Supporting Information

Elusive π-Helical Peptide Foldamers Spotted by Chiroptical Studies

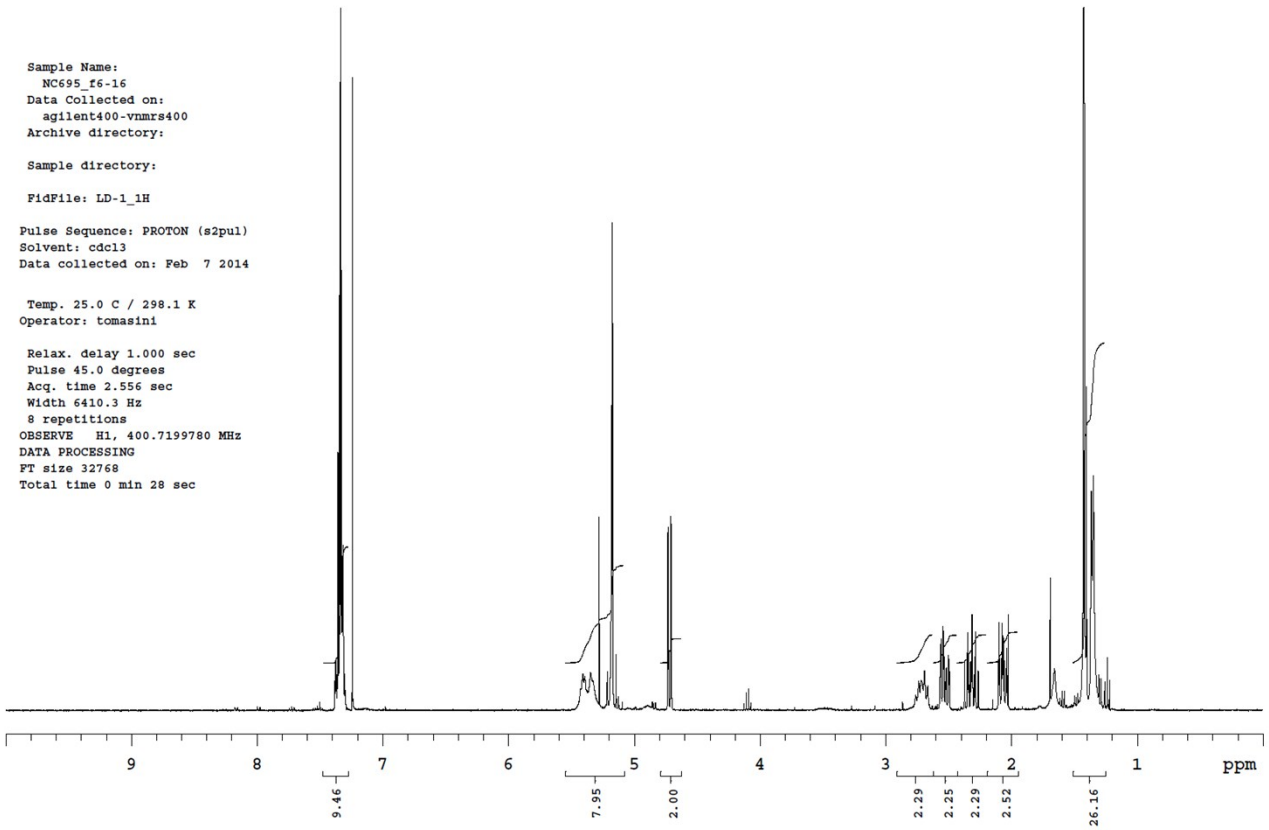
Sergio Di Silvio,[a] Fabio Bologna,[a] Lorenzo Milli,[a] Demetra Giuri,[a] Nicola Zanna,[a] Nicola Castellucci,[a] Magda Monari,[a] Matteo Calvaresi,[a] Marcin Górecki,[b,c] Gaetano Angelici,[c] Claudia Tomasini*[a] and Gennaro Pescitelli*[c]

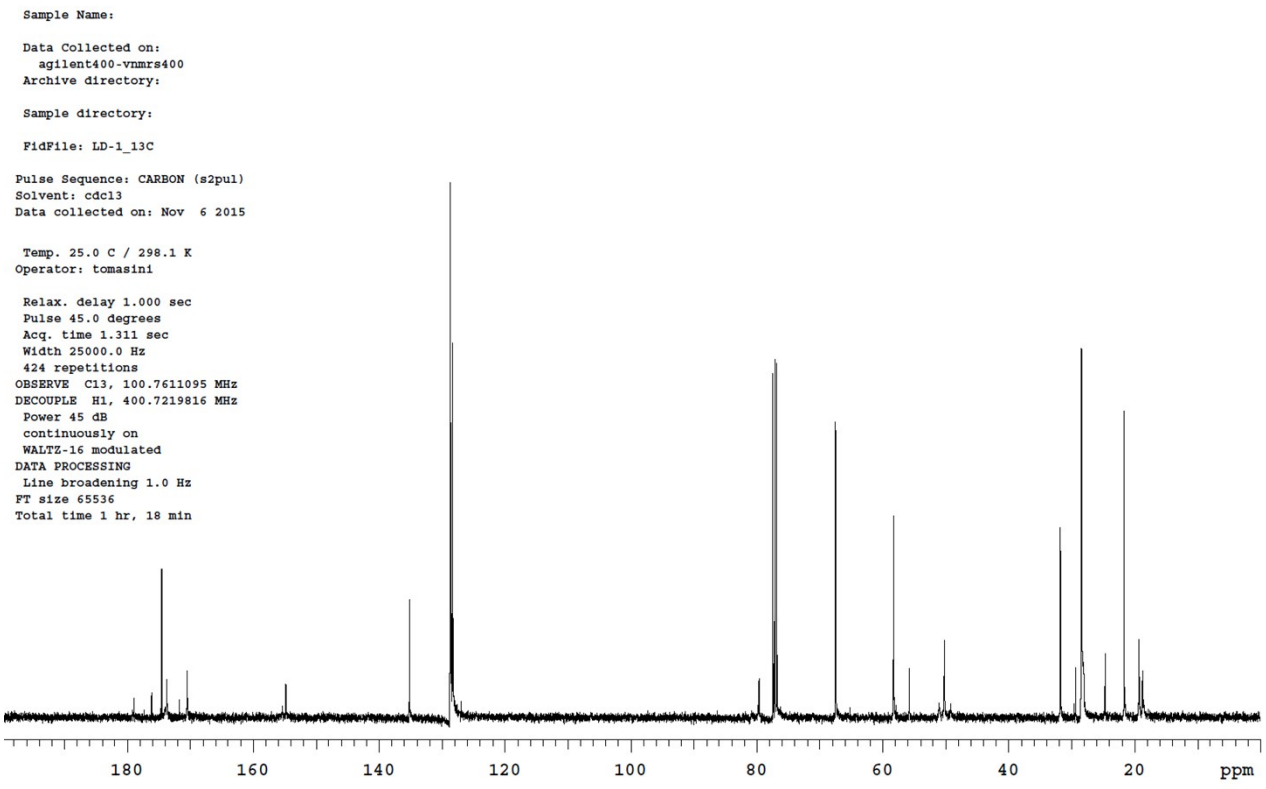
a. Dipartimento di Chimica "G. Ciamician" - Alma Mater Studiorum Università di Bologna -Via F. Selmi 2, 40126 Bologna – Italy
b. Dipartimento di Chimica e Chimica Industriale - Università di Pisa - Via Moruzzi 13 - 56124 Pisa – Italy
c. Institute of Organic Chemistry, Polish Academy of Sciences, Warsaw 01-224, -Poland

*Corresponding authors' e-mail: claudia.tomasini@unibo.it; gennaro.pescitelli@unipi.it

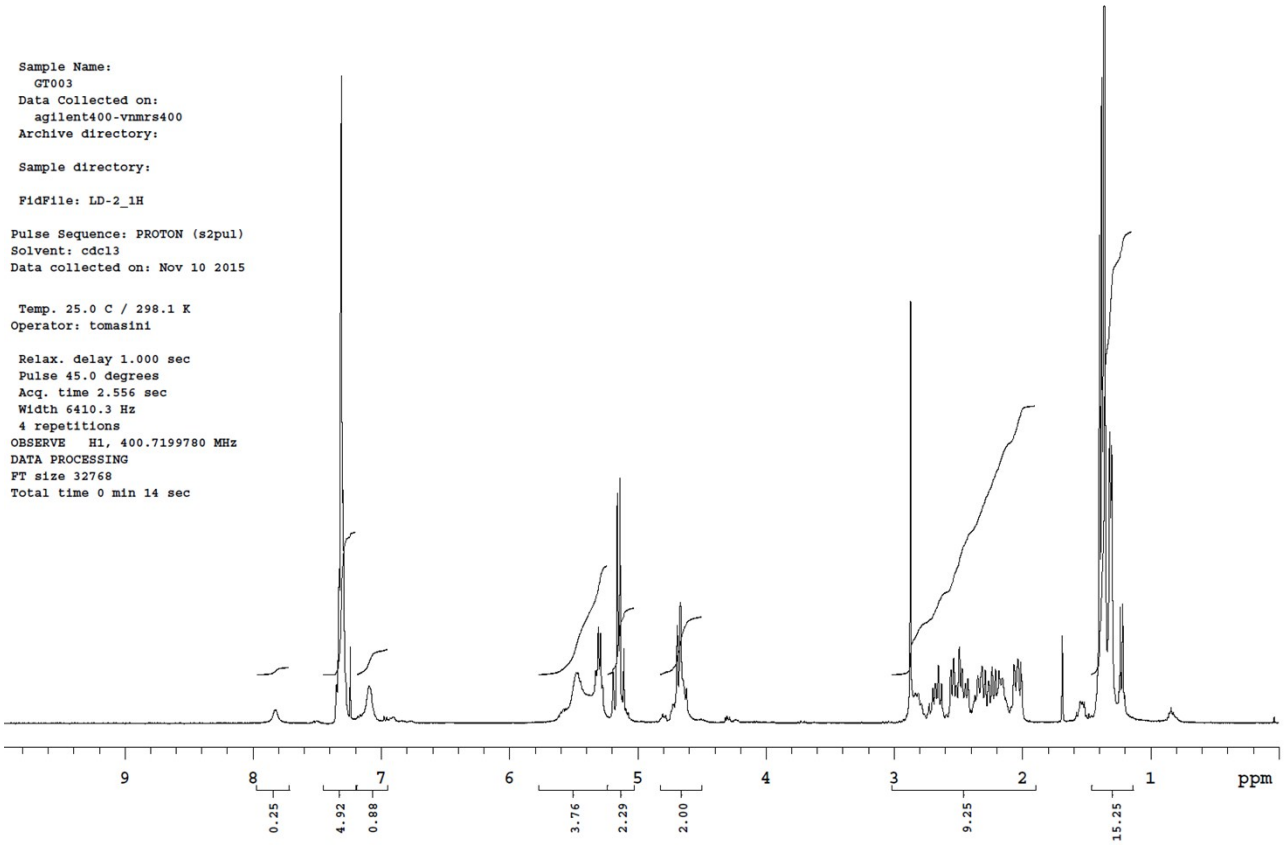
Table of Contents

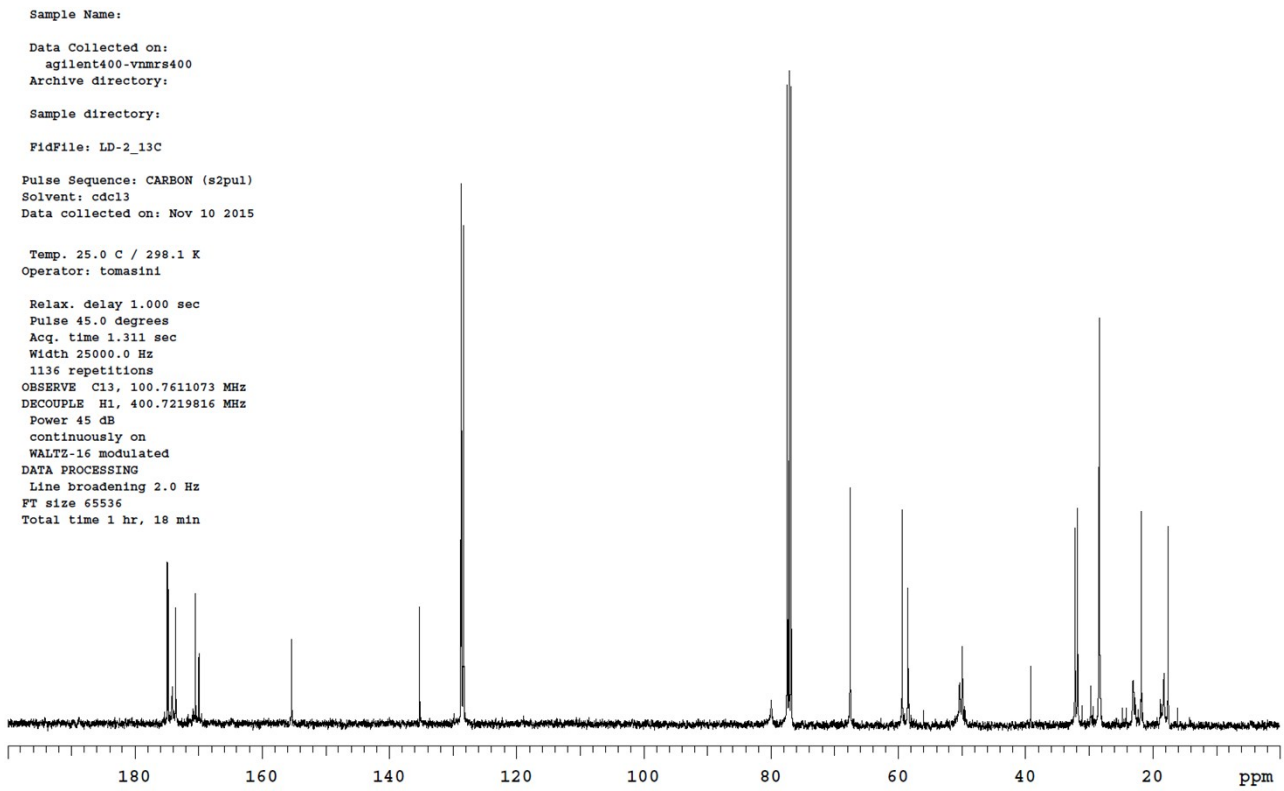
Boc-L-Ala-D-pGlu-OBn LD-1



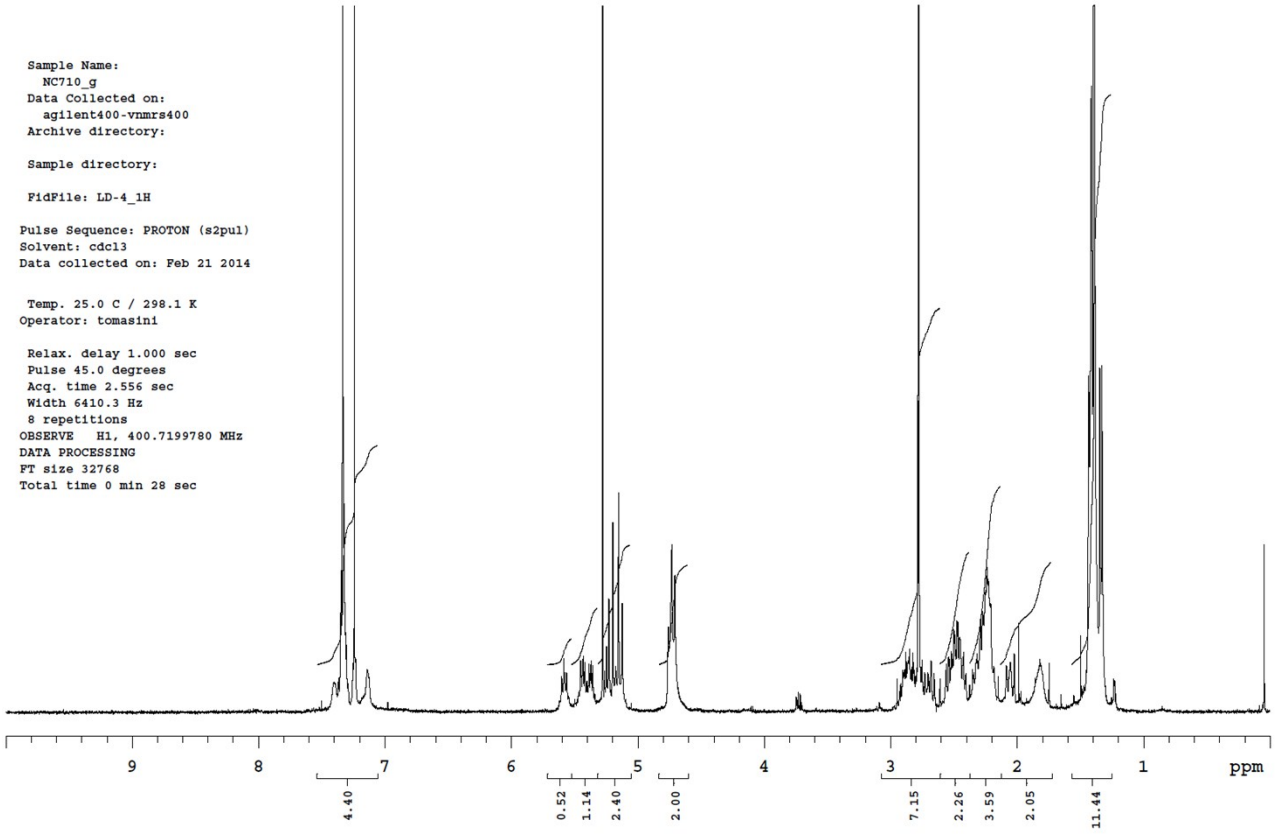


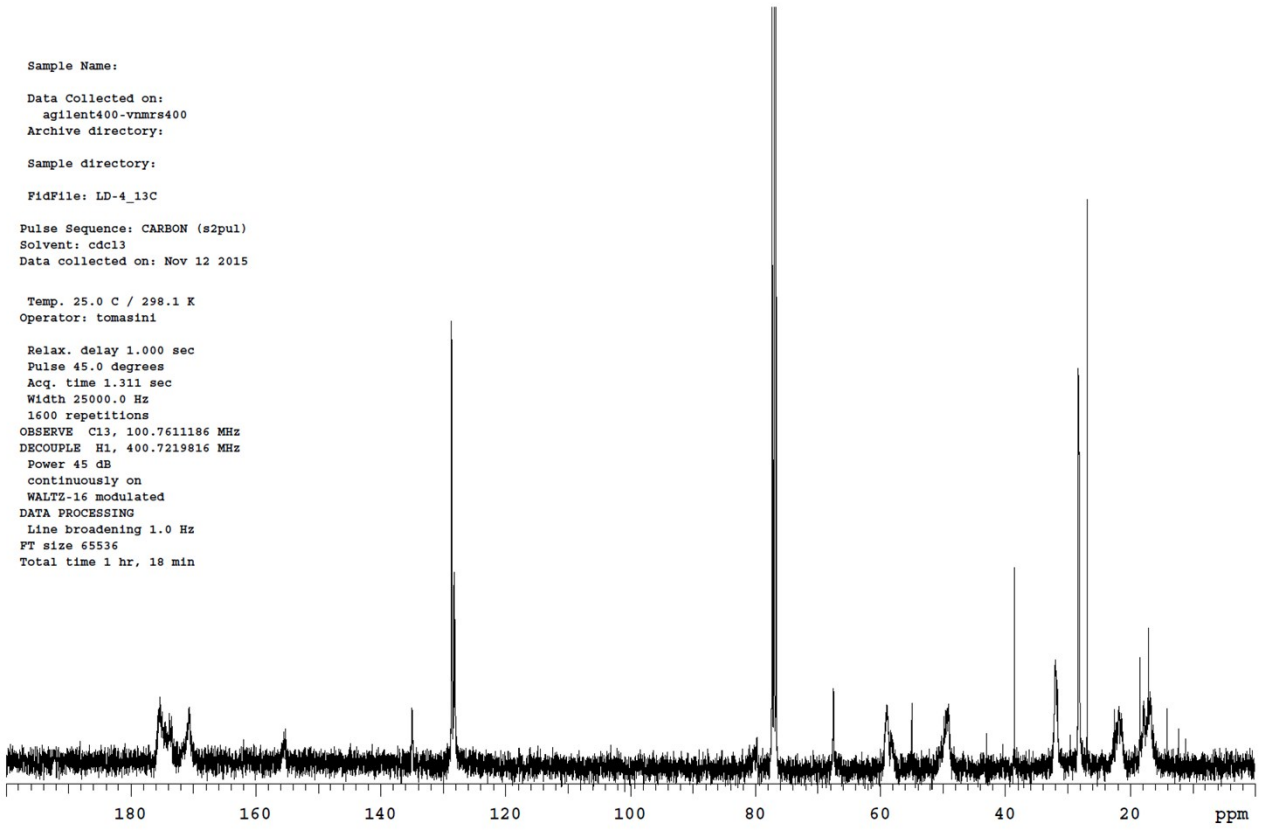
Boc-(L-Ala-D-pGlu)$_2$-OBn LD-2



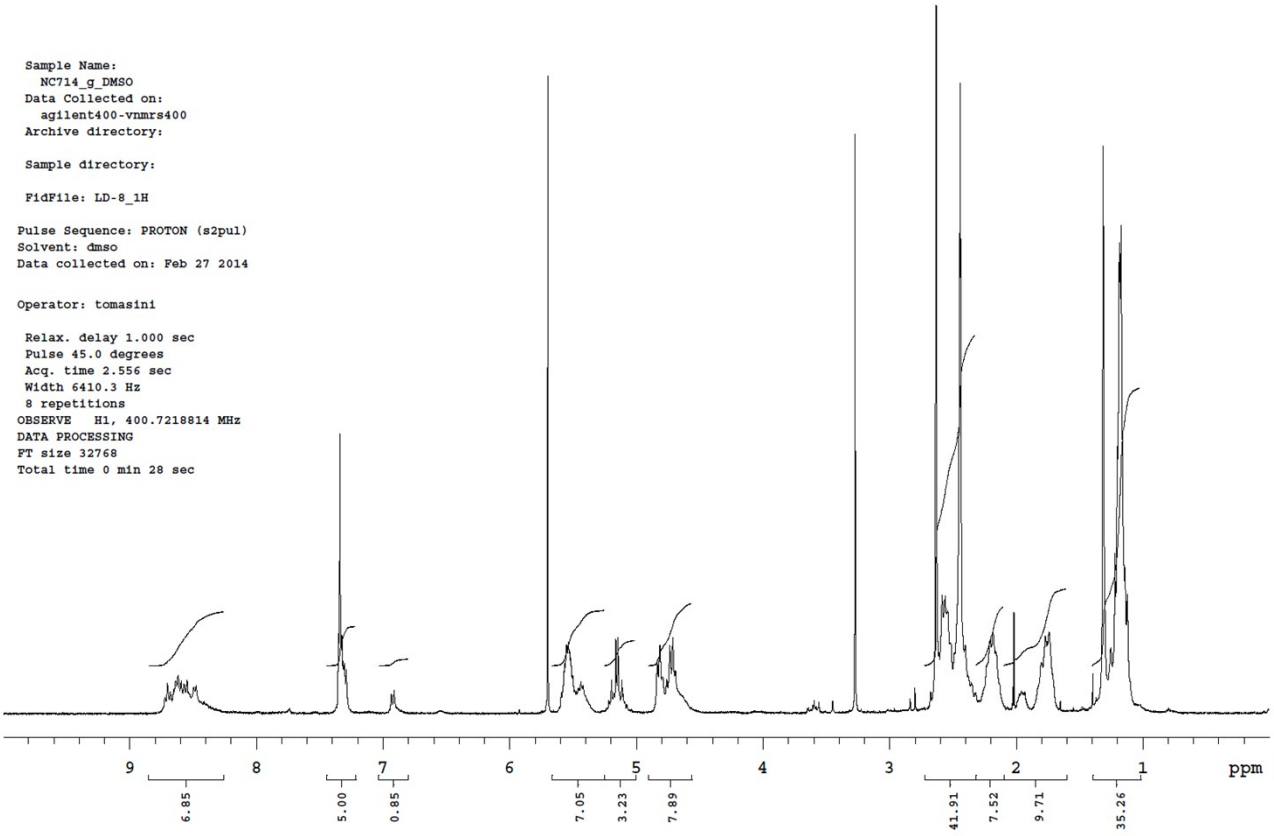


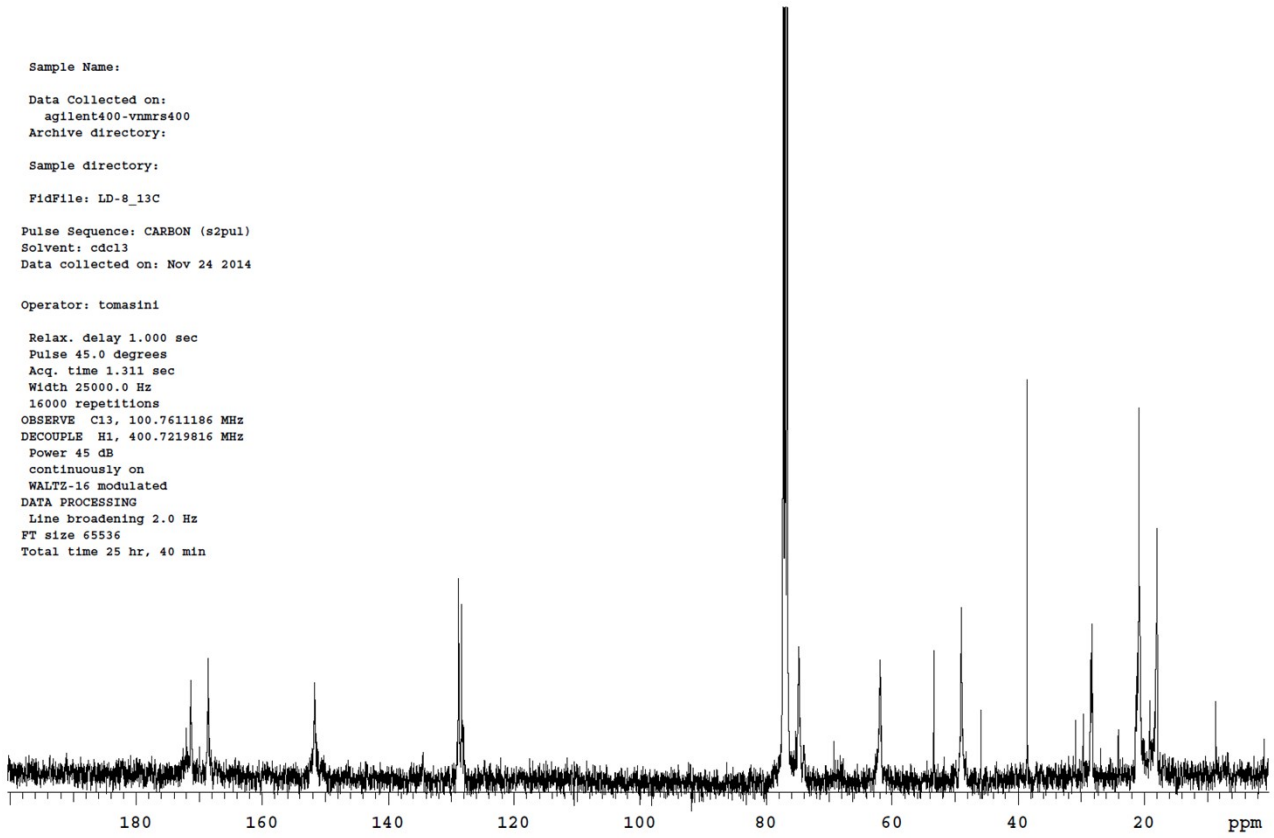
Boc-(L-Ala-D-pGlu)$_4$-OBn LD-4



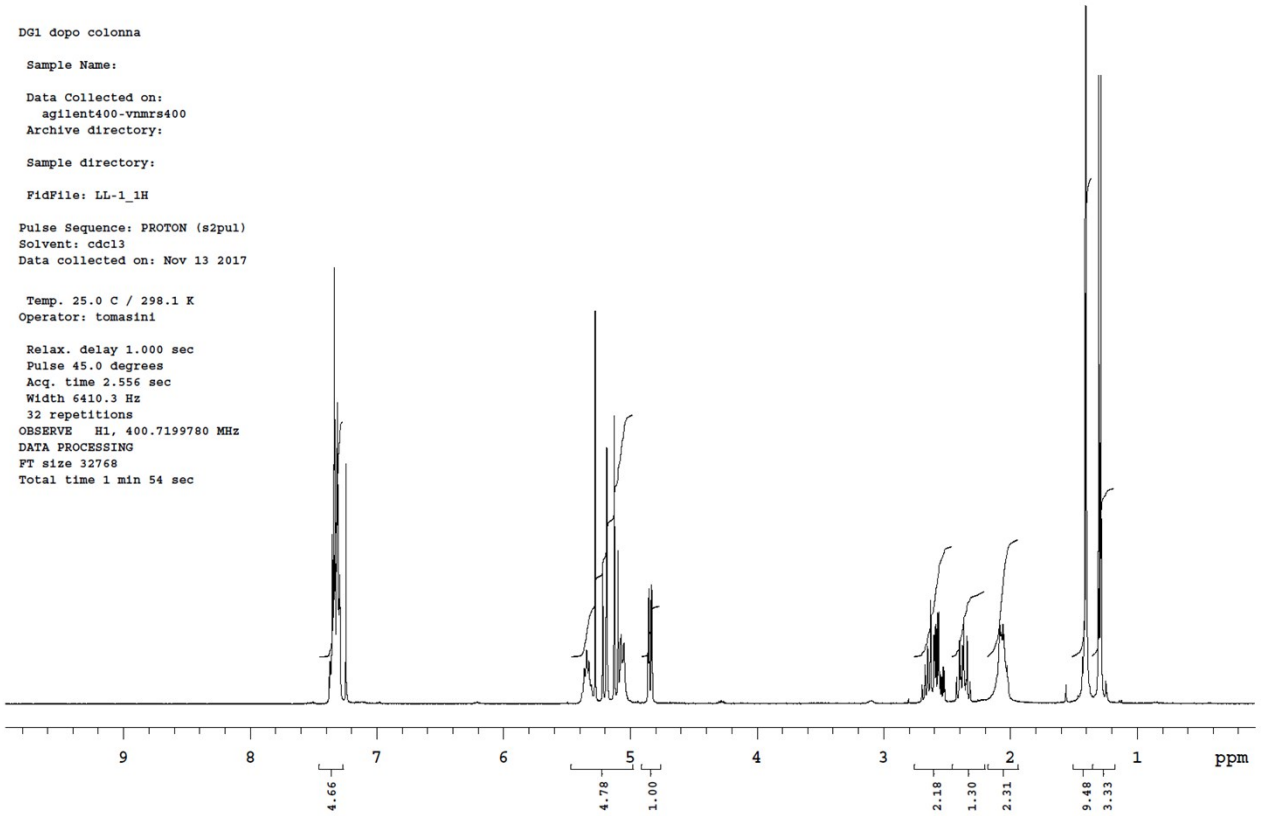


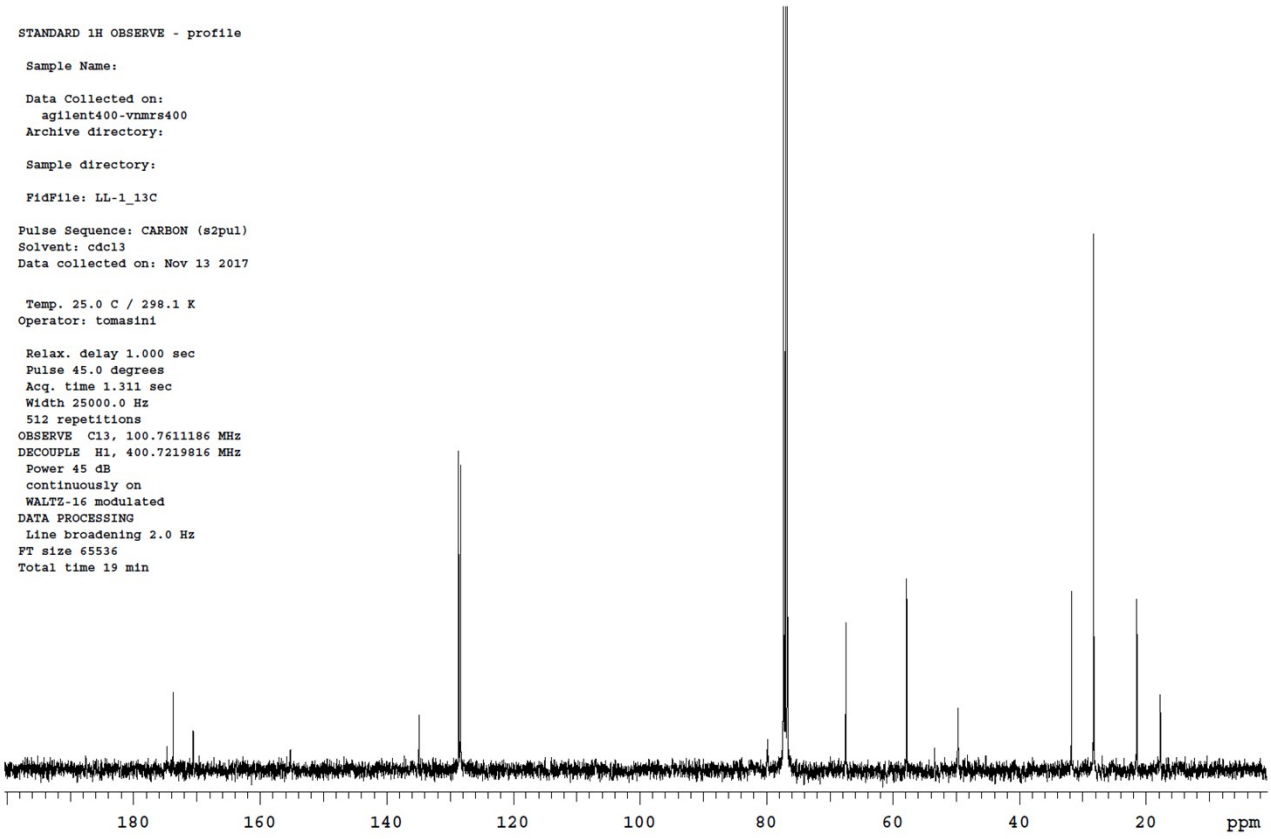
Boc-(L-Ala-D-pGlu)$_8$-OBn LD-8



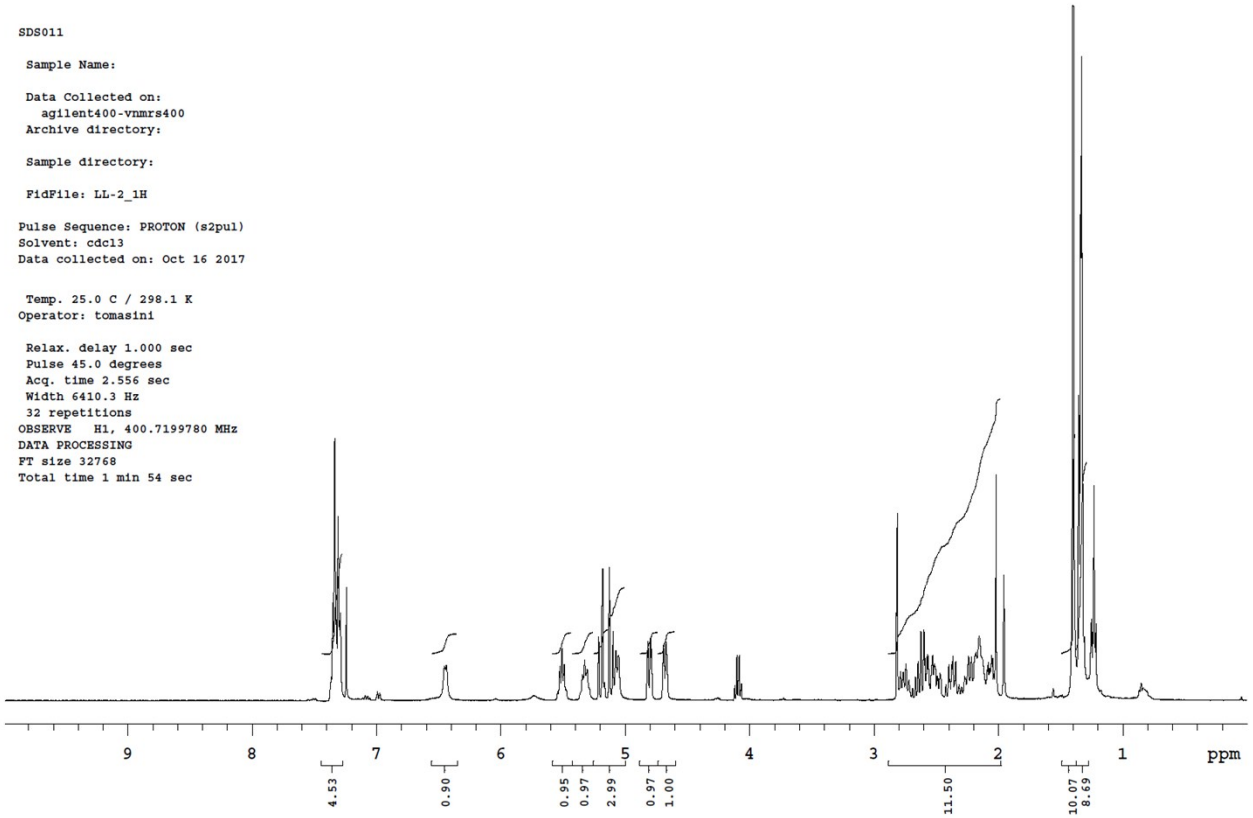


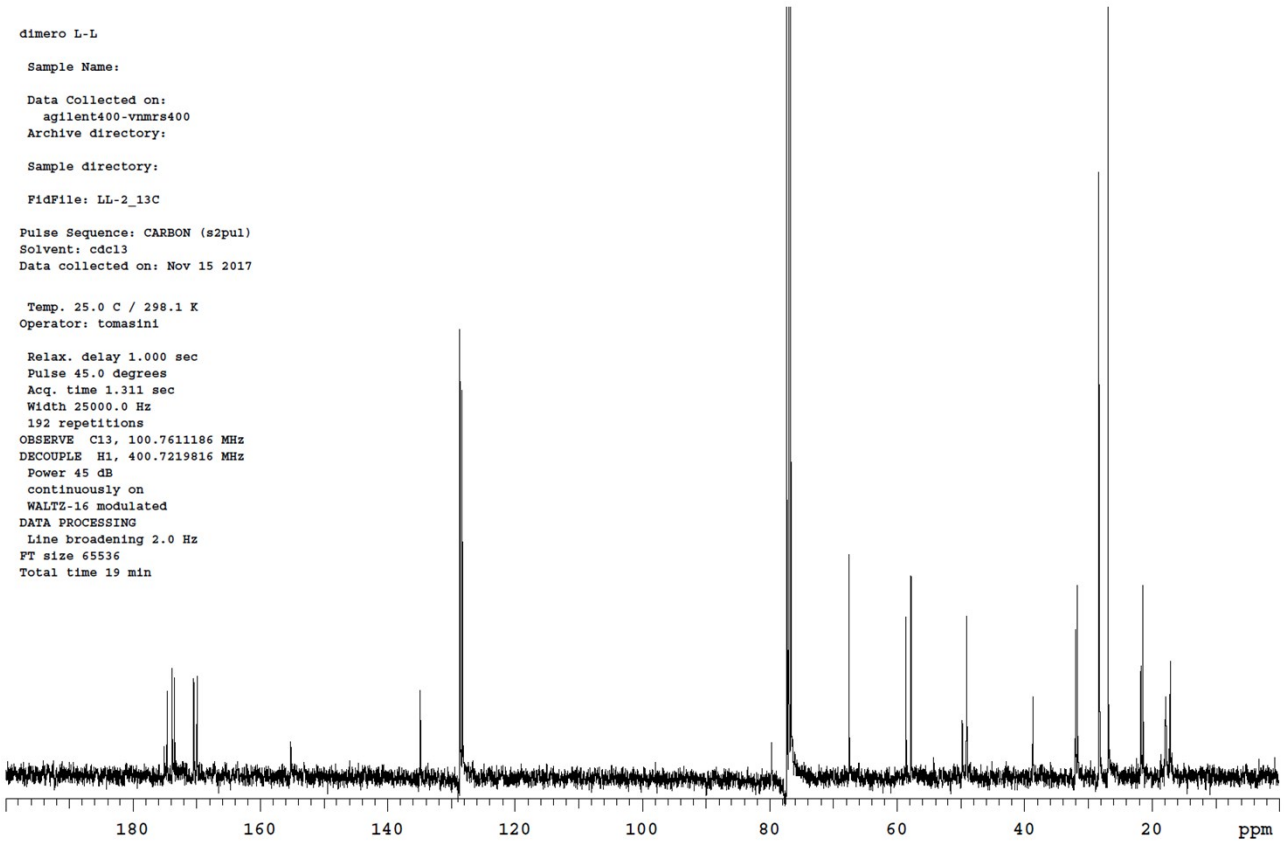
Boc-L-Ala-L-pGlu-OBn LL-1


DG1 dopo colonna
Sample Name:
Data Collected on:
agilent400-vnmrs400
Archive directory:
Sample directory:
FidFile: LL-1_1H
Pulse Sequence: PROTON (s2pul)
Solvent: cdcl3
Data collected on: Nov 13 2017
Temp. 25.0 C / 298.1 K
Operator: tomasini
Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.556 sec
Width 6410.3 Hz
32 repetitions
OBSERVE H1, 400.7199780 MHz
DATA PROCESSING
FT size 32768
Total time 1 min 54 sec
9
8
7
6
5
4
3
2
1
ppm
4.66
4.78
1.00
2.18
1.30
2.31
9.48
3.33

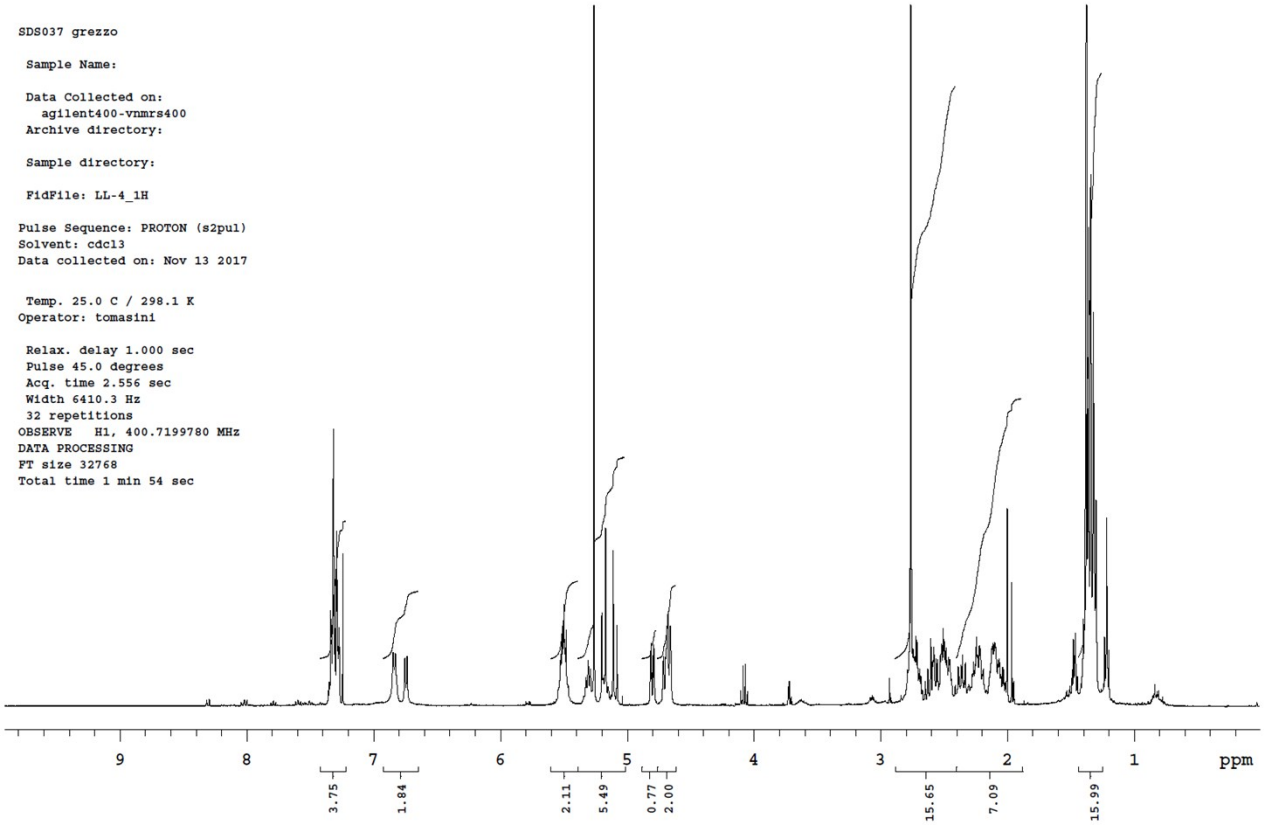

STANDARD 1H OBSERVE - profile
Sample Name:
Data Collected on:
agilent400-vnmrs400
Archive directory:
Sample directory:
FidFile: LL-1_13C
Pulse Sequence: CARBON (s2pul)
Solvent: cdcl3
Data collected on: Nov 13 2017
Temp. 25.0 C / 298.1 K
Operator: tomasini
Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.311 sec
Width 25000.0 Hz
512 repetitions
OBSERVE C13, 100.7611186 MHz
DECOUPLE H1, 400.7219816 MHz
Power 45 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 2.0 Hz
FT size 65536
Total time 19 min
180
160
140
120
100
80
60
40
20
ppm

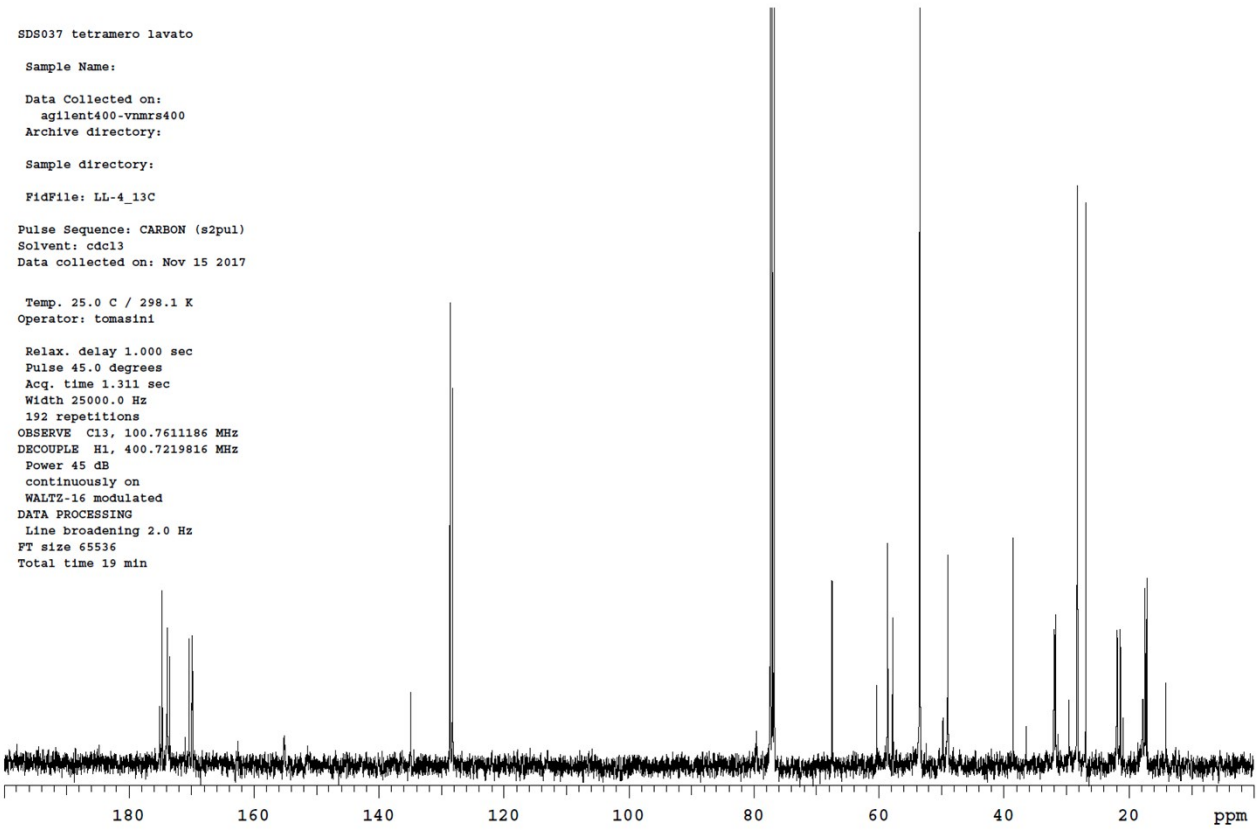
Boc-(L-Ala-L-pGlu)$_2$-OBn LL-2



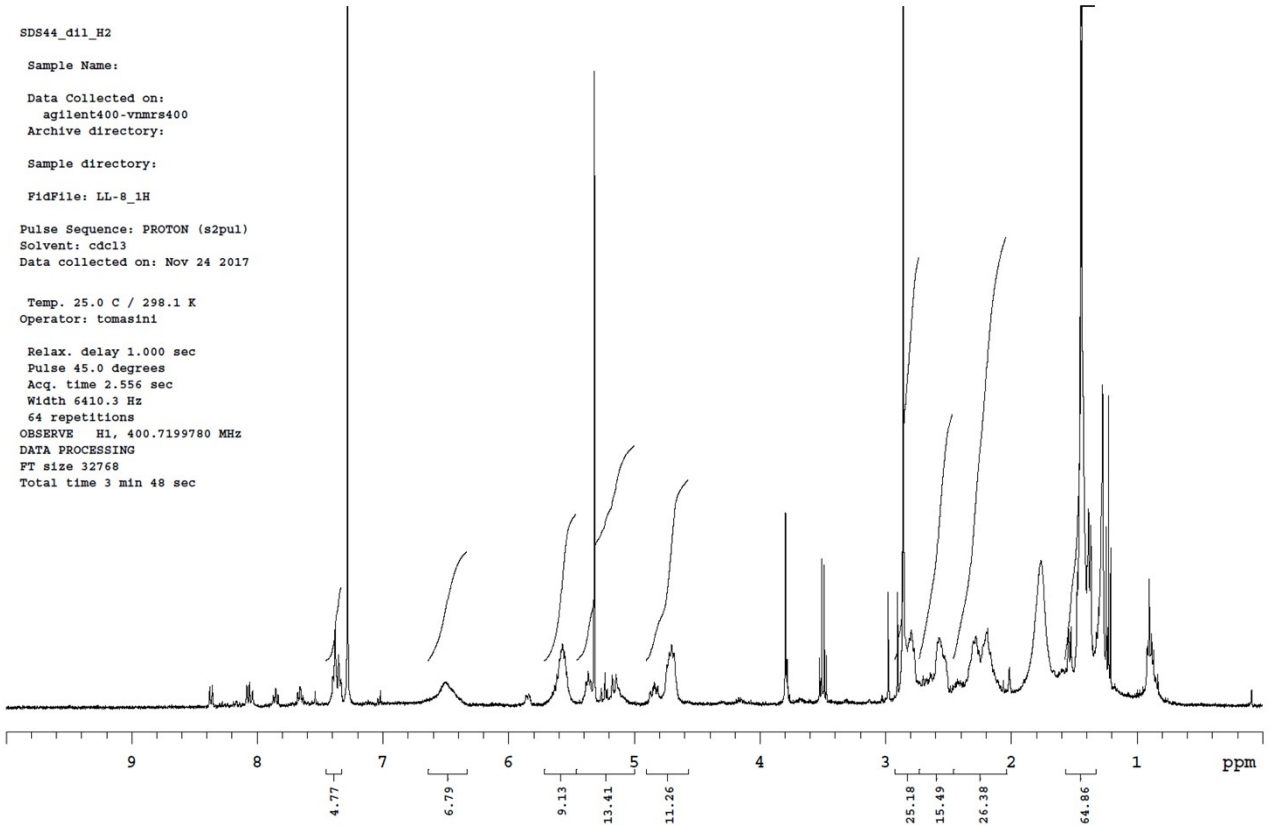


Boc-(L-Ala-L-pGlu)$_4$-OBn LL-4





Boc-(L-Ala-L-pGlu)$_8$-OBn LL-8


SDS44_d11_H2
Sample Name:
Data Collected on:
agilent400-vnmrs400
Archive directory:
Sample directory:
FidFile: LL-8_1H
Pulse Sequence: PROTON (s2pul)
Solvent: cdcl3
Data collected on: Nov 24 2017
Temp. 25.0 C / 298.1 K
Operator: tomasini
Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 2.556 sec
Width 6410.3 Hz
64 repetitions
OBSERVE H1, 400.7199780 MHz
DATA PROCESSING
FT size 32768
Total time 3 min 48 sec
9
8
7
6
5
4
3
2
1
ppm
4.77
6.79
9.13
13.41
11.26
25.18
15.49
26.38
64.86

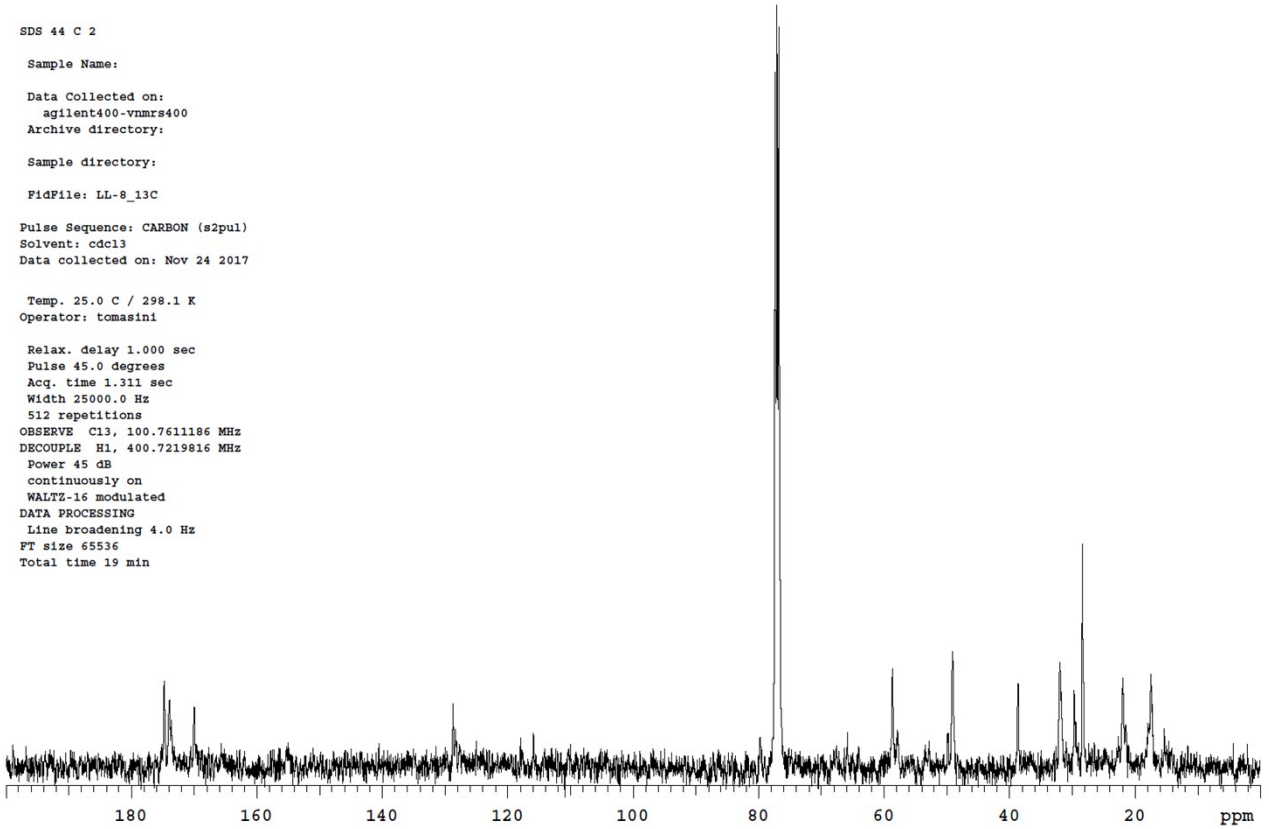

SDS 44 C 2
Sample Name:
Data Collected on:
agilent400-vnmrs400
Archive directory:
Sample directory:
FidFile: LL-8_13C
Pulse Sequence: CARBON (s2pul)
Solvent: cdcl3
Data collected on: Nov 24 2017
Temp. 25.0 C / 298.1 K
Operator: tomasini
Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.311 sec
Width 25000.0 Hz
512 repetitions
OBSERVE C13, 100.7611186 MHz
DECOUPLE H1, 400.7219816 MHz
Power 45 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 4.0 Hz
FT size 65536
Total time 19 min
180
160
140
120
100
80
60
40
20
ppm

Table S1. Crystal data and structure refinement for **LL-1** and **LD-1.**

Compound	**LL-1**	**LD-1**
Formula	$C_{20}H_{26}N_2O_6$	$C_{20}H_{26}N_2O_6$
Fw	390.43	390.43
Crystal symmetry	Tetragonal	Monoclinic
Space group	*P4₃*	*P2₁/c*
a, Å	10.5551(8)	18.3138(10)
b, Å	10.5551(8)	10.0679(6)
c, Å	18.9321(14)	11.1281(6)
α	90	90
β	90	102.629(2)
γ	90	90
Cell volume, Å^3	2109.2(4)	2002.2(2)
Z	4	4
D_C, Mg m^{-3}	1.229	1.295
μ(Mo-K_α), mm^{-1}	0.091	0.096
F(000)	832	832
Crystal size/ mm	0.30 x 0.25 x 0.22	0.33 x 0.30 x 0.25
θ limits, °	2.209 - 28.311	2.279 - 28.453
Reflections collected	32476	30842
Unique obs. Reflections [$F_o > 4\sigma(F_o)$]	5178 [R(int) = 0.0399]	5043 [R(int) = 0.0566]
Goodness-of-fit-on F^2	1.034	1.131
R_1 (F)[a], wR_2 (F^2) [$I > 2\sigma(I)$]	0.0654, 0.1686	0.0687, 0.1522
Largest diff. peak and hole, e. Å^{-3}	0.226 and -0.220	0.552 and -0.675

[a] $R_1 = \Sigma||F_o|-|F_c||/\Sigma|F_o|$.[b] $wR_2 = [\Sigma w(F_o^2-F_c^2)^2/\Sigma w(F_o^2)^2]^{1/2}$ where $w = 1/[\sigma^2(F_o^2) + (aP)^2 + bP]$ where $P = (F_o^2 + F_c^2)/3$.

Table S2. Hydrogen bonds for **LD-1** [Å and deg] measured on the X-ray structure of the racemate.

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(2)-H(2N)...O(5)#1	0.84(5)	2.16(6)	2.973(4)	161(4)
C(14)-H(14)...O(3)	0.98	2.20	2.868(5)	124.5
C(10)-H(11B)...O(4)#2	0.97	2.53	3.380(6)	145.8
C(17)-H(17B)...O(5)	0.96	2.57	3.085(7)	113.8
C(18)-H(18A)...O(5)	0.96	2.41	3.000(8)	119.7

Symmetry transformations used to generate equivalent atoms:
#1 y,-x+1,z+1/4 #2 -y+1,x,z-1/4

Table S3. Relevant torsion angles in the X-ray structures of **LL-1** (enantiopure form) and **LD-1** (in the racemate).

LL-1	ϕ	ψ	ω
L-Ala	-	139.03	-173.94
L-pGlu	-69.01	-	-

LD-1	ϕ	ψ	ω
L-Ala	-	-172.12	-163.90
D-pGlu	-74.90	-	-

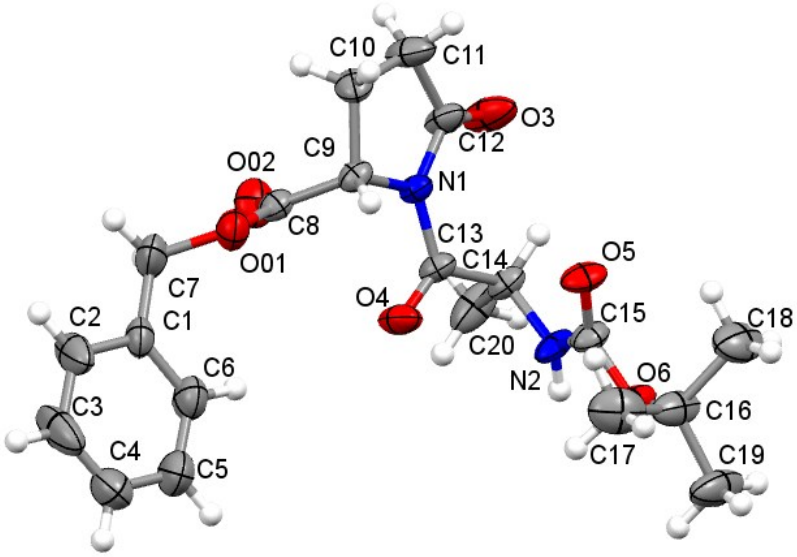


Figure S1. ORTEP drawing of **LL-1**. Thermal ellipsoids are drawn at the 50% of the probability level.

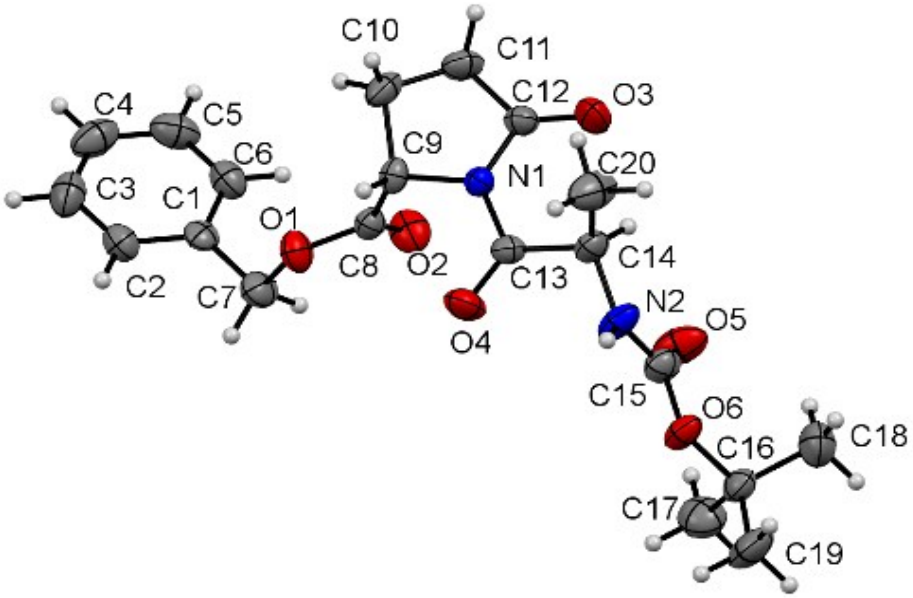


Figure S2. ORTEP drawing of **LD-1** in the racemate. Thermal ellipsoids are drawn at the 50% of the probability level.

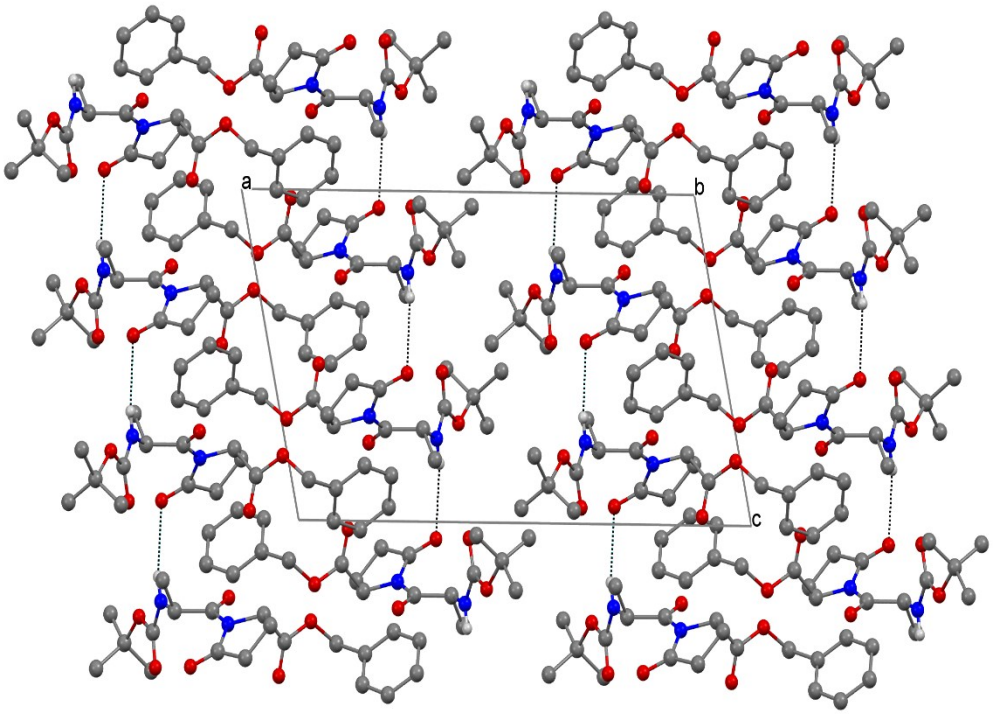


Figure S3. View down the b axis of the crystal packing of **LD-1**. Black dotted lines indicate the C=O···H-N hydrogen bond. All hydrogen bonds except the amidic ones have been omitted for sake of clarity.

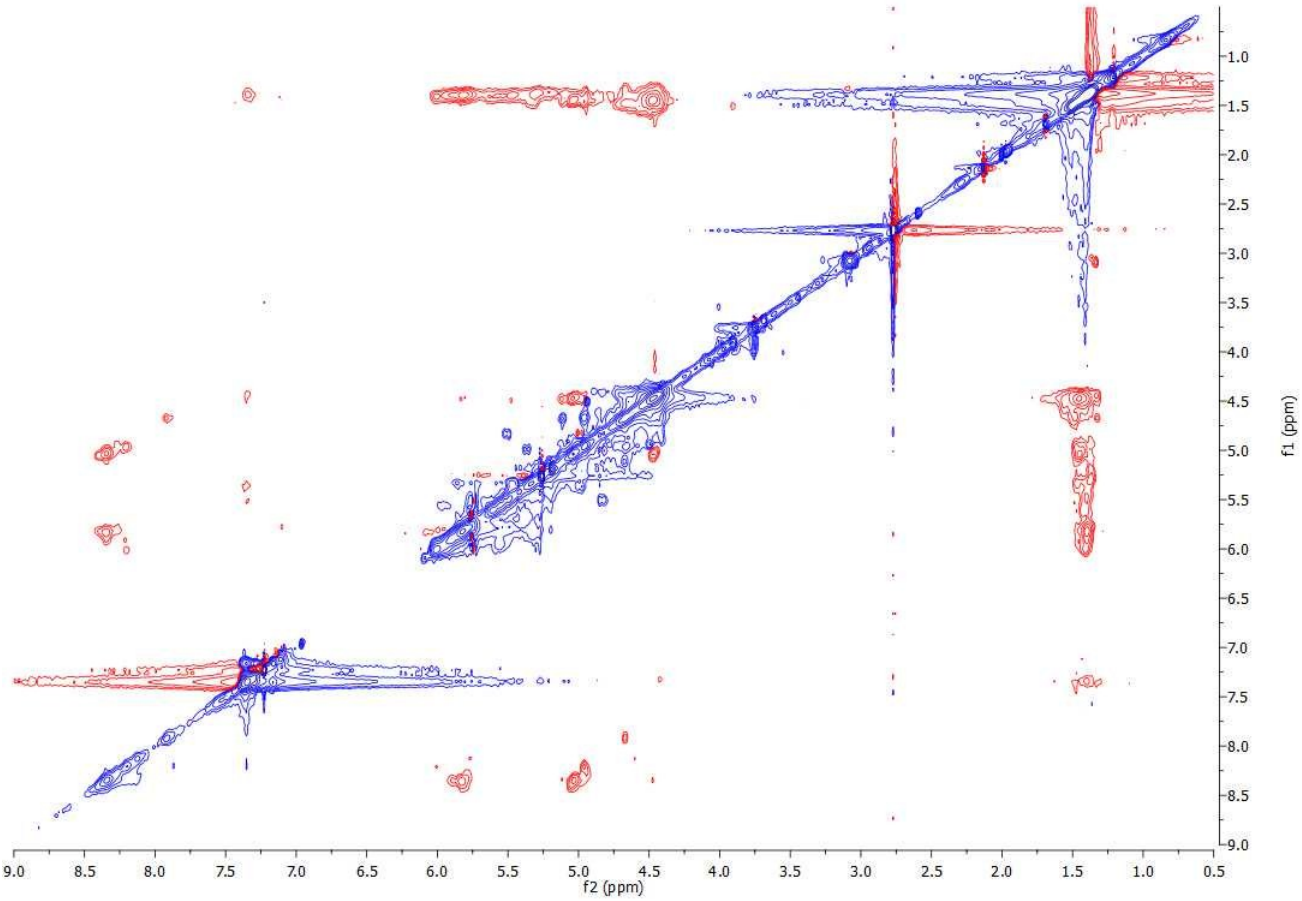


Figure S4. ROESY experiment performed on a 10^{-2} M solution of **LD-8** in $CDCl_3$.

Computational Section

Molecular dynamics simulations. Molecular dynamics simulations were carried out with the AMBER 16.0 suite of programs.[S1] The peptide is parametrized using the Amber ff14SB force field[S2] (the standard RESP procedure is carried out to assign charges to atoms by Antechamber). A linear conformation of the peptide was built and immersed in a solvent box of explicit acetonitrile molecules. Periodic boundary conditions were used. An equilibration protocol was applied, resulting in an unconstrained well-tempered NPT ensemble at target conditions. A Langevin thermostat was used to set a constant temperature at 300 K and 1 atm. Particle Mesh Ewald[S3] summation was used throughout (cut off radius of 10 Å for the direct space sum). Bonds involving H atoms were constrained using the SHAKE algorithm,[S4] and a time step of 2 fs was applied in all runs. Overall sampling time for MD production was 100 ns. Snapshot structures were saved into individual trajectory files every 5000 time steps, that is, every 10 ps of molecular dynamics, for a total of 100000 snapshots. MD simulations were carried out using pmemd. VMD was used to visualize the trajectory.[S5] Two-dimensional free-energy profiles for the peptide in explicit solvent were obtained as a function of PCA1 and PCA2. The program "ptraj"[S6] in the AMBER package was used in the PCA. The free energy values are given as normalized energy. The energy landscape of the peptide is visualized by means of free-energy functions, which are projected as contour lines onto a two-dimensional space formed by the PCA1/PCA2 axes. These coordinates are derived from a principal component analysis.[S7] The normalized free-energy change associated with the passage between two different states of a system in thermodynamic equilibrium is given by $\Delta G = -RT (\ln p1/p2)$. Here, R is the ideal gas constant, T is the absolute temperature, and pi is the probability of finding the system in state i. The two-dimensional space defined by the PCA1 and PCA2 axes has been divided into a grid and the free energy has been calculated for each bin of the grid on the basis of the previous equation. To obtain the p values, the trajectory at ambient temperature was projected onto the PCA1/PCA2 space, and p corresponds to the number of times the trajectory "visits" a given bin. Torsion angles Φ and Ψ in the **LD-8** structure (Table S4, Figure S8) are calculated as the average values during the MD simulation.

QM calculations- All calculations were run with Gaussian16 (Revision A.03. Wallingford, CT; 2016) with default grids and convergence criteria. The input structures for all calculations were obtained by DFT re-optimization, at B3LYP/6-311G(d,p) level of theory, of representative energy minima picked up from molecular dynamics simulations, as explained in the main text. Excited states calculations were run with time-dependent DFT (TDDFT) at CAM-B3LYP/def2-TZVP level of theory including a continuum solvent model (IEF-PCM) for acetonitrile, including 24 roots for **LD/LL-1**, 36 roots for **LD/LL-2**, 50 roots for **LD/LL-3**, and 64 roots for **LD/LL-4**, respectively. Frequency calculations were run at the same level B3LYP/6-311G(d,p2) used for

geometry optimizations *in vacuo*; all calculated frequencies were real for all structures considered. Geometry and frequency calculations were repeated at GD3-B3LYP/6-311G(d,p) *in vacuo*, that is, including Grimme's empirical dispersion of GD3 type.[S8] Transition density plots were generated with the program Multiwfn (v. 3.3.8),[S9] using an isovalue of 0.0004. Calculated ECD and VCD spectra were generated with the program SpecDis v.1.71.[S10]

References

S1) D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (**2016**), AMBER 2016, University of California, San Francisco.

S2) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.

S3) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

S4) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327-341.

S5) Humphrey, W.; Dalke, A.; Schulten, K. J. Mol. Graph. **1996**, *14*, 33–38.

S6) Roe, D. R.; Cheatham III, T. E. *J. Chem. Theory Comput.* **2013**, *9*, 3084-3095.

S7) Daidone, I.; Amadei, A. *WIREs Comput. Mol. Sci.* **2012**, *2*, 762-770.

S8) Grimme, S.; Antony, J.; Ehrlich S.; H. Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

S9) Lu, T.; Chen, F. *J. Comput. Chem.* **2012**, *33*, 580–592.

S10) Bruhn, T.; Schaumlöffel, A.; Hemberger, Y.; Pescitelli, G. SpecDis Version 1.71, Berlin, Germany, **2017**, Https://Specdis-Software.Jimdo.Com/. *SpecDis version 1.71, Berlin, Ger. 2017, https//specdis-software.jimdo.com/*.

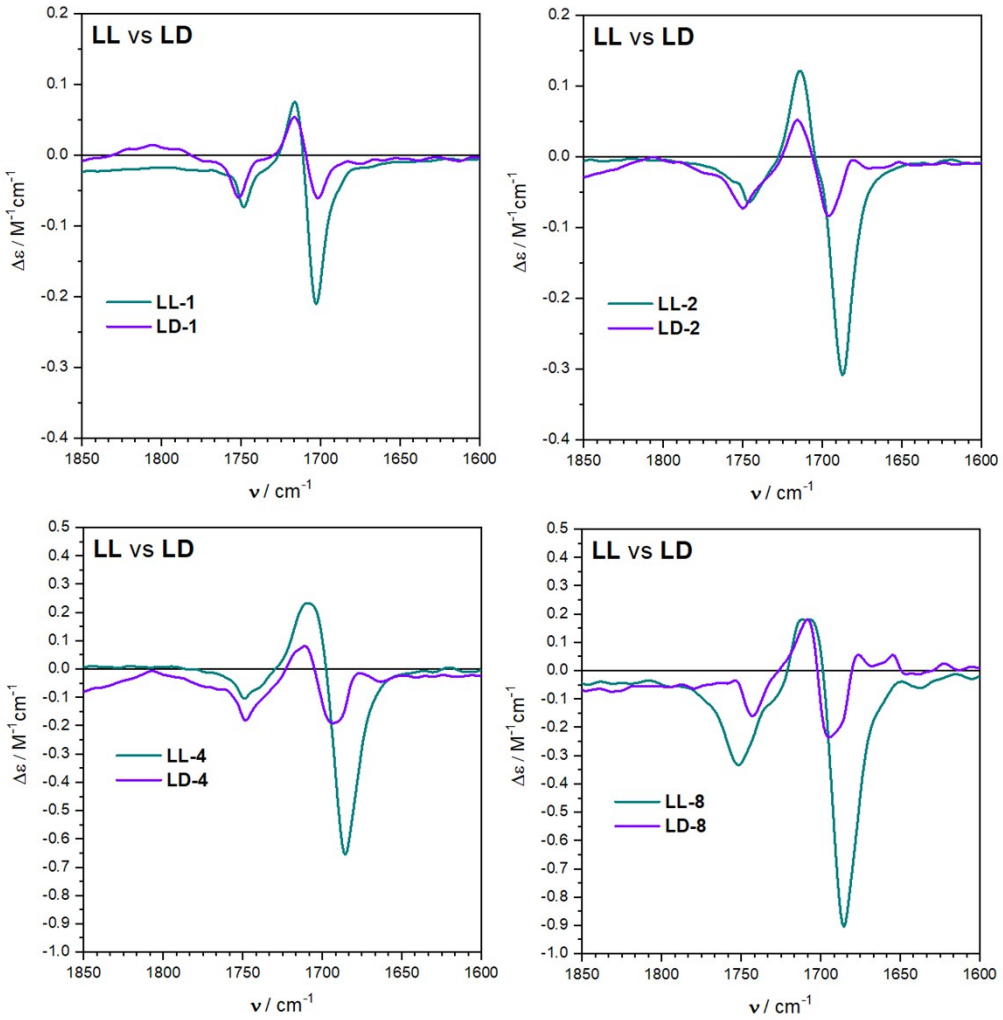


Figure S5. Comparison between the VCD spectra of the **LD**- and **LL**-series recorded for the same oligomer length.

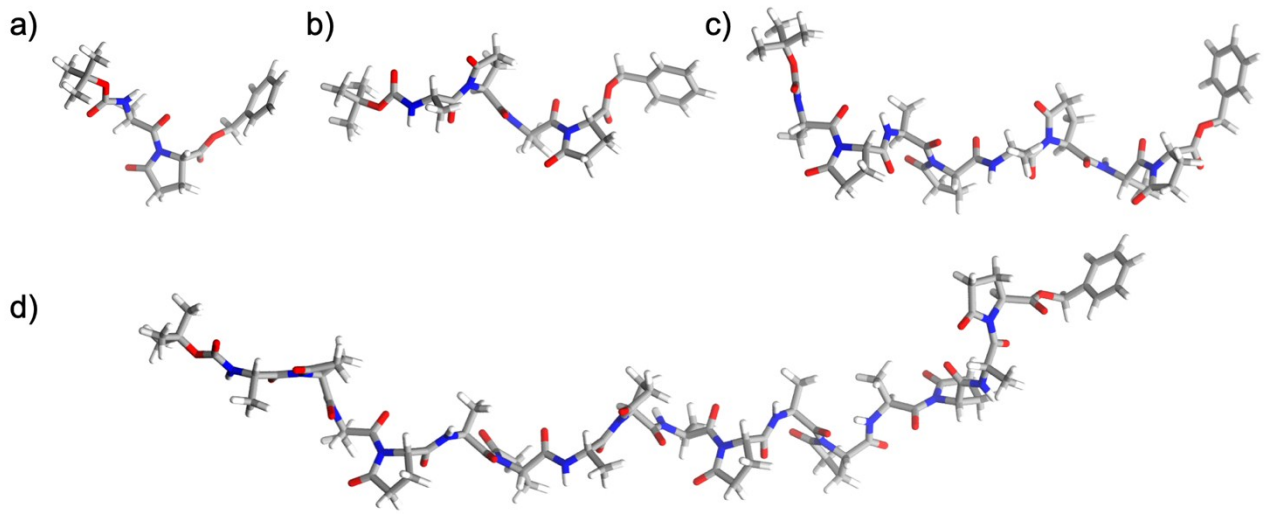


Figure S6. Representative structure of a) **LL-1**, b) **LL-2**, c) **LL-4**, d) **LL-8,** corresponding to the minima of the free energy landscape in Figure 8a-d

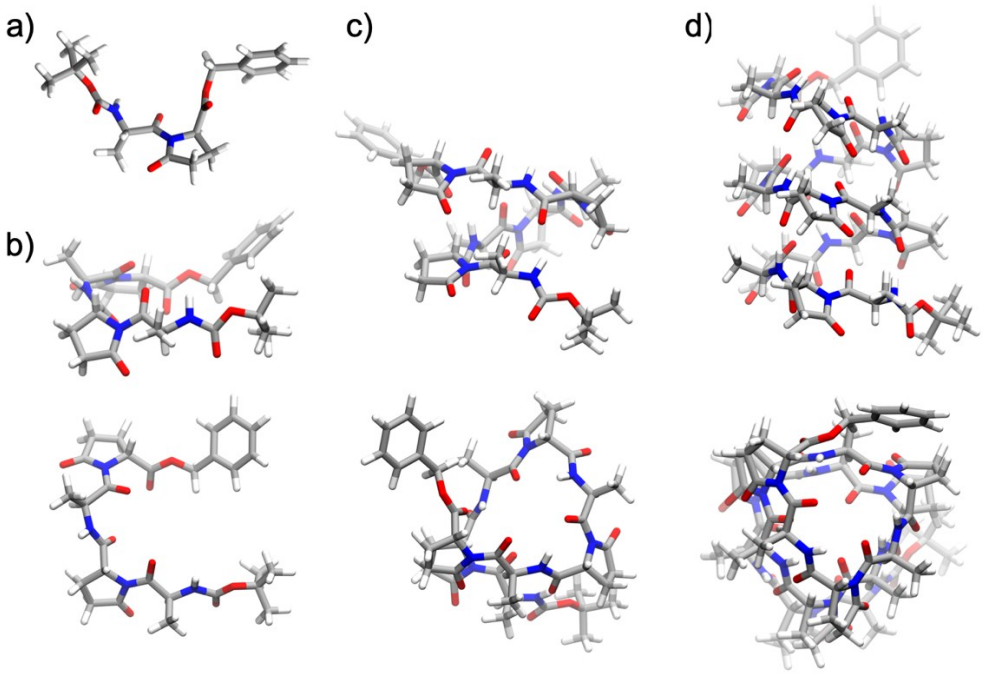


Figure S7. Representative structure of a) **LD-1**, b) **LD-2**, c) **LD-4**, d) **LD-8,** corresponding to the minima of the free energy landscape in Figure 88-h. In b) , c) and d) both front view and top view perspectives are showed.

Table S4. Torsion angles Φ and Ψ in the **LD-8** structure obtained by MD simulations.

Residue	1 Ala (N-term)	2 pGlu	3 Ala	4 pGlu	5 Ala	6 pGlu	7 Ala	8 pGlu
Phi	-	71.22	-119.31	73.62	-136.08	77.17	-143.22	70.83
Psi	152.39	-134.51	155.27	-115.07	156.21	-111.55	158.27	-116.42

Residue	9 Ala	10 pGlu	11 Ala	12 pGlu	13 Ala	14 pGlu	15 Ala	16 pGlu (C-term)
Phi	-140.6	75.28	-147.75	76.18	-149.74	76.43	-147.66	77.32
Psi	159.31	-111.61	158.315	-111.27	156.99	-115.86	154.66	-

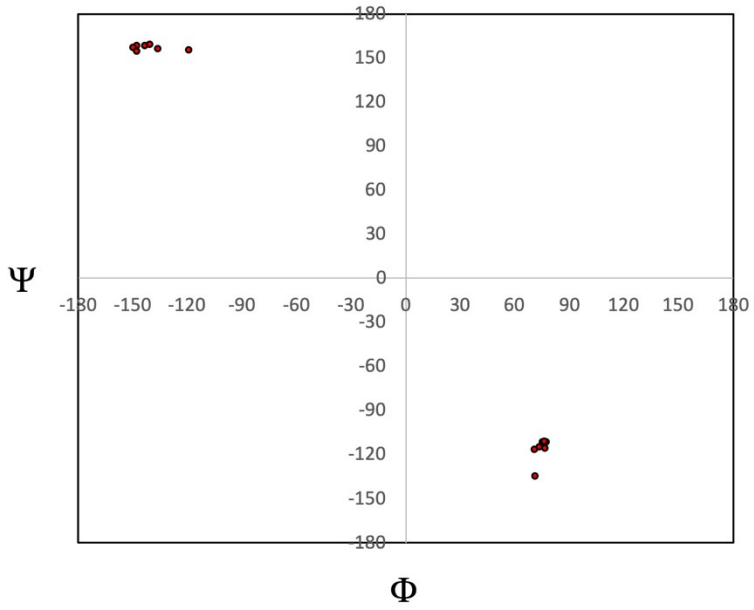


Figure S8. Ramachandran plot of **LD-8** structure obtained by MD simulations.

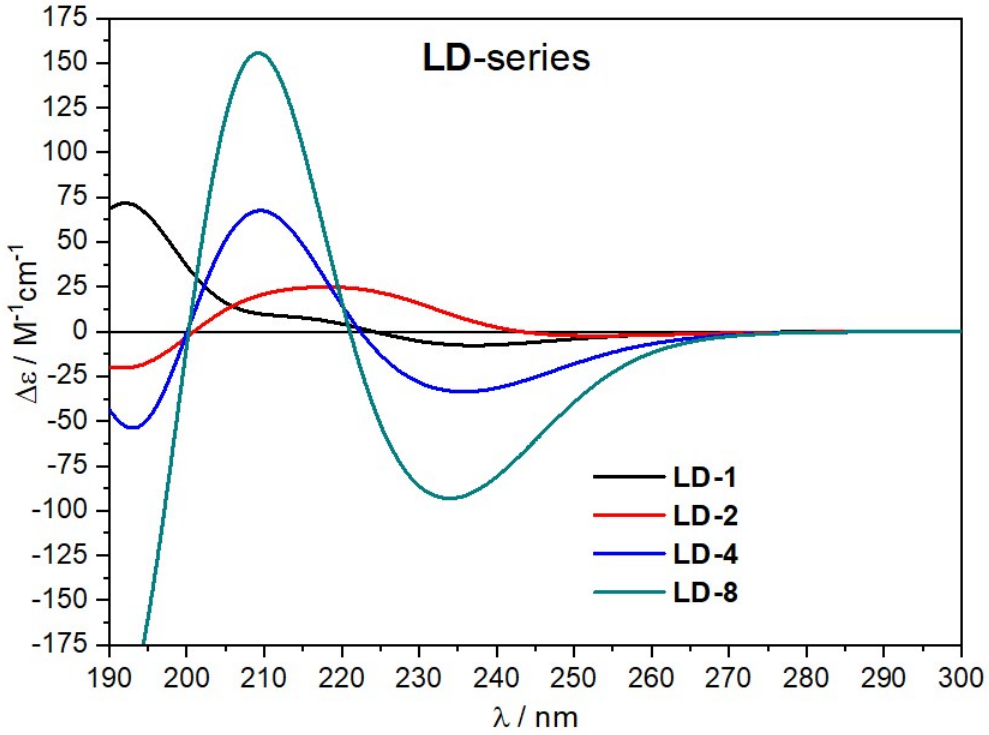


Figure S9A. Calculated ECD spectra for the **LD** series at the CAM-B3LYP/def2-TZVP//B3LYP/6-311G(d,p) level of theory with PCM for acetonitrile, using the global energy minima from MD simulations.

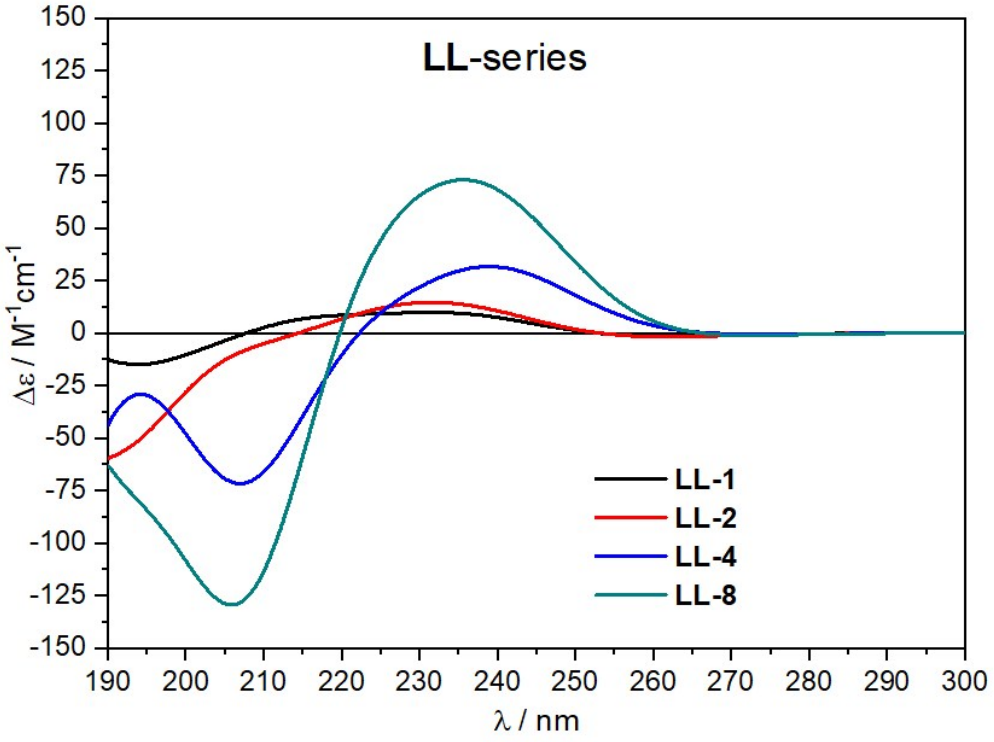


Figure S9B. Calculated ECD spectra for the **LL** series at the CAM-B3LYP/def2-TZVP//B3LYP/6-311G(d,p) level of theory with PCM for acetonitrile, using the global energy minima from MD simulations.

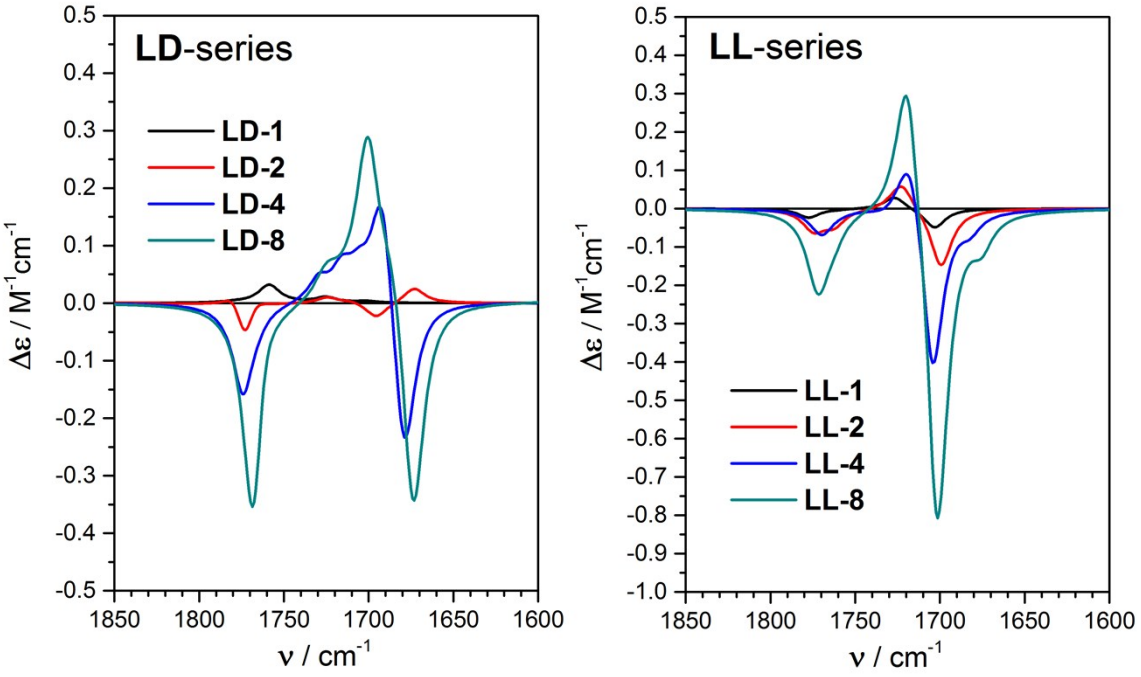


Figure S10. Calculated VCD spectra for the **LD** and **LL** series. Calculations run at the B3LYP/6-311G(d,p) level of theory *in vacuo*; frequency scaling factor 0.975; spectra plotted as sums of Lorentzians with 8 cm^{-1} full width at half height.

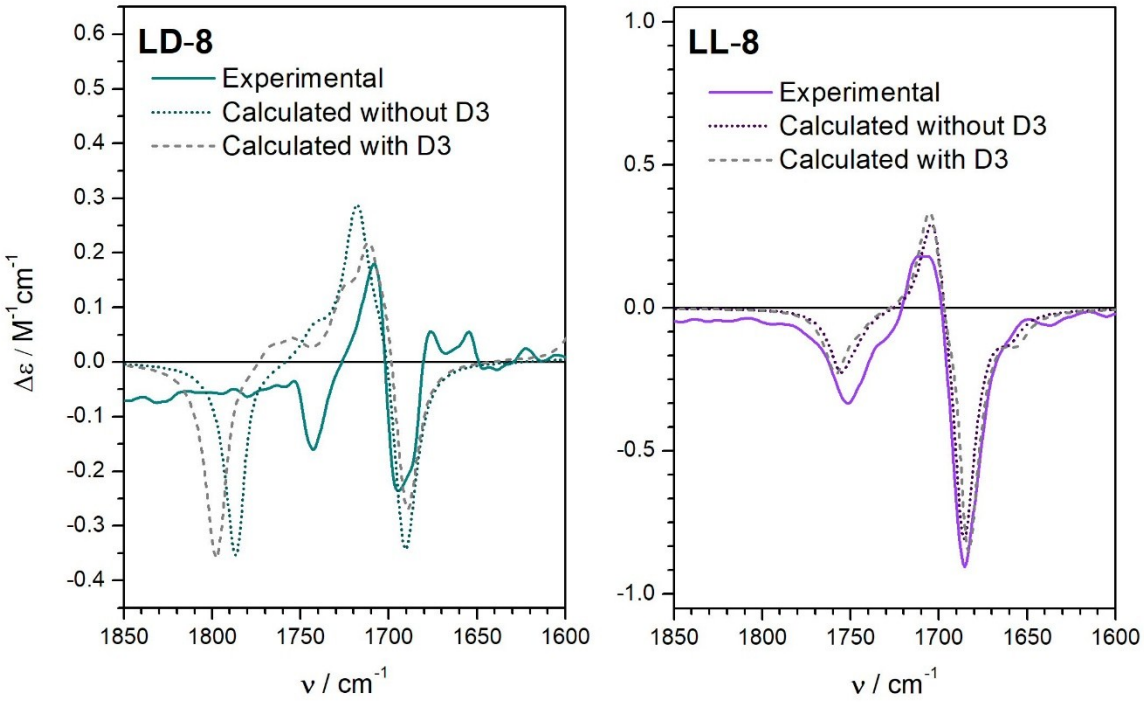


Figure S11. Comparison between experimental and B3LYP/6-311G(d,p) calculated VCD spectra of **LL-8** and **LD-8** without (color dotted lines) or with (grey dashed lines) Grimme's GD3 dispersion correction. Frequency scaling factors: B3LYP for **LD-8**, 0.966; D3-B3LYP for **LD-8,** 0.99; B3LYP and D3-B3LYP for **LL-8,** 0.985. Spectra plotted as sums of Lorentzians with 8 cm^{-1} full width at half height.