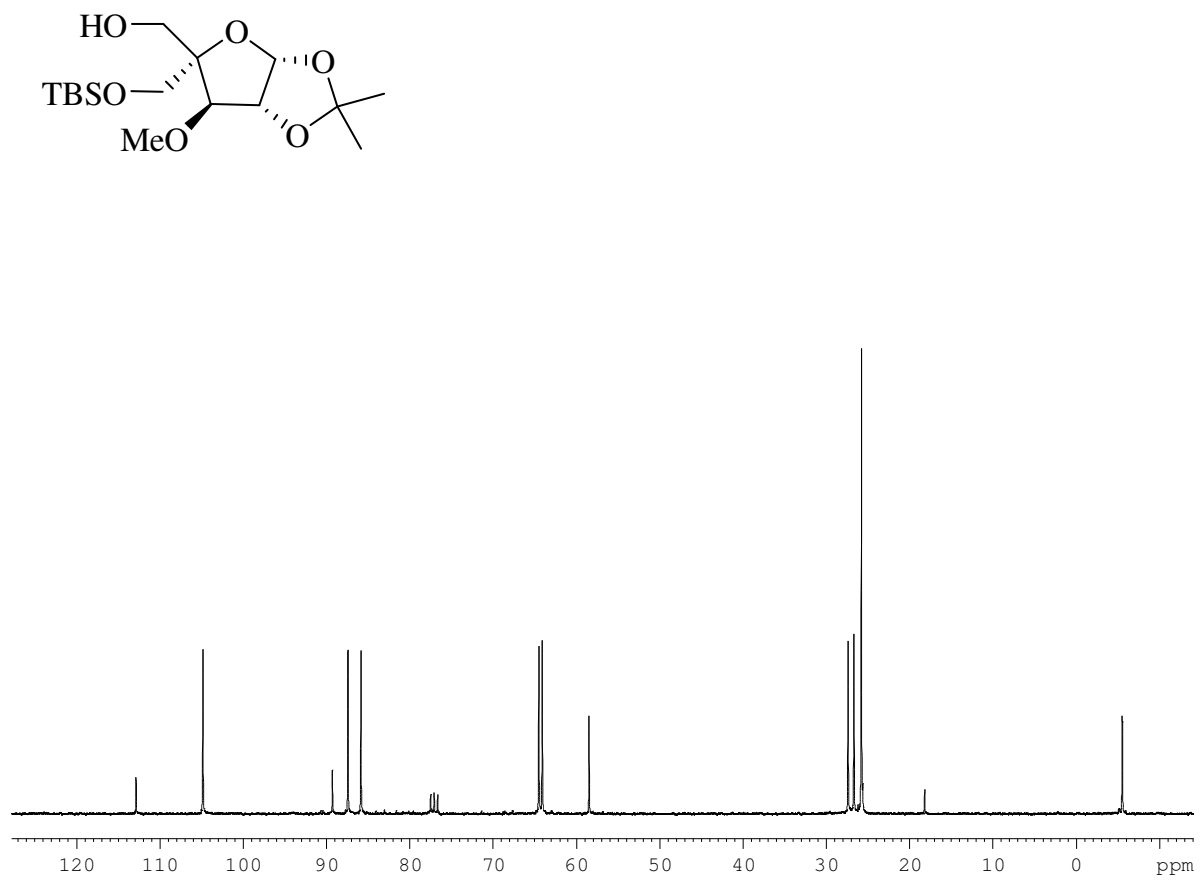


Figure S5. ^{13}C NMR Spectrum of **7** (75 MHz, CDCl_3 , 295K)

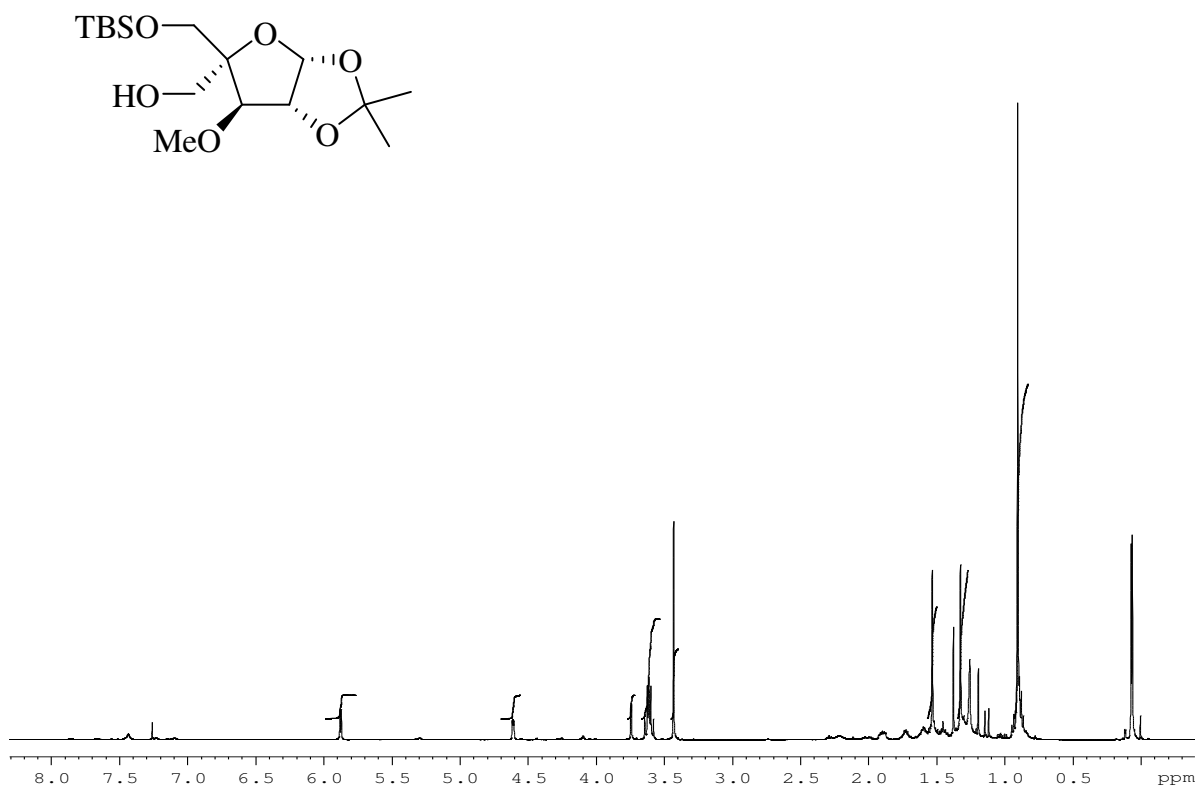


Figure S6. ^1H NMR Spectrum of **8** (300 MHz, CDCl_3 , 295K).

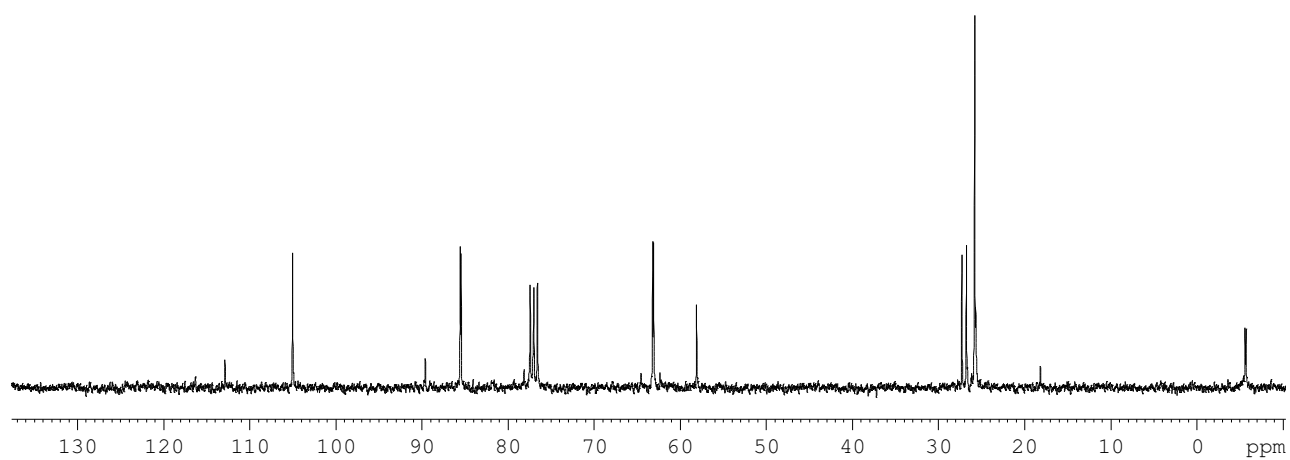
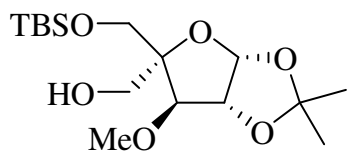


Figure S7. ^{13}C NMR Spectrum of **8** (75 MHz, CDCl_3 , 295K).

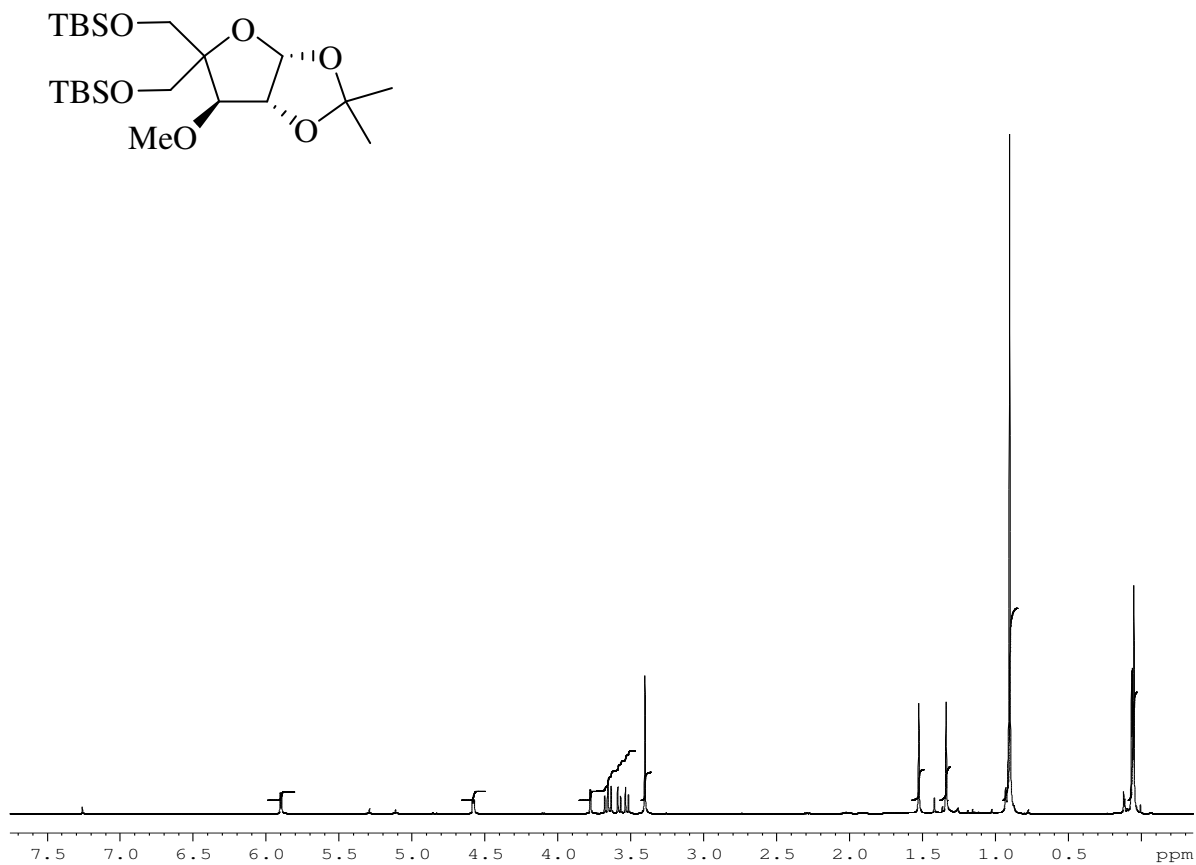


Figure S8. ^1H NMR Spectrum of **9** (500 MHz, CDCl_3 , 295K).

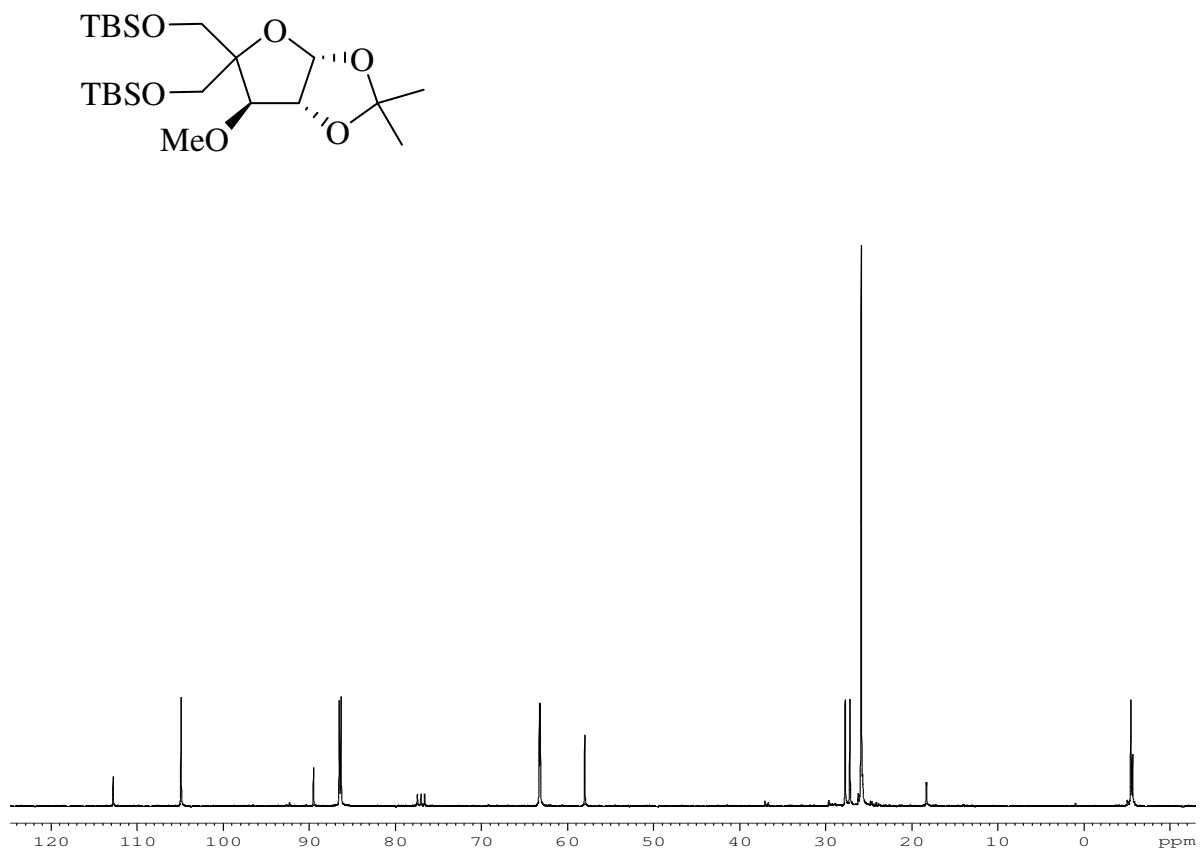


Figure S9. ^{13}C NMR Spectrum of **9** (75 MHz, CDCl_3 , 295K).

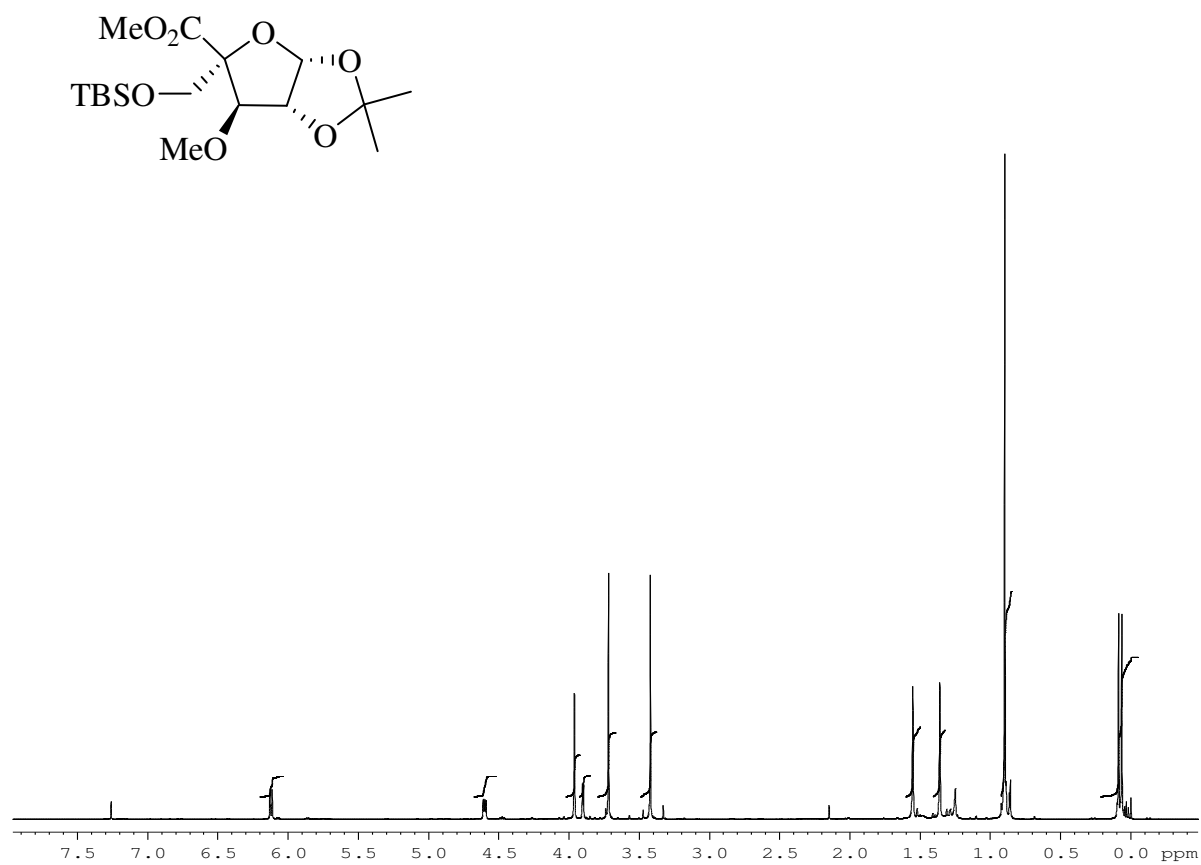


Figure S10. ^1H NMR Spectrum of **12** (500 MHz, CDCl_3 , 295K).

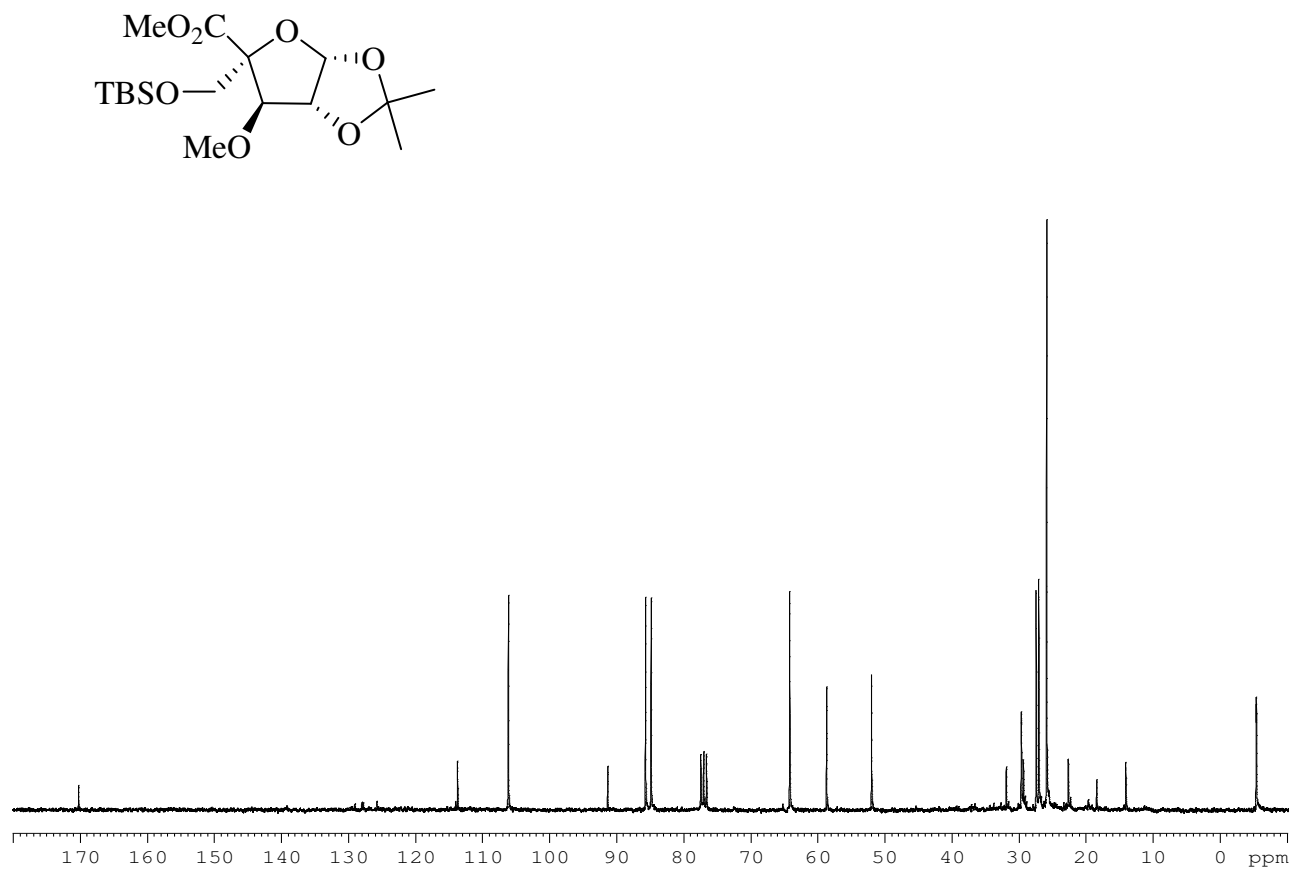


Figure S11. ^{13}C NMR Spectrum of **12** (75 MHz, CDCl_3 , 295K).

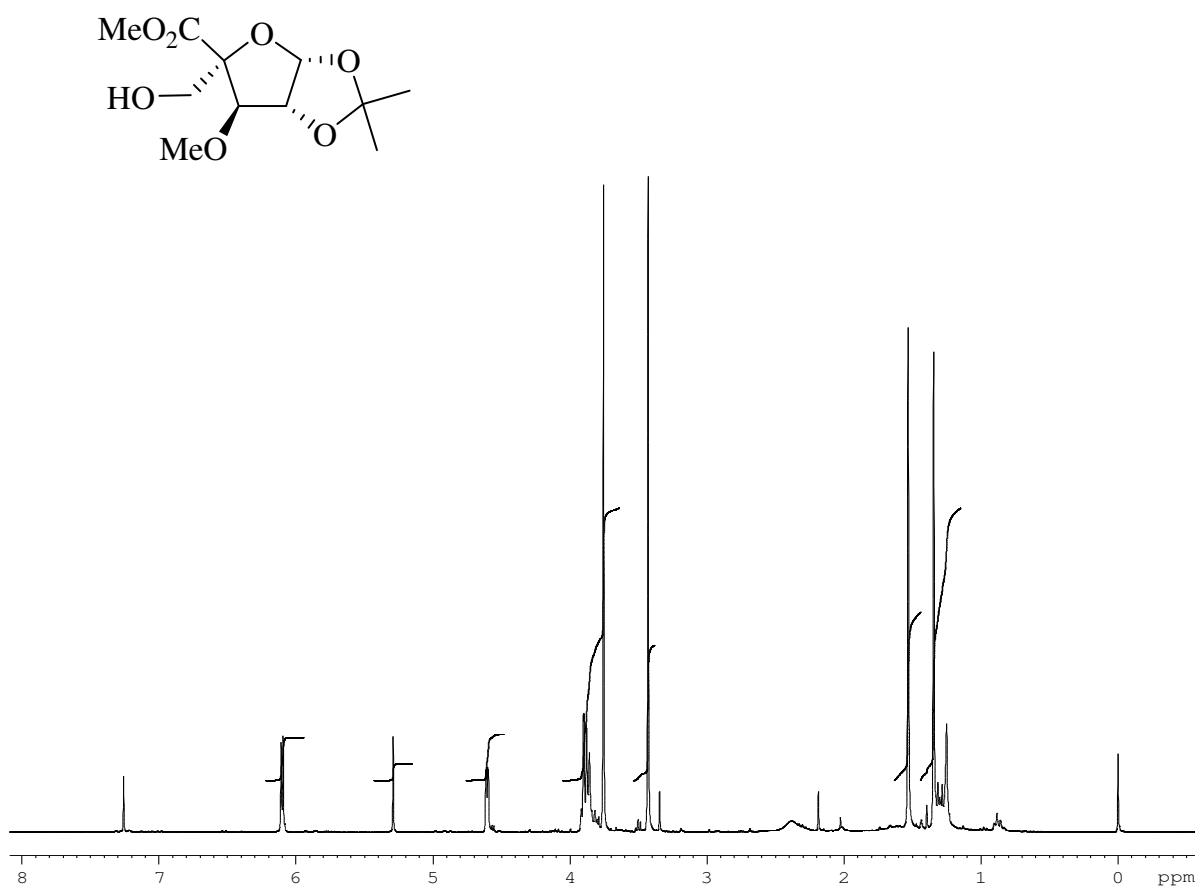


Figure S12. ^1H NMR Spectrum of **13** (500 MHz, CDCl_3 , 295K).

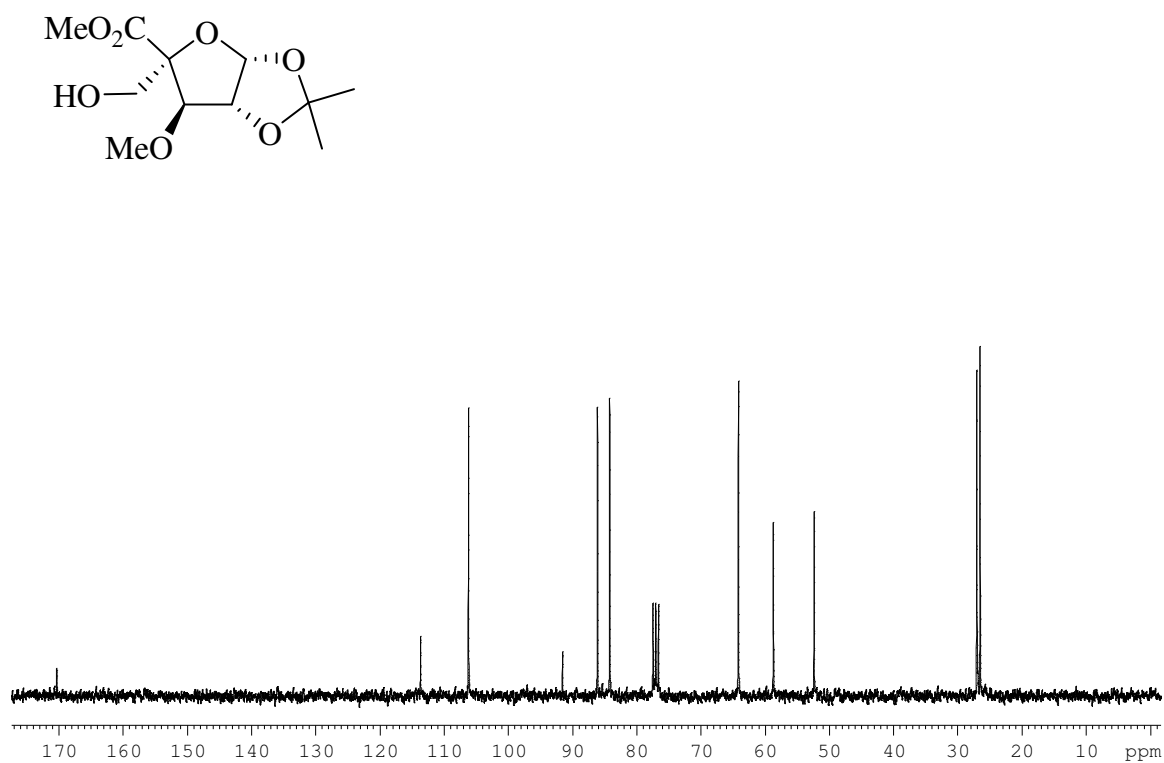


Figure S13. ^{13}C NMR Spectrum of **13** (75 MHz, CDCl_3 , 295K).

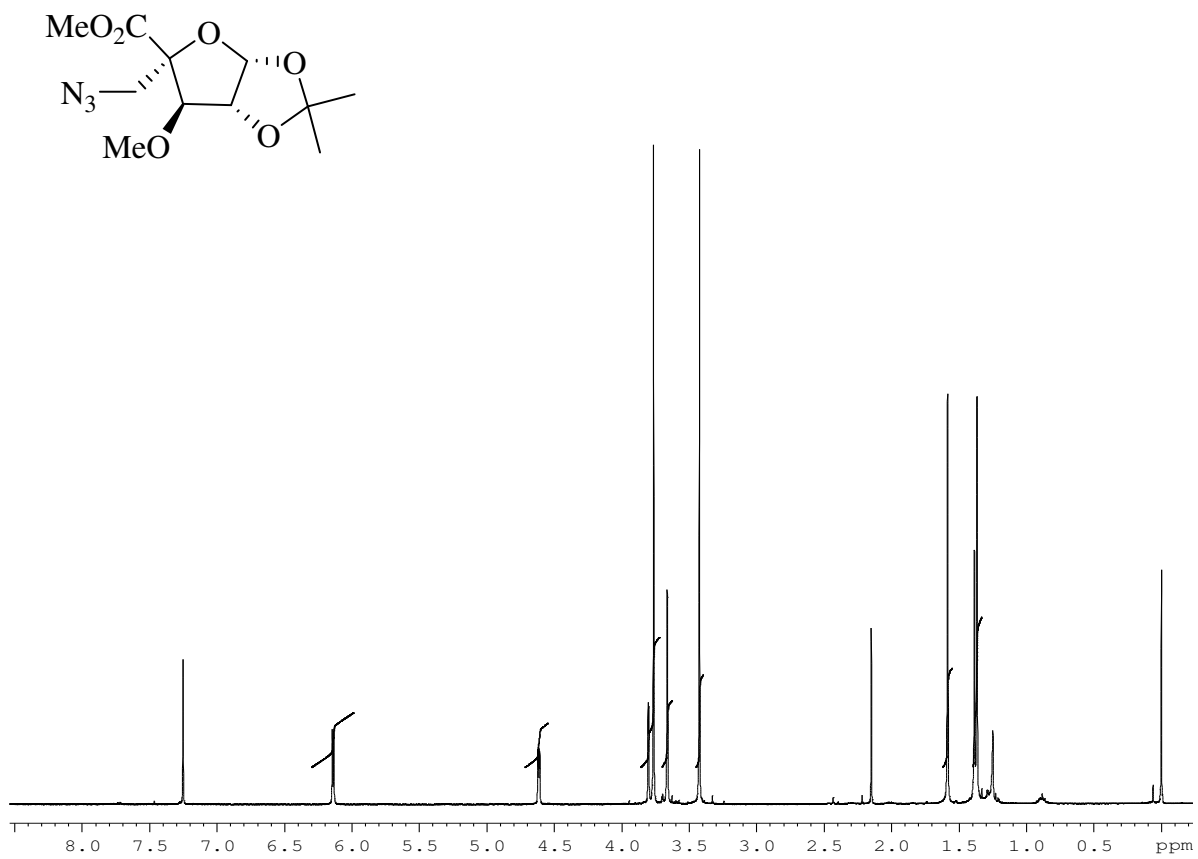


Figure S14. ^1H NMR Spectrum of **14** (400 MHz, CDCl_3 , 295K).

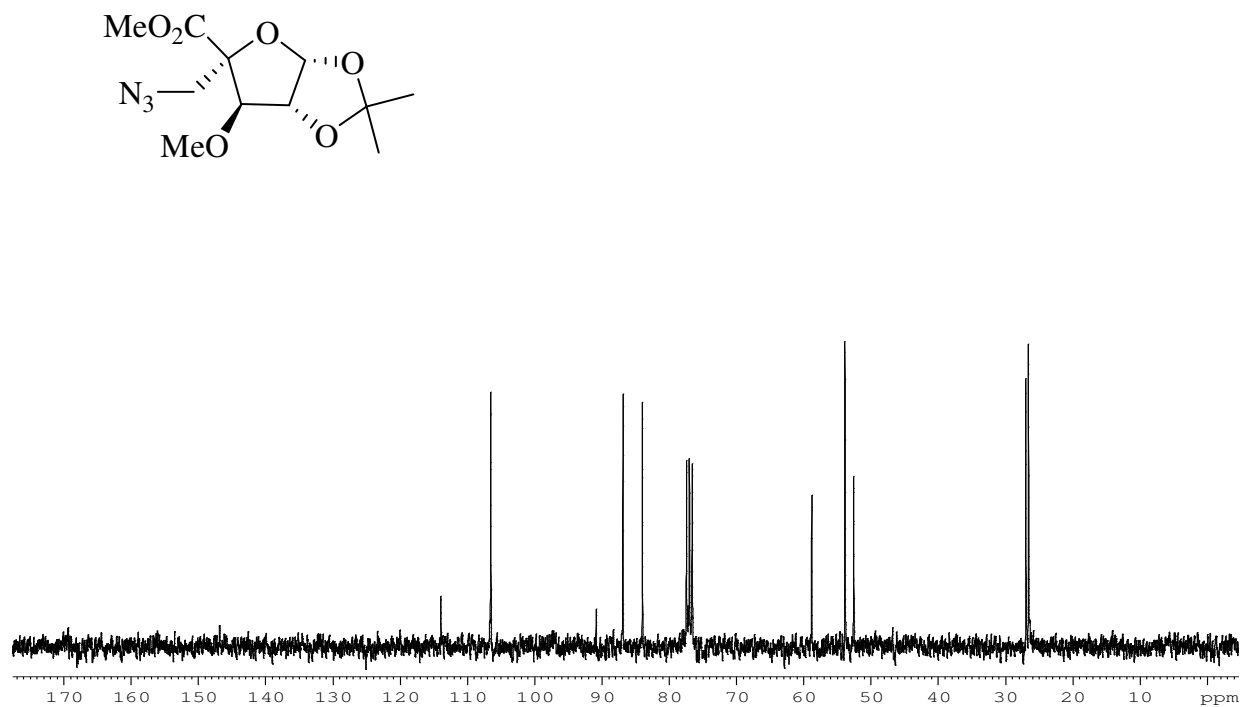


Figure S15. ^{13}C NMR Spectrum of **14** (75 MHz, CDCl_3 , 295K).

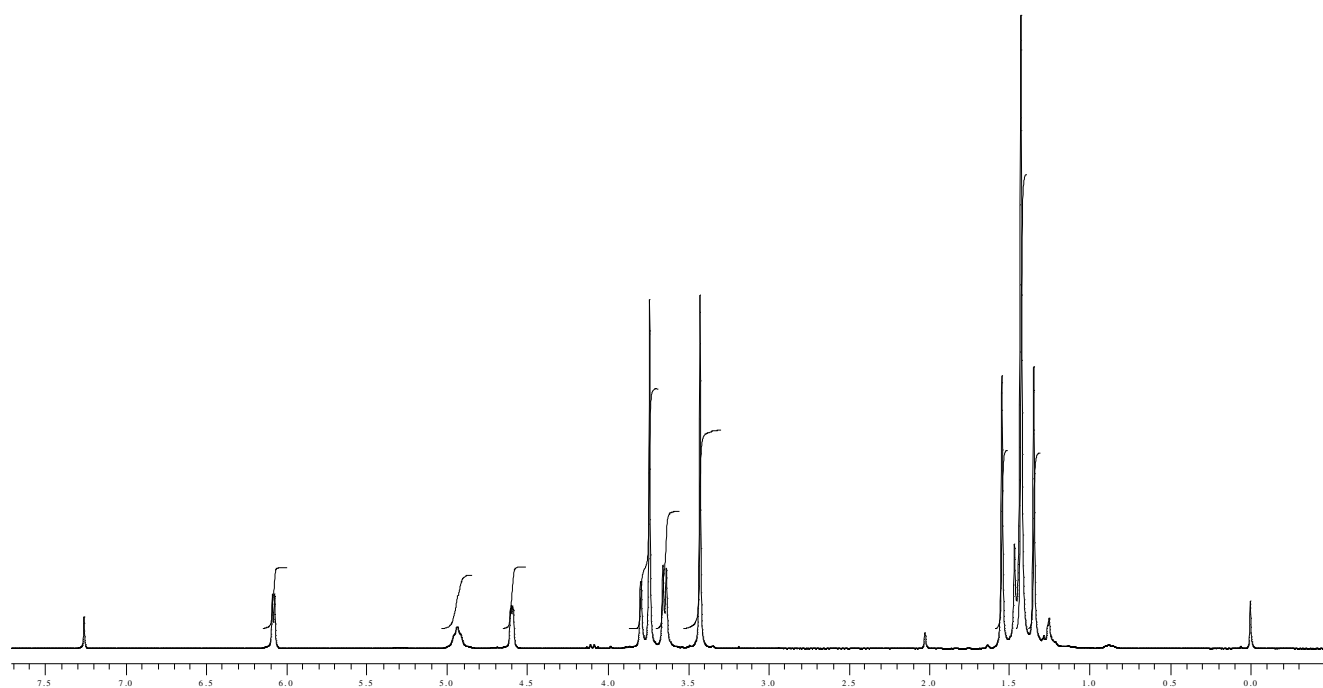
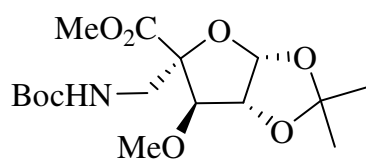


Figure S16. ^1H NMR Spectrum of **1** (300 MHz, CDCl_3 , 295K).

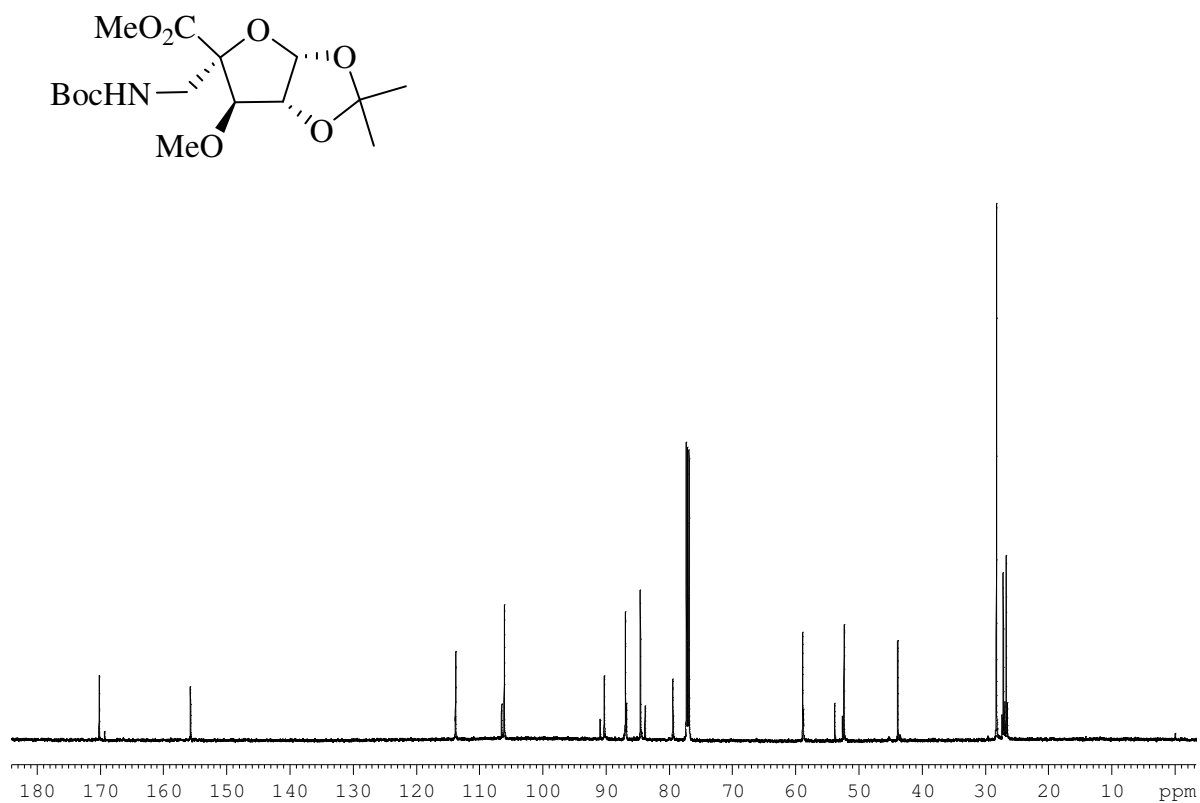


Figure S17. ^{13}C NMR Spectrum of **1** (75 MHz, CDCl_3 , 295K).

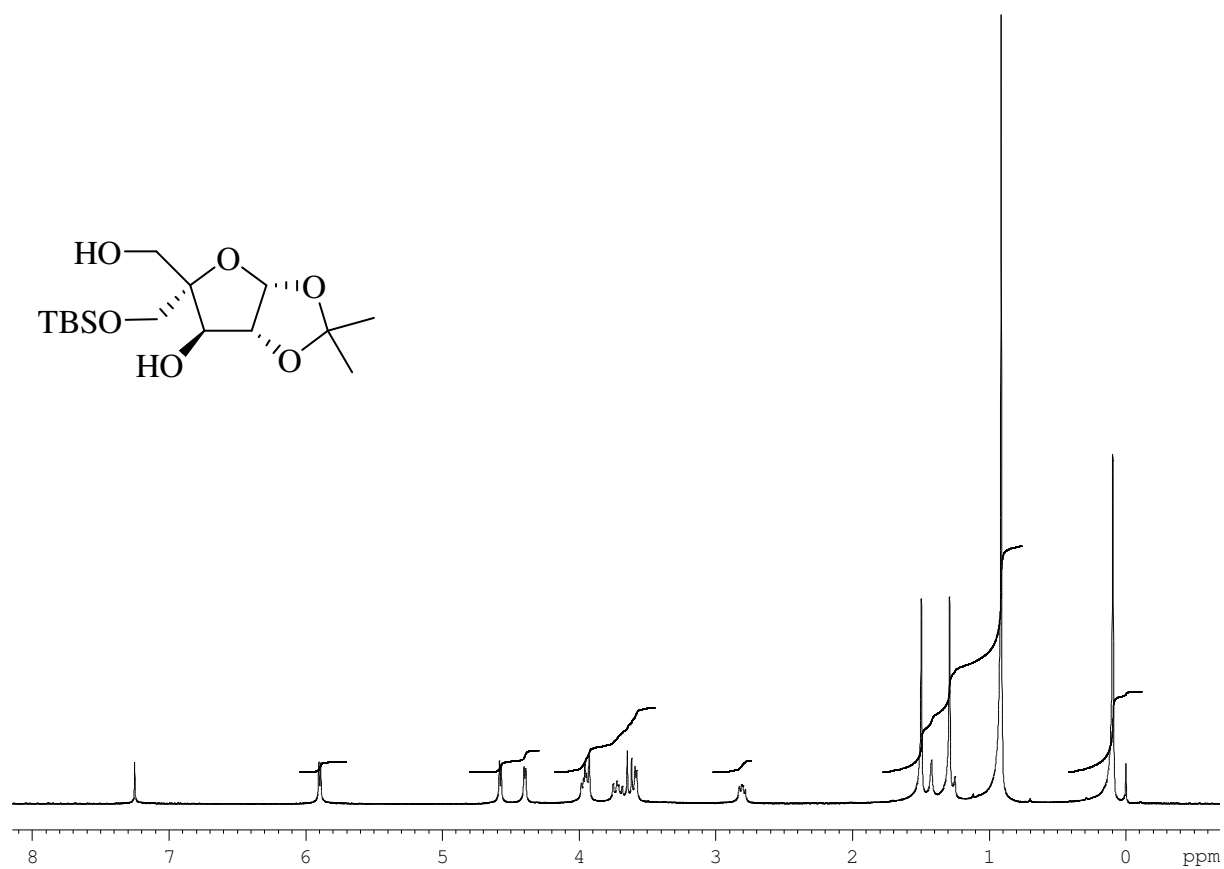


Figure S18. ^1H NMR Spectrum of **16** (300 MHz, CDCl_3 , 295K).

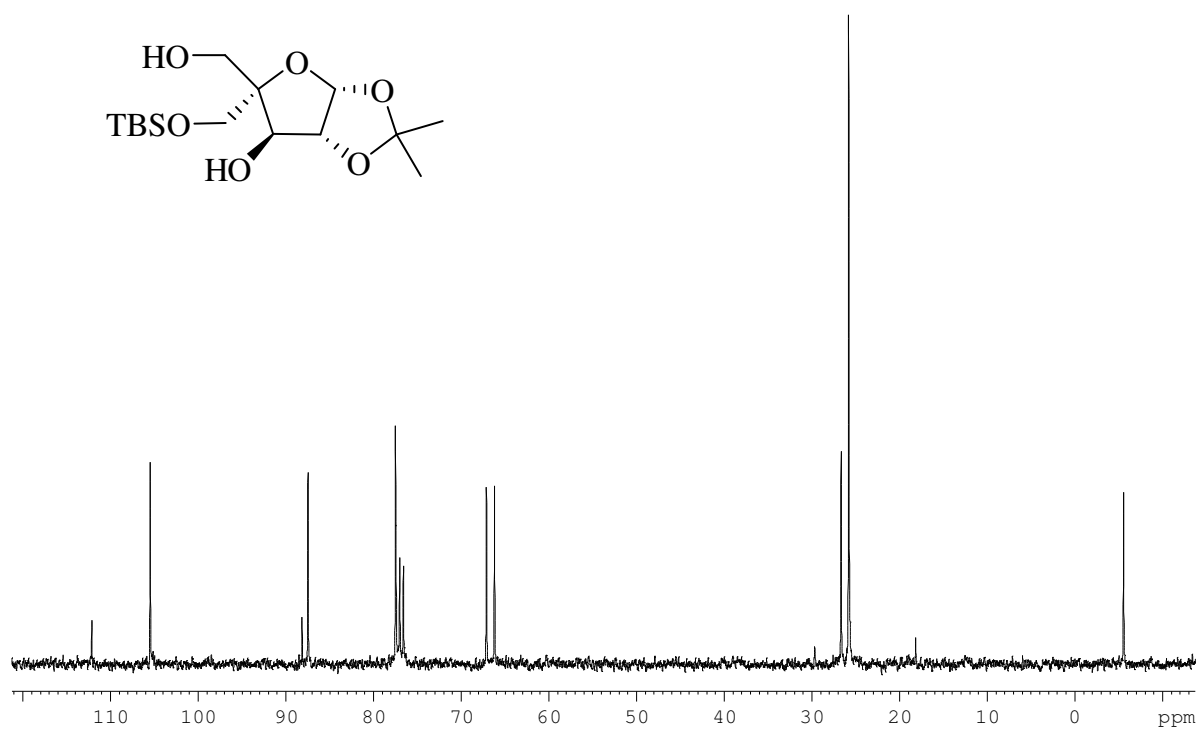


Figure S19. ^{13}C NMR Spectrum of **16** (75 MHz, CDCl_3 , 295K).

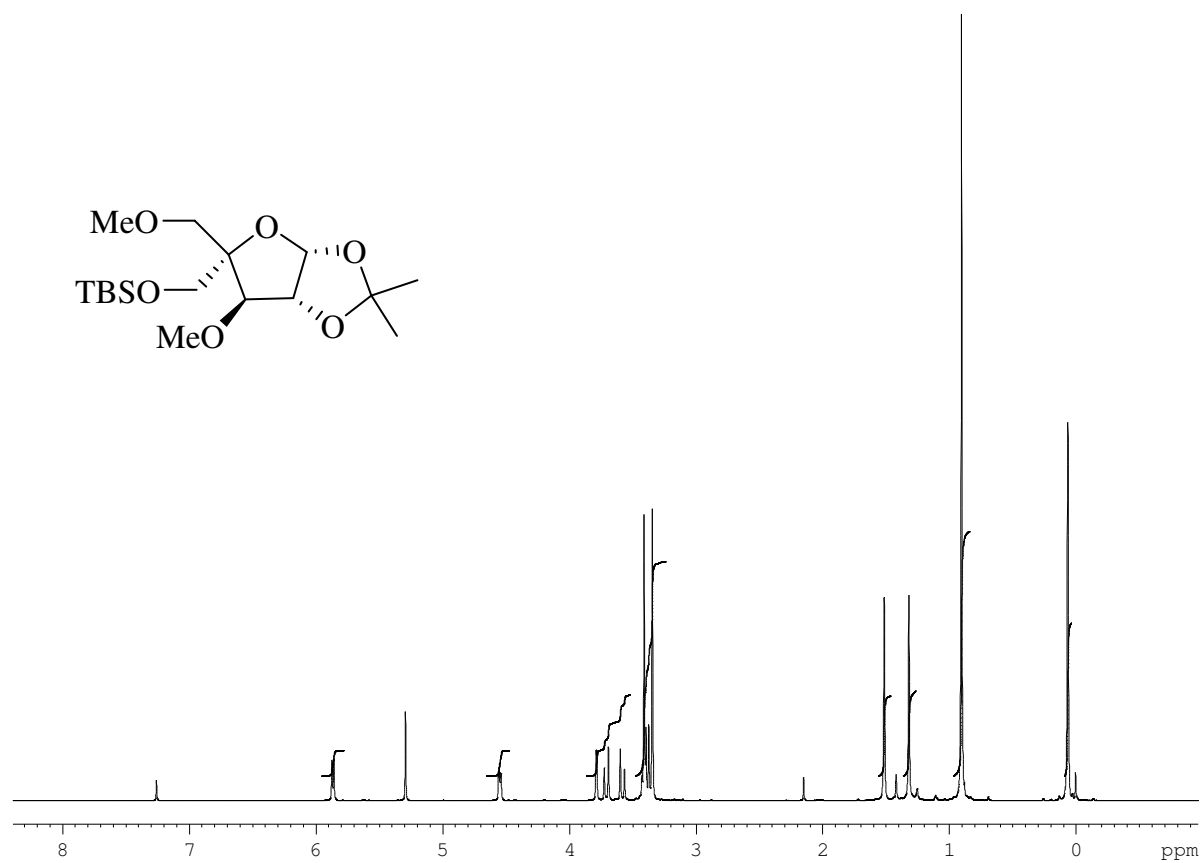


Figure S20. ^1H NMR Spectrum of **17** (300 MHz, CDCl_3 , 295K).

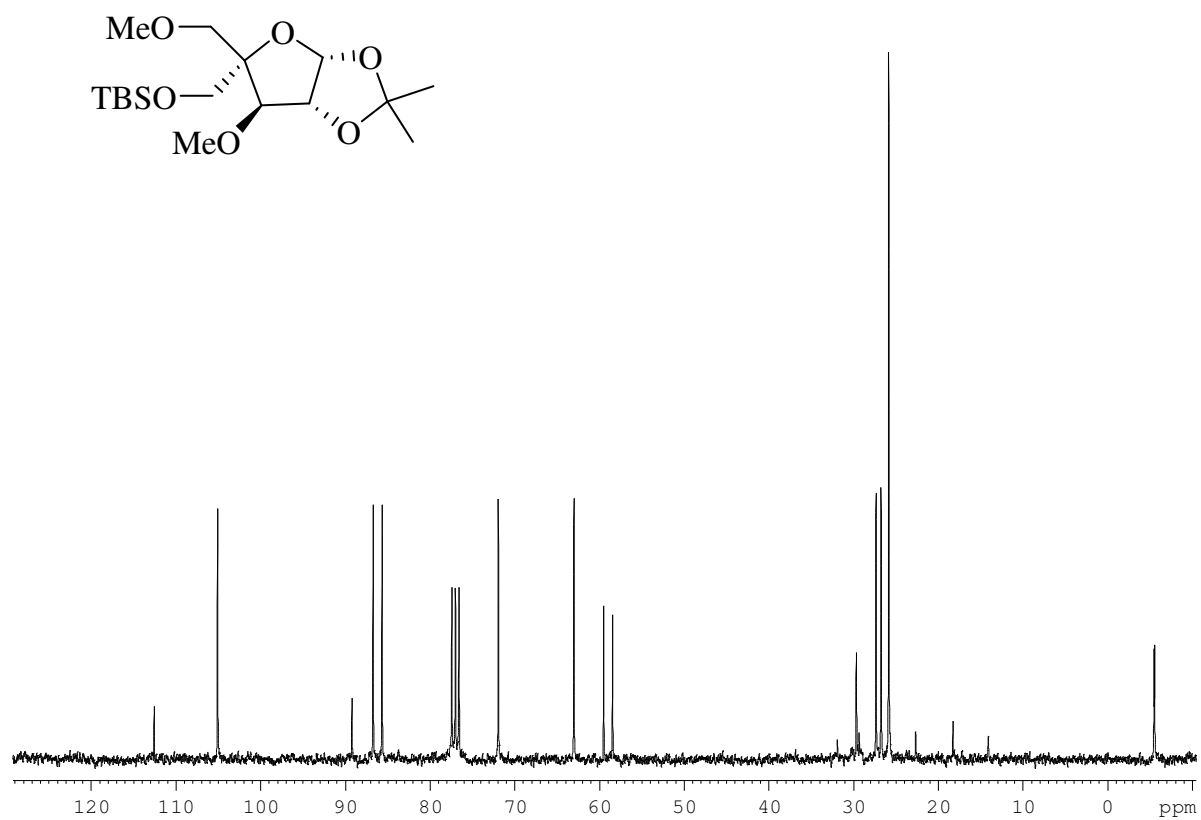


Figure S21. ^{13}C NMR Spectrum of **17** (75 MHz, CDCl_3 , 295K).

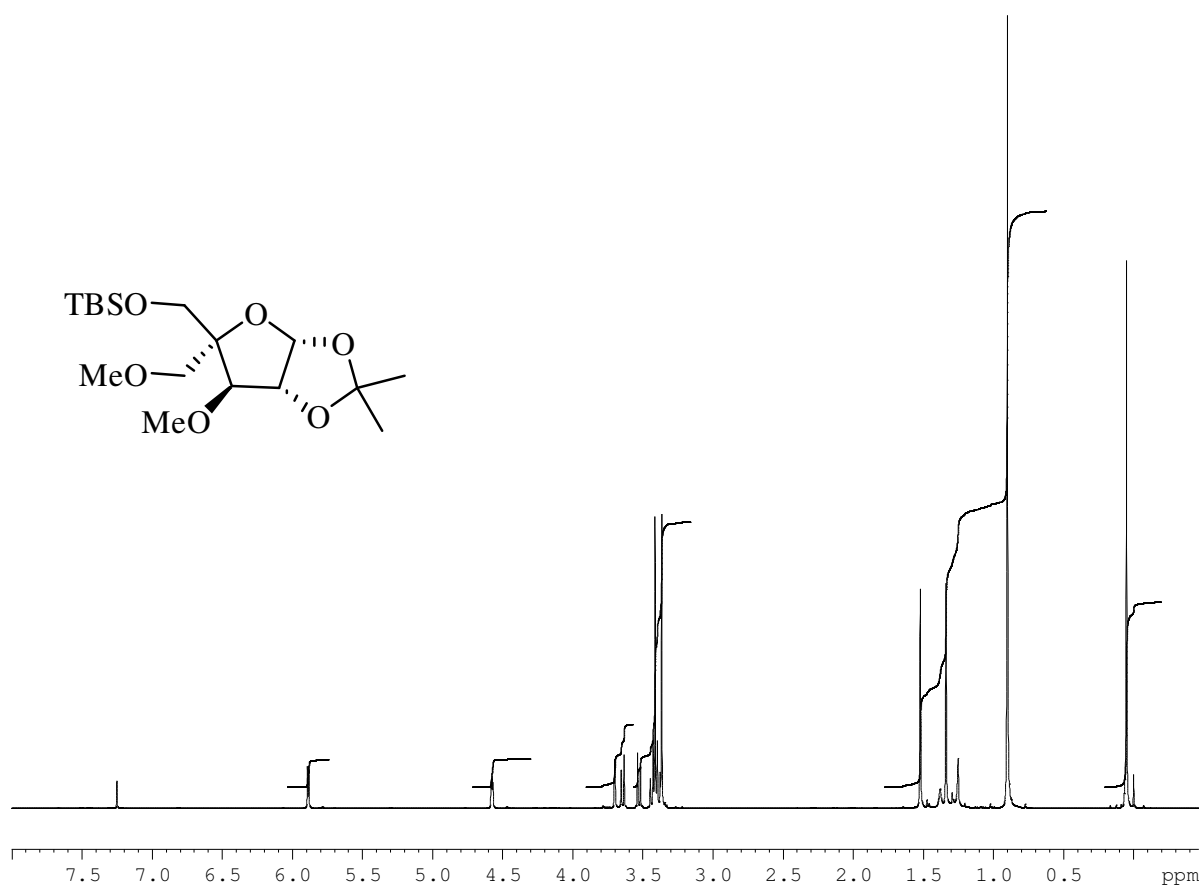


Figure S22. ^1H NMR Spectrum of **18** (300 MHz, CDCl_3 , 295K).

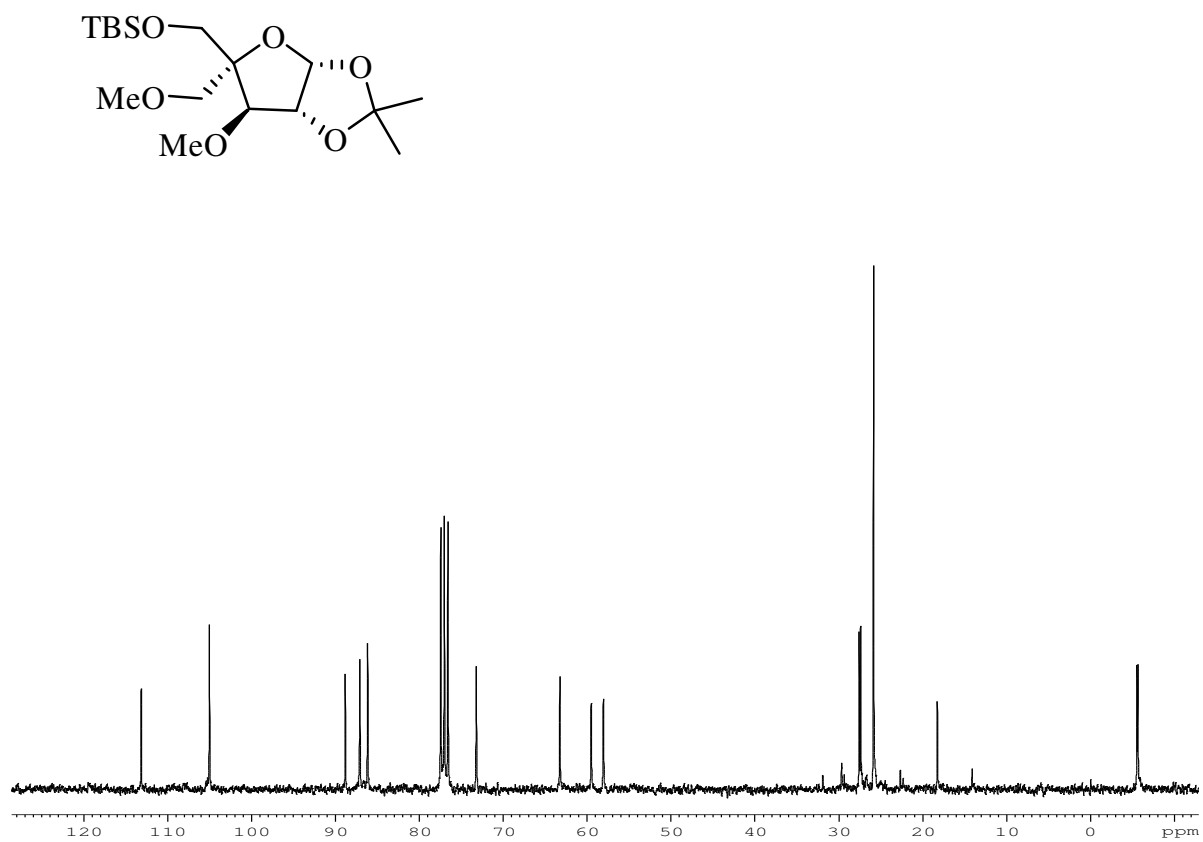


Figure S23. ^{13}C NMR Spectrum of **18** (75 MHz, CDCl_3 , 295 K).

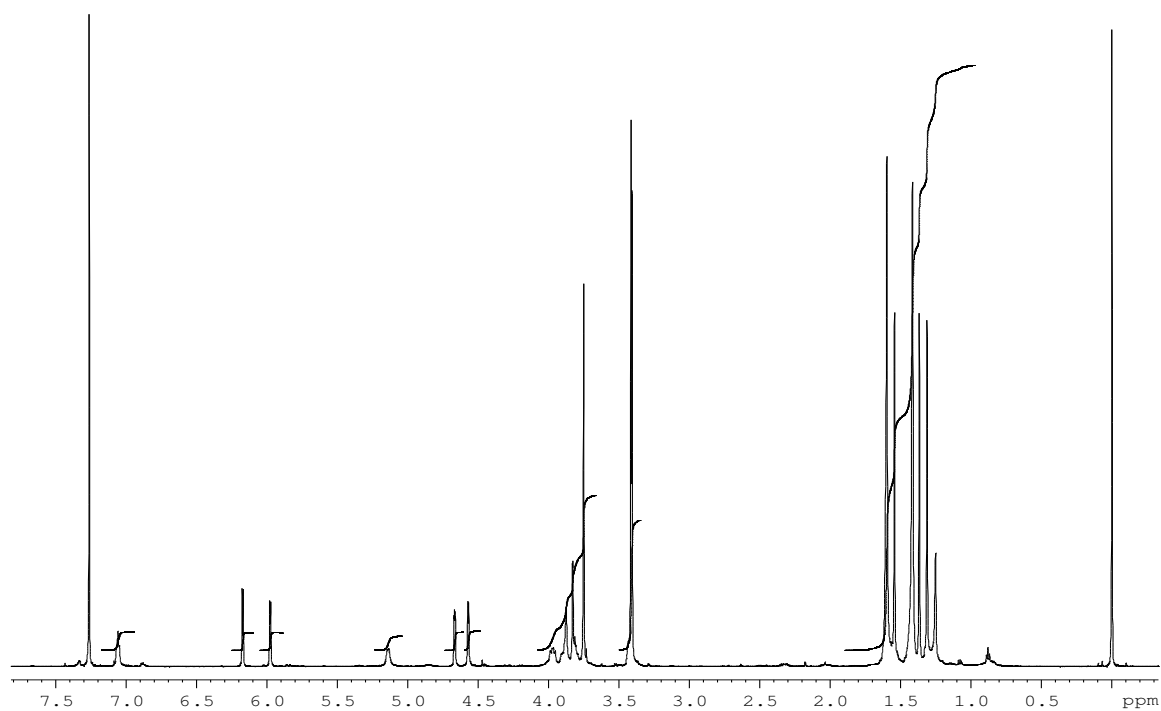
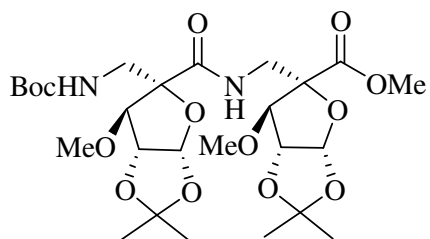


Figure S24. ^1H NMR Spectrum of peptide **2** (600 MHz, CDCl_3 , 288K).

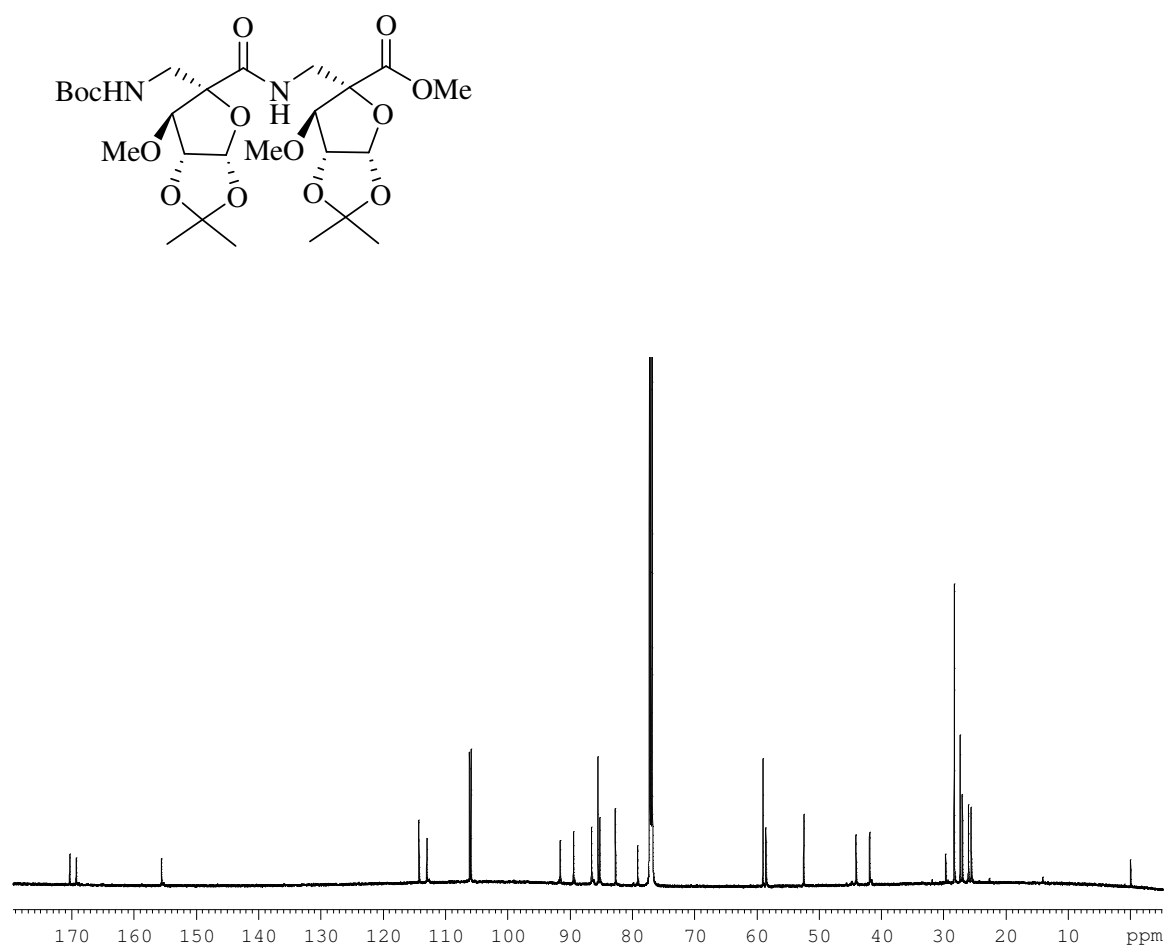


Figure S25. ^{13}C NMR Spectrum of peptide **2** (150 MHz, CDCl_3 , 288K).

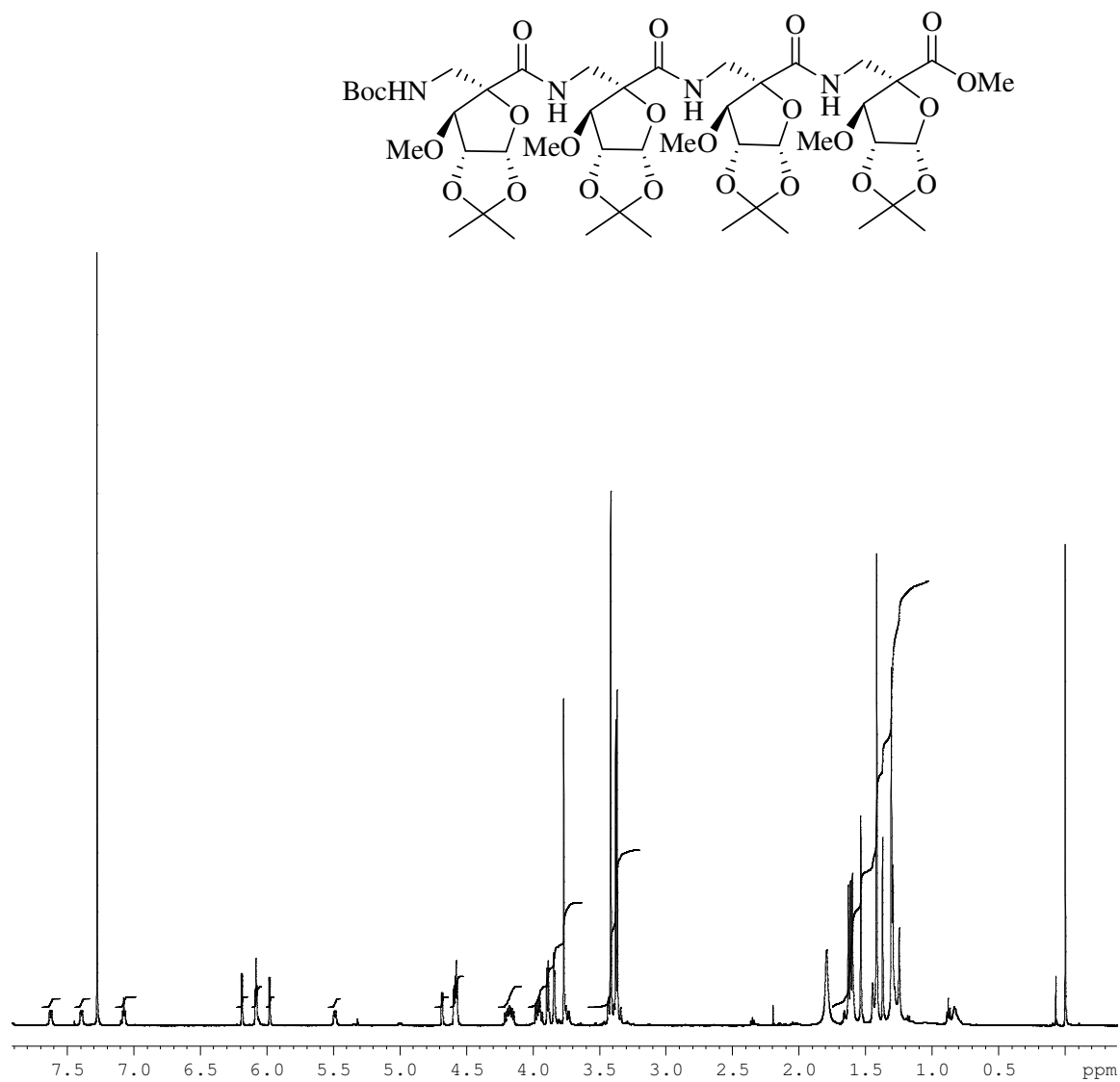


Figure S26. ¹H NMR Spectrum of peptide **3** (600 MHz, CDCl₃, 278K).

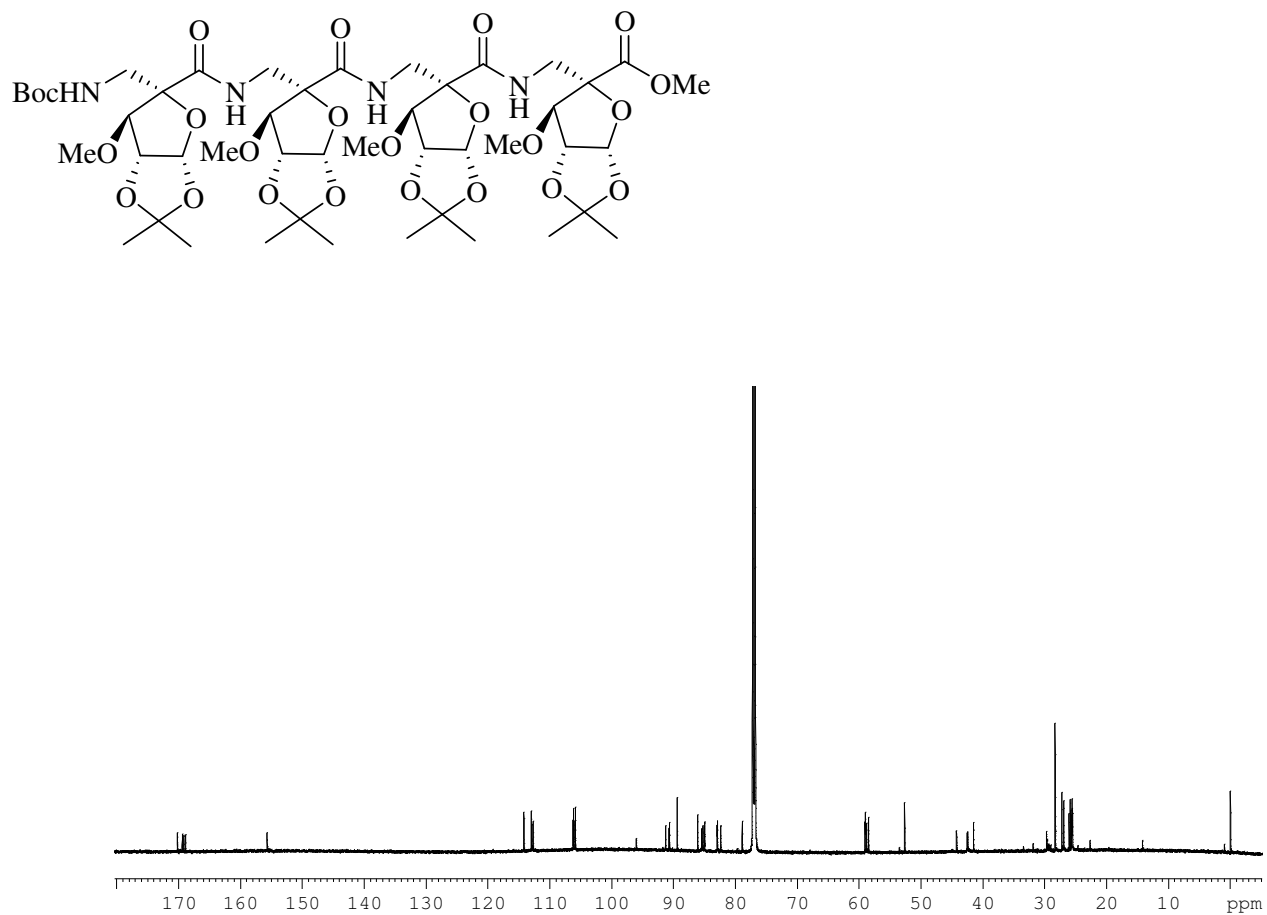


Figure S27. ^{13}C NMR Spectrum of peptide 3 (150 MHz, CDCl_3 , 278K).

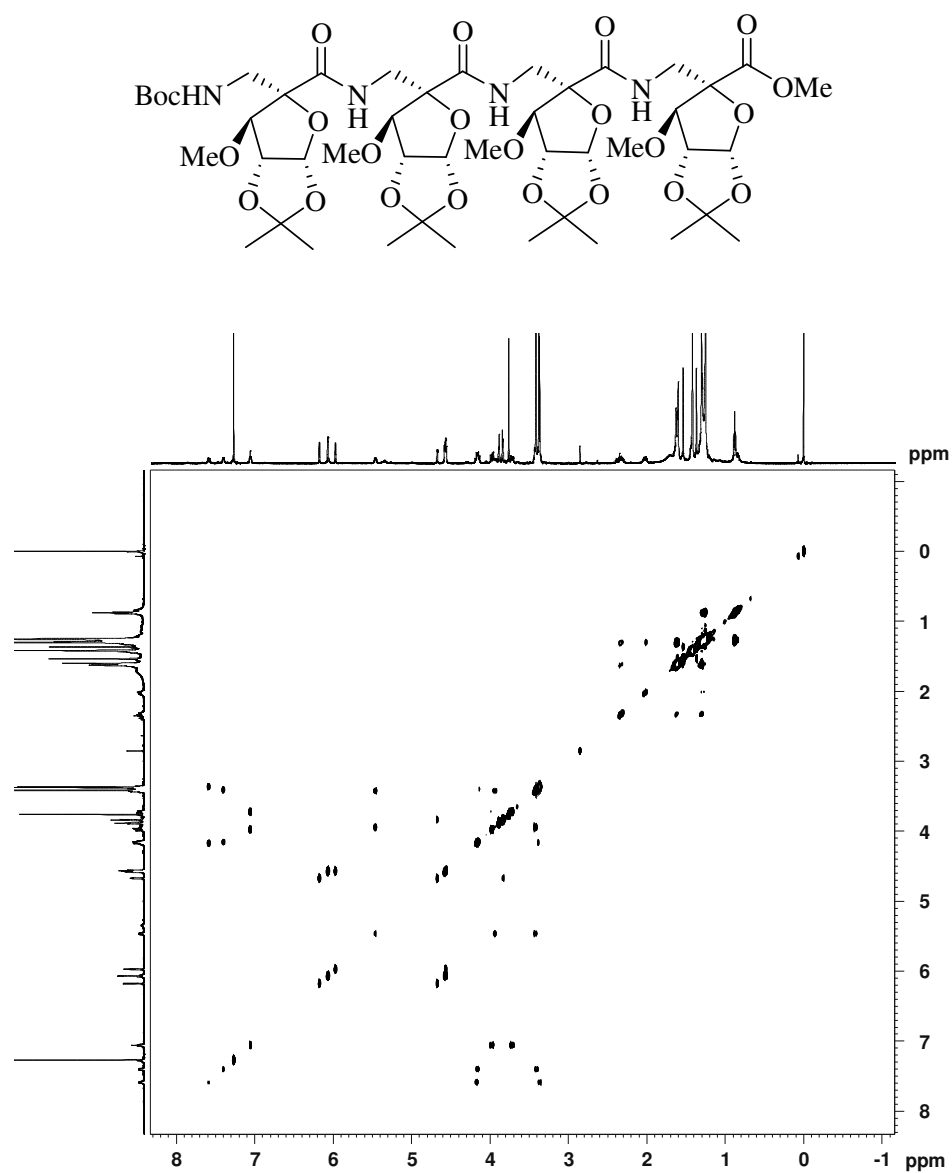


Figure S28. TOCSY Spectrum of peptide **3** (600 MHz, CDCl₃, 278K).

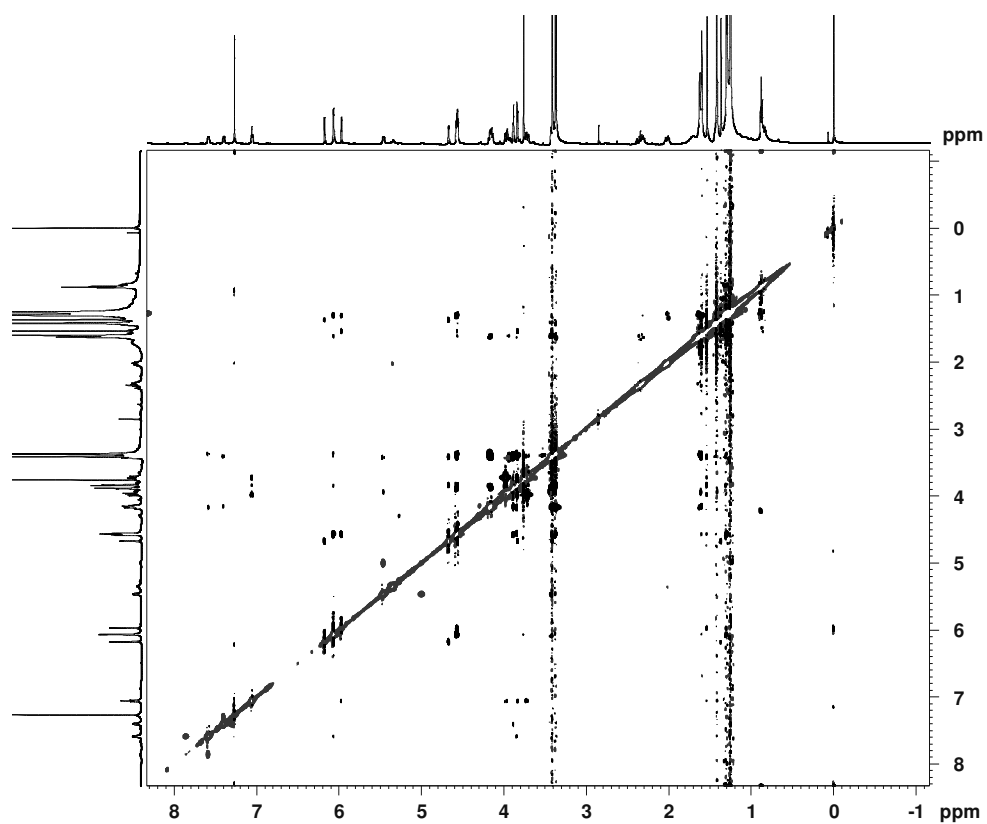
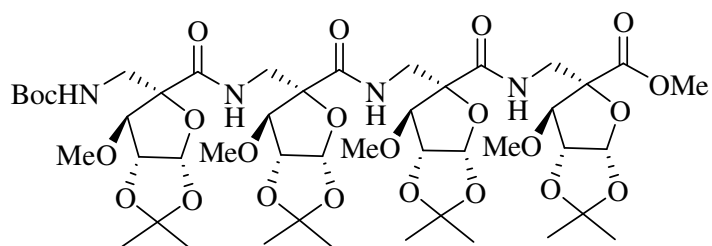


Figure S29. ROESY Spectrum of peptide **3** (600 MHz, CDCl_3 , 278K).

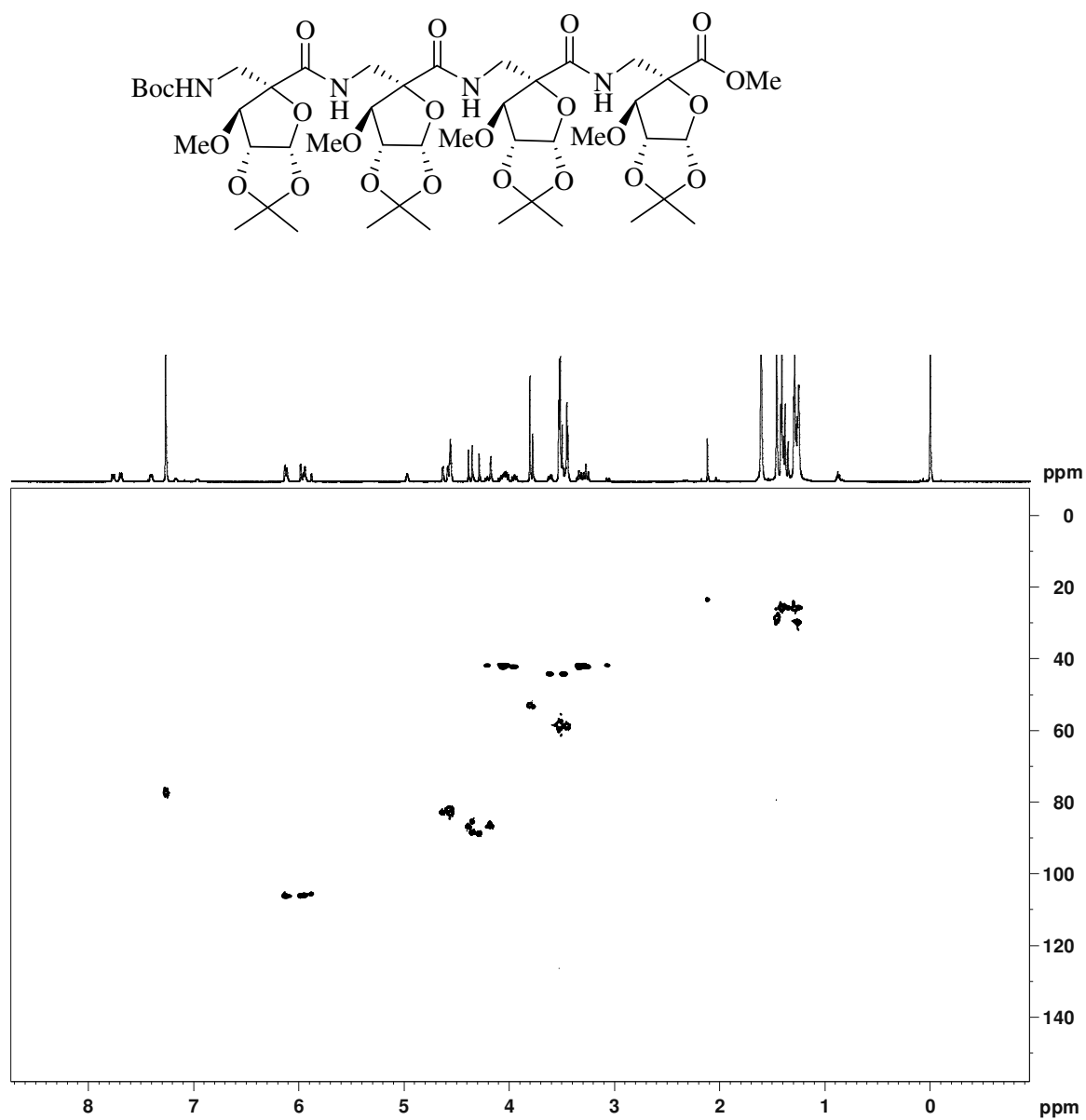
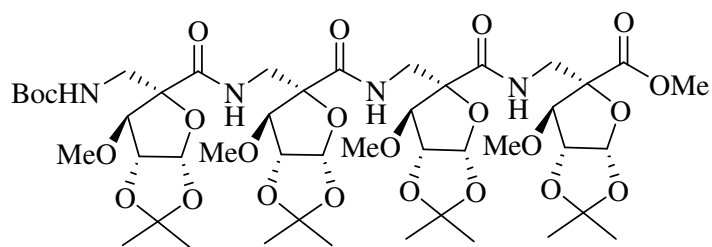
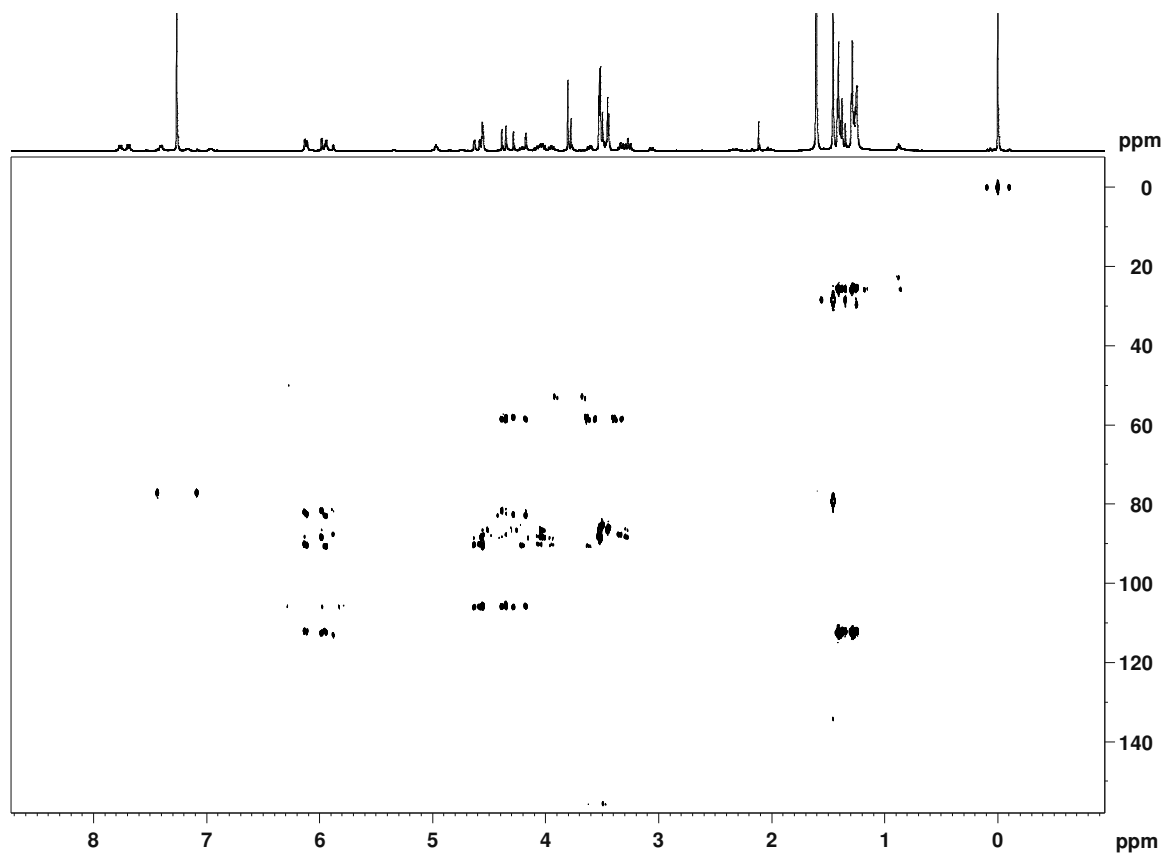


Figure S30. HSQC Spectrum of peptide **3** (600 MHz, CDCl₃, 278K).



A



B

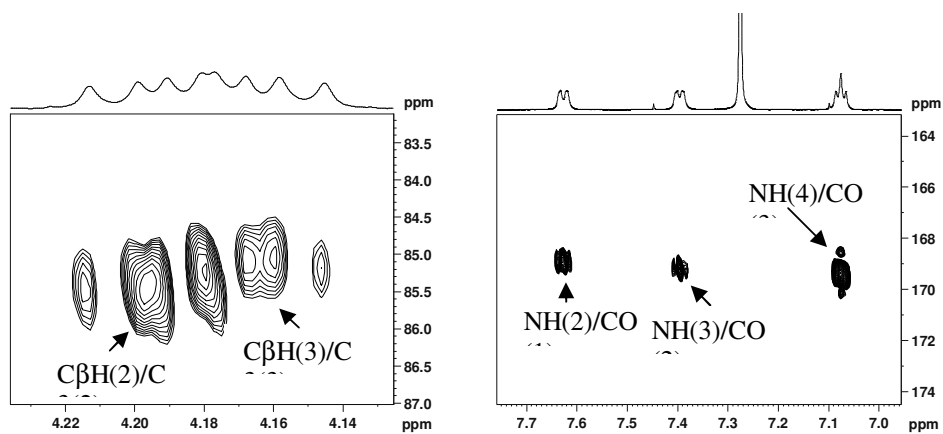


Figure S31. (A) HMBC Spectrum and (B) expansion of peptide **3** (600 MHz, CDCl_3 , 278K).

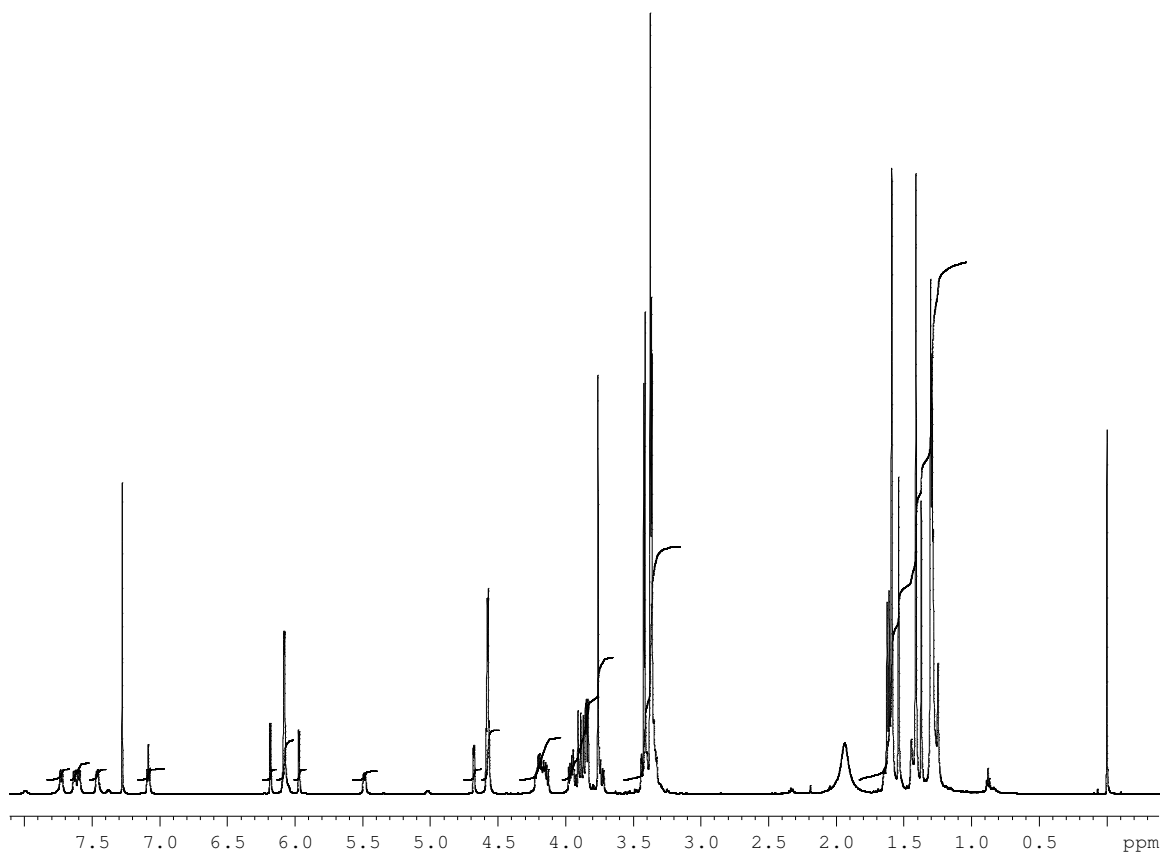
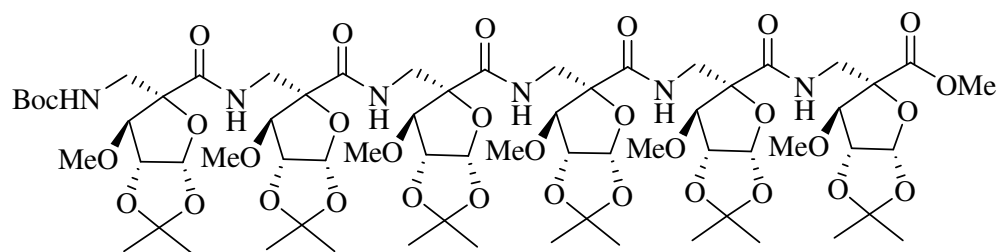


Figure S32. ^1H NMR Spectrum of peptide **4** (600 MHz, CDCl_3 , 278K).

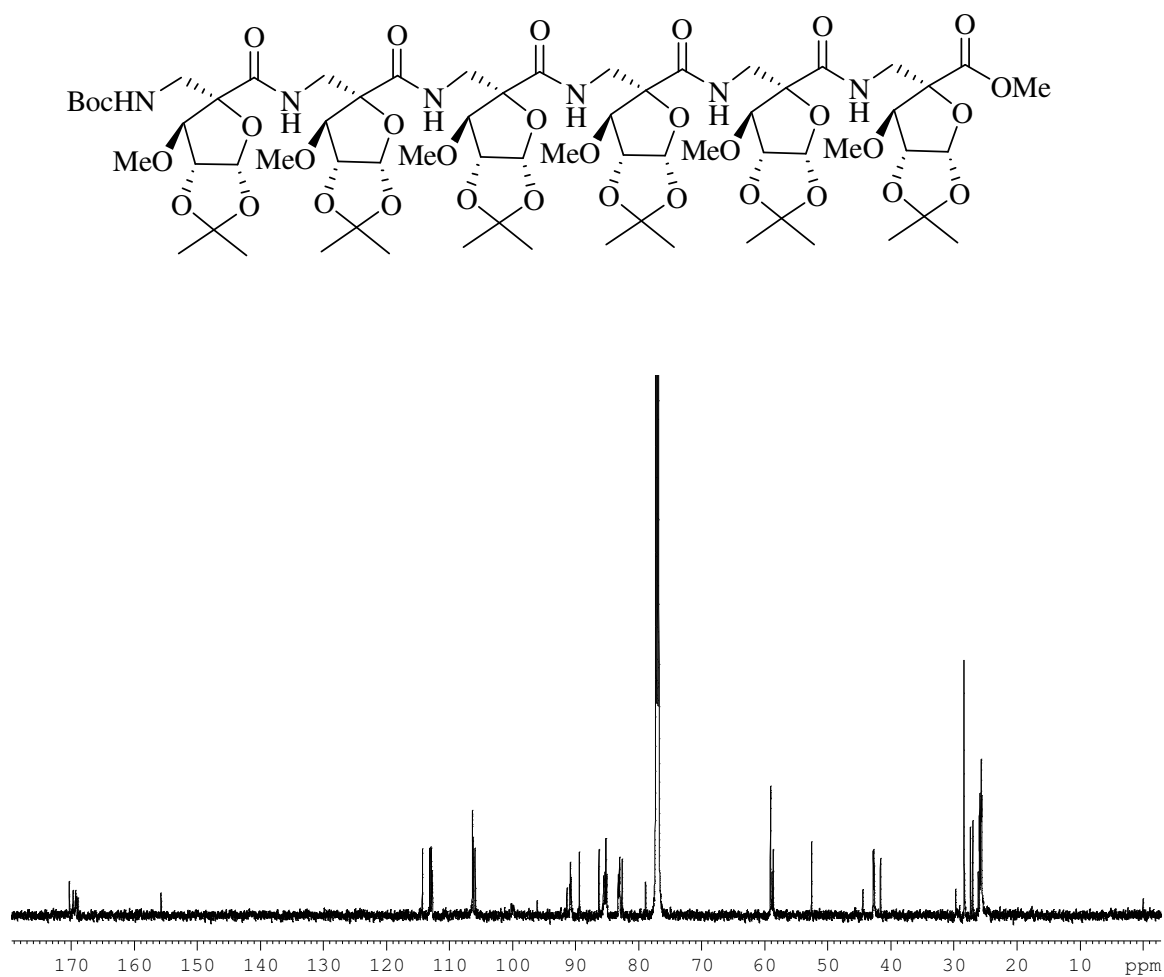


Figure S33. ^{13}C NMR Spectrum of peptide 4 (150 MHz, CDCl_3 , 278K).

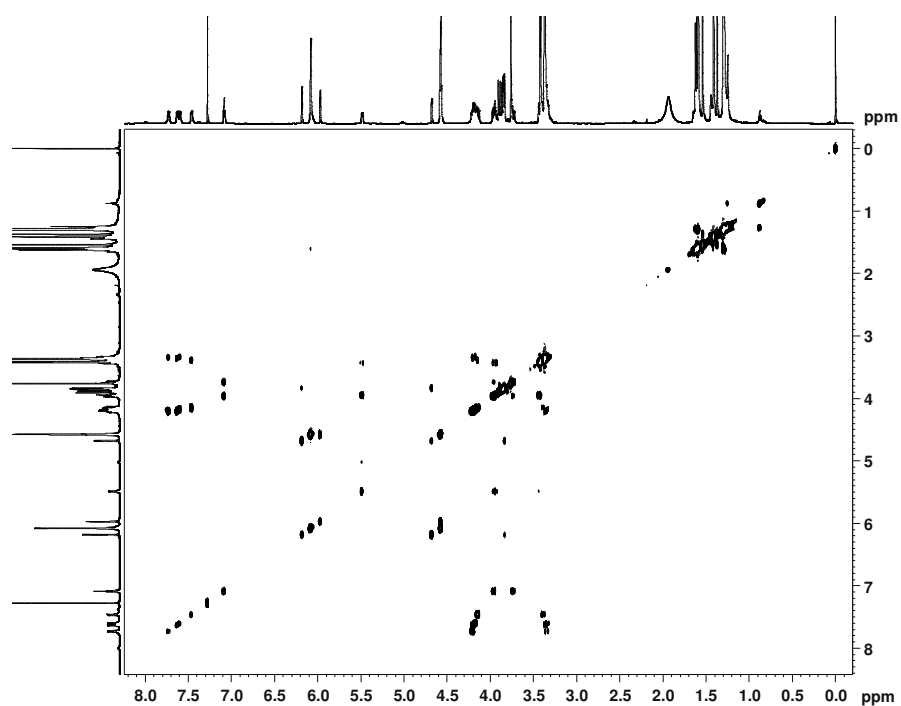
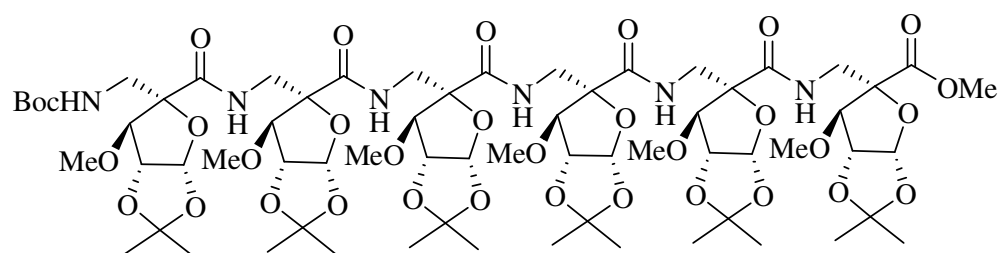


Figure S34. TOCSY Spectrum of peptide **4** (600 MHz, CDCl₃, 278K).

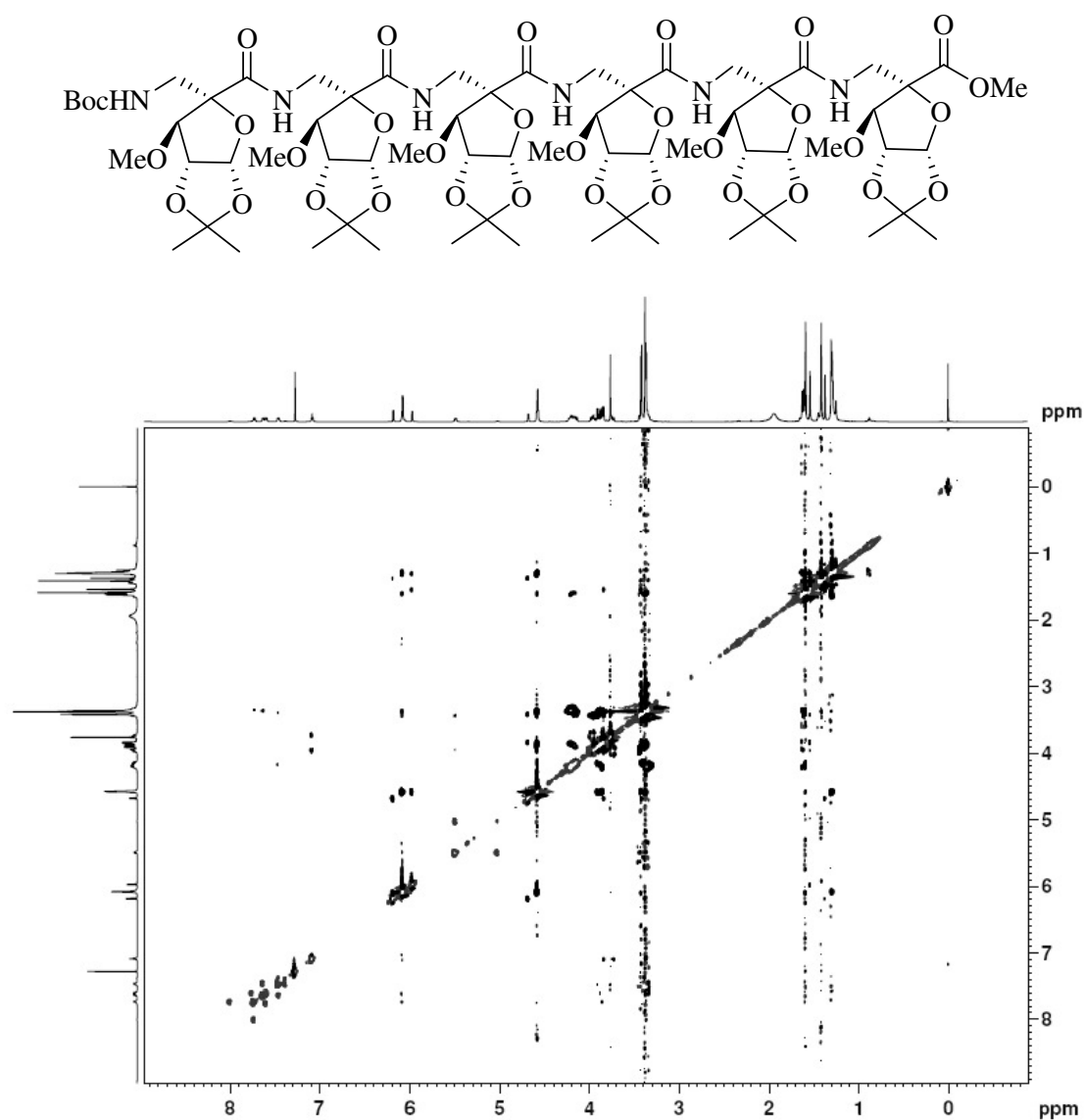


Figure S35. ROESY Spectrum of peptide **4** (600 MHz, CDCl_3 , 278K).

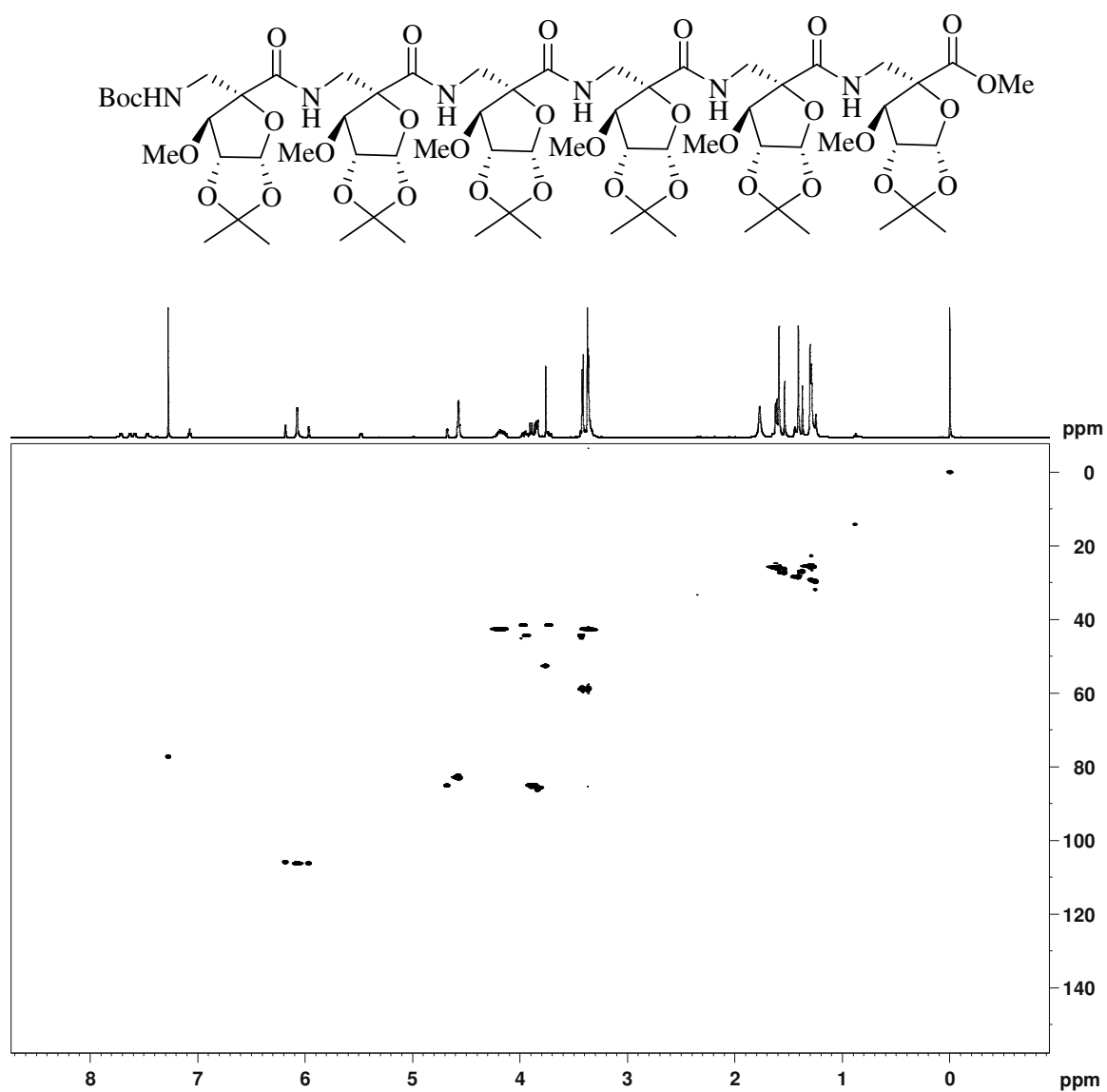


Figure S36. HSQC Spectrum of peptide 4 (600 MHz, CDCl₃, 278K).

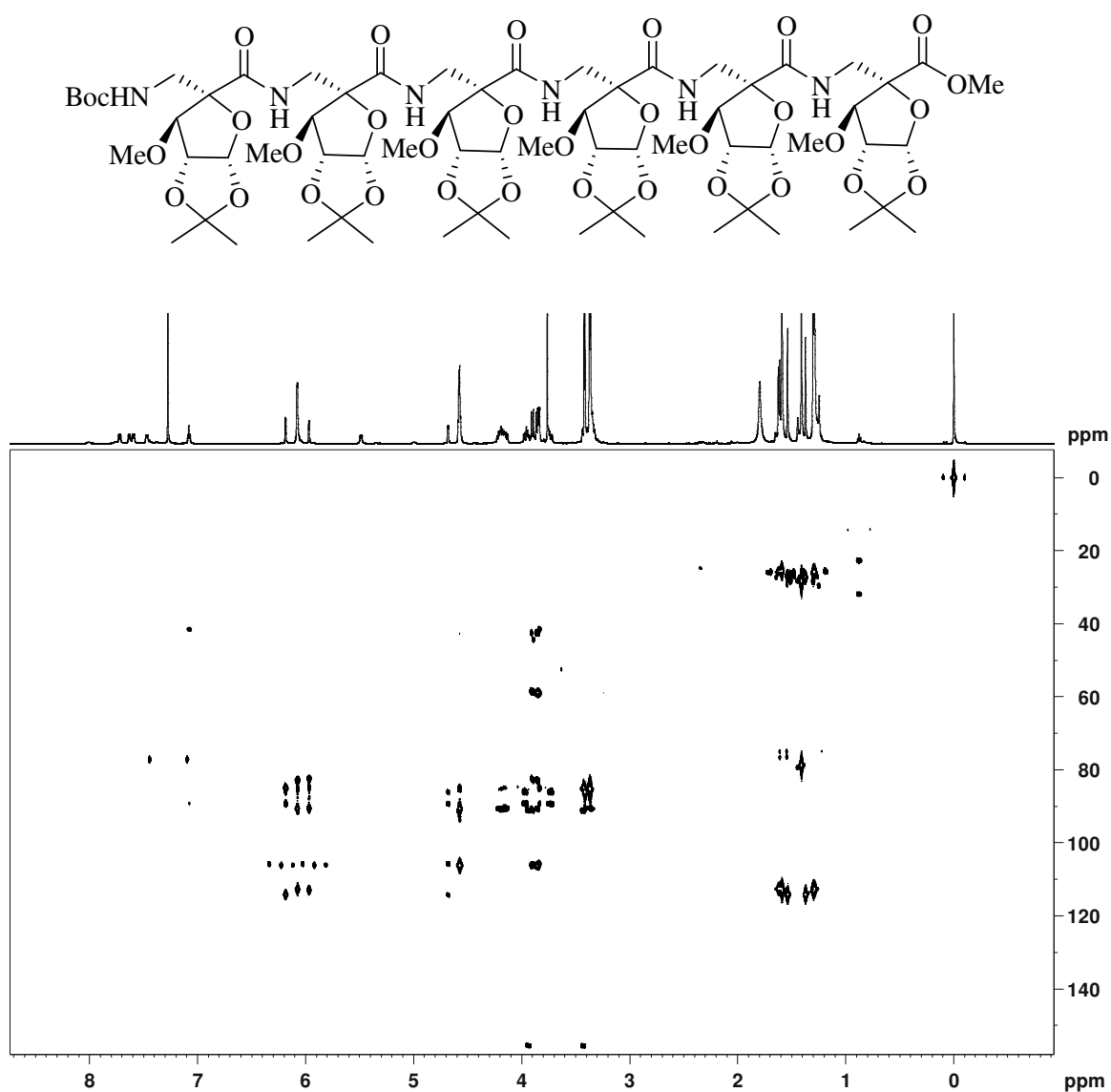


Figure S37. HMBC Spectrum of peptide **4** (600 MHz, CDCl₃, 278K).

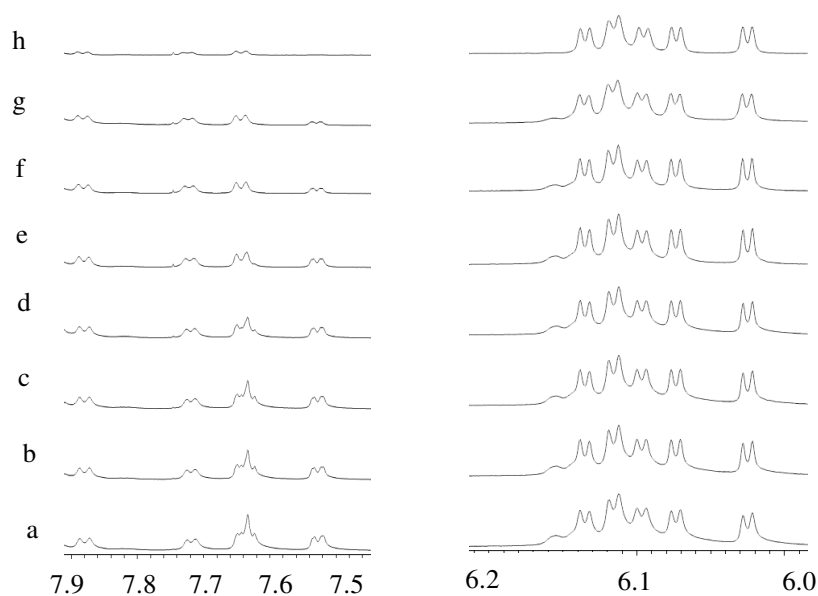
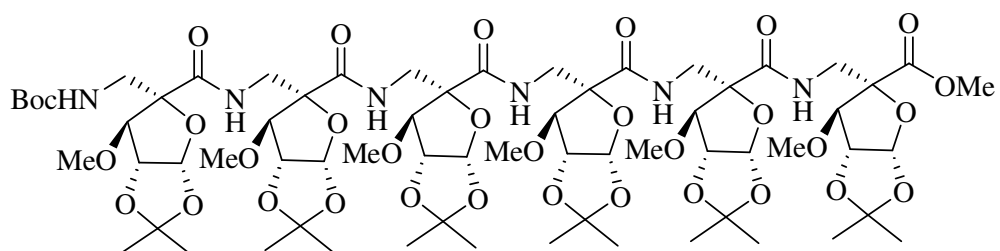


Figure S38. ^1H NMR Spectrum of peptide **4** (600 MHz, CD_3OD , 288K) as a function of time: a) 0 h; b) 1 h; c) 2 h; d) 8 h; e) 21 h; f) 45 h; g) 69 h; h) 117 h.

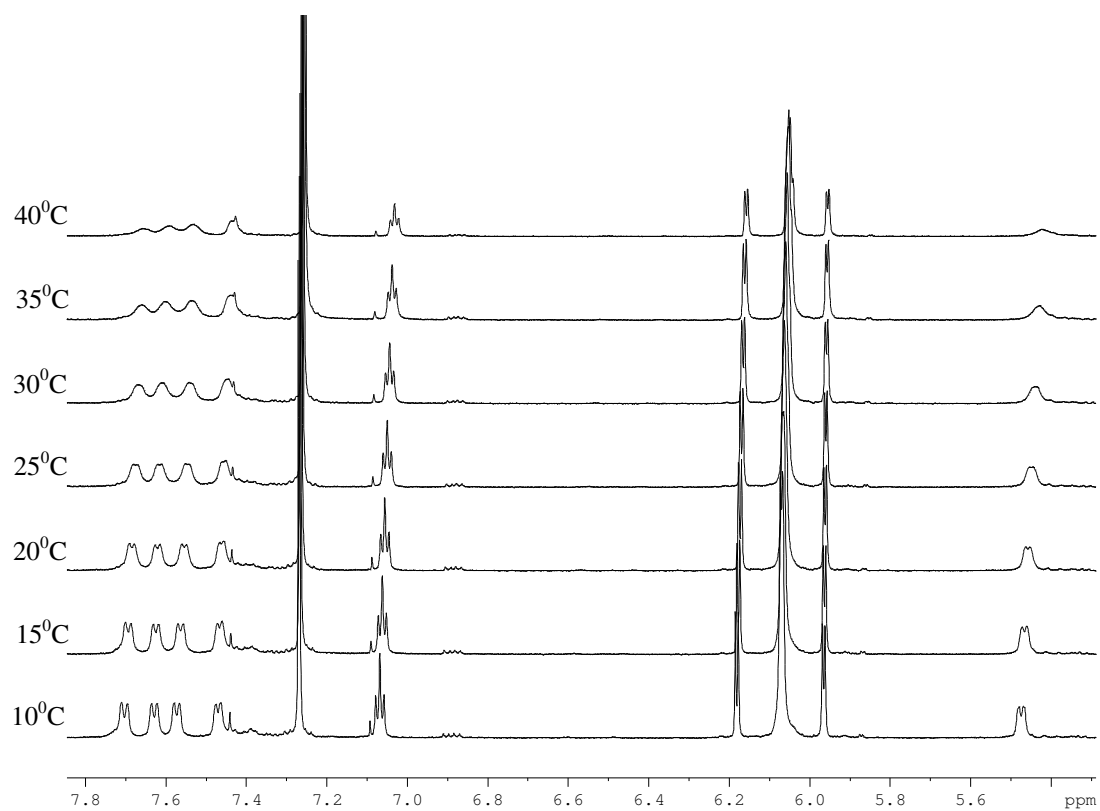


Figure S39. NMR spectra of peptide **4** in CDCl₃ as a function of temperature.

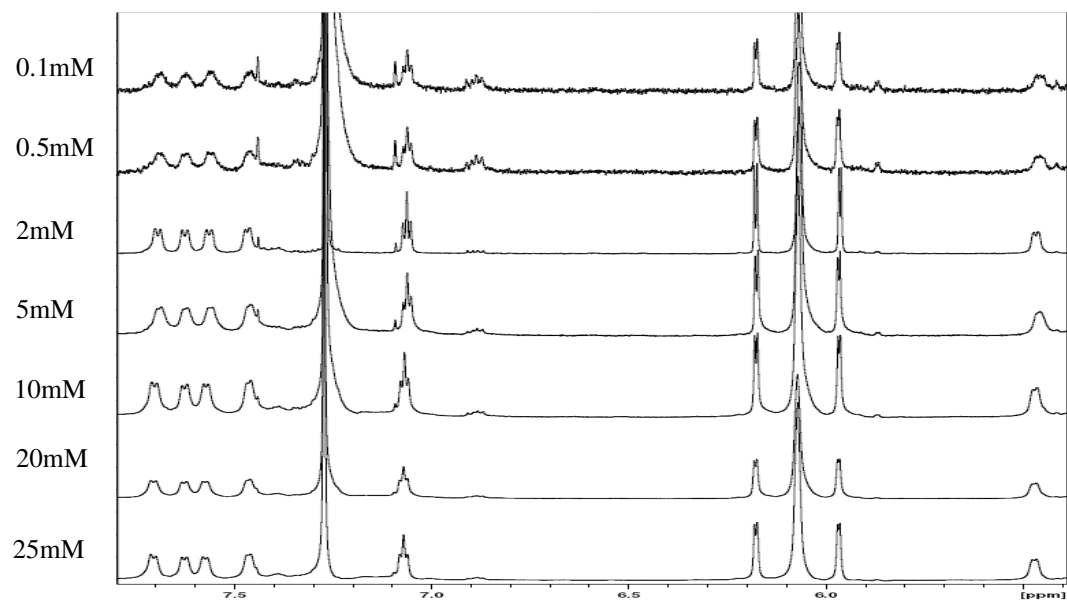


Figure S40. NMR spectra of peptide **4** as a function of concentration

Molecular Dynamics (MD): Model building and Restrained Molecular dynamics simulations on **3** and **4** were carried out using the Insight-II (97.0)/Discover program on a Silicon Graphics O2 workstation. The CVFF force field with default parameters was used throughout the simulations using a distance dependent dielectric constant with $\epsilon = 4.7$ (dielectric constant of deuterated chloroform). The distance constraints were derived from the volume integrals obtained from the ROESY spectra using a two-spin approximation and a reference distance of 1.8 Å for the geminal proton's C β H and C β' H in backbone. The upper and lower bound of the distance constraints have been obtained by enhancing and reducing the derived distance by 10% (See Tables S2 for peptide **3** and Table S3 for peptide **4**). The dihedral angle constraints used for both peptides **3** and **4** are listed in the Table S1. Backbone dihedral angles for the last residue were not constrained, as there was inadequate support from the NMR data. Distance and the dihedral angle constraints were applied with a force constant of 15 kcal/ Å and 5 kcal/ radian in the form of flat bottom potential. First, minimizations were done with steepest descent, followed by conjugate gradient methods for a maximum of 1000 iterations each or RMS deviation of 0.001 kcal/ mol, whichever was earlier. The energy-minimized structures were then subjected to MD simulations at a temperature of 300 K. The molecules were initially equilibrated for 1 ps and subsequently subjected to a 2 ns simulated annealing protocol. Starting from 300 K they were heated to 1500 K in four steps increasing the temperature by 300 K and simulating for 2.5 ps at each step, and then subsequently cooled back to 300 K in 4 steps decreasing the temperature by 300 K in each step again simulating for 2.5 ps at each step. After this, a structure was saved and the above process was repeated 100 times. The 100 structures generated so were energy minimized with the above

protocol. From these 100 energy minimized structures, only twenty of the best possible structures were superimposed for display. For peptide **3**, the backbone and heavy atoms RMSD, are 0.56Å and 1.29Å respectively. The corresponding values for peptide **4** are 1.37 Å and 1.44 Å.

Table S1. Dihedral angle constraints used for peptide **3** and **4** (except for the residue at the C-terminal)

Dihedral angle	(°)
ϕ	-120 ± 20
θ	-60 ± 20
ψ	180 ± 20
C2H-C2-C3-C3H	-90 ± 20

Table S2. Distance constraints used in the MD calculations for peptide 3

Residue	Atom	Residue	Atom	Upper Bound	Lower Bound
1	NH	1	C3H	3.58	2.93
2	NH	2	C3H	3.47	2.84
2	NH	1	C1H	3.70	3.03
2	C β H _(pro-R)	2	C3H	2.60	2.13
3	NH	2	C1H	3.74	3.06
3	NH	3	C3H	3.47	2.84
3	C β H _(pro-R)	3	C3H	2.61	2.14
4	NH	3	C1H	3.54	2.89
4	NH	4	C3H	3.49	2.85
4	C β H _(pro-R)	4	C3H	2.71	2.20

Table S3. Distance constraints used in the MD calculations for peptide 4

Residue	Atom	Residue	Atom	Upper Bound	Lower Bound
1	NH	1	C3H	3.56	2.91
2	NH	2	C3H	3.49	2.85
2	NH	1	C1H	3.92	3.21
3	NH	2	C1H	3.71	3.03
3	NH	3	C3H	3.40	2.79
3	C β H _(pro-R)	3	C3H	2.74	2.24
4	NH	3	C1H	3.81	3.12
4	NH	4	C3H	3.55	2.91
4	C β H _(pro-R)	4	C3H	2.79	2.28
5	NH	4	C1H	3.92	3.21
5	NH	5	C3H	3.54	2.89
6	NH	5	C1H	3.82	3.13
6	NH	6	C3H	3.37	2.76
6	C β H _(pro-R)	6	C3H	2.72	2.22

CD Spectra of peptides 3 and 4. The CD spectra of 0.2 mM solution in methanol for both **3** and **4** show a narrow and another broad shallow negative maxima at ~192 nm and at about 215 nm, respectively, alongwith a maxima for **4** at ~ 195, while the $[\theta]$ values are zero at about 194 and 200 nm for **4**. Though the molecular ellipticities θ per residue are rather small, still suggesting there is secondary structures for these oligomers.

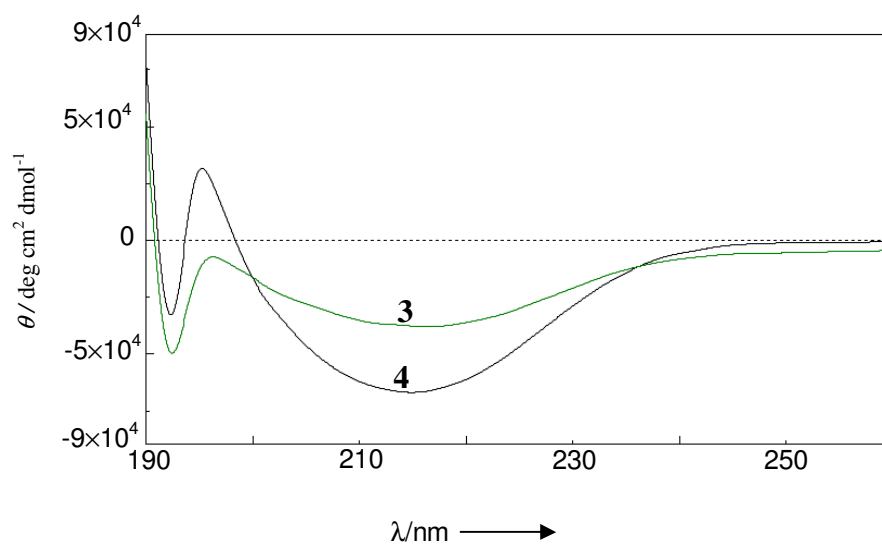


Figure S41. CD spectra for peptides **3** and **4** in MeOH.