

Supplementary Information for

Bottom-up design of peptide shapes in water using oligomers of *N*-methyl-L/D-alanine

Jumpei Morimoto^{†*a}, Marin Yokomine^{†a}, Yota Shiratori^{†b}, Takumi Ueda^{cd}, Takayuki Nakamuro^e, Kiyofumi Takaba^f, Saori Maki-Yonekura^f, Koji Umezawa^{gh}, Koichiro Miyanishiⁱ, Yasuhiro Fukuda^a, Takumu Watanabe^a, Mayuko Suga^a, Ayumi Inayoshi^a, Takuya Yoshida^c, Wataru Mizukamiⁱ, Koh Takeuchi^d, Koji Yonekura^{fjk}, Eiichi Nakamura^e, Shinsuke Sando^{*ab}

^a Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

^b Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

^c Graduate School of Pharmaceutical Sciences, Osaka University, 1-6, Yamadaoka, Suita, Osaka, Japan.

^d Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

^e Department of Chemistry, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

^f RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan.

^g Department of Biomedical Engineering, Graduate School of Science and Technology, Shinshu University, 8304 Minami-minowa, Kami-ina, Nagano 399-4598, Japan.

^h Institute for Biomedical Sciences, Interdisciplinary Cluster for Cutting Edge Research, Shinshu University, Matsumoto, Nagano 390-8621, Japan.

ⁱ Center for Quantum Information and Quantum Biology, Osaka University, 1-2 Machikaneyama, Toyonaka, Osaka 560-0043, Japan

^j Advanced Electron Microscope Development Unit, RIKEN-JEOL Collaboration Center, RIKEN Baton Zone Program, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan.

^k Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai, 2-1-1 Katahira, Aoba, Miyagi 980-8577, Japan.

[†] These authors contributed equally.

* Corresponding authors

Table of Contents

Abbreviations	4
Materials and Methods	4
General remarks.....	4
Synthesis of acetyl- <i>N</i> -methyl-L-alanine dimethylamide (1).....	5
Synthesis of acetyl- <i>N</i> -methyl-L-alanine piperazine, <i>N</i> -Boc (2).....	6
Synthesis of acetyl- <i>N</i> -methyl-L-alanyl- <i>N</i> -methyl-L-alanine dimethylamide (4).....	7
Synthesis of acetyl- <i>N</i> -methyl-D-alanyl- <i>N</i> -methyl-L-alanine dimethylamide (5).....	8
Synthesis of acetyl- <i>N</i> -methyl-L-alanyl- <i>N</i> -methyl-L-alanine piperazine, Boc (6).....	8
Synthesis of acetyl- <i>N</i> -methyl-D-alanyl- <i>N</i> -methyl-L-alanine piperazine, Boc (7).....	9
Synthesis of LLDDLL hexamer (8) for NMR analysis.	9
Synthesis of LLDDLL hexamer (9) for the crystallographic analysis.....	10
Synthesis of LLLLDL hexamer (10) for NMR analysis.....	10
Synthesis of LLLLLL and DDDDDD hexamers (S5 and S6) for CD spectroscopic analysis.	10
Synthesis of LLDDLLDDLL decamer for SMART-EM analysis.	11
Synthesis of LLLLLLLLLL decamer for SMART-EM analysis.....	12
Synthesis of functionalized LLDDLL hexamer (11).	12
Generation of Ramachandran-type plots of acetyl- <i>N</i> -methyl-L/D-alanine dimethylamide.	12
DFT-optimized structure of acetyl- <i>N</i> -methyl-L-alanine dimethylamide.....	13
DFT-optimized structures of oligo- <i>N</i> -methylalanines.....	13
Conformational search of dimers and hexamers by multicanonical MD simulations.....	13
NMR measurements.	14
Crystallization of an NMA monomer and dimers.	15
X-ray crystallography.	15
XFEL analysis of LLDDLL crystals.	15
Circular dichroism studies.	16
SMART-EM.....	16
Supplementary Figures and Tables.....	18
Fig. S1 ¹ H-NMR spectrum of L-NMA monomer.....	18
Fig. S2 HSQC spectrum of L-NMA monomer.....	19
Fig. S3 HMBC spectrum of L-NMA monomer.....	20
Fig. S4 EASY-ROESY spectrum of L-NMA monomer.	21
Table S1 NOE summary of acetyl- <i>N</i> -methyl-L-alanine dimethylamide.....	22
Table S2 Chemical shift table of acetyl- <i>N</i> -methyl-L-alanine dimethylamide.....	23
Table S3 Representative conformers of the 10 clusters from McMD simulations of LLDDLL and LLLLDL ..	24
Fig. S5 ¹ H-NMR spectrum of LL dimer.....	25
Fig. S6 HSQC spectrum of LL dimer.....	26
Fig. S7 HMBC spectrum of LL dimer.....	27
Fig. S8 EASY-ROESY spectrum of LL dimer.....	28
Table S4 NOE summary of LL dimer.....	29

Table S5 Chemical shift table of LL dimer.....	30
Fig. S9 ¹ H-NMR spectrum of DL dimer.	31
Fig. S10 HSQC spectrum of DL dimer.	32
Fig. S11 HMBC spectrum of DL dimer.	33
Fig. S12 ROESY spectrum of DL dimer.	34
Table S6 NOE summary of DL dimer	35
Table S7 Chemical shift table of DL dimer	36
Fig. S13 A list of possible peptide shapes realized by the sequences of LLDL DL, LDLDLL, and LDLDLD. .	37
Table S8 Representative conformers of the 10 clusters from McMD simulations of LLDDLL and LLLLDL ..	38
Fig. S14 HSQC spectrum of LLDDLL hexamer.....	39
Fig. S15 HMBC spectrum of LLDDLL hexamer.....	40
Fig. S16 ROESY spectrum of LLDDLL hexamer.	41
Table S9 NOE summary of LLDDLL hexamer.....	42
Table S10 Chemical shift table of LLDDLL hexamer.....	43
Table S11 Summary of XFEL analysis of LLDDLL	44
Fig. S17 HSQC spectrum of LLLLDL hexamer.	45
Fig. S18 HMBC spectrum of LLLLDL hexamer.	46
Fig. S19 ROESY spectrum of LLLLDL hexamer.....	47
Table S12 NOE summaries of LLLLDL hexamer.....	48
Table S13 Chemical shift table of LLLLDL hexamer	49
Fig. S20 Circular dichroism (CD) spectra of oligo-NMAs.	50
Fig. S21 The overlays of the TEM images of LLDDLLDDLL and molecular models from MD simulations...	51
Fig. S22 The overlays of the TEM images of LLLLLLLLLL and molecular models from MD simulations.....	51
Fig. S23 HSQC spectrum of functionalized LLDDLL hexamer.	52
Fig. S24 HMBC spectrum of functionalized LLDDLL hexamer.....	53
Fig. S25 ROESY spectrum of functionalized LLDDLL hexamer.	54
Table S14 NOE summaries of functionalized LLDDLL hexamer	55
Table S15 Chemical shift table of functionalized LLDDLL hexamer.....	56
Fig. S26 UV chromatograms of the purified monomer and oligomers of <i>N</i> -methylalanine.	59
Video S1 Video imaging of LLDDLLDDLL captured by OneView-IS.....	60
Video S2 Video imaging of LLLLLLLLLL captured by OneView-IS.....	60
References	61

Abbreviations

Boc, *tert*-butoxycarbonyl; Cbz, benzyloxycarbonyl; DMT-MM, 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride; DMF, *N,N*-dimethylformamide; DCM, dichloromethane; DIPEA, *N,N*-diisopropylethylamine; THF, tetrahydrofuran; TMP, 2,4,6-trimethylpyridine; HFIP, 1,1,1,3,3,3-hexafluoro-2-propanol; COMU, ethyl 2-cyano-2-((dimethyliminio)(morpholino)methoxyimino)acetate hexafluorophosphate; HATU, 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate; HOAt, 3*H*-1,2,3-triazolo[4,5-*b*]pyridin-3-ol; TFA, trifluoroacetic acid; HOBt, 1*H*-benzo[*d*][1,2,3]triazol-1-ol; EDC·HCl, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride; NMP, *N*-methyl-2-pyrrolidone; DIC, *N,N'*-diisopropylcarbodiimide; oxyma, ethyl cyano(hydroxyimino)acetate.

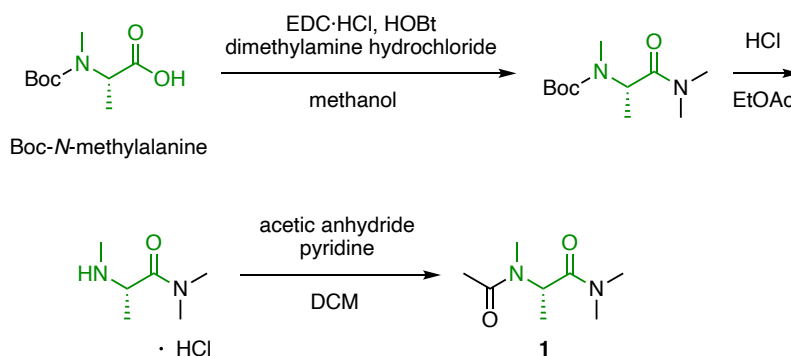
Materials and Methods

General remarks. Chemicals and solvents used in this study were purchased from commercial suppliers and used without further purification. Flash column chromatography was performed with an Isolera Spektra instrument equipped with a Biotage® Sfär Silica HC D 10 g cartridge. Preparative high performance liquid chromatography (HPLC) was performed on a Prominence HPLC system (Shimadzu) with a 5C18-AR-II column (Nacalai tesque, 10 mm I.D.×150 mm, 34350-41). Liquid chromatography-mass spectrometry (LC-MS) and ultra performance liquid chromatography (UPLC) was performed on an ACQUITY UPLC HClass/SQD2 (Waters) using ACQUITY UPLC BEH C₁₈ 1.7 mm Column (Waters, 2.1 mm × 50 mm). HRMS data were obtained using microTOF II (Bruker Daltonics). All quantum chemical calculations were carried out with the Gaussian16 package. A black powder of aggregate of conical carbon nanotubes (CNTs) was purchased from NEC Co. (carbon nanohorn, Lot No. 181-3-2; assay: 90%) and used for the preparation of water-miscible amino-CNT as previously published¹. Distilled water was further purified with Millipore Milli-Q (Milli-Q Reference, Merck). Multiplicities of ¹H NMR signals are reported as follows: s, singlet; d, doublet; q, quartet; m, multiplet.

Principal moment inertia (PMI) analysis. The PMI analysis was conducted according to the literature². Main chain atoms (N, C_α, and C=O) of each compound were extracted and used for the analysis. For miniature proteins, six consecutive amino acid residues were comprehensively extracted from each protein and analyzed. The model structures of NMA hexamers were generated by the following procedures. First, the most stable conformer of L-NMA residue ($\phi = -135^\circ$, $\psi = 75^\circ$) and D-NMA residue ($\phi = 135^\circ$, $\psi = -75^\circ$) on the Ramachandran-type plot was connected to generate each hexamer structure and each hexamer was capped with acetyl group at N-terminus and dimethylamide group at C-terminus. The generated conformers were energy minimized for 20 steps with the constraints of backbone dihedral angles (ϕ , ψ , and ω) by molecular mechanics calculations to relax the structures. The geometry of the conformers was optimized by DFT calculations at B3LYP/6-

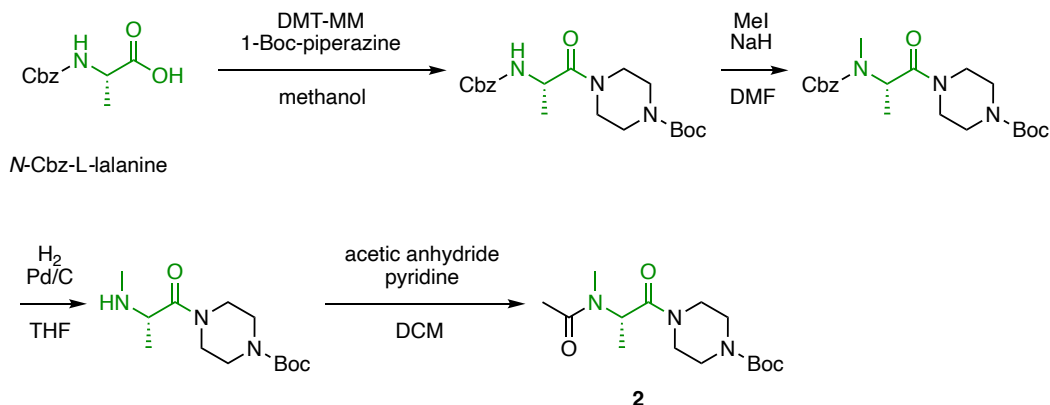
311G(d) level. The DFT calculations were conducted with the solvation effects of water using the CPCM model. Considering that the geometry optimization by DFT calculations did not significantly change the monomer conformation in the hexamers (the average differences of ϕ and ψ values between the conformations before and after the geometry optimization were less than 30°), the geometry optimization step was omitted for generating the model structures of NMA decamers.

Synthesis of acetyl-*N*-methy-L-alanine dimethylamide (**1**).



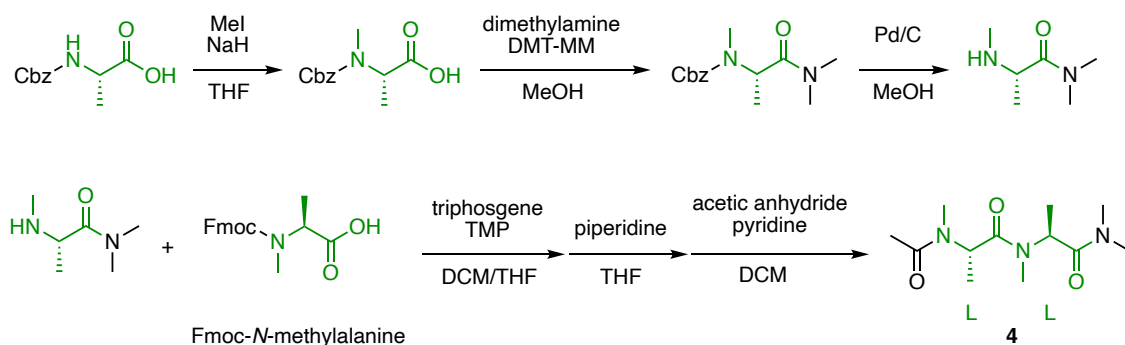
To a 25-mL round-bottom flask containing a magnetic stirring bar were added *N*-Boc-*N*-methyl-L-alanine (0.30 g, 1.5 mmol), and HOBT (0.22 g, 1.6 mmol, 1.1 equiv.) and 2.95 mL MeCN. To the solution was added DIPEA (0.97 mL, 5.6 mmol, 3.7 equiv.), dimethylamine hydrochloride (0.18 g, 2.2 mmol, 1.5 equiv.), and EDC·HCl (0.31 g, 1.6 mmol, 1.1 equiv.). The solution was stirred at room temperature overnight. After evaporation of the solvent, the residue was dissolved in water and DCM, and the product was extracted with DCM twice. The organic phase was washed with sat. Na_2CO_3 aq. twice and brine once. The collected organic phase was dried over Na_2SO_4 . The solvent was removed by evaporation. The product was used for the next reaction without further purification. In a 100-mL round-bottom flask containing a magnetic stirring bar, the product was dissolved in 2 mL of ethyl acetate and cooled on ice. 2 mL of 4 M HCl in ethyl acetate was added and the solution was allowed to reach room temperature. The solution was stirred for 30 min. The solvent was removed under reduced pressure. The product was dissolved in 4.6 mL of DCM and cooled on ice. Pyridine (0.74 mL, 9.1 mmol) and acetic anhydride (0.43 mL, 4.6 mmol) were added and the solution was allowed to reach room temperature. The reaction mixture was stirred for 30 min at room temperature. The solvent was removed under reduced pressure and the residue was subjected to silica gel column chromatography (DCM/MeOH) to give **1** (0.11 g, 0.64 mmol, 43%). $^1\text{H}/^{13}\text{C}$ NMR (D_2O): Chemical shifts are summarized in Table S3. HRMS (ESI-TOF MS) m/z : $[\text{M}+\text{Na}]^+$ Calcd for $\text{C}_8\text{H}_{16}\text{N}_2\text{NaO}_2^+$ 195.1104; Found 195.1109.

Synthesis of acetyl-*N*-methyl-L-alanine piperazine, *N*-Boc (2).



To a 50-mL round-bottom flask containing a magnetic stirring bar were added *N*-Cbz-L-alanine (1.1 g, 5.0 mmol), and 1-Boc-piperazine (1.4 g, 7.5 mmol, 1.5 equiv.) and 5.0 mL MeOH. To the solution was added DMT-MM (1.8 g, 5.5 mmol, 1.1 equiv.), and the solution was stirred at room temperature overnight. After evaporation of the solvent, the residue was dissolved in water and DCM, and the product was extracted with DCM three times. The collected organic phase was dried over Na₂SO₄. The solvent was removed by evaporation and the crude product was used for the next reaction without further purification. To a 100-mL round-bottom flask containing a magnetic stirring bar were added the product, iodomethane (0.63 mL, 10 mmol, 2.0 equiv.), and DMF (35 mL). After the solution was cooled on ice, sodium hydride (60% in oil, 0.40 g [0.24 g], 10 mmol, 2.0 equiv.) was added portion wise under stirring. The solution was allowed to reach room temperature and stirred for 4 h. The reaction mixture was diluted with ethyl acetate and washed with NH₄Cl aq. and brine. The organic phase was dried over Na₂SO₄. The solvent was removed under reduced pressure and the crude product was used for the next reaction without further purification. The product and palladium 10% on carbon (0.25 g), were suspended in MeOH under a nitrogen atmosphere in a round-bottom flask. The flask was charged with hydrogen gas while cooling at -78°C and the reaction mixture was stirred at room temperature overnight. The reaction mixture was filtered through celite using ethyl acetate as eluent. The solvent was removed under reduced pressure. As far as TLC confirmed, the product was almost pure, so the crude product was used for the next reaction without further purification. To a 100-mL round-bottom flask containing a magnetic stirring bar were added the product and DCM (20 mL). Acetic anhydride (2.4 mL, 25 mmol, 5.0 equiv.) and pyridine (4.0 mL, 50 mmol, 10 equiv.) were added to the solution. The reaction mixture was stirred for 30 min at room temperature and the solvent was removed under reduced pressure. The residue was purified by silica gel column chromatography (DCM/MeOH) to give **2** (0.69 g, 2.2 mmol, 44%). ¹H NMR (CDCl₃, 400 MHz): δ 5.54 (q, J = 6.8 Hz, 1H), 3.23–3.69 (m, 8H), 2.90 (s, 3H), 2.10 (s, 3H), 1.46 (s, 9H), 1.27 (d, J = 6.8 Hz, 3H). ¹³C NMR (CDCl₃, 100 MHz): δ 169.7, 169.3, 154.0, 79.8, 47.7, 44.7, 41.6, 30.5, 27.9, 21.5, 14.0. HRMS (ESI-TOF MS) m/z : [M+Na]⁺ Calcd for C₁₅H₂₇N₃NaO₄⁺ 336.1894; Found 336.1910.

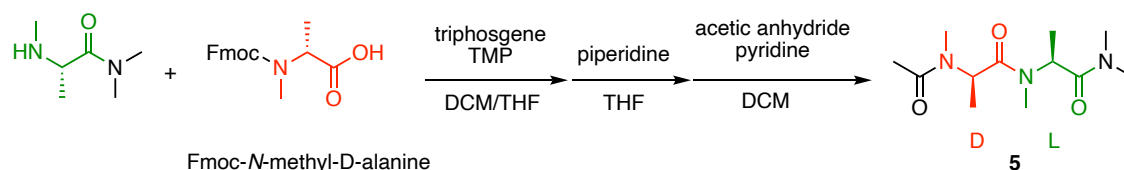
Synthesis of acetyl-*N*-methyl-L-alanyl-*N*-methyl-L-alanine dimethylamide (4).



N-Cbz-L-alanine (1.1 g, 5.0 mmol) and THF (25 mL) were added to a 100-mL round-bottom flask containing a magnetic stirring bar. Iodomethane (2.5 mL, 40 mmol, 8.0 equiv.) was added to the solution. After cooling on ice, sodium hydride (60% in oil, 0.60 g [0.36 g], 15 mmol, 3.0 equiv.) was added portion wise under stirring. The resulting mixture was stirred at 60 °C overnight. The reaction mixture was diluted with water and ethyl acetate. The product was extracted with water twice and the aqueous layer was acidified with 2 M HCl aq. to pH 3–4. The product was extracted with ethyl acetate twice and the collected organic layer was dried over Na₂SO₄. After removing the solvent under reduced pressure, the product was used for the next reaction without further purification. Cbz-*N*-methyl-L-alanine was dissolved in THF (5.0 mL). 2 M dimethylamine in THF (3.8 mL, 7.5 mmol, 1.5 equiv.) was added to the solution. DMT-MM (1.8 g, 5.5 mmol, 1.1 equiv.) was added to the solution and the reaction mixture was stirred at room temperature overnight. The solvent was removed under reduced pressure and the residue was diluted with water and DCM. The product was extracted with DCM twice and the collected organic phase was dried over Na₂SO₄. After removing the solvent under reduced pressure, the product was purified on silica gel chromatography. Palladium on carbon (150 mg) suspended in 10 mL of MeOH was applied to Cbz-*N*-methyl-L-alanine dimethylamide and the flask was charged with hydrogen gas. The reaction mixture was stirred at room temperature overnight. After the palladium on carbon was removed through celite, the solvent was removed under reduced pressure to give *N*-methyl-L-alanine dimethylamide (0.28 g, 44% in three steps). Fmoc-*N*-methyl-L-alanine (0.20 g, 0.60 mmol, 1.2 equiv.) and triphosgene (60 mg, 0.20 mmol, 0.4 equiv.) were dissolved in 2.5 mL of THF. TMP (0.20 mL, 1.5 mmol, 3.0 equiv.) was added. *N*-Methyl-L-alanine dimethylamide (65 mg, 0.50 mmol, 1.0 equiv.) dissolved in 2.5 mL DCM was added to the THF solution, and the reaction mixture was stirred for 2 h at room temperature. Water was added to quench the reaction and the product was extracted using ethyl acetate three times. The collected organic phase was dried over Na₂SO₄. After removing the solvent under reduced pressure, the product was dissolved in 5.0 mL of THF and piperidine (0.44 mL, 4.5 mmol, 9.0 equiv.) was added. The reaction mixture was stirred for 1 h at room temperature. After removing the solvent under reduced pressure, the residue was purified by silica gel column chromatography. The yield was 79 mg (73%). The purified product was dissolved in 2.0 mL of DCM, and acetic anhydride (0.34 mL, 3.7 mmol, and 10 equiv.) and pyridine (0.15 mL, 1.8 mmol, 5.0 equiv.) were added to the reaction mixture. The reaction mixture was stirred for 3 h at room temperature and the solvent was removed under reduced pressure. The residue was purified by silica gel column chromatography to give **4** (69

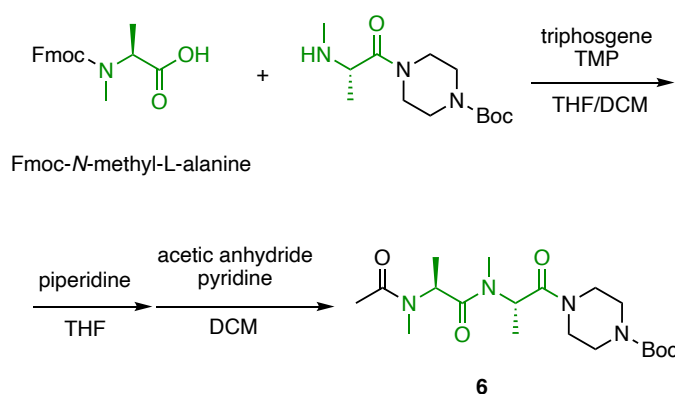
mg, 0.27 mmol, 75%). $^1\text{H}/^{13}\text{C}$ NMR (D_2O): Chemical shifts are summarized in Table S6. HRMS (ESI-TOF MS) m/z : $[\text{M}+\text{Na}]^+$ Calcd for $\text{C}_{12}\text{H}_{23}\text{N}_3\text{NaO}_3^+$ 280.1632; Found 280.1634.

Synthesis of acetyl-*N*-methyl-D-alanyl-*N*-methyl-L-alanine dimethylamide (5).



Fmoc-*N*-methyl-D-alanine (0.20 g, 0.60 mmol, 1.2 equiv.) and triphosgene (60 mg, 0.20 mmol, 0.4 equiv.) were dissolved in 2.5 mL of THF. TMP (0.20 mL, 1.5 mmol, 3.0 equiv.) was added to the solution under stirring. *N*-Methyl-L-alanine dimethylamide (65 mg, 0.50 mmol, 1.0 equiv.) dissolved in 2.5 mL DCM was added, and the reaction mixture was stirred for 2 h at room temperature. Water was added to quench the reaction and the product was extracted using ethyl acetate three times. The collected organic phase was dried over Na_2SO_4 . After removing the solvent under reduced pressure, the product was dissolved in 5.0 mL of THF and piperidine (0.44 mL, 4.5 mmol, 9.0 equiv.) was added. The reaction mixture was stirred for 1 h at room temperature. After removing the solvent under reduced pressure, the residue was purified by silica gel column chromatography. The yield was 81 mg (75%). The purified product was dissolved in 2.0 mL of DCM. Acetic anhydride (0.34 mL, 3.7 mmol, 10 equiv.) and pyridine (0.15 mL, 1.8 mmol, 5.0 equiv.) were added to the DCM solution. The reaction mixture was stirred for 3 h at room temperature and the solvent was removed under reduced pressure. The residue was purified by silica gel column chromatography to give **5** (75 mg, 0.29 mmol, 81%). $^1\text{H}/^{13}\text{C}$ NMR (D_2O): Chemical shifts are summarized in Table S8. HRMS (ESI-TOF MS) m/z : $[\text{M}+\text{Na}]^+$ Calcd for $\text{C}_{12}\text{H}_{23}\text{N}_3\text{NaO}_3^+$ 280.1632; Found 280.1631.

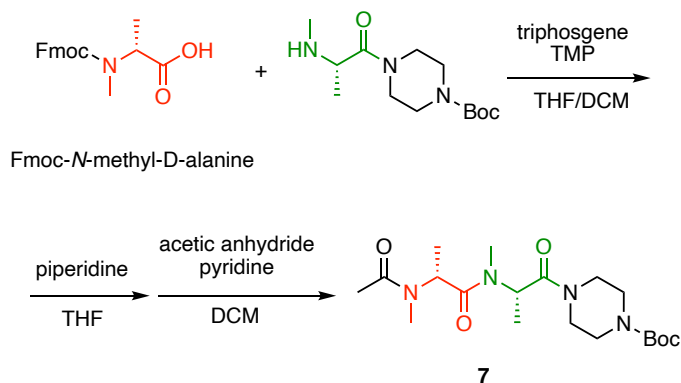
Synthesis of acetyl-*N*-methyl-L-alanyl-*N*-methyl-L-alanine piperazine, Boc (6).



Fmoc-*N*-methyl-L-alanine (0.38 g, 1.2 mmol, 1.2 equiv.) and triphosgene were dissolved in 2.5 mL of THF. To the solution was added TMP (0.40 mL, 3.0 mmol, 3.0 equiv.). *N*-Methyl-L-alanine piperazine, *N*-Boc (0.28 g, 1.0 mmol, 1.0 equiv.) dissolved in DCM (0.5 mL) and THF (2.0 mL) was added, and the reaction mixture was stirred for 3 h at room temperature. NH_4Cl aq. was added to quench the reaction and the product was extracted using ethyl acetate three times. The collected organic phase was dried over Na_2SO_4 . After removing the solvent under reduced pressure, the product

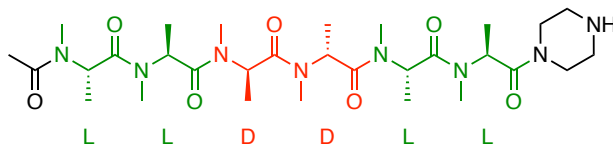
was dissolved in 5.0 mL of THF and piperidine (0.89 mL, 9.0 mmol, 9.0 equiv.) was added. The reaction mixture was stirred for 1 h at room temperature. After removing the solvent under reduced pressure, the residue was purified by silica gel column chromatography (0.21 g, 0.59 mmol, 59%). The product was dissolved in 3.0 mL of DCM, and acetic anhydride (0.56 mL, 5.9 mmol, 10 equiv.) and pyridine (0.24 mL, 3.0 mmol, 5.0 equiv.) were added to the reaction mixture. The reaction mixture was stirred for 30 min and the solvent was removed under reduced pressure. The residue was purified by silica gel column chromatography to give **6** (0.22 g, 0.55 mmol, 93%). ¹H NMR (CDCl₃, 400 MHz): δ 5.45–5.54 (m, 2H), 3.21–3.69 (m, 8H), 2.93 (s, 3H), 2.89 (s, 3H), 2.11 (s, 3H), 1.46 (s, 9H), 1.27 (d, *J* = 6.8 Hz, 3H), 1.26 (d, *J* = 7.2 Hz, 3H). ¹³C NMR (CDCl₃, 100 MHz): δ 171.0, 170.4, 169.2, 154.4, 80.3, 48.9, 48.7, 44.9, 41.9, 30.9, 29.7, 28.2, 21.8, 14.5, 14.3. HRMS (ESI-TOF MS) *m/z*: [M+Na]⁺ Calcd for C₁₉H₃₄N₄NaO₅⁺ 421.2421; Found 421.2441.

Synthesis of acetyl-*N*-methyl-D-alanyl-*N*-methyl-L-alanine piperazine, Boc (**7**).



7 was synthesized in the same manner as **6**. Fmoc-*N*-methyl-D-alanine was used instead of Fmoc-*N*-methyl-L-alanine. Starting from 1.0 mmol of *N*-methyl-L-alanine piperazine, 0.90 mmol of **7** (0.25 g) was obtained. (Yield was 90%.) ¹H NMR (CDCl₃, 400 MHz): δ 5.38–5.52 (m, 2H), 3.28–3.68 (m, 8H), 2.96 (s, 6H), 2.10 (s, 3H), 1.45 (s, 9H), 1.28 (d, *J* = 6.8 Hz, 3H), 1.28 (d, *J* = 6.8 Hz, 3H). ¹³C NMR (CDCl₃, 100 MHz): δ 171.5, 170.6, 169.7, 154.4, 80.2, 48.9, 48.6, 45.0, 41.9, 31.2, 29.9, 28.3, 21.9, 14.5, 14.3. HRMS (ESI-TOF MS) *m/z*: [M+Na]⁺ Calcd for C₁₉H₃₄N₄NaO₅⁺ 421.2421; Found 421.2411.

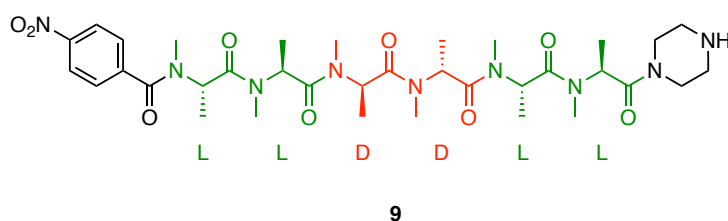
Synthesis of LLDDLL hexamer (**8**) for NMR analysis.



The hexamer **8** was synthesized on resin. Trityl chloride resin (1.55 mmol/g, 64 mg, 0.1 mmol) was swollen in THF for 10 min in a fritted syringe. After removing THF, piperazine (20 equiv.) in THF was added to the resin and the syringe was shaken for 2 h. After the resin was washed with THF and DCM three times each, the resin was incubated with DCM/MeOH/DIPEA = 17/2/1 for 15 min. The resin was washed with DCM five times and DMF three times. Fmoc-*N*-methyl-L-alanine (3.0 equiv.,

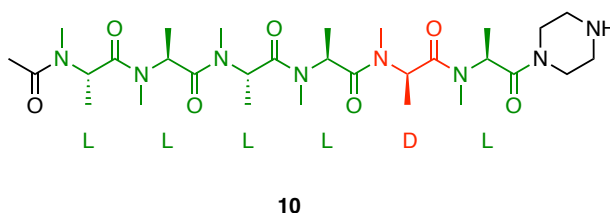
0.2 M), COMU (3.0 equiv., 0.2 M), and DIPEA (6.0 equiv., 0.4 M) were dissolved in DMF and applied to the resin. The resin was incubated for 2 h with continuous shaking. The resin was washed with DMF three times. The resin was incubated with 20% piperidine/DMF twice (3 min and 12 min), and the resin was washed with DMF three times. The coupling and deprotection were repeated to synthesize the hexamer on resin. The hexamer was capped using acetic anhydride by incubating the resin with a solution of acetic anhydride (10 equiv., 0.5 M) and DIPEA (20 equiv., 1.0 M) in DMF for 2 h. After the resin was washed with DMF and DCM three times each, the resin was treated with 30% HFIP/DCM for 45 min three times to cleave the hexamer from the resin. After removing the solvent, the crude product was purified using HPLC. Yield was 78%. HRMS (ESI-TOF MS) m/z : $[M+Na]^+$ Calcd for $C_{30}H_{54}N_8NaO_7^+$ 661.4008; Found 661.4010.

Synthesis of LLDDLL hexamer (9) for the crystallographic analysis.



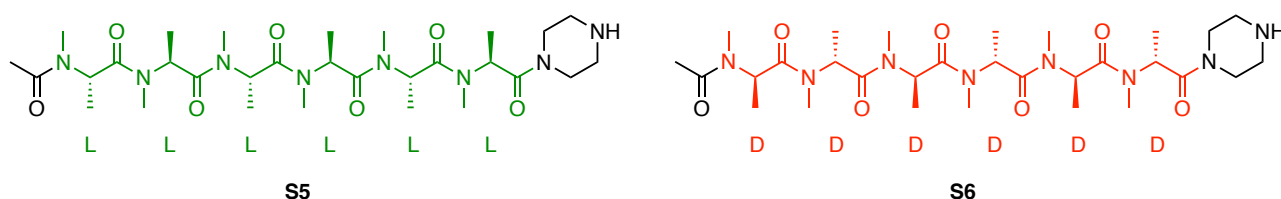
The hexamer **9** was synthesized on resin in a similar manner with **8**. After the synthesis of the NMA hexamer on resin, 4-nitrobenzoic acid (4.0 equiv., 0.2 M), COMU (4.0 equiv., 0.2 M), and DIPEA (8.0 equiv., 0.4 M) in DMF were applied to the resin and the reaction mixture was stirred for 1 h. The resin was washed with DMF three times. The resin was treated with 30% HFIP/DCM for 45 min three times to cleave the hexamer from the resin. After removing the solvent, the crude product was purified using HPLC. The yield was 89%. HRMS (ESI-TOF MS) m/z : $[M+Na]^+$ Calcd for $C_{35}H_{55}N_9NaO_9^+$ 768.4015; Found 768.4035.

Synthesis of LLLLDL hexamer (10) for NMR analysis.



The oligomer **10** was synthesized on resin in a similar manner with **8**. Yield was 24%. HRMS (ESI-TOF MS) m/z : $[M+Na]^+$ Calcd for $C_{30}H_{54}N_8NaO_7^+$ 661.4008; Found 661.4011.

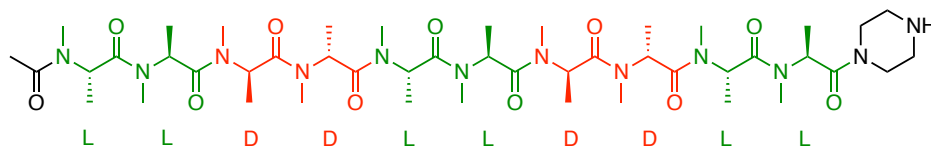
Synthesis of LLLLLL and DDDDDD hexamers (S5 and S6) for CD spectroscopic analysis.



The oligomers **S5** and **S6** were synthesized on resin in a similar manner to **8**. Yield was 78% and 95% for **S5** and **S6**, respectively. ESI-MS m/z : **S5** $[M+H]^+$ Calcd for $C_{30}H_{55}N_8O_7^+$ 639.4; Found 639.7; **S6** $[M+H]^+$ Calcd for $C_{30}H_{55}N_8O_7^+$ 639.4; Found 639.6.

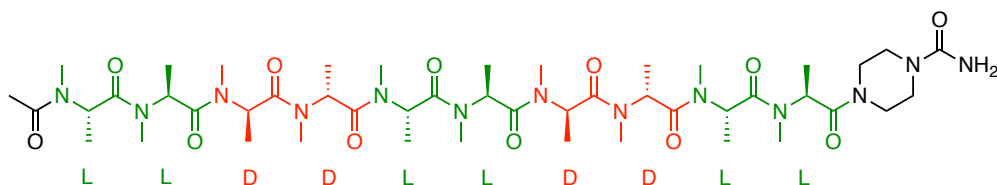
Synthesis of LLDDLLDDLL decamer for SMART-EM analysis.

Solid phase synthesis of the decamer with C-terminal piperazine



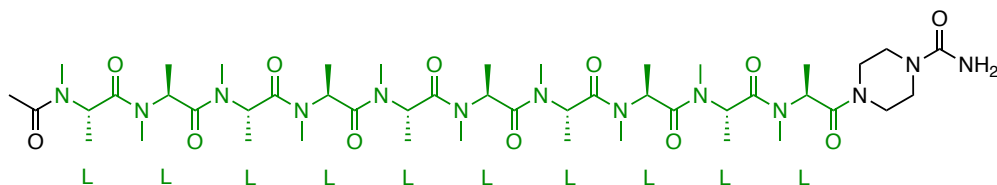
The oligomer with C-terminal piperazine was synthesized on resin. 0.13 g of trityl chloride resin (1.55 mmol/g, 0.20 mmol) was swollen in DCM for 10 min in a fritted syringe. After removing DCM, piperazine (0.35 g, 4.0 mmol, 20 equiv.) in DCM was added to the resin and the syringe was shaken for 2 h. After the resin was washed with DCM three times each, the resin was incubated with DCM/MeOH/DIPEA = 17/2/1 for 15 min. The resin was washed with DCM/MeOH/DIPEA = 17/2/1 twice, DCM three times and DMF three times. Fmoc-*N*-methyl-L-alanine (0.26 g, 0.80 mmol, 4.0 equiv.), COMU (0.34 g, 0.80 mmol, 4.0 equiv.), and DIPEA (0.27 mL, 1.6 mmol, 8.0 equiv.) were dissolved in 4.0 mL of DMF and applied to the resin. The resin was incubated for 1.5 h with continuous shaking. The resin was washed with DMF three times. The resin loaded Fmoc group was quantified according to the previous report. The loaded percentage was determined to be 61%. The following oligomer synthesis was conducted using an automated peptide synthesizer Syro I. 20 μ mol of the oligomer was used for the synthesis. On the synthesizer, Fmoc deprotection was conducted by incubating the resin with 20% piperidine/DMF twice (3 min and 12 min), and the resin was washed with DMF six times. A coupling reaction was conducted by incubating the resin with 4 equiv. of Fmoc-*N*-methyl-L/D-alanine (0.17 M), 4 equiv. of HATU (0.17 M), 4 equiv. of HOAt (0.17 M), and 8 equiv. of DIPEA (0.34 M) in NMP for 45 min. After the coupling reaction, the resin was washed with DMF three times. The cycle of coupling reaction and washing with DMF was repeated once. The coupling and deprotection were repeated until the Fmoc group of the *N*-terminal residue was deprotected on the synthesizer. After deprotection of the *N*-terminal residue, acetic anhydride (0.20 mmol, 10 equiv.) and pyridine (0.40 mmol, 20 equiv.) in 0.40 mL of DMF was applied to the resin and the reaction vessel was shaken for 1.5 h. The resin was washed with DMF and DCM three times each. The resin was treated with 30% HFIP/DCM for 20 min five times to cleave the oligomer from the resin. The resin was washed with MeOH three times, and all the cleaved peptide solution and washing solutions were combined. After removing the solvent, the crude product was purified using HPLC. The yield was 14 mg (64%). The product is considered to be obtained as TFA adduct.

C-terminal carbamoylation



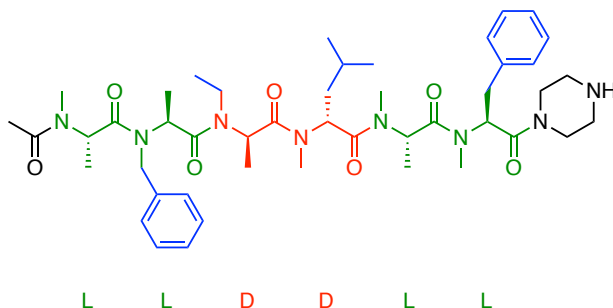
6.0 mg of the oligomer with C-terminal piperazine (5.5 μmol) was mixed with 1.5 equiv. of DIPEA, 2.5 equiv. of phenyl carbamate. The mixture was incubated at 70 $^{\circ}\text{C}$ overnight. The product was purified using HPLC to give 1.0 mg of the objective compound. Yield was 19%. ESI-MS m/z : $[\text{M}+\text{Na}]^+$ Calcd for $\text{C}_{47}\text{H}_{84}\text{N}_{13}\text{O}_{12}^+$ 1022.6; Found 1023.1.

Synthesis of LLLLLLLLLL decamer for SMART-EM analysis.



The LLLLLLLLLL with C-terminal piperidine was synthesized on resin and carbamoylated in solution in the same manner as LLDDLDDLL. The yield was 24%. ESI-MS m/z : $[\text{M}+\text{Na}]^+$ Calcd for $\text{C}_{47}\text{H}_{84}\text{N}_{13}\text{O}_{12}^+$ 1022.6; Found 1023.4.

Synthesis of functionalized LLDDL hexamer (11).



The functionalized LLDDL hexamer with C-terminal piperazine was synthesized on resin in a similar manner with non-functionalized LLDDL hexamer. The functional groups were introduced according to the previously reported methods³⁻⁵. The *N*-ethyl group was introduced by Fukuyama-Mitsunobu reaction condition, and the *N*-benzyl group was introduced by reductive amination. The functional groups on α -carbons were introduced using *N*-methyl-D-leucine and *N*-methyl-L-phenylalanine instead of *N*-methylalanine. The coupling reaction of amino acids were conducted using DIC/Oxyma as coupling reagents at 60 $^{\circ}\text{C}$ for 3 h. Double coupling was conducted when necessary. The yield was 16%. HRMS (ESI-TOF MS) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{46}\text{H}_{71}\text{N}_8\text{O}_7^+$ 847.5440; Found 847.5429.

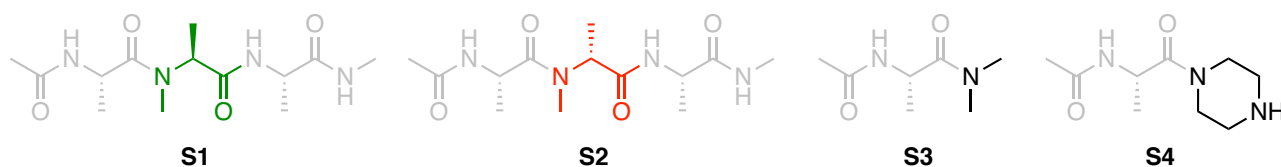
Generation of Ramachandran-type plots of acetyl-*N*-methyl-L/D-alanine dimethylamide. First, a series of conformers of acetyl-*N*-methyl-L-alanine dimethylamide was generated by combinatorically fixing ϕ and ψ at every 15 $^{\circ}$ from -180° to 180° . 625 conformers were generated in total. Each conformer was optimized at the B3LYP/6-31G(d) level of theory. The ϕ and ψ angles were fixed to the initial values during the optimization. The single point energy of the geometry-optimized conformation was calculated at the B3LYP/6-31G(d) level of theory, and the energies of all the conformers were plotted to generate the Ramachandran-type plot of acetyl-*N*-methyl-L-alanine dimethylamide. The Ramachandran-type plot of acetyl-*N*-methyl-D-alanine dimethylamide was

generated in the same manner. All the DFT calculations were conducted with solvent water modeled with polarizable continuum model.

DFT-optimized structure of acetyl-*N*-methyl-L-alanine dimethylamide. The most stable conformation of acetyl-*N*-methyl-L-alanine dimethylamide on the Ramachandran-type plot was further optimized at the B3LYP/6-311G(d) level of theory with the solvation effects of water using the CPCM model without fixing ϕ and ψ angles.

DFT-optimized structures of oligo-*N*-methylalanines. Initial conformations of *N*-methylalanine hexamers with acetylated N-terminus and dimethylamidated C-terminus were prepared using dihedral angles (ϕ , ψ , ω) of $(-135^\circ, 75^\circ, 180^\circ)$ for acetyl-*N*-methyl-L-alanine residue and $(135^\circ, -75^\circ, 180^\circ)$ for acetyl-*N*-methyl-D-alanine residue. The conformations were optimized by DFT calculations at the B3LYP/6-31G(d) level of theory with the solvation effects of water using the CPCM model. The conformations were further optimized at the B3LYP/6-311G(d) level of theory with the solvation effects of water using the CPCM model.

Conformational search of dimers and hexamers by multicanonical MD simulations. The empirical model of dimers for MD simulations were constructed by connecting monomeric residues. The residue model was built from the monomer whose terminals were capped with acetyl-alanyl and alanyl-(*N*-methyl) amine groups (S1 for L-NMA and S2 for D-NMA). The capped monomer was optimized with B3LYP/6-31G(d,p) by GAMESS⁶. By using residuegen in AmberTools, the partial atomic charges for the residue were assigned. The force field other than the partial charges was taken from Generalized Amber Force Field (GAFF). Similarly, the C-terminal residue of dimethylamide was modeled from acetyl-alanyl (*N,N*-dimethyl) amine (S3). Another C-terminal residue of piperazine group was built from acetyl-alanyl piperazine (S4). Finally, the dimer models were generated by connecting the acetyl group, two L- or D-residues and dimethylamide, and the hexamer models were generated by connecting the acetyl group, six L- or D-residues and piperazine. The LEaP in AmberTools were used.



The trivial-trajectory-parallelized multicanonical MD (McMD) simulations⁷ were performed in the temperature range from 260 K to 1500 K with 100 multiple trajectories per a simulation system. This temperature range has previously been shown to be sufficient for capturing *cis/trans* interconversions at *N*-methyl amide bonds⁸. The system contains the hexamer immersed in TIP3P water molecules within a truncated octahedron. The periodic boundary condition was used, and the long-range electrostatic interaction was calculated with Particle-Mesh Ewald method. The time step

was 2 fs with SHAKE algorithm to keep covalent bonds with hydrogen atom. After training the multicanonical weight until a constant-sampling probability was achieved across the temperature range, the production McMD run was conducted for the total of 1 μ s. The snapshots were saved every 20 ps to be analyzed equilibrium state and potential of mean force at 300 K by the reweighting method⁷. We implemented the McMD program within Amber22 and used it for the above simulations. The dihedral angles and structural clustering were calculated by AmberTools2022⁹. Ten representative structures of each peptide were selected by clustering the trajectories from the MD simulations.

Analysis of conformational dynamics of NMA decamers in vacuum by conventional canonical MD simulations. The molecular model for decamer was built by connecting the monomeric residues. The residue modeling was described above. The two decamers, acetyl-LLDDLDDLL-dimethylamide and acetyl-LLLLLLLLLL-dimethylamide were modeled.

The initial dihedral angles were restrained in stable region on Ramachandran-type plot. The stable ϕ and ψ values of L-NMA were in degrees $[-150, -120]$ and $[60, 90]$, respectively. On the other hand, in D-NMA, the stable ϕ and ψ values were in degrees $[120, 150]$ and $[-90, -60]$. At heating and equilibration process, a harmonic-restraint potential energy was added to keep the dihedral angles within the stable region.

MD simulations were conducted in vacuum condition at two temperatures: 200 and 300 K. The temperature was controlled by Langevin thermostat with γ_{ln} of 1.0 ps^{-1} . The time step was 2 fs with SHAKE. The non-covalent interaction was calculated with cutoff length 999 Å, apparently with no cutoff. The system was heated from 1 K to the target temperature for 1 ns, and it was equilibrated for 10 ns at the target temperature. After the equilibration, the production run was performed for 1 μ s without the restraint. The snapshots were saved every 100 ps. The MD simulation was repeated five times with different initial velocities. Then, total of 5 μ s was simulated and 50,000 snapshots were saved. The saved snapshots were used for structure clustering analysis, and the centroid structures in clusters were used as representatives for EM image fitting.

NMR measurements. Lyophilized peptides were directly dissolved in D₂O for NMR measurements. NMR spectra were recorded on a Bruker AVANCE-III HD 800, 600, or 500 spectrometer (Bruker, Billerica, MA, USA), equipped with a cryogenic probe at 298 K, unless otherwise stated. One-dimensional ¹H- and ¹³C-NMR experiments were performed using standard pulse sequences and phase cycling (zgpr and zgpg30). The inter-scan delays were set to 1.0 or 2.0 sec. Two-dimensional ROESY and EASY-ROESY spectra were recorded using the previously reported pulse program (roesyetgp and roesyadjsphpr) and parameters with a mixing time of 250 ms. The spectral widths were set to 7,211 or 6,010 Hz for both dimensions, and 2,048×2,048 or 1,024×1,024 complex points were recorded. Two-dimensional ¹H-¹³C HSQC and HMBC experiments were performed using standard pulse sequences and phase cycling (hsqcetgpsp, hmbcgpndqf, and hmbcgplpndqf). In the HSQC experiments, the spectral widths were set to 6,009–8,012 Hz and 14,084–20,119 Hz for the ¹H and ¹³C dimensions, respectively, and 1,024×512 complex points were recorded. In the HMBC

experiments, the spectral widths were set to 5,000–8,012 Hz and 25,149–40,242 Hz for the ^1H and ^{13}C dimensions, respectively, and 2,048–4,096 \times 1,024–2,048 complex points were recorded. The inter-scan delays were set to 1.0 or 1.5 sec in all two-dimensional experiments. All the spectra were processed and analyzed by the Topspin 3.5 or 4.1 software (Bruker). Sodium trimethylsilylpropanesulfonate was used as an external standard for the chemical shifts of each compound in both ^1H and ^{13}C dimensions.

Crystallization of an NMA monomer and dimers. NMA monomer and dimers were dissolved in a mixture of ethyl acetate and hexane. The solution was left in a glass vial capped with a plastic cap with small holes to let solvent slowly evaporate until crystals appeared in solution.

X-ray crystallography. A crystal was mounted with mineral oil on a loop-type mount and set on VariMax Dual (Rigaku). The X-ray diffraction data was measured at $-180\text{ }^\circ\text{C}$ using Cu K α radiation ($\lambda = 1.5418\text{ \AA}$). The structure was solved using Olex2 software (Rigaku) equipped with SHELX and PLATON. The non-hydrogen atoms were refined anisotropically. Hydrogen atoms were placed on ideal positions.

XFEL analysis of LLDDLL crystals. The 3D structure of the hexamer was determined by the serial XFEL crystallography¹⁰. Crystals of hexamer were obtained in a 1:1 mixture of hexane and ethyl acetate as micrometer-sized grains. The grains were suspended in Cargille Type B immersion oil (Cargille Laboratories) and spread over a polyimide plate with an area of $4 \times 4\text{ mm}^2$. The XFEL measurement was performed at Beamline 2, SACLA XFEL facility. The photon energy of the XFEL and the beam size were adjusted to 15.0 keV and $\sim 1\text{ }\mu\text{m}$, respectively. The sample plate was scanned over the entire plane with XZ-translation and phi-rotation movements, being exposed every $10\text{ }\mu\text{m}$ by XFEL pulses. The diffraction patterns were recorded on an MX300-HS detector (Rayonix) placed 55 mm downward from the sample plane. All data collection was performed at room temperature. Total 729,635 images were collected. To obtain the reference cell parameters for the effective indexing of the serial diffraction images, rotational electron diffraction was also measured. The crystals from the same batch as those for the XFEL analysis were spread over a copper TEM grid covered with a holey carbon film (Quantifoil R1/4). After drying up the solution, a mixture of hexane and ethyl acetate, the grid was immersed in liquid nitrogen, and transferred into a CRYO ARM 300 electron microscope (JEOL) operated at an accelerating voltage of 300 kV and maintained at a specimen temperature of 93 K. Diffraction patterns were recorded on a DE64 detector (Direct Electron) by illuminating crystals with a parallel electron beam during rotating the sample stage from -68° to 68° ¹¹. The rotational series were iteratively collected for several tens of different crystal grains. The electron diffraction patterns were processed with DIALS¹² for indexing the diffraction spots. The common lattice parameters, monoclinic, $a = 16.7\text{ \AA}$, $b = 6.0\text{ \AA}$, $c = 40.3\text{ \AA}$, $\beta = 92^\circ$, found in the rotational series were provided for processing serial XFEL diffraction images with CrystFEL suite¹³. Thus 26,709 images (3.7% of the total collected images) were successfully indexed from extracted 140,435 images (19.2%) showing diffraction spots. The integrated and merged diffraction intensities

were then used for ab initio phasing with SHELXT¹⁴ and the initial structure was refined with SHELXL¹⁵. All non-hydrogen atoms were modeled with anisotropic displacement parameters and hydrogen atoms were generated using a riding model. Structure determination was also possible by the direct method from the electron diffraction data alone, but its data statistics against the model and geometries of atom positions were poorer than those from the XFEL data.

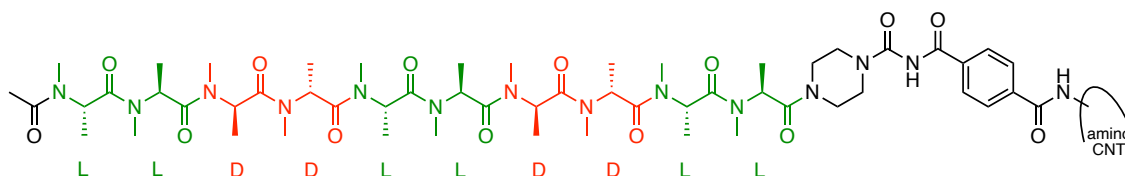
Circular dichroism studies. A 100 μM solution of oligo-NMA in 10 mM phosphate buffer was used for the measurements. CD spectra were acquired at 25 °C using a CD spectrometer (JASCO, J-1500) equipped with a 1 mm pathlength quartz cuvette (JASCO, 204J). The data pitch was set to 0.1 nm. The scanning speed was 100 nm/min, and each spectrum was obtained by averaging three scans. A baseline spectrum was recorded using 10 mM phosphate buffer and subtracted from the sample spectra. All data points were baseline-corrected, converted to molar ellipticity ($\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$), and plotted.

SMART-EM.

General Remarks for TEM measurements. A TEM grid precoated with a lacy microgrid (NS-C15, pore size 1.5–8 μm) was purchased from Okenshoji Co., Ltd. TEM simulation was performed using a multi-slice procedure implemented in BioNet elbis[®] software. The fishhook for capturing peptides was prepared based on a previous study.¹⁶

We used a Cs aberration-corrected JEOL JEM-ARM200F transmission electron microscope equipped with a complementary metal-oxide-semiconductor camera (Gatan OneView). Dynamic imaging of the peptides using single-molecule atomic-resolution time-resolved electron microscopy (SMART-EM) at angstrom resolution was conducted with an electron dose rate (EDR) between 10^6 and $10^7\text{ e}^- \text{nm}^{-2} \text{s}^{-1}$ and an acceleration voltage of 80 kV at a frame rate of 50 frame per second (fps) at 2,000,000 magnifications. We first surveyed the whole CNT aggregates on the grid at $\times 100,000$ magnification to find suitable areas for careful analysis. To analyze them in detail, we increased the magnification to $\times 2,000,000$ and started the video recording. Defocus value was set manually to approximately -10 nm (under-focus), a typical value that offers the best compromise between image contrast and richness of molecular structural information. The images were recorded in a .dm4 format in the Gatan DigitalMicrograph software. The images taken in a .dm4 format were transformed into 8-bit or 32-bit .tiff format files by Gatan DigitalMicrograph and Fiji software.¹⁷ All images were processed using a bandpass filter (filtering structures smaller than 2 pixels and larger than 40 pixels, tolerance of direction: 5%) and linearly adjusted with brightness and contrast. The initial structures of peptides were obtained from computational calculations (shown above), and the models were constructed on Materials Studio software.

KAT ligation of peptides on CNT to prepare peptide@CNT.¹⁸



KAT@CNT (0.5 mg) was dispersed in a 1:1 mixture of THF and aqueous buffer of pH 7.4 ($1 \times$ PBS solution, NIPPON GENE). A decamer of LLDDLLDDLL (0.2 μ mol, 0.8 mM, 250 μ L) and 1,3,5-trichloroisocyanuric acid (TCAA, 110 μ L, 2.2 mM, 1.2 equiv. to LLDDLLDDLL) were added and the resulting mixture stirred at 40 $^{\circ}$ C for 22 h (4 mL solution in total). The solution was directly filtrated through a PTFE membrane filter (ADVANTEC, pore size: 100 nm) and washed with dried-MeOH (1 mL \times 3) to obtain 0.42 mg of a black powder after vacuum drying (60 Pa) for 1 h (peptide@CNT). LLLLLLLLLL was treated according to this procedure.

The black powder of peptide@CNT forming a large agglomerate was wetted with MeOH (2 mL/mg) under nitrogen and gently ground for 3 minutes in an agate mortar so as to break the agglomerate into individual CNT aggregates for TEM analysis. The dispersion was filtered through a Kiriya 5A filter (pore sizes: 7 μ m), and the filtrate (10 μ L) was drop casted onto a TEM microgrid placed on a paper that absorbed excess MeOH. The resulting TEM grid was placed in vacuo (60 Pa) at 298 K for 1 h.

For comparisons between TEM images and structures under vacuum, the viewing angle was optimized from the coordinates obtained in the MD experiments under vacuum at 200–300 K and superimposed on the TEM images (Figure S20–21).

Supplementary Figures and Tables

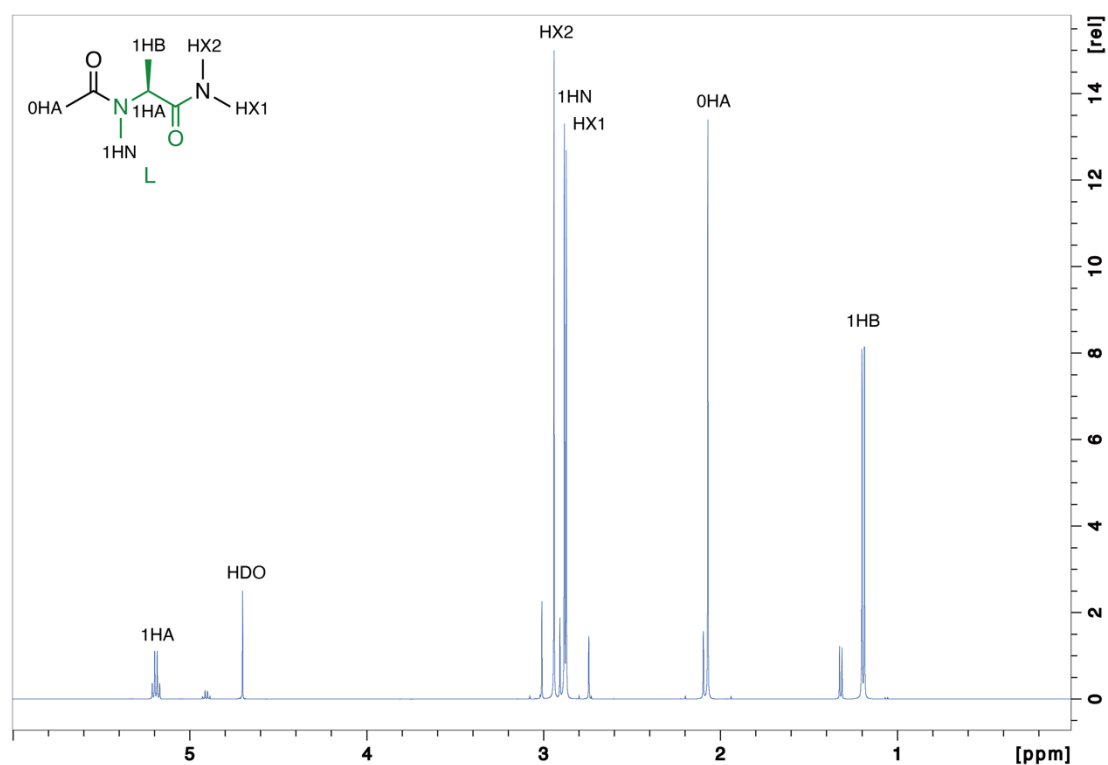


Fig. S1 ^1H -NMR spectrum of L-NMA monomer. The ^1H -NMR spectrum was recorded in D_2O . The name of the proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are not labeled.

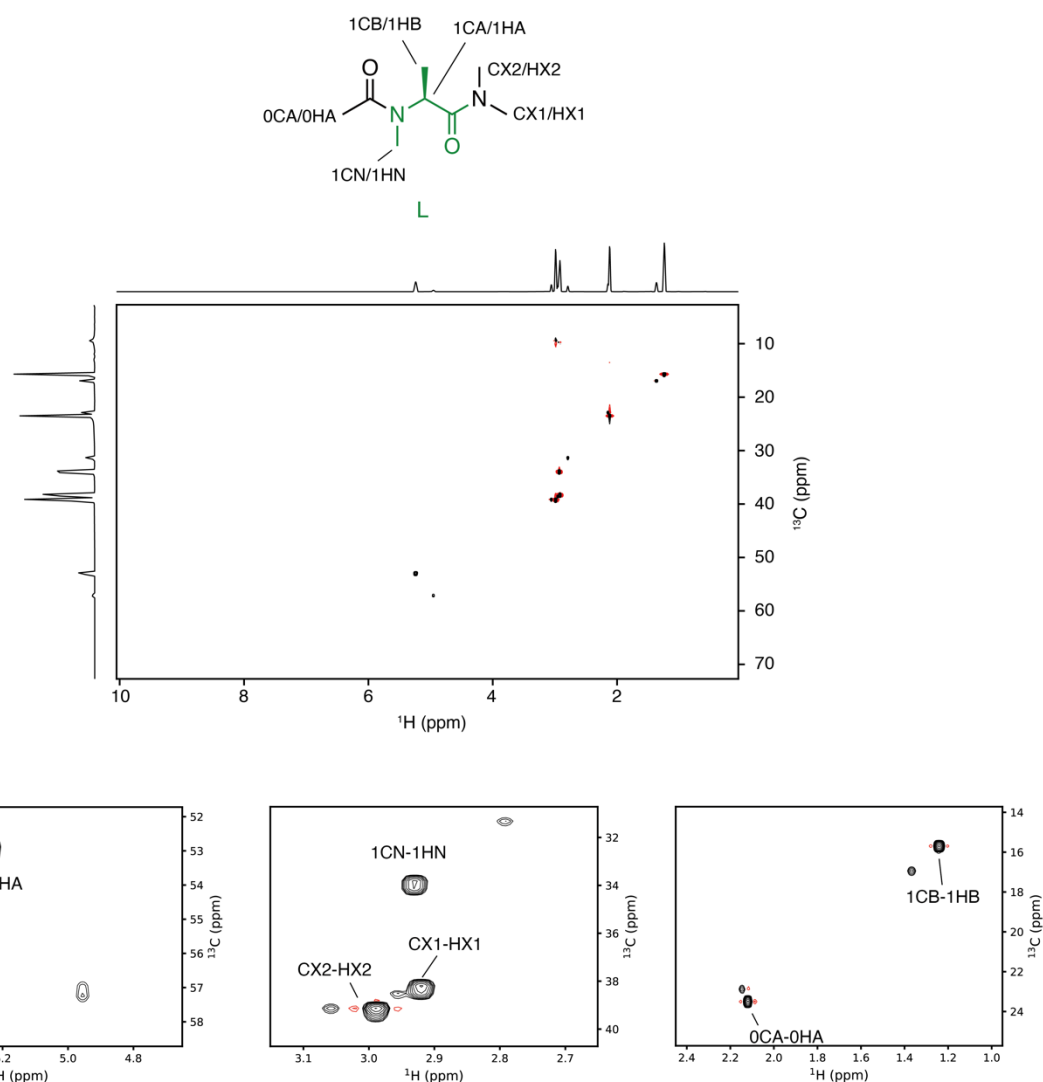


Fig. S2 HSQC spectrum of L-NMA monomer. The HSQC spectrum was recorded in D₂O. Black and red contours indicate positive and negative values, respectively. Enlarged views of the HSQC spectrum are shown at the bottom. The name of the carbon/proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are not labeled.

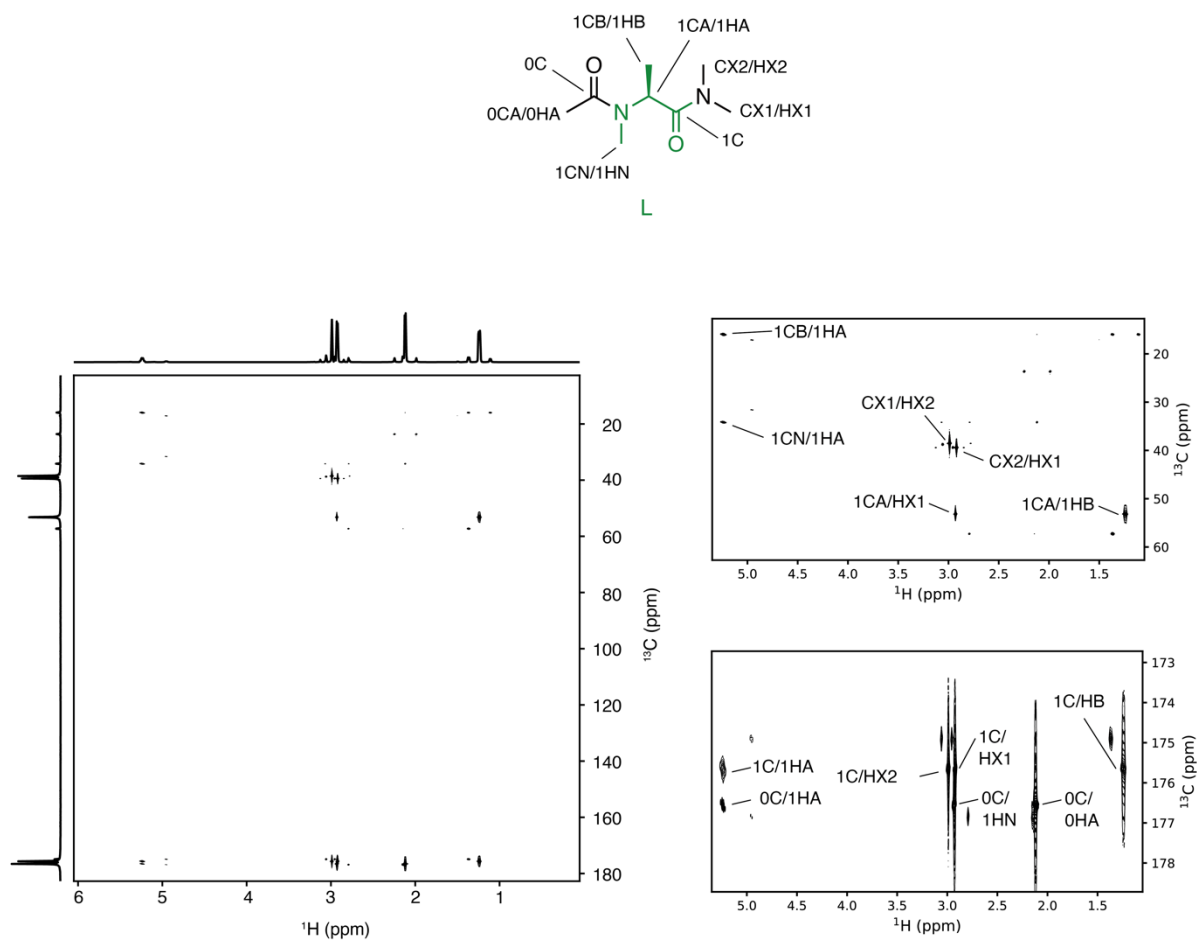


Fig. S3 HMBC spectrum of L-NMA monomer. The HMBC spectrum was recorded in D₂O. Enlarged views of the HMBC spectrum are shown on the right. The name of the carbon/proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are not labeled.

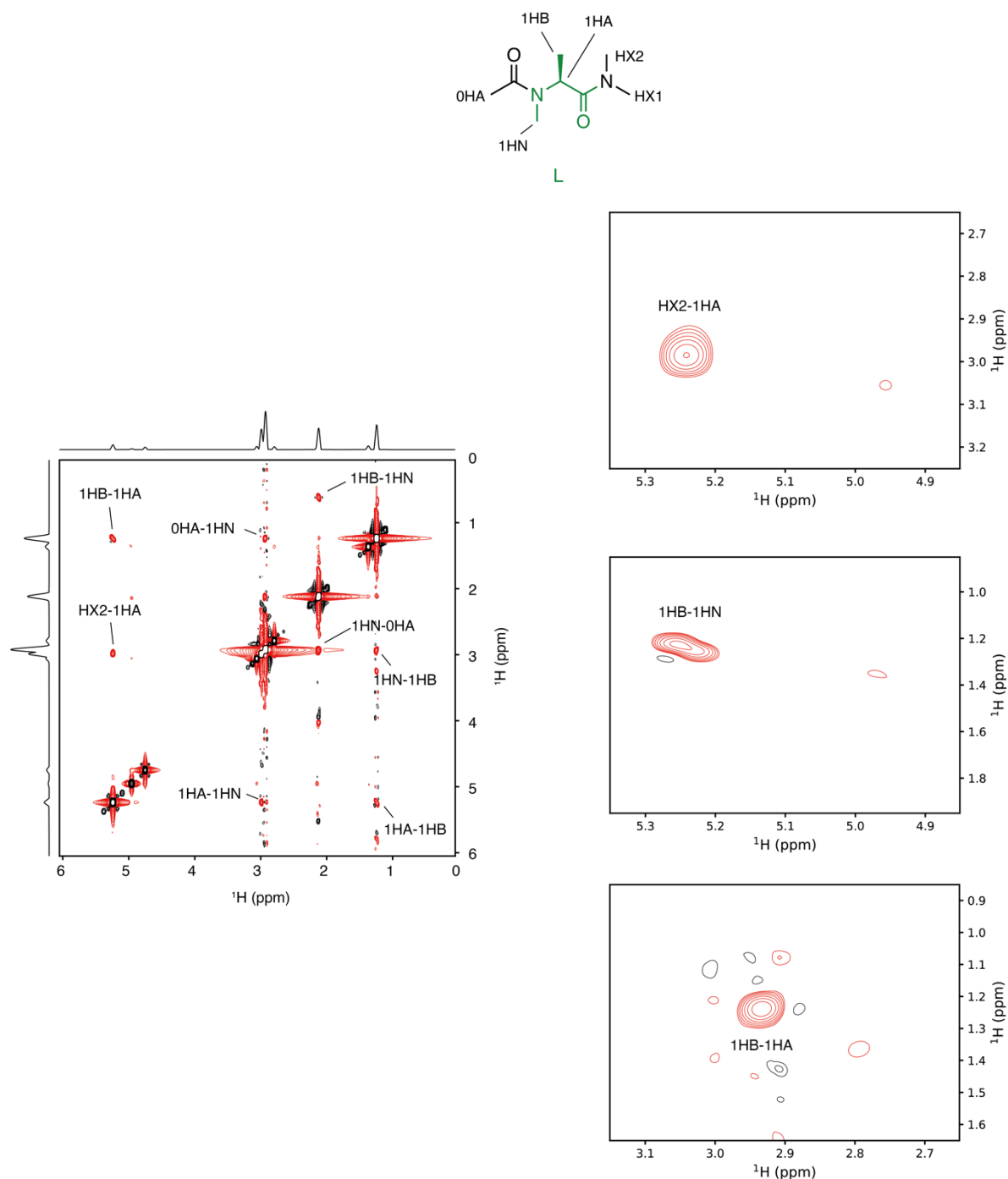


Fig. S4 EASY-ROESY spectrum of L-NMA monomer. The ROESY spectrum was recorded in D_2O . Black and red contours indicate positive and negative values, respectively. Enlarged views of the ROESY spectrum are shown on the right. The name of the proton is labeled on the chemical structure and near each peak. The minor peaks which are derived from conformers with a *cis* amide bond are not labeled.

Table S1 NOE summary of acetyl-*N*-methyl-L-alanine dimethylamide

Atom 1 *	Atom 2 *	Intensity **
1HA	HX2	s
1HN	1HB	m
1HA	1HB	s

* HA, α -proton; HX, C-terminal dimethylamide protons; HN, *N*-methyl protons; HB, β -protons.

** m, medium; s, strong.

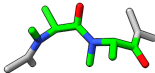
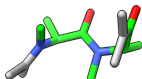
Table S2 Chemical shift table of acetyl-*N*-methyl-L-alanine dimethylamide

Residue No.	Atom *	Nuclei	Chemical shift (ppm)
0	C	¹³ C	176.6
0	CA	¹³ C	23.8
0	HA	¹ H	2.12
1	C	¹³ C	175.7
1	CA	¹³ C	53.2
1	CB	¹³ C	16.1
1	CN	¹³ C	34.2
C-term	CX1	¹³ C	38.5
C-term	CX2	¹³ C	39.4
1	HA	¹ H	5.24
1	HB	¹ H	1.24
1	HN	¹ H	2.94
C-term	HX1	¹ H	2.92
C-term	HX2	¹ H	2.99

* See the atom names on the chemical structure in Fig. S4.

Table S3 Representative conformers of the 10 clusters from McMD simulations of LLDDL and LLLDL

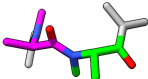
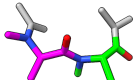
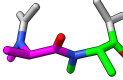
LL dimer

Cluster	Representative conformer ^a	Population (%)	Cluster	Representative conformer	Population (%)
1 ^b		88.0	6		0.7
2		5.6	7		0.4
3		2.6	8		0.3
4		1.4	9		0.1
5		0.8	10		0.1

^a Representative conformer from the population of 5% or more are shown in the table.

^b Cluster 1 satisfies all the backbone interproton distances with the observed NOE signals.

DL dimer

Cluster	Representative conformer ^a	Population (%)	Cluster	Representative conformer	Population (%)
1 ^b		65.0	6		1.5
2		12.4	7		0.9
3		11.9	8		0.3
4		4.3	9		0.3
5		3.2	10		0.2

^a Representative conformer from the population of 5% or more are shown in the table.

^b Cluster 1 satisfies all the backbone interproton distances with the observed NOE signals.

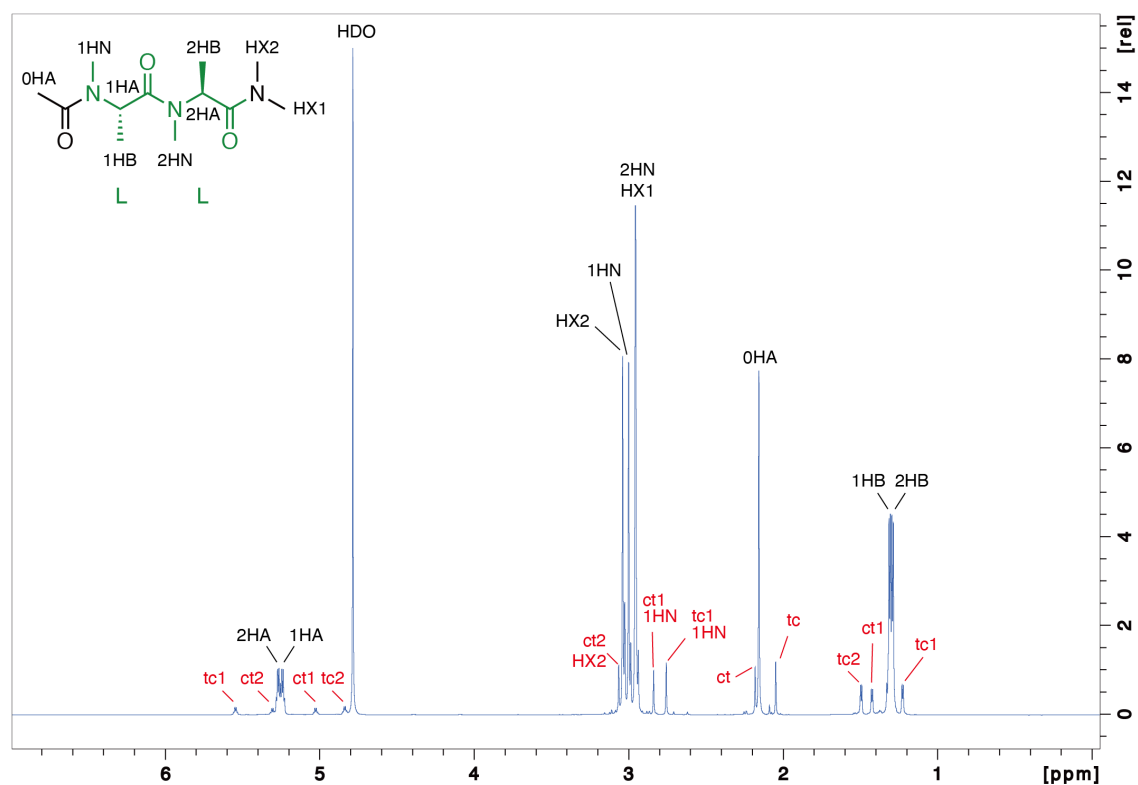


Fig. S5 ^1H -NMR spectrum of LL dimer. The ^1H -NMR spectrum was recorded in D_2O . The name of the proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are labeled in red text.

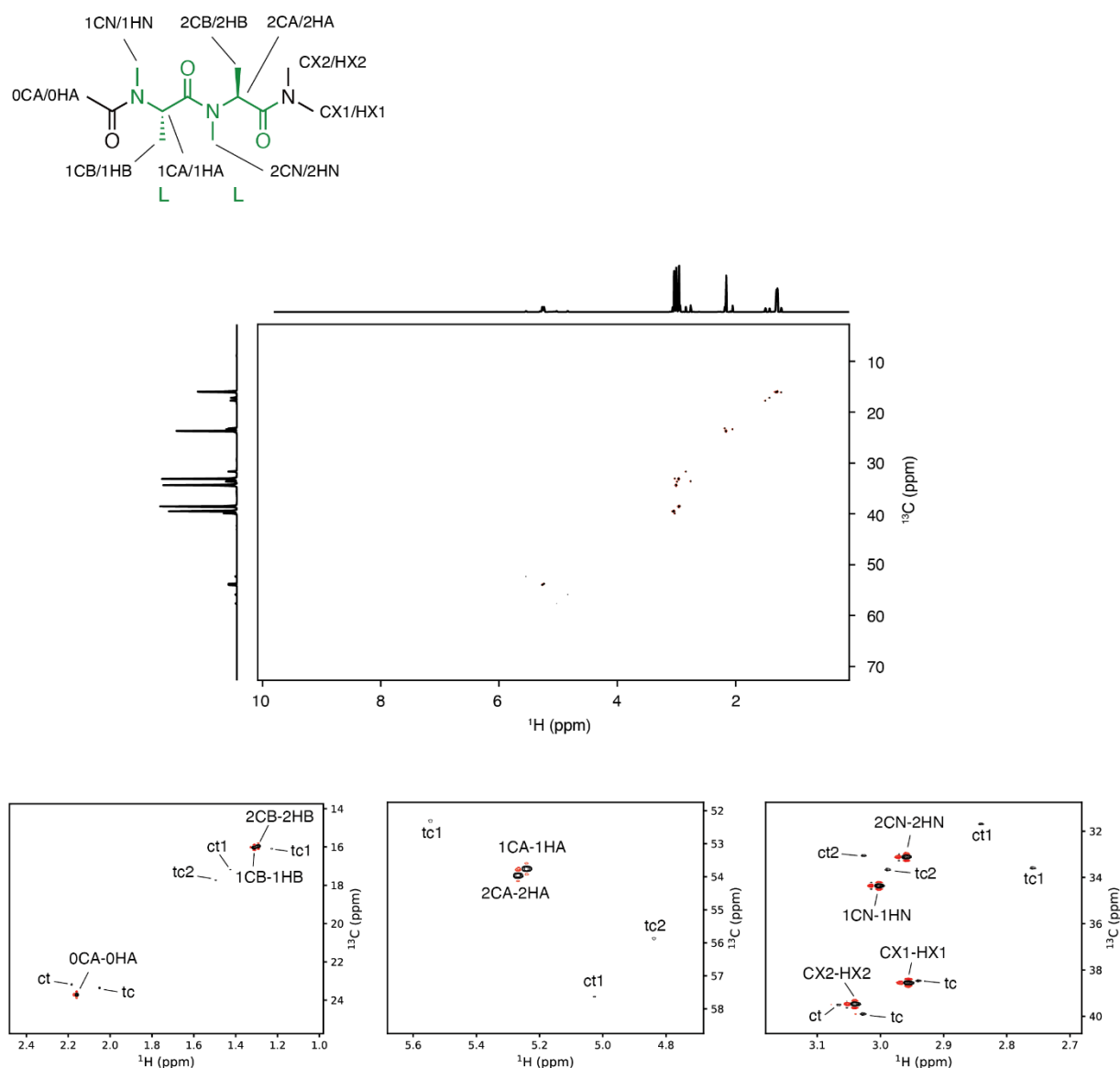


Fig. S6 HSQC spectrum of LL dimer. The HSQC spectrum was recorded in D_2O . Black and red contours indicate positive and negative values, respectively. Enlarged views of the HSQC spectrum are shown at the bottom. The name of the carbon/proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are labeled as ct and tc, indicating the position of the *cis* (c) and *trans* (t) amide bonds within the dipeptide. The numbers 1 and 2 following these labels denote that the signal corresponds to the N-terminal and C-terminal residues, respectively. The population of each conformer was calculated to be $tt = 78\%$, $tc = 12\%$, $ct = 10\%$, and $cc = 0\%$ (not detected) ($t = \text{trans}$, $c = \text{cis}$).

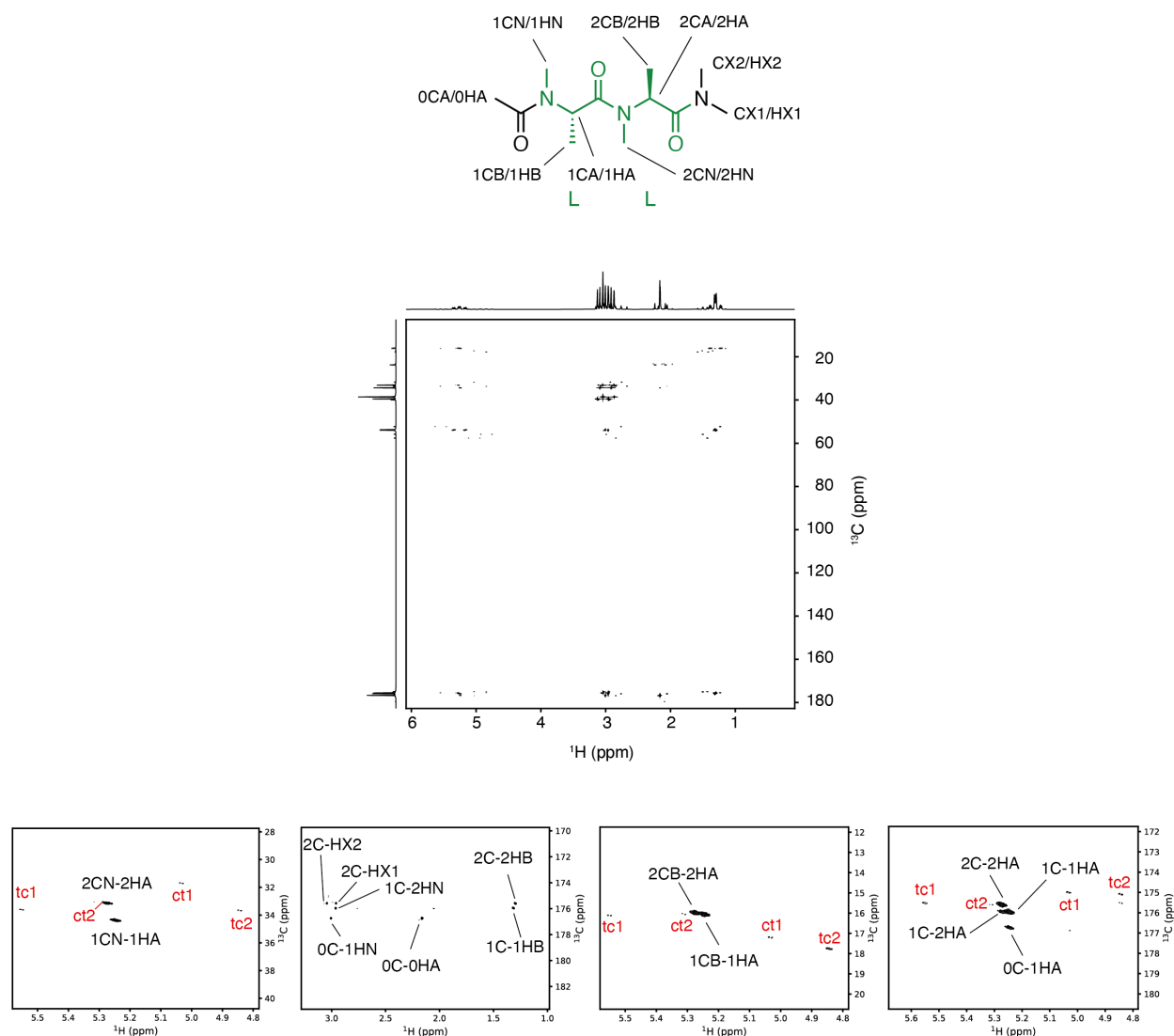


Fig. S7 HMBC spectrum of LL dimer. The HMBC spectrum was recorded in D₂O. Enlarged views of the HMBC spectrum are shown at the bottom. The name of the carbon/proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are labeled in red text.

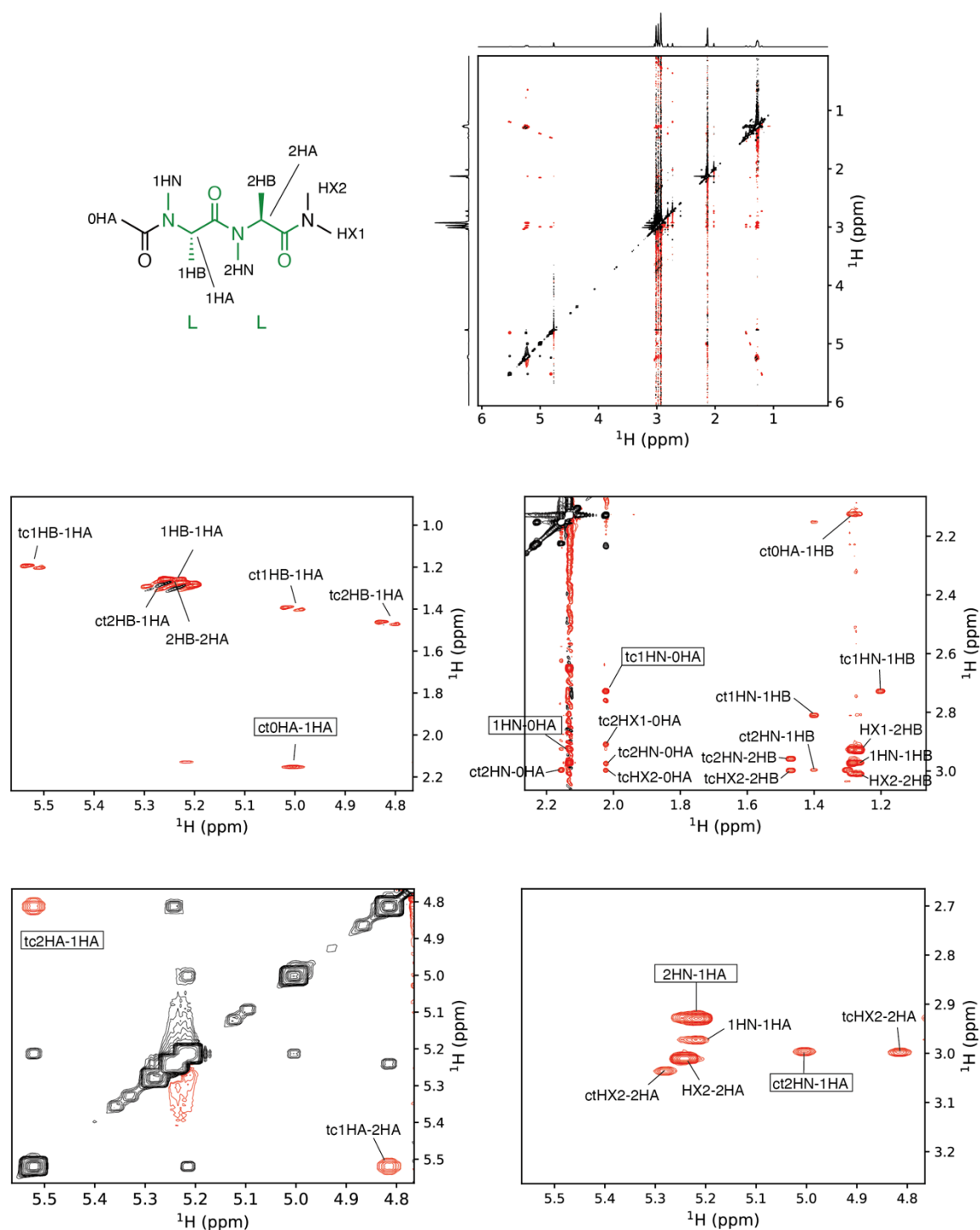


Fig. S8 EASY-ROESY spectrum of LL dimer. The ROESY spectrum was recorded in D₂O. Black and red contours indicate positive and negative values, respectively. Enlarged views of the ROESY spectrum are shown on the middle and bottom panels. The name of the proton is labeled on the chemical structure and near each peak. *Trans* and *cis* geometries of the amide bonds were determined based on NOE signals between NH and HA (for *trans*), and HA and HA (for *cis*) of neighboring residues. The corresponding NOE signals are highlighted with black boxes.

Table S4 NOE summary of LL dimer

Atom 1 *	Atom 2 *	Intensity **
1HA	2HN	s
1HN	1HB	m
1HA	1HB	s
2HA	HX2	s
2HN	2HB	m
2HA	2HB	s
0HA	1HN	w
1HA	1HN	w

* HA, α -proton; HX, C-terminal dimethylamide protons; HN, *N*-methyl protons; HB, β -protons.

** m, medium; s, strong; w, weak.

Table S5 Chemical shift table of LL dimer

Residue No.	Atom *	Nuclei	Chemical shift (ppm)
0	C	¹³ C	176.7
0	CA	¹³ C	23.8
0	HA	¹ H	2.16
1	C	¹³ C	176
1	CA	¹³ C	53.8
1	CB	¹³ C	16.1
1	CN	¹³ C	34.4
1	HA	¹ H	5.24
1	HB	¹ H	1.31
1	HN	¹ H	3
2	C	¹³ C	175.6
2	CA	¹³ C	54
2	CB	¹³ C	16
2	CN	¹³ C	33.2
C-term	CX1	¹³ C	38.6
C-term	CX2	¹³ C	39.5
2	HA	¹ H	5.27
2	HB	¹ H	1.29
2	HN	¹ H	2.96
C-term	HX1	¹ H	2.95
C-term	HX2	¹ H	3.04

* See the atom names on the chemical structure in Fig. S8.

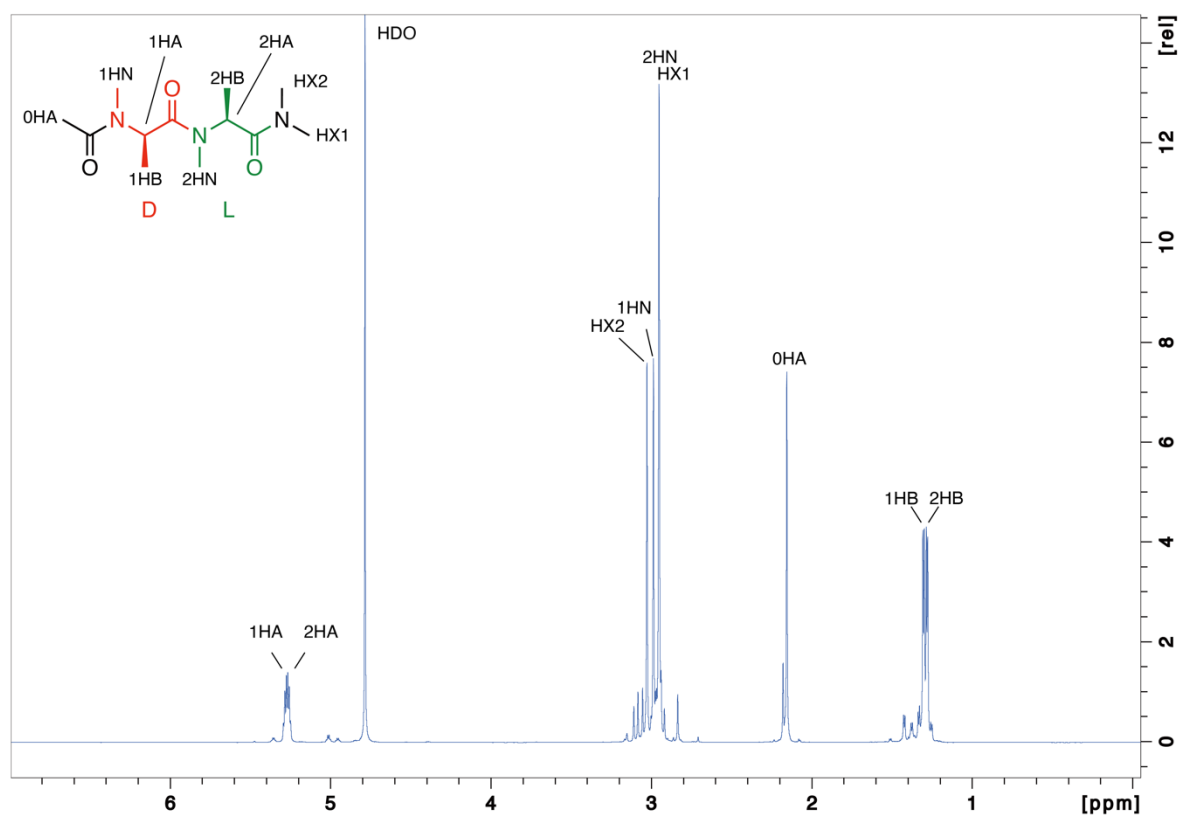


Fig. S9 ^1H -NMR spectrum of DL dimer. The ^1H -NMR spectrum was recorded in D_2O . The name of the proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are not labeled.

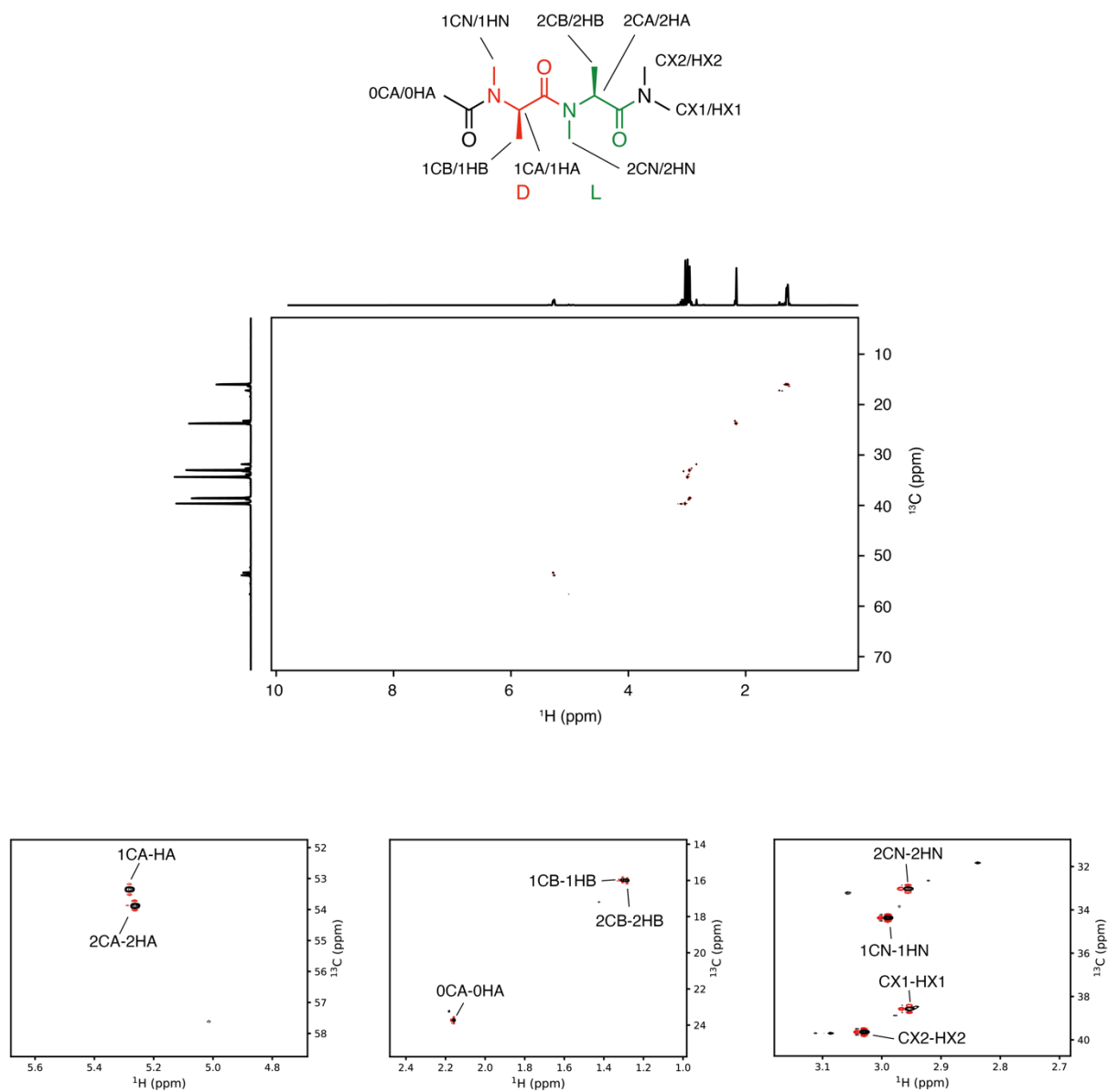


Fig. S10 HSQC spectrum of DL dimer. The HSQC spectrum was recorded in D_2O . Black and red contours indicate positive and negative values, respectively. Enlarged views of the HSQC spectrum are shown at the bottom. The name of the carbon/proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are not labeled.

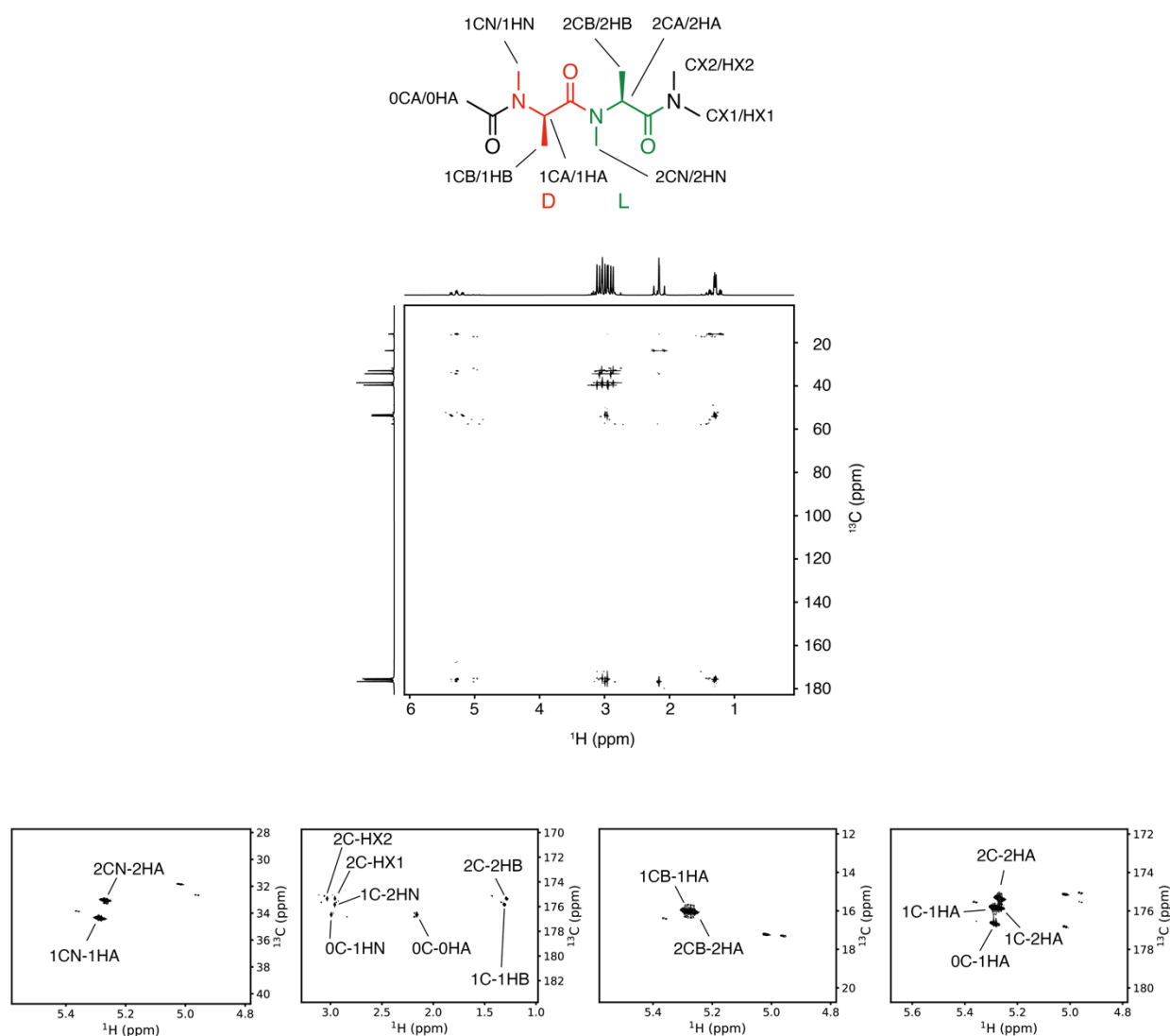


Fig. S11 HMBC spectrum of DL dimer. The HMBC spectrum was recorded in D₂O. Enlarged views of the HMBC spectrum are shown at the bottom. The name of the carbon/proton is labeled on the chemical structure and above each peak. The minor peaks which correspond to conformers with a *cis* amide bond are not labeled.

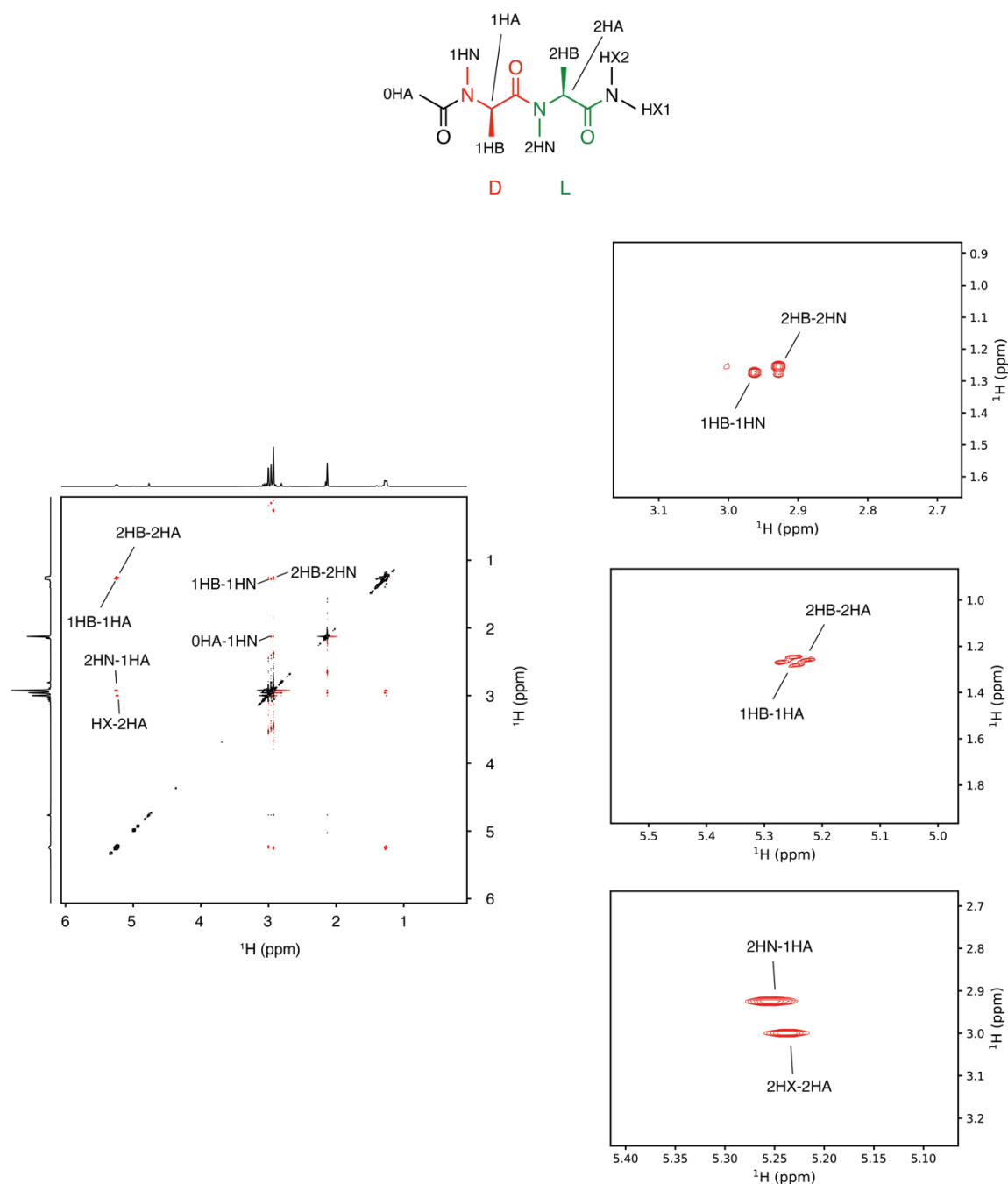


Fig. S12 ROESY spectrum of DL dimer. The EASY-ROESY spectrum was recorded in D₂O. Black and red contours indicate positive and negative values, respectively. Enlarged views of the ROESY spectrum are shown on the right. The name of the proton is labeled on the chemical structure and near each peak. The minor peaks which are derived from conformers with a *cis* amide bond are not labeled. The population of each conformer was calculated to be *tt* = 82%, *tc* = 6.7%, *ct* = 11.4%, and *cc* = 0% (not detected) (*t* = *trans*, *c* = *cis*).

Table S6 NOE summary of DL dimer

Atom 1 *	Atom 2 *	Intensity **
1HA	2HN	s
1HN	1HB	m
1HA	1HB	s
2HA	HX2	s
2HN	2HB	m
2HA	2HB	s
0HA	1HN	w
1HA	1HN	w
2HA	2HN	w

* HA, α -proton; HX, C-terminal dimethylamide protons; HN, *N*-methyl protons; HB, β -protons.

** m, medium; s, strong; w, weak.

Table S7 Chemical shift table of DL dimer

Residue No.	Atom *	Nuclei	Chemical shift (ppm)
0	C	¹³ C	176.7
0	CA	¹³ C	23.8
0	HA	¹ H	2.16
1	C	¹³ C	175.8
1	CA	¹³ C	53.4
1	CB	¹³ C	16
1	CN	¹³ C	34.4
1	HA	¹ H	5.28
1	HB	¹ H	1.3
1	HN	¹ H	2.99
2	C	¹³ C	175.4
2	CA	¹³ C	53.9
2	CB	¹³ C	16.1
2	CN	¹³ C	33.1
C-term	CX1	¹³ C	38.6
C-term	CX2	¹³ C	39.7
2	HA	¹ H	5.26
2	HB	¹ H	1.28
2	HN	¹ H	2.95
C-term	HX1	¹ H	2.95
C-term	HX2	¹ H	3.02

* See the atom names on the chemical structure in Fig. S12

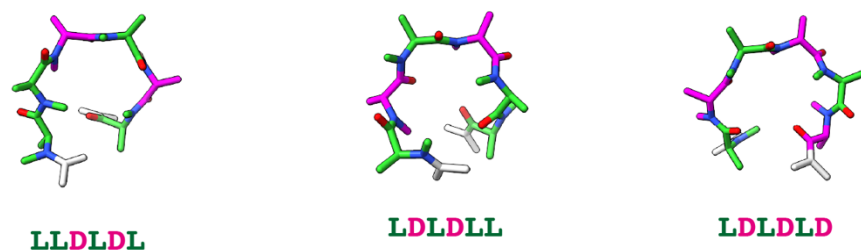
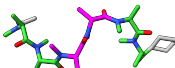
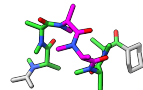
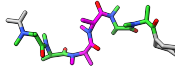
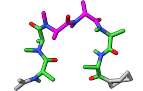
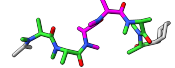


Fig. S13 A list of possible peptide shapes realized by the sequences of LLDLDL, LDLDLL, and LDLDLD. The initial conformations of the oligomers were constructed by linking the stable conformations of L-NMA and D-NMA residues with trans amide bonds. However, these initial conformations exhibited intramolecular atomic clashes. To resolve this, molecular mechanics calculations were performed with the MMFF94 force field within Avogadro19 (version 1.2.0). The minimization step produced conformations having no intramolecular clashes. The resulting structures were further optimized through DFT calculations to obtain the model structures.

Table S8 Representative conformers of the 10 clusters from McMD simulations of LLDDL and LLLDL

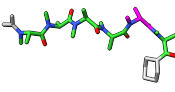
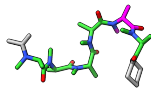
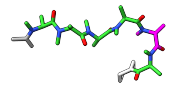
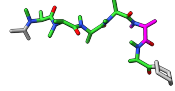
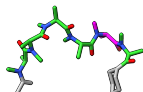
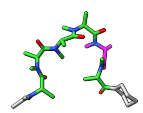
LLDDL hexamer

Cluster	Representative conformer ^a	Population (%)	Cluster	Representative conformer	Population (%)
1 ^b		59.0	6		0.7
2		20.7	7		0.5
3		7.0	8		0.4
4		6.0	9		0.4
5		5.2	10		0.1

^a Representative conformer from the population of 5% or more are shown in the table.

^b Cluster 1 satisfies all the backbone interproton distances with the observed NOE signals.

LLLLDL hexamer

Cluster	Representative conformer ^a	Population (%)	Cluster	Representative conformer	Population (%)
1 ^b		31.9	6		5.4
2 ^b		28.8	7		2.5
3		11.7	8		2.0
4		9.2	9		1.4
5		6.3	10		0.7

^a Representative conformer from the population of 5% or more are shown in the table.

^b Cluster 1 does not satisfy all the backbone interproton distances with the observed NOE signals while cluster 2 does.

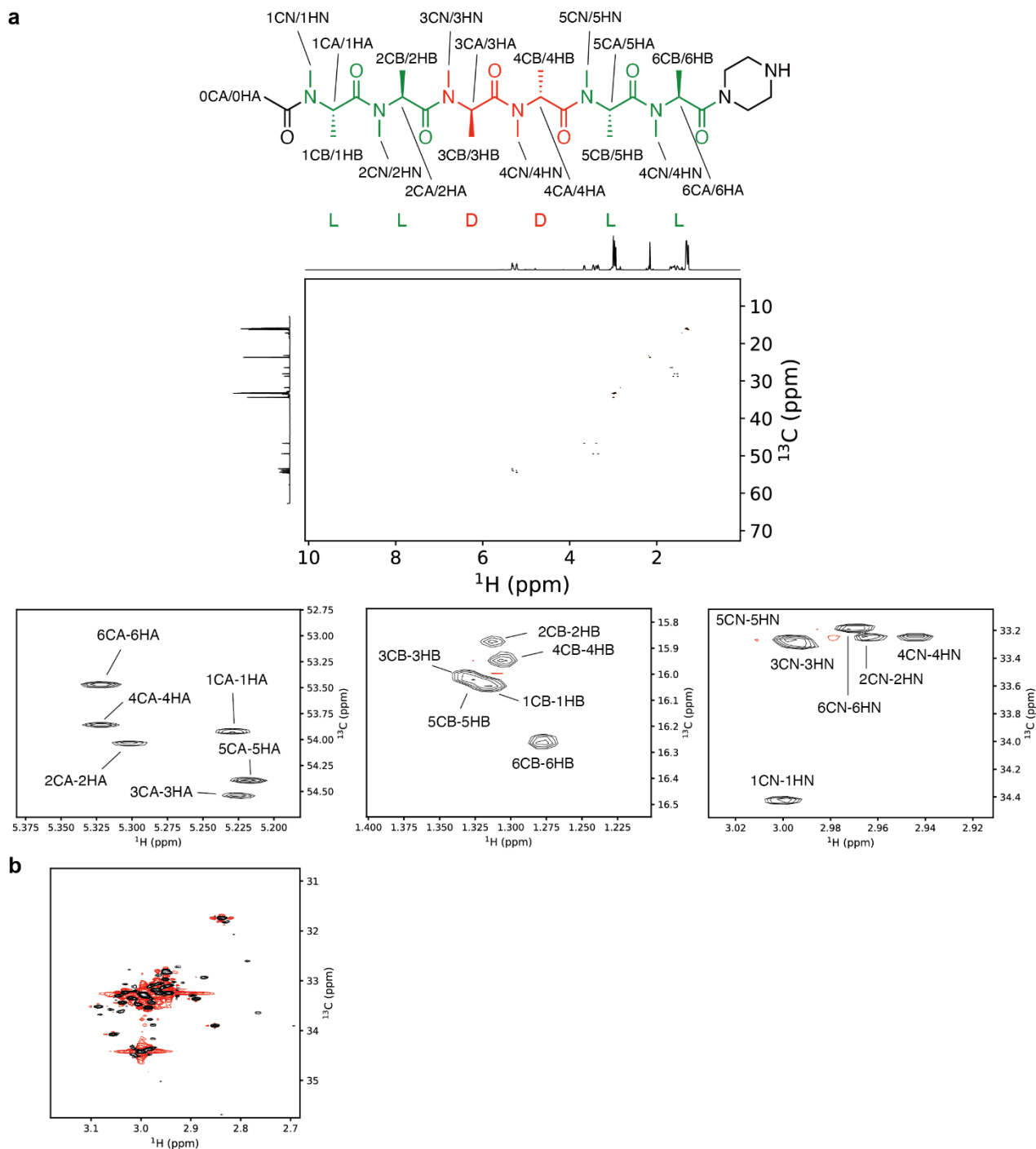


Fig. S14 HSQC spectrum of LLDDLL hexamer. The HSQC spectrum was recorded in D₂O. Black and red contours indicate positive and negative values, respectively. (a) The structure and the HSQC spectrum of the hexamer. Enlarged views of the HSQC spectrum are shown below the entire spectrum. The name of the carbon/proton are labeled on the chemical structure and around each peak. (b) An enlarged view of the region corresponding to *N*-substituted methyl group. The spectrum was used for calculating the population of all-*trans* state. The basal contour level was set to be 50 times lower than that of the other spectra. The population of the all-*trans* state (66%) was calculated as the ratio of the total positive peak volume of the resonances from the *N*-substituted methyl groups in the all-*trans* state to the total positive peak volume of all resonances from the *N*-substituted methyl groups.

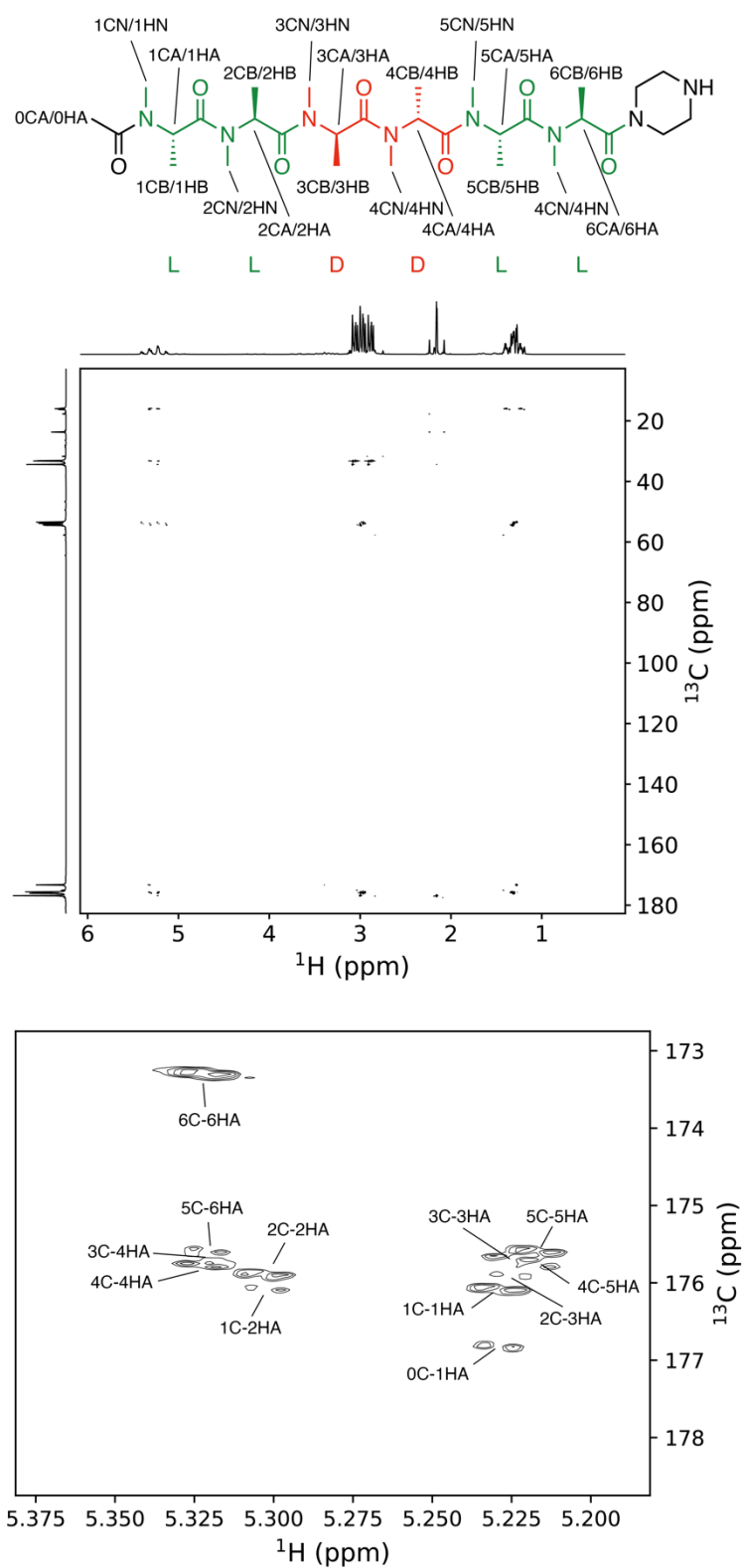


Fig. S15 HMBC spectrum of LLDDLL hexamer. The HMBC spectrum was recorded in D_2O . An enlarged view of the HMBC spectrum is shown at the bottom. The name of the carbon/proton are labeled on the chemical structure and above each peak.

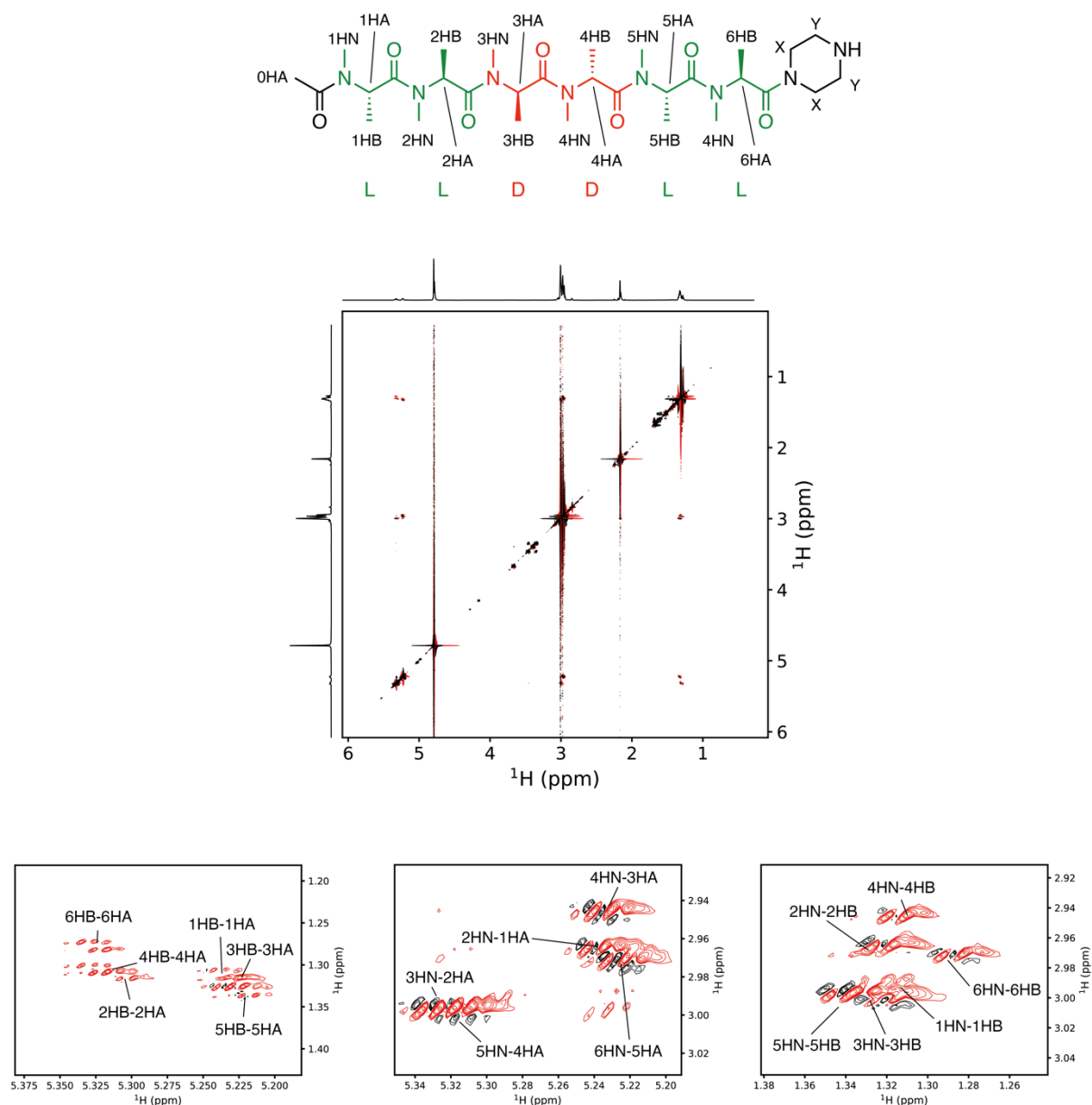


Fig. S16 ROESY spectrum of LLDDLL hexamer. The ROESY spectrum was recorded in D₂O. Black and red contours indicate positive and negative values, respectively. Enlarged views of the ROESY spectrum are shown at the bottom. The name of the proton is labeled on the chemical structure and around each peak.

Table S9 NOE summary of LLDDL hexamer

Atom 1 *	Atom 2 *	Intensity **
1HN	1HA	w
1HN	1HB	m
1HA	1HB	s
1HA	2HN	s
2HN	2HA	w
2HN	2HB	m
2HA	2HB	s
2HA	3HN	s
3HN	3HA	w
3HN	3HB	m
3HA	3HB	s
3HA	4HN	s
4HN	4HA	w
4HN	4HB	m
4HA	4HB	s
4HA	5HN	s
5HN	5HA	w
5HN	5HB	m
5HA	5HB	s
5HA	6HN	s
6HN	6HA	w
6HN	6HB	m
6HA	6HB	s

* HA, α -proton; HN, *N*-methyl protons; HB, β -protons.

** m, medium; s, strong; w, weak.

Table S10 Chemical shift table of LLDDLL hexamer

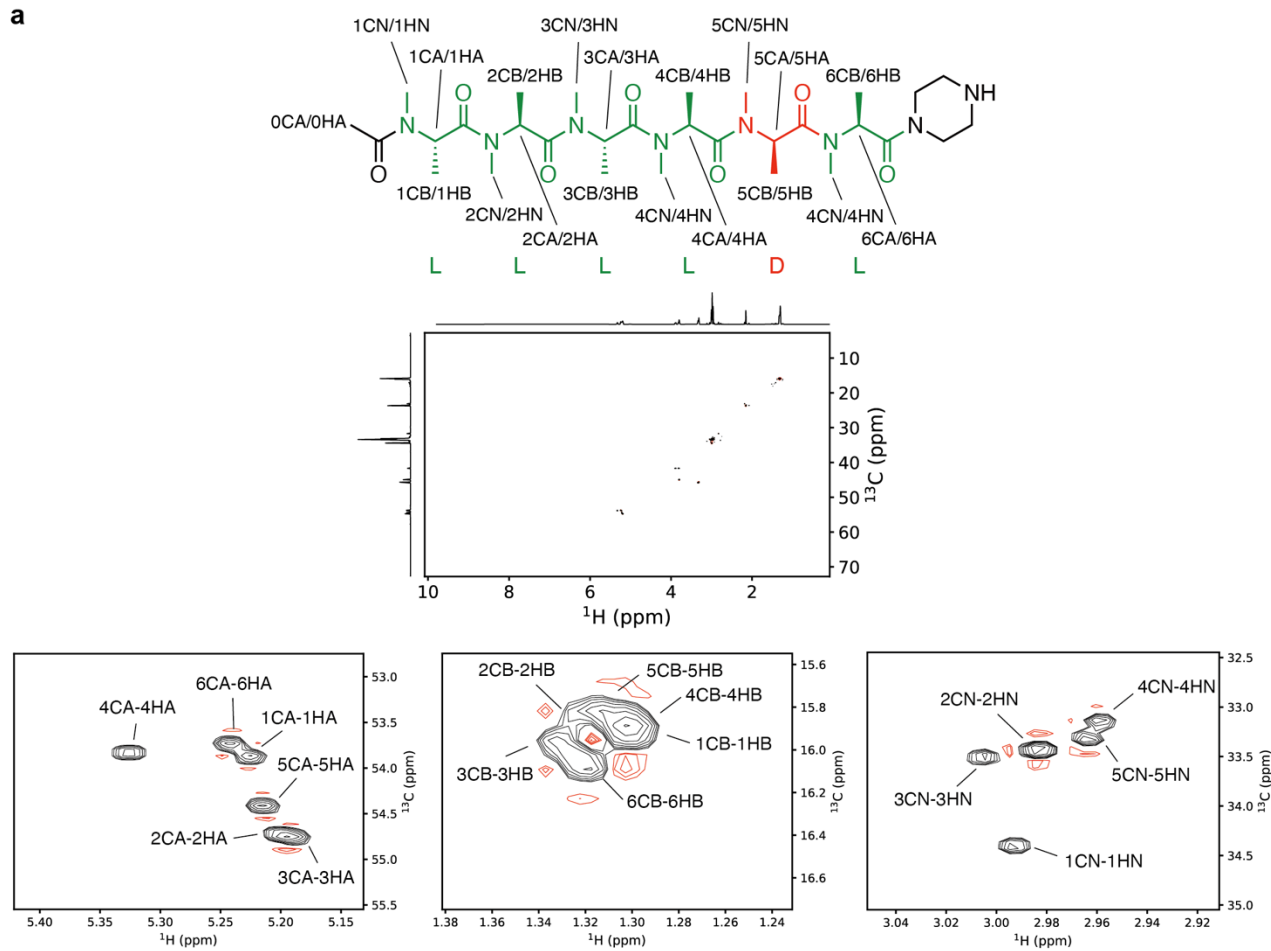
Residue No.	Atom *	Nuclei	Chemical shift (ppm)	Residue No.	Atom *	Nuclei	Chemical shift (ppm)
0	C	¹³ C	176.8	5	C	¹³ C	175.6
0	CA	¹³ C	23.7	5	CA	¹³ C	54.4
0	HA	¹ H	2.16	5	CB	¹³ C	16
1	C	¹³ C	176.1	5	CN	¹³ C	33.3
1	CA	¹³ C	53.9	5	HA	¹ H	5.22
1	CB	¹³ C	16.1	5	HB	¹ H	1.34
1	CN	¹³ C	34.5	5	HN	¹ H	3
1	HA	¹ H	5.23	6	C	¹³ C	173.3
1	HB	¹ H	1.32	6	CA	¹³ C	53.5
1	HN	¹ H	3	6	CB	¹³ C	16.3
2	C	¹³ C	175.9	6	CN	¹³ C	33.2
2	CA	¹³ C	54	6	CX	¹³ C	49.5
2	CB	¹³ C	15.9	6	CY	¹³ C	46.7
2	CN	¹³ C	33.3	6	HA	¹ H	5.33
2	HA	¹ H	5.31	6	HB	¹ H	1.28
2	HB	¹ H	1.32	6	HN	¹ H	2.98
2	HN	¹ H	2.97	6	HX1	¹ H	3.35
3	C	¹³ C	175.7	6	HX2	¹ H	3.46
3	CA	¹³ C	54.5	6	HY1	¹ H	3.4
3	CB	¹³ C	16	6	HY2	¹ H	3.67
3	CN	¹³ C	33.3				
3	HA	¹ H	5.23				
3	HB	¹ H	1.33				
3	HN	¹ H	3				
4	C	¹³ C	175.8				
4	CA	¹³ C	53.9				
4	CB	¹³ C	16				
4	CN	¹³ C	33.3				
4	HA	¹ H	5.32				
4	HB	¹ H	1.31				
4	HN	¹ H	2.95				

* See the atom names on the chemical structure in Fig. S15.

Table S11 Summary of XFEL analysis of LLDDL

	LLDDL
Chemical formula	C ₃₅ H ₅₅ N ₉ O ₉ ·H ₂ O
Formula weight	763.89
Crystal system, space group	Monoclinic, <i>P</i> 2 ₁
Temperature (K)	Room temperature
<i>a</i> , <i>b</i> , <i>c</i> (Å)	16.67(3), 6.012(9), 40.37(7)
β (°)	91.8(11)
<i>V</i> (Å ³)	4044(12)
<i>Z</i>	4
<i>D</i> _{calc} (g/cm ³)	1.255
Radiation type	XFEL
Photon energy (keV)	15.0
No. of collected images	729,635
No. of hit images	140,435
No. of indexed images	26,709
<i>R</i> _{split} (%)	7.84
<i>R</i> ₁ (<i>F</i> >4σ(<i>F</i>)), <i>wR</i> ₂ (all <i>F</i>)	0.1069, 0.3311
No. of reflections	6450
No. of parameters	971
CCDC number	2251714

a



b

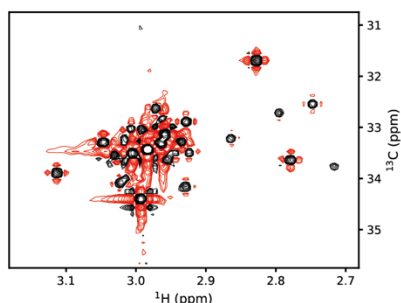


Fig. S17 HSQC spectrum of LLLLDL hexamer. The HSQC spectrum was recorded in D₂O. Black and red contours indicate positive and negative values, respectively. (a) The structure and the HSQC spectrum of the hexamer. Enlarged views of the HSQC spectrum are shown below the entire spectrum. The name of the carbon/proton is labeled on the chemical structure and around each peak. (b) An enlarged view of the region corresponding to *N*-substituted methyl group. The spectrum was used for calculating the population of all-*trans* state. The basal contour level was set to be 50 times lower than that of the other spectra. The population of the all-*trans* state (74%) was calculated as the ratio of the total positive peak volume of the resonances from the *N*-substituted methyl groups in the all-*trans* state to the total positive peak volume of all resonances from the *N*-substituted methyl groups.

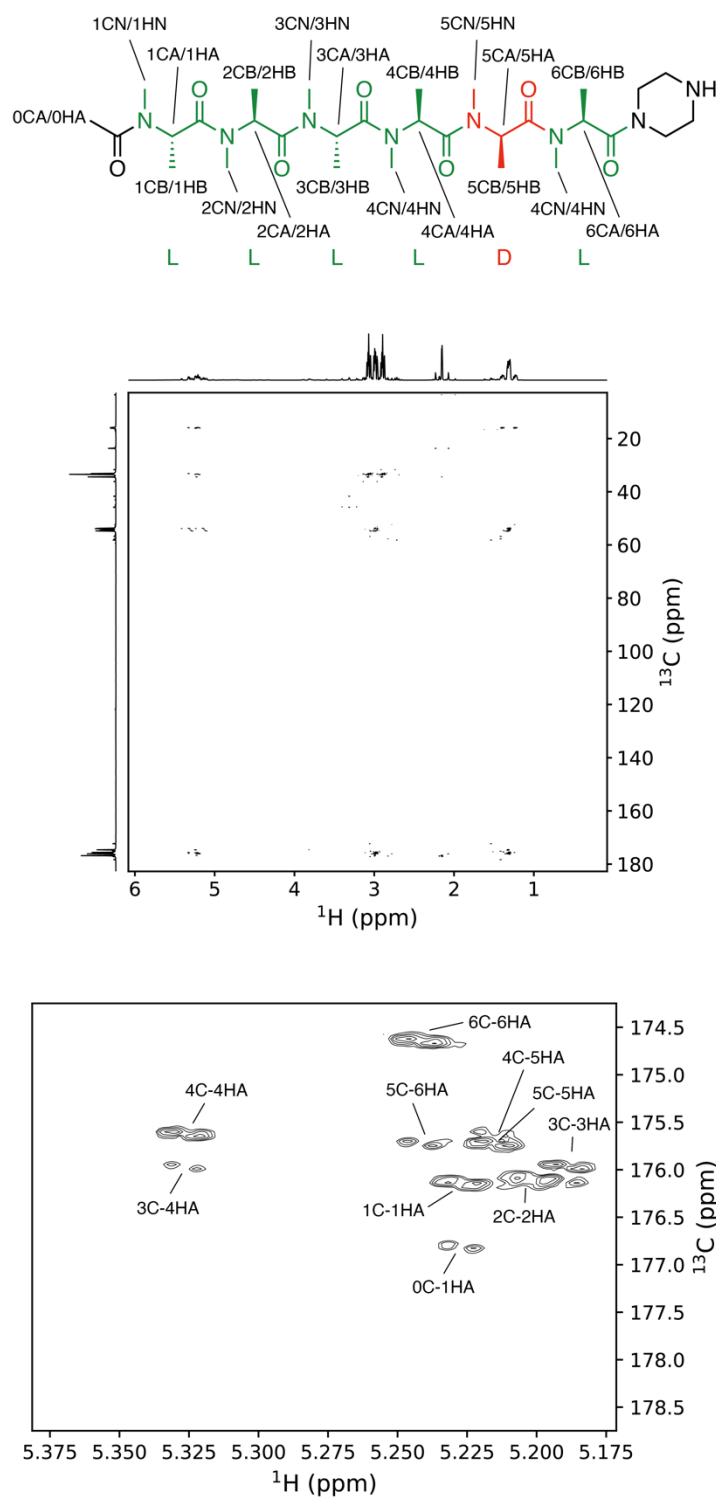


Fig. S18 HMBC spectrum of LLLLDL hexamer. The HMBC spectrum was recorded in D_2O . An enlarged view of the HMBC spectrum is shown at the bottom. The name of the carbon/proton is labeled on the chemical structure and above each peak.

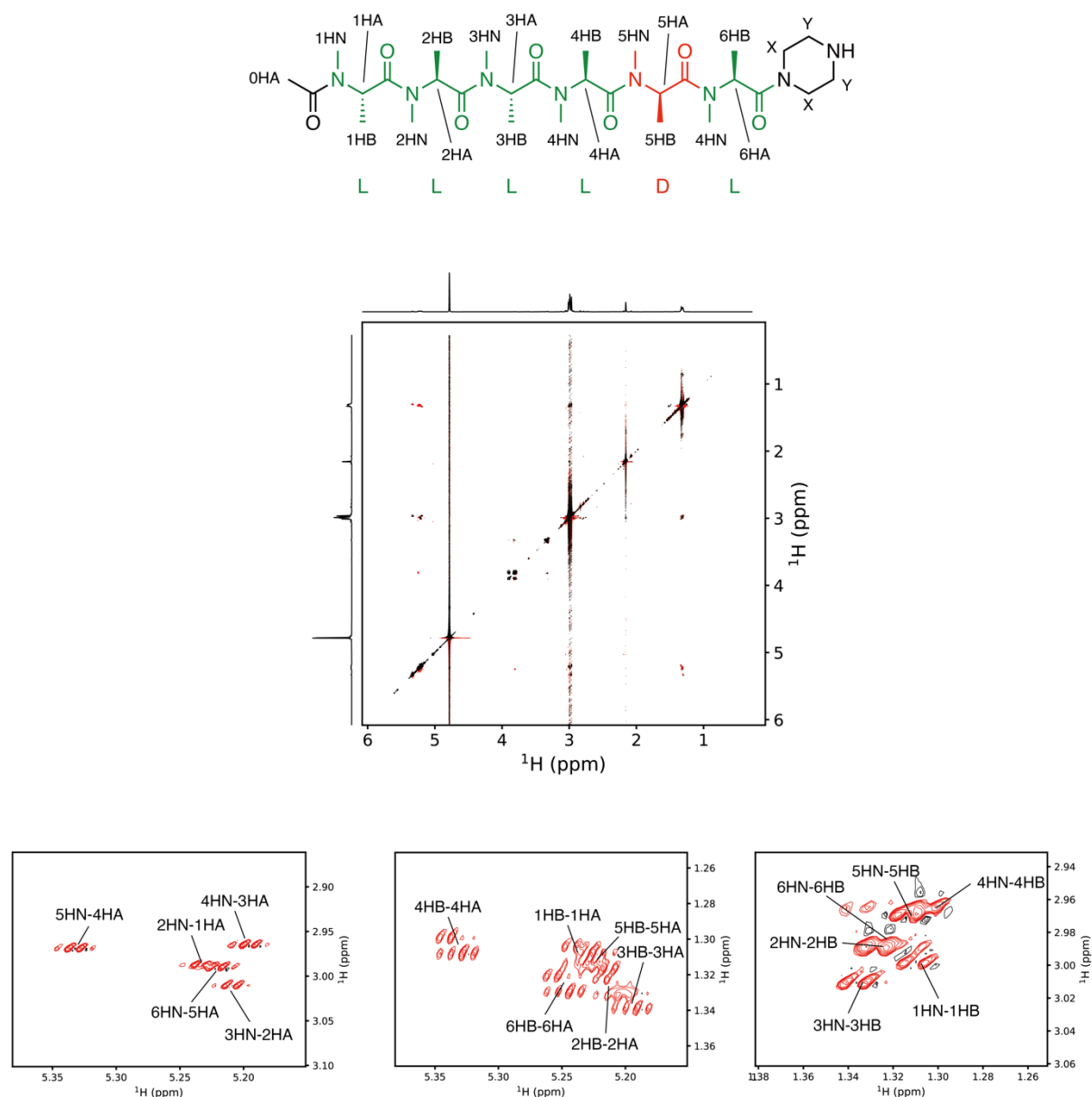


Fig. S19 ROESY spectrum of LLLLDL hexamer. The ROESY spectrum was recorded in D_2O . Black and red contours indicate positive and negative values, respectively. Enlarged views of the ROESY spectrum are shown at the bottom. The name of the proton is labeled on the chemical structure and around each peak.

Table S12 NOE summaries of LLLLDL hexamer

Atom 1 *	Atom 2 *	Intensity **
1HN	1HB	m
1HA	1HB	s
1HA	2HN	s
2HN	2HB	m
2HA	2HB	s
2HA	3HN	s
3HN	3HB	m
3HA	3HB	s
3HA	4HN	s
4HN	4HB	m
4HA	4HB	s
4HA	5HN	s
5HN	5HB	m
5HA	5HB	s
5HA	6HN	s
6HN	6HB	m
6HA	6HB	s

* HA, α -proton; HN, *N*-methyl protons; HB, β -protons.

** m, medium; s, strong.

Table S13 Chemical shift table of LLLLDL hexamer

Residue No.	Atom *	Nuclei	Chemical shift (ppm)	Residue No.	Atom *	Nuclei	Chemical shift (ppm)
0	C	¹³ C	176.8	5	C	¹³ C	175.7
0	CA	¹³ C	23.7	5	CA	¹³ C	54.4
0	HA	¹ H	2.16	5	CB	¹³ C	15.9
1	C	¹³ C	176.1	5	CN	¹³ C	33.4
1	CA	¹³ C	53.9	5	HA	¹ H	5.22
1	CB	¹³ C	16	5	HB	¹ H	1.31
1	CN	¹³ C	34.4	5	HN	¹ H	2.97
1	HA	¹ H	5.23	6	C	¹³ C	174.6
1	HB	¹ H	1.31	6	CA	¹³ C	53.8
1	HN	¹ H	3	6	CB	¹³ C	16.1
2	C	¹³ C	176.1	6	CN	¹³ C	33.5
2	CA	¹³ C	54.8	6	CX	¹³ C	41.7
2	CB	¹³ C	15.9	6	CX1	¹³ C	41.7
2	CN	¹³ C	33.5	6	CX2	¹³ C	45
2	HA	¹ H	5.21	6	CY	¹³ C	45.7
2	HB	¹ H	1.33	6	CY1	¹³ C	45.8
2	HN	¹ H	2.99	6	CY2	¹³ C	5.25
3	C	¹³ C	176	6	HA	¹ H	1.33
3	CA	¹³ C	54.8	6	HB	¹ H	2.99
3	CB	¹³ C	16	6	HN	¹ H	3.81
3	CN	¹³ C	33.5	6	HX11	¹ H	3.89
3	HA	¹ H	5.19	6	HX12	¹ H	3.8
3	HB	¹ H	1.33	6	HX2	¹ H	3.31
3	HN	¹ H	3.01	6	HY1	¹ H	3.33
4	C	¹³ C	175.6	6	HY2	¹ H	175.7
4	CA	¹³ C	53.9				
4	CB	¹³ C	15.9				
4	CN	¹³ C	33.2				
4	HA	¹ H	5.33				
4	HB	¹ H	1.3				
4	HN	¹ H	2.96				

* See the atom names on the chemical structure in Figure S18.

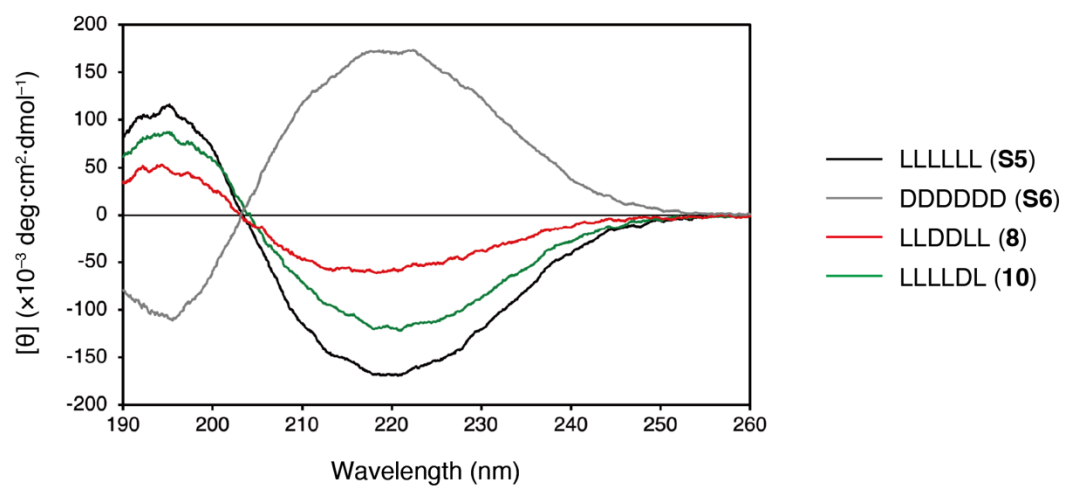


Fig. S20 Circular dichroism (CD) spectra of oligo-NMAs. The measurements were conducted using 100 μM oligo-NMAs in 10 mM phosphate buffer (pH 7.2).

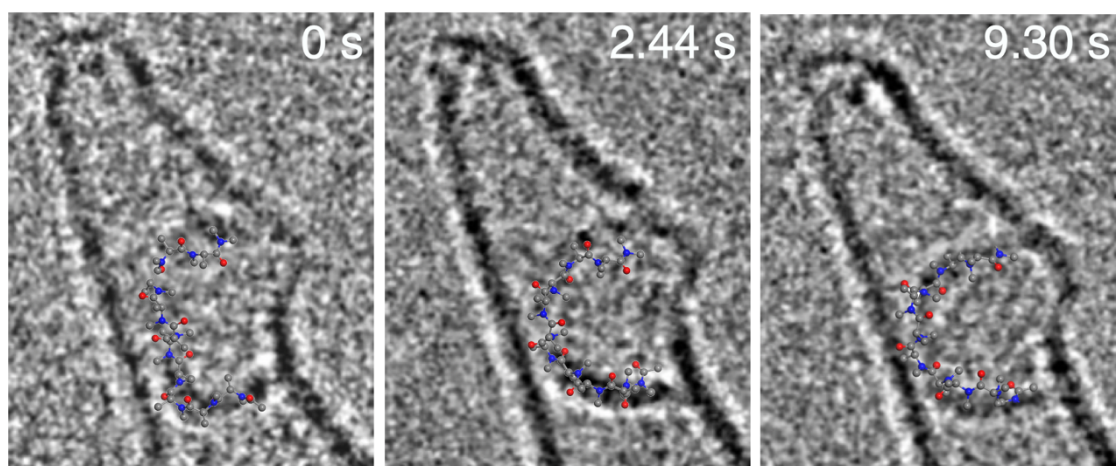


Fig. S21 The overlays of the TEM images of LLDDLDDLL and molecular models from MD simulations.

Scale bar is 1 nm.

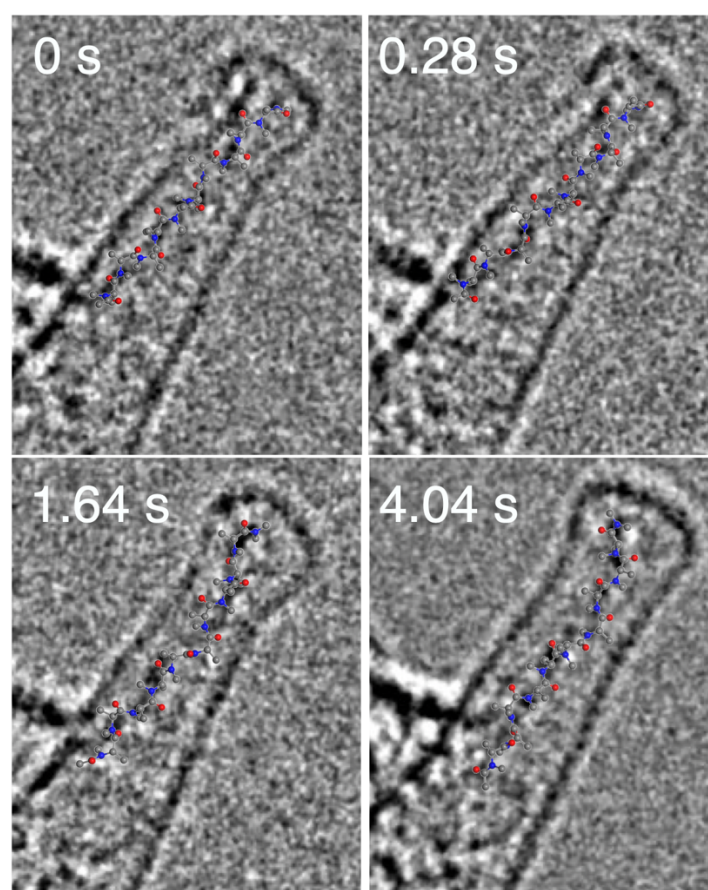
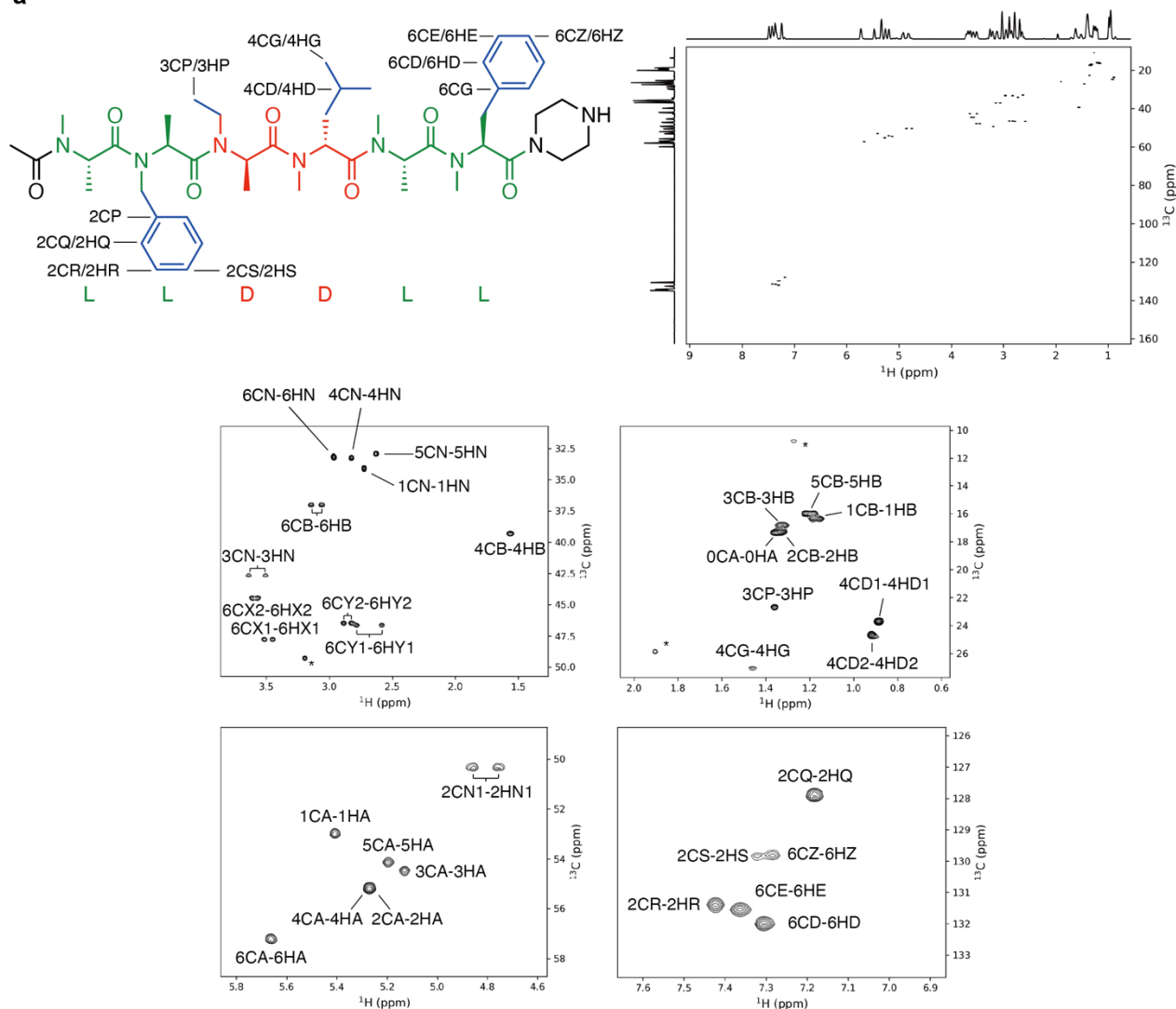


Fig. S22 The overlays of the TEM images of LLLLLLLLLL and molecular models from MD simulations.

Scale bar is 1 nm.

a



b

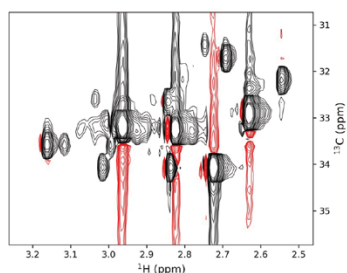


Fig. S23 HSQC spectrum of functionalized LLDDLL hexamer. The HSQC spectrum was recorded in D₂O. Black and red contours indicate positive and negative values, respectively. (a) The structure and the HSQC spectrum of the hexamer. Enlarged views of the HSQC spectrum are shown below the entire spectrum. The name of the carbon/proton is labeled on the chemical structure and around each peak. The name of the backbone carbon/proton are the same with Figure S14. (b) An enlarged view of the region corresponding to *N*-substituted methyl group. The spectrum was used for calculating the population of all-*trans* state. The basal contour level was set to be 50 times lower than that of the other spectra. The population of the all-*trans* state (77%) was calculated as the ratio of the total positive peak volume of the resonances from the *N*-substituted methyl groups in the all-*trans* state to the total positive peak volume of all resonances from the *N*-substituted methyl groups.

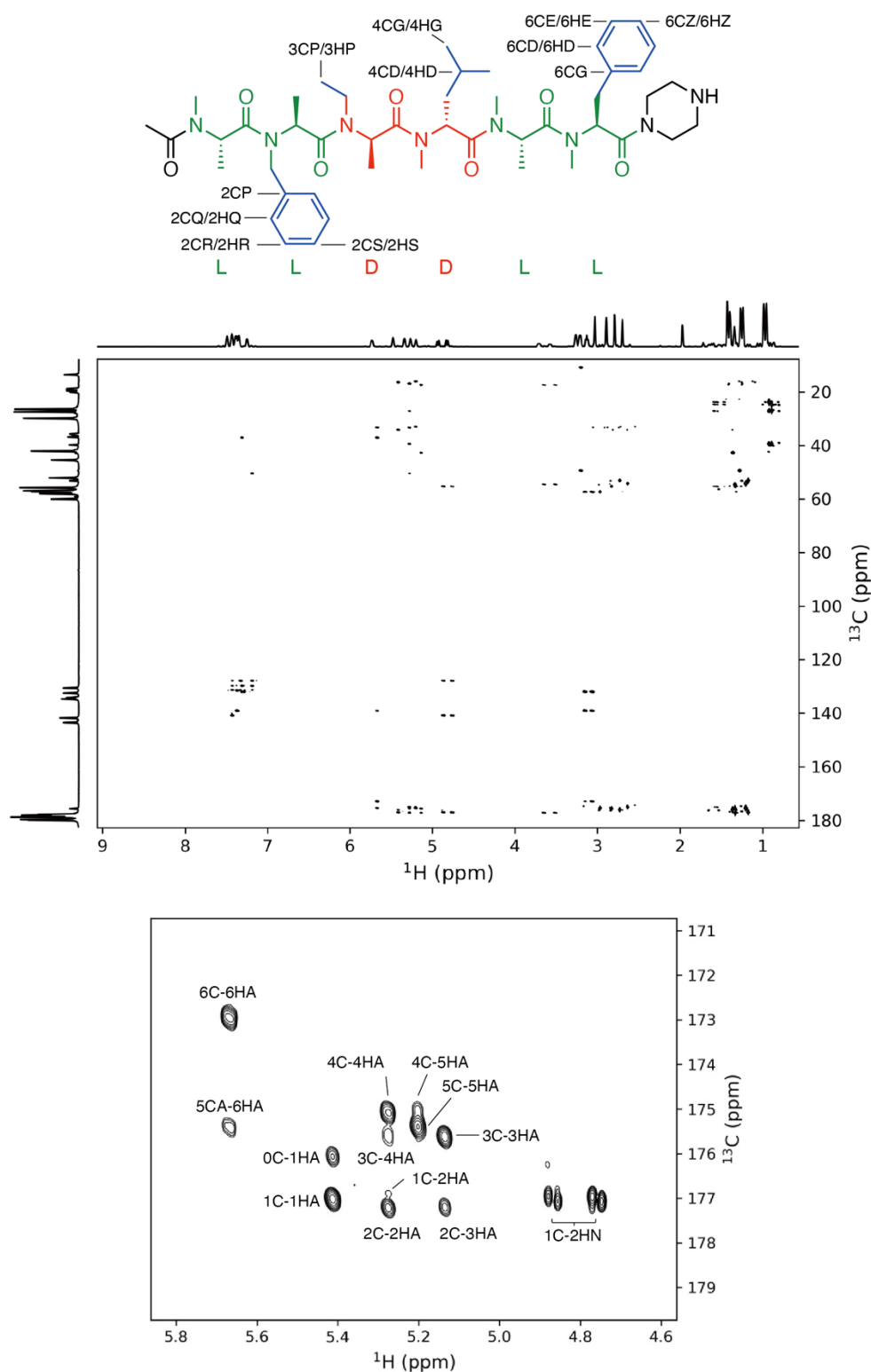


Fig. S24 HMBC spectrum of functionalized LLDDLL hexamer. The HMBC spectrum was recorded in D₂O. An enlarged view of the HMBC spectrum is shown at the bottom. The name of the carbon/proton is labeled on the chemical structure and above each peak.

Table S14 NOE summaries of functionalized LLDDLL hexamer

Atom 1 *	Atom 2 *	Intensity **	Atom 1 *	Atom 2 *	Intensity **
0HA	1HN	s	4HD2	4HA	m
1HA	2HQ	w	4HD2	4HN	w
1HB	1HA	s	4HG	4HN	m
1HB	1HN	s	4HN	3HA	s
1HN	1HA	w	4HN	4HA	w
1HN	2HQ	m	5HB	5HA	s
1HN	2HR	w	5HB	5HN	s
2HA	2HQ	m	5HB	6HN	w
2HB	2HA	s	5HN	4HA	s
2HB	2HN1	m	5HN	5HA	w
2HB	2HN2	w	5HN	6HD	w
2HB	2HQ	s	5HN	6HE	w
2HB	2HR	w	6HA	6HD	s
2HB	2HS	w	6HB1	6HD	m
2HN1	2HQ	m	6HB2	6HD	m
2HN2	1HA	s	6HN	5HA	s
2HN2	2HQ	m	6HN	6HA	w
2HQ	2HR	s	6HN	6HD	w
2HQ	2HS	m	6HN	6HB1	s
3HB	3HA	s	6HN	6HB2	m
3HB	3HN1	s	6HX11	6HA	s
3HB	3HN2	m	6HX11	6HX22	m
3HB	4HN	m	6HX12	6HA	s
3HN1	2HA	m	6HX21	6HA	w
3HN1	3HA	w	6HX22	6HA	w
3HN2	2HA	m	6HY11	6HX11	w
3HN2	3HA	w	6HY11	6HX12	w
3HP	3HN1	m	6HY11	6HX21	w
3HP	3HN2	w	6HY11	6HX22	w
4HB	4HA	m	6HY12	6HX11	w
4HB	4HN	s	6HY12	6HX12	w
4HB	5HN	m	6HY21	6HD	w
4HD1	4HA	s	6HY21	6HE	w
4HD1	5HN	w	6HY22	6HX11	w

* The name of the proton is defined on the chemical structure shown in Fig. S22.

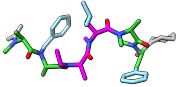
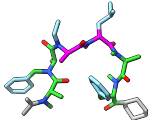
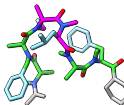
** w, weak; m, medium; s, strong.

Table S15 Chemical shift table of functionalized LLDDLL hexamer

Residue No.	Atom *	Nuclei	Chemical shift (ppm)	Residue No.	Atom *	Nuclei	Chemical shift (ppm)
0	C	¹³ C	176.1	4	HD12	¹ H	0.89
0	CA	¹³ C	17.5	4	HD22	¹ H	0.92
0	HA	¹ H	1.36	4	HN	¹ H	2.83
1	C	¹³ C	177	5	C	¹³ C	175.4
1	CA	¹³ C	53.1	5	CA	¹³ C	54.2
1	CB	¹³ C	16.5	5	CB	¹³ C	16.2
1	CN	¹³ C	34.2	5	CN	¹³ C	33
1	HA	¹ H	5.41	5	HA	¹ H	5.2
1	HB	¹ H	1.17	5	HB	¹ H	1.2
1	HN	¹ H	2.73	5	HN	¹ H	2.63
2	C	¹³ C	177.2	6	C	¹³ C	173
2	CA	¹³ C	55.3	6	CA	¹³ C	57.3
2	CB	¹³ C	17	6	CB	¹³ C	37.2
2	CN	¹³ C	50.4	6	CG	¹³ C	139.1
2	CP	¹³ C	140.9	6	CD	¹³ C	132
2	CQ	¹³ C	128	6	CE	¹³ C	131.6
2	CR	¹³ C	131.4	6	CO	¹³ C	33.3
2	CS	¹³ C	130.3	6	CX1	¹³ C	47.9
2	HA	¹ H	5.28	6	CX2	¹³ C	44.5
2	HB	¹ H	1.33	6	CY1	¹³ C	46.7
2	HN1	¹ H	4.76	6	CY2	¹³ C	46.6
2	HN2	¹ H	4.86	6	CZ	¹³ C	129.8
2	HQ	¹ H	7.19	6	HA	¹ H	5.67
2	HR	¹ H	7.43	6	HB1	¹ H	3.07
2	HS	¹ H	7.32	6	HB2	¹ H	3.15
3	C	¹³ C	175.6	6	HD	¹ H	7.31
3	CA	¹³ C	54.6	6	HE	¹ H	7.37
3	CB	¹³ C	17.5	6	HN	¹ H	2.97
3	CN	¹³ C	42.8	6	HX11	¹ H	3.46
3	CP	¹³ C	22.8	6	HX12	¹ H	3.52
3	HA	¹ H	5.14	6	HX21	¹ H	3.57
3	HB	¹ H	1.33	6	HX22	¹ H	3.61
3	HN1	¹ H	3.51	6	HY11	¹ H	2.59
3	HN2	¹ H	3.64	6	HY12	¹ H	2.78
3	HP	¹ H	1.36	6	HY21	¹ H	2.83
4	C	¹³ C	175.1	6	HY22	¹ H	2.89
4	CA	¹³ C	55.3	6	HZ	¹ H	7.28
4	CB	¹³ C	39.4				
4	CG	¹³ C	27.2				
4	CD1	¹³ C	23.9				
4	CD2	¹³ C	24.8				
4	CN	¹³ C	33.4				
4	HA	¹ H	5.27				
4	HB1	¹ H	1.57				
4	HB2	¹ H	1.57				
4	HG	¹ H	1.47				

* See the atom names on the chemical structure in Fig. S22.

Table S16 Representative conformers of the 10 clusters from McMD simulations of functionalized LLDDL

Cluster	Representative conformer ^a	Population (%)	Cluster	Representative conformer	Population (%)
1		70.9	6		0.8
2		16.8	7		0.7
3		6.5	8		0.3
4		1.8	9		0.3
5		1.7	10		0.1

^a Representative conformer from the population of 5% or more are shown in the table.

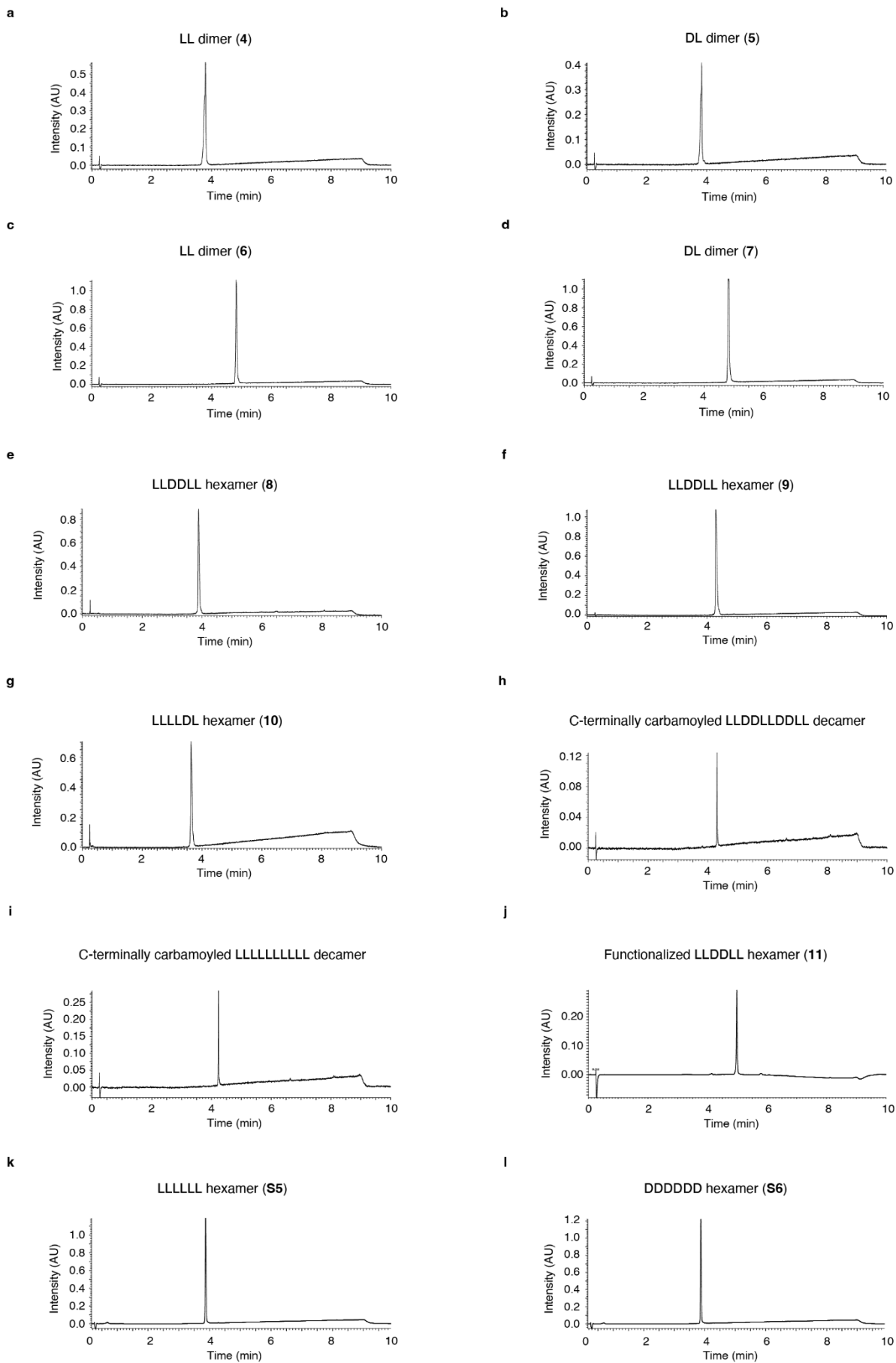


Fig. S26 UV chromatograms of the purified monomer and oligomers of *N*-methylalanine. UV chromatograms of the purified compounds. For (a)–(g), (k), and (l), the following linear gradient was used: 0–2 min, B 1%; 2–7 min, B 1–95%; 7–8 min, B 95%; 8–10 min, B 1%. For (h)–(j), the following linear gradient was used: 0–1 min, B 5%; 1–8 min, B 5–95%; 8–10 min, B 5%.

Legends for Supporting Video

Video S1 Video imaging of LLDDLLDDLL captured by OneView-IS

This is the video of LLDDLLDDLL captured by OneView-IS (Figure 4f and Figure S19). The experimental conditions were an acceleration voltage of 80 kV, EDR of $3.8 \times 10^6 \text{ e}^- \text{ nm}^{-2} \text{ s}^{-1}$, and an exposure time of 20 milliseconds for each frame. The playback speed of the video is the original video recording. Scale bar is 1 nm.

Video S2 Video imaging of LLLLLLLLLL captured by OneView-IS

This is the video of LLLLLLLLLL captured by OneView-IS (Figure 4g and Figure S20). The experimental conditions were an acceleration voltage of 80 kV, EDR of $4.9 \times 10^6 \text{ e}^- \text{ nm}^{-2} \text{ s}^{-1}$, and an exposure time of 20 milliseconds for each frame. The playback speed of the video is the original video recording. Scale bar is 1 nm.

References

- 1 H. Isobe, T. Tanaka, R. Maeda, E. Noiri, N. Solin, M. Yudasaka, S. Iijima and E. Nakamura, *Angew. Chem., Int. Ed.*, 2006, **45**, 6676–6680.
- 2 W. H. B. Sauer and M. K. Schwarz, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 987–1003.
- 3 J. Morimoto, Y. Fukuda, D. Kuroda, T. Watanabe, F. Yoshida, M. Asada, T. Nakamura, A. Senoo, S. Nagatoishi, K. Tsumoto and S. Sando, *J. Am. Chem. Soc.*, 2019, **141**, 14612–14623.
- 4 Y. Fukuda, M. Yokomine, D. Kuroda, K. Tsumoto, J. Morimoto and S. Sando, *Chem. Sci.*, 2021, **12**, 13292–13300.
- 5 M. Yokomine, J. Morimoto, Y. Fukuda, Y. Shiratori, D. Kuroda, T. Ueda, K. Takeuchi, K. Tsumoto and S. Sando, *Angew. Chem. Int. Ed.*, 2022, **61**, e202200119.
- 6 G. M. J. Barca, C. Berton, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Włoch, P. Xu, F. Zahariev and M. S. Gordon, *J. Chem. Phys.*, 2020, **152**, 154102.
- 7 J. Ikebe, K. Umezawa, N. Kamiya, T. Sugihara, Y. Yonezawa, Y. Takano, H. Nakamura and H. Higo, *J. Comput. Chem.*, 2011, **32**, 1286–1297.
- 8 S. Ono, M. R. Naylor, C. E. Townsend, C. Okumura, O. Okada, H. W. Lee and R. S. Lokey, *J. Chem. Inf. Model*, 2021, **61**, 5601–5613.
- 9 D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, G. A. Cisneros, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, M. C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O’Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C. L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, J. Wang, J. Wang, H. Wei, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D. M. York, S. Zhao and P. A. Kollman, *Amber 2022, University of California, San Francisco*.
- 10 K. Takaba, S. Maki-Yonekura, I. Inoue, K. Tono, T. Hamaguchi, K. Kawakami, H. Naitow, T. Ishikawa, M. Yabashi and K. Yonekura, *Nat. Chem.*, 2023, **15**, 491–497.
- 11 K. Takaba, S. Maki-Yonekura, S. Inoue, T. Hasegawa and K. Yonekura, *Front Mol Biosci*, 2021, **7**, 1–12.
- 12 M. T. B. Clabbers, T. Gruene, J. M. Parkhurst, J. P. Abrahams and D. G. Waterman, *Acta Crystallogr. D Struct. Biol.*, 2018, **74**, 506–518.
- 13 T. A. White, V. Mariani, W. Brehm, O. Yefanov, A. Barty, K. R. Beyerlein, F. Chervinskii, L. Galli, C. Gati, T. Nakane, A. Tolstikova, K. Yamashita, C. H. Yoon, K. Diederichs and H. N. Chapman, *J. Appl. Crystallogr.*, 2016, **49**, 680–689.
- 14 G. M. Sheldrick, *Acta Crystallogr. A*, 2015, **71**, 3–8.
- 15 G. M. Sheldrick, *Acta Crystallogr. C Struct. Chem.*, 2015, **71**, 3–8.

- 16 T. Nakamuro, K. Kamei, K. Sun, J. W. Bode, K. Harano and E. Nakamura, *J. Am. Chem. Soc.*, 2022, **144**, 13612–13622.
- 17 J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak and A. Cardona, *Nat. Methods*, 2012, **9**, 676–682.
- 18 A. O. Galvez, C. P. Schaack, H. Noda and J. W. Bode, *J. Am. Chem. Soc.*, 2017, **139**, 1826–1829.
- 19 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminform.*, 2012, **4**, 1–17.