Supporting information

Linear Consecutive Hexaoxazoles as G4 Ligands Inducing Chair-type Anti-parallel Topology of Telomeric G-Quadruplex

Shogo Sasaki,[a] Yue Ma,[b] Takumi Ishizuka,[c] Hong-Liang Bao,[c] Takatsugu Hirokawa,[d, e, f] Yan Xu,[c] Masayuki Tera,[a] Kazuo Nagasawa[a,*]

a Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, Japan
E-mail: knaga@cc.tuat.ac.jp

b Institute of Global Innovation Research (GIR), Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, Japan

c Division of Chemistry, Department of Medical Sciences, Faculty of Medicine, University of Miyazaki, 5200 Kihara, Kiyotake, Miyazaki 889-1692, Japan

d Transborder Medical Research Center, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, 305-8575, Japan

e Division of Biomedical Science, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8575, Japan

f Cellular and Molecular Biotechnology Research Institute, National Institute of Advanced Industrial Science and Technology, 2-4-7 Aomi, Koto-ward, Tokyo 135-0064, Japan

Table of Contents:

Supplementary Methods

FRET melting assay

The FRET melting assay was performed with an excitation wavelength of 470–505 nm and a detection wavelength of 523–543 nm using the Thermal Cycler Dice Real Time System III (Takara). The dual fluorescently labeled oligonucleotides Flu-telo21 was used in this protocol. The donor fluorophore was 6-carboxyfluorescein (FAM) and the acceptor fluorophore was 6-carboxytetramethyl- rhodamine (TAMRA). All purified nucleotides (Sigma Genosys) were dissolved as stock solutions (100 μM) in MilliQ water to be used without further purification. Further dilutions of the oligonucleotides were performed with 60 mM potassium cacodylate buffer (pH 7.4), and FRET experiments were carried out with a 0.4 μM oligonucleotide solution. Dual-labeled DNA was annealed by heating at 99 °C for 5 min, and then slowly cooled to room temperature. Ligands were prepared as DMSO stock solutions (10 mM) and diluted to 1 mM using DMSO, and then diluted to 100 μM using 60 mM potassium cacodylate buffer (pH 7.4). Next, the annealed DNA (20 μL) and the compound solution (20 μL) were distributed across 96-well plates (Takara), with a total volume of 40 μL, with the labeled oligonucleotide (0.2 μM) and the compound (1.0 μM). The plates were incubated at 25 °C for 12 h. Subsequent experiments used the following temperature procedure in RT-PCR, finishing as follows: 25 °C for 20 min, and then a stepwise increase of 1 °C every minute from 25 °C until 99 °C. During the procedures, we measured the FAM fluorescence after each step. The change in the melting temperature at 1.0 μM compound concentration—ΔT_m (1.0 μM)—was calculated from at least three experiments by subtraction of the blank from the averaged melting temperature of each compound.

Circular dichroism (CD) spectrometry

CD spectra were recorded on a J-720 spectropolarimeter (JASCO, Tokyo, Japan) using a quartz cell of 1 mm optional path length and an instrument scanning speed of 500 nm min-1 with a response time of 1 s, and over a wavelength range of 220–320 nm. All purified nucleotides (Sigma Genosys) used in this protocol (Table S1) were dissolved as a 1.0 mM stock solution in MilliQ water to be used without further purification. Further dilution of the nucleotide was 50 mM Tris-HCl buffer (with 100 mM KCl, 100 mM NaCl) from 1 mM stock solutions to give a concentration of 10 μM. The solution was annealed by heating at 95 °C for 5 min, and then slowly cooled to room temperature, and then titrated into the oligonucleotide samples up to 0-5 equivalents using

ligands, and incubated overnight. Finally, CD spectra are representative of five averaged scans taken at 25 °C.

2-Ap-based fluorescence quenching analysis

The Fluorescent of 2-aminopurine quenching analysis was performed with an excitation wavelength of 300 nm and a detection wavelength of 320-500 nm using the FP-8600 (JASCO, Tokyo, Japan). The labeled oligonucleotides 22AG-Ap (table S1) were used in this protocol, and were dissolved as stock solutions (100 μM) in MilliQ water to be used without further purification. Further dilutions of the oligonucleotides were performed with 50 mM Tris-HCl buffer (pH 7.4) with 100 mM KCl, and this experiment was carried out with a 2.0 μM oligonucleotide solution. 22AG-Ap was annealed by heating at 95 °C for 5 min, and then slowly cooled to room temperature. The ligands were prepared as DMSO stock solutions (10 mM) and diluted to 100 μM using 50 mM Tris-HCl buffer (pH 7.4) with 100 mM KCl. The compound solution (100 μM, 10 μL) was added to the annealing DNA solution with a total volume 1mL, with the labeled oligonucleotide (2.0 μM) and the compound (4.0 μM). The solution was incubated at 25 °C for 12 h. Finally, we measured the fluorescent of aminopurine at least three times at 25 °C, and the spectra are representative of three averaged scans taken.

Computational analysis

For initial coordinates of L2H2-2M2EA-6LCO (**2**), ionization and energy minimization were performed with the OPLS3e force field in the LigPrep script in the Maestro (Schrödinger, LLC, New York, NY, USA). These minimized structures were employed as input structures for docking simulations. The human telomeric G4 structures with two types of topologies, chair-type (PDB ID: 6JKN) and basket-type (PDB ID: 143D), were refined for docking simulations using constrained energy refinements in the OPLS3e force field (Schrödinger LLC). Docking simulations were performed using the Glide XP[1] docking program (Schrödinger, LLC, New York, NY, USA). Finally, ligand binding free energy of the all poses was calculated and re-ranked using the MM-GBSA (Schrödinger, LLC, New York, NY, USA).

Oligonucleotides Sequences

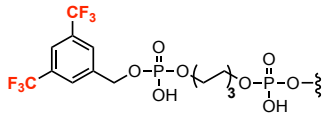
Table S1. Sequences of oligonucleotides used in this paper

Oligonucleotides	Sequence	G4 structure (with K^+)
Flu-telo21	FAM-d[GGG TTA GGG TTA GGG TTA GGG]-TAMRA	Mixture
22AG	d[AGG GTT AGG GTT AGG GTT AGG G]	Mixture
23TAG	d[TAG GGT TAG GGT TAG GGT TAG GG]	Hybrid-1
24TTA	d[TTA GGG TTA GGG TTA GGG TTA GGG]	Mixture
25TAG	d[TAG GGT TAG GGT TAG GGT TAG GGT A]	Hybrid-2
^{19}F-22AG [a]	^{19}F-d[AGG GTT AGG GTT AGG GTT AGG G]	Mixture
22AG-Ap	d[AGG GTT AGG GTT *Ap*GG GTT AGG G]	Mixture
dsDNA	d[TAT AGC TAT ATT TTT TTA TAG CTA TA]	-

a: DNA oligonucleotide bearing 3,5-bis(trifluoro methyl)phenyl moiety at the 5' terminal

^{19}F-22AG : ^{19}F- 5'-AGG GTT AGG GTT AGG GTT AGG G-3'

^{19}F =



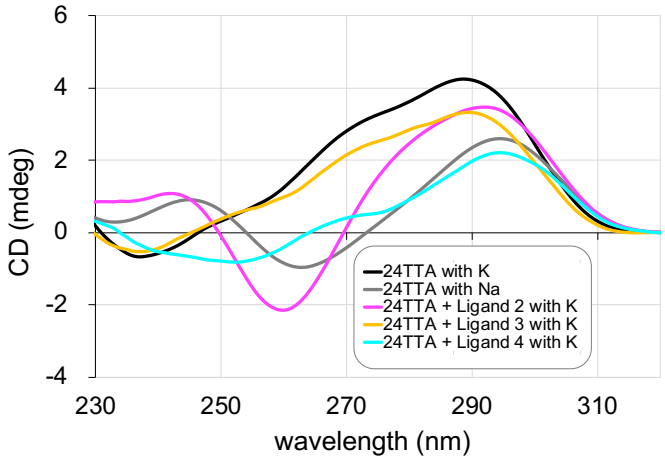


Fig. S1 The CD spectra of telomeric DNA (24TTA; 2 μM) with Na^+ (gray) and K^+ (black), and plus 10 μM of L2H2-2M2EA-6LCO (**2**; magenta), L2A2-2M2EAc-6LCO (**3**; yellow) and L2G2-2M2EG-6LCO (**4**; cyan) with K^+, respectively.

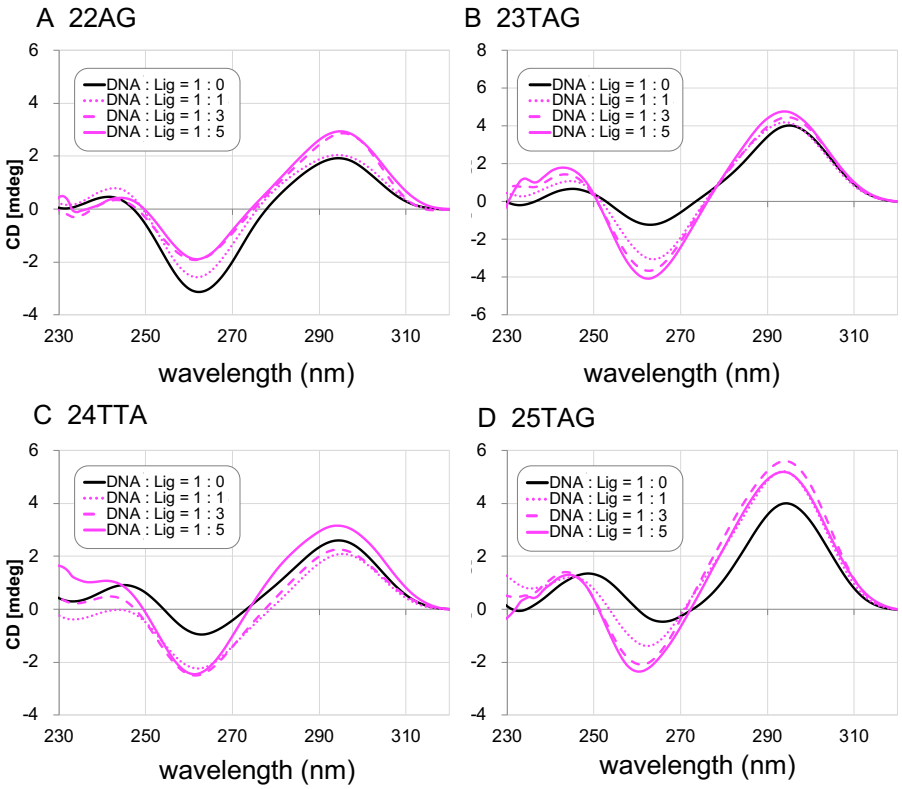


Fig. S2 CD spectra of telomeric DNA (2 μM) in 50 mM Tris-HCl with 100 mM NaCl (pH 7.5) in the presence of 0-5 equivalents of L2H2-2M2EA-6LCO (**2**); (A) 22AG, (B) 23TAG; (C) 24TTA, (D) 25TAG, respectively.

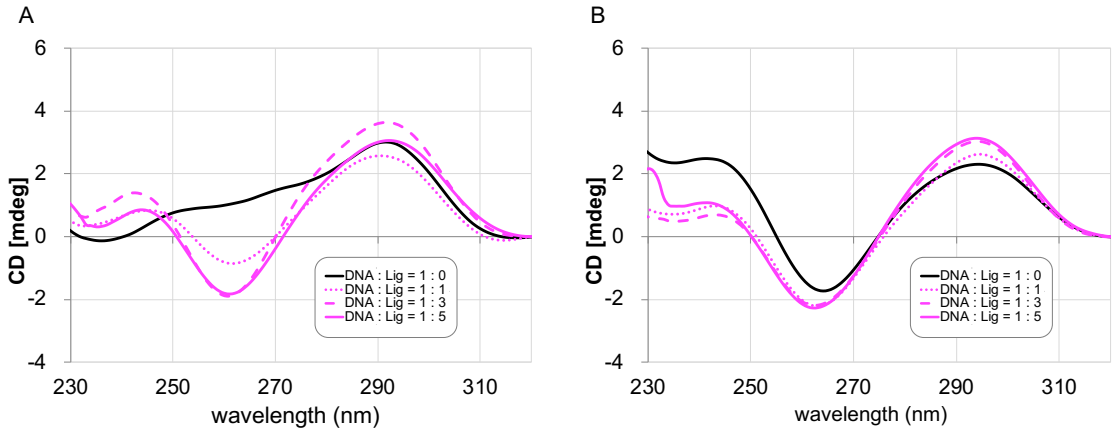


Fig. S3 CD spectra of 22AG-Ap (2 μM) in 50 mM Tris-HCl with (A) 100 mM KCl or (B) 100 mM NaCl (pH 7.5) in the presence of 0-5 equivalents of L2H2-2M2EA-6LCO (**2**).

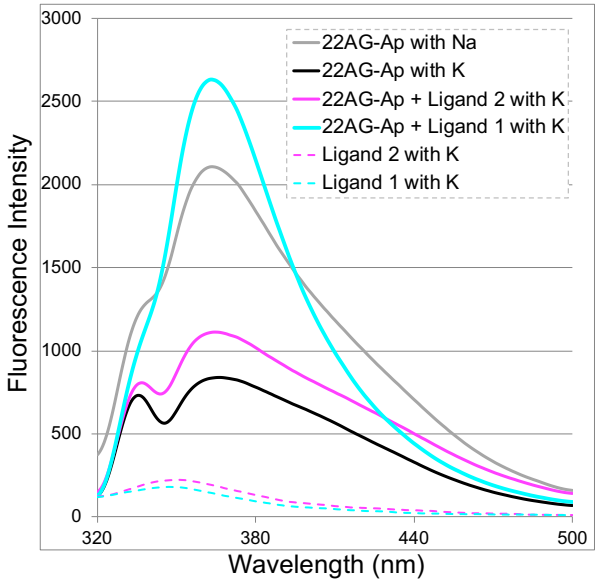


Fig. S4 2-Ap-based fluorescence quenching analysis in 22AG. Fluorescence spectra of 22AG-Ap (2 μM) with Na^+ (gray), or K^+ plus 4 μM ligand **2** (magenta) and ligand **1** (cyan), or 4 μM ligand **2** (dotted magenta) and ligand **1** (dotted cyan).

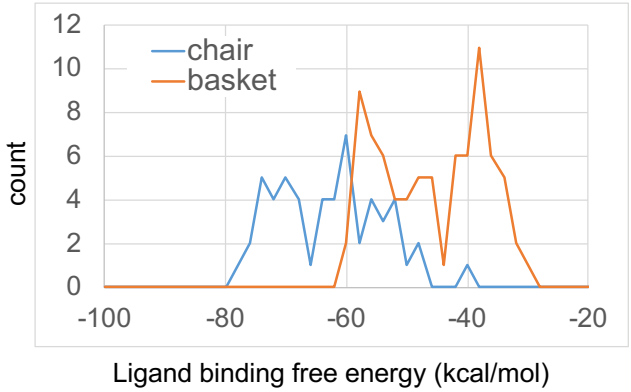


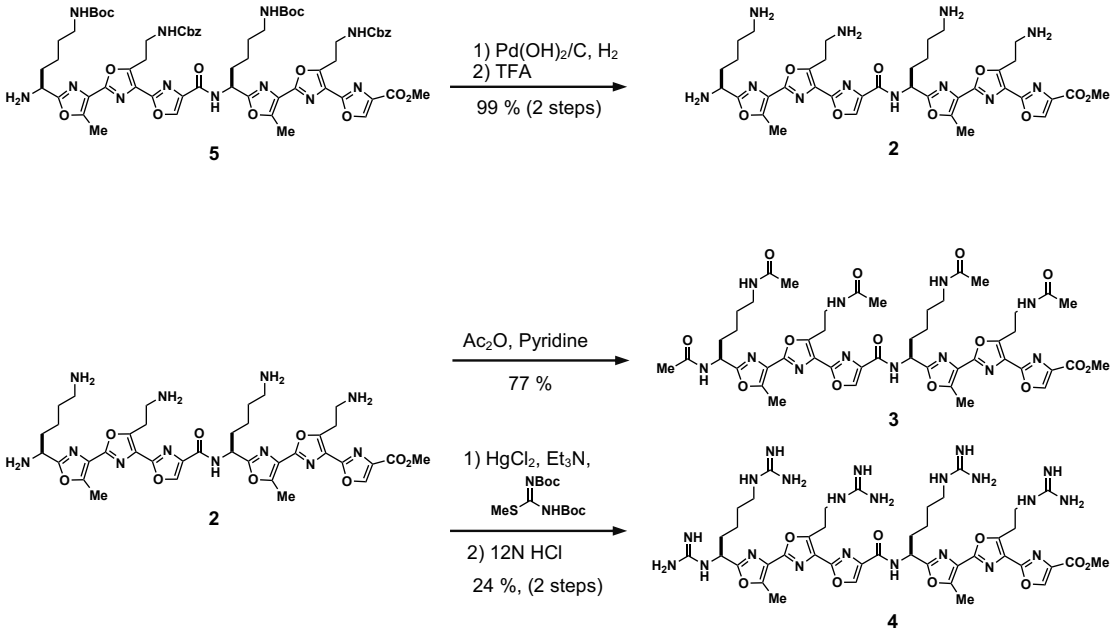
Fig. S5 The ligand binding free energy of the complex of a chaie-type anti-parallel G4 (cyan) and a basket-type anti-parallel G4 (orange) with **2**.

Synthesis and characterization

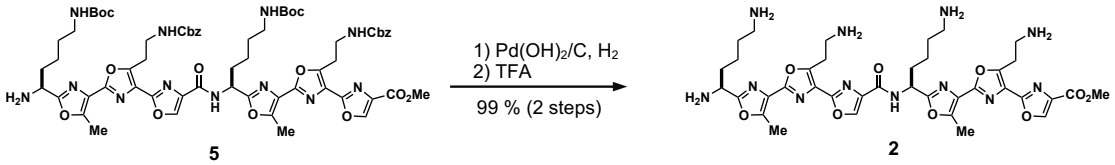
Instrumentation

Optical rotations were measured on a JASCO P-2200 polarimeter. ^{1}H and ^{13}C NMR spectra were recorded on JEOL JNM-ECA 500. The spectra are referenced internally according to the residue solvent signals of DMSO-d_6 (^{1}H NMR; d = 2.50 ppm, ^{13}C NMR; d = 40 ppm). Date for ^{1}H NMR are recorded as follow; chemical shift (δ, ppm), multiplicity (s, singlet; t, triplet; m, multiplet; br. broad), integration, coupling constant (Hz). Date for ^{13}C NMR are recorded in terms of chemical shift (δ, ppm). Mass spectra were recorded on JEOL JMS-T100LC spectrometer with ESI-MS mode using MeOH as solvent.

Synthesis of the compounds **2**, **3** and **4**

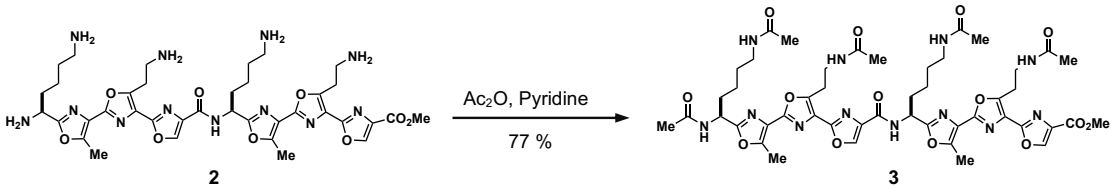


Scheme S1. Synthesis of L2H2-2M2EA-6LCO (**2**), L2A2-2M2EAc-6LCO (**3**) and L2G2 -2M2EG-6LCO (**4**)



5 2

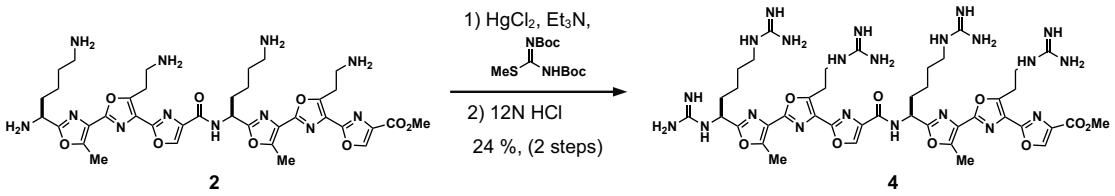
L2H2-2M2EA-6LCO (**2**): To a solution of **5**[2] (30 mg, 26.3 μmol) in THF-MeOH (1:1, 1mL) was added $Pd(OH)_2/C$ (20 mg, 67 wt%) and the reaction mixture was stirred at room temperature under hydrogen atmosphere. After stirring for 30 min, the mixture was filtered through a pad of Celite® and eluted with MeOH, and the solution was concentrated *in vacuo*. To a solution of the crude compound in CH_2Cl_2 (1 mL) was added TFA (0.5 mL) at 0 ºC. The reaction mixture was stirred for 30 min at this temperature, and then concentrated *in vacuo* to give **2** (21mg, 99 %); $[a]^{25}_D$ = -1.2 (c 1.2 MeOH); ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.20-8.81 (m, 4H), 8.81 (s, 1H), 8.31-8.08 (m, 6H), 8.08-7.78 (m, 6H), 4.50-4.60 (dt, J = 6.0, 14.5 Hz, 1H), 4.62 (t, J = 7.0 Hz,1H), 3.85 (s, 3H), 3.64-3.49 (m, 4H), 2.85-2.74 (m, 4H), 2.72 (s, 3H), 2.68 (s, 3H), 2.16-1.94 (m, 4H), 1.68-1.29 (m, 8H); ^{13}C NMR (125 MHz, DMSO-d_6) d 163.1, 161.5, 160.3, 159.6, 159.2, 159.0, 156.0, 155.6, 155.4, 155.0, 152.5, 151.5, 150.2, 150.0, 145.9, 137.0, 133.8, 126.9, 126.5, 124.7, 124.3, 118.7, 116.3, 52.5, 47.9, 47.0, 39.1, 38.9, 37.5, 37.2, 31.7, 31.4, 27.1, 26.8, 24.2, 23.0, 21.9, 11.9; HRMS (ESI, $[M+H]^+$) calcd for C37H49N12O9 805.3745, found 805.3719.



2 3

L2A2-2M2EAc-6LCO (**3**): To a solution of **2** (10.5 mg, 13.0 μmol) in pyridine (1 mL) was added acetic anhydride (1 mL) at 0 ºC. After stirring for 1h, the mixture was concentrated *in vacuo*. The residue was purified by silica gel column ($CHCl_3$/MeOH = 100:1 to 5:1) to give **3** (10.1 mg, 77 %); $[a]^{25}_D$ = -279.4 (c 1.0 CH_3Cl); ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.96 (s, 1H), 8.86 (d, J = 8.5 Hz, 1H), 8.76 (s, 1H), 8.52 (d, J = 8.0 Hz, 1H), 8.09-8.01 (m, 2H), 7.84-7.79 (m, 2H), 5.21 (dt, J = 6.0, 14.5 Hz, 1H), 4.93 (dt, J = 6.0, 14.0 Hz, 1H), 3.46-3.35 (m, 4H), 3.05-2.97 (m, 4H), 2.67 (s, 3H), 1.87 (s, 3H), 1.77 (s, 3H), 1.76 (s, 3H), 1.71 (s, 3H), 1.69 (s, 3H) 1.50-1.19 (m, 12H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 170.3, 169.8, 169.7, 169.4, 163.8, 163.3, 161.5,

160.3, 158.5, 158.3, 158.0, 157.8, 156.3, 155.6, 155.2, 152.4, 151.2, 150.8, 145.8, 136.8, 133.7, 126.3, 125.8, 124.4, 119.1, 116.6, 79.7, 70.3, 52.5, 47.1, 47.0, 38.8, 38.7, 37.4, 37.2, 32.8, 31.9, 29.6, 29.2, 26.4, 23.3, 23.1, 22.9, 22.9; HRMS (ESI, $[M+Na]^+$) calcd for C47H58N14O12Na 1037.4093, found 1037.4070.



L2G2-2M2EG-6LCO (**4**): To a solution of the amine **2** (30.8 mg, 38.3 μmol) in DMF (1 mL) was added Et_3N (64.6 μL, 460 μmol), 1,3-bis(tert-butoxycarbonyl)-2-methyl-2-thiopsudourea (111 mg, 383 μmol), and $HgCl_2$ (62.4 mg, 230 μmol), and the mixture was stirred for 1 h at room temperature. The reaction mixture was filtered through a pad of Celite® and eluted with ethyl acetate. To the filtrates were added H_2O, and the organic layer was extracted with ethyl acetate. The extracts were dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was purified by flash column chromatography on silica gel ($CHCl_3$/ethyl acetate = 8:1). To a solution of the compound (18.4 mg, 9.13 μmol) in CH_2Cl_2 (1 mL) was added 12N HCl (0.5 mL) at 0 ºC and the mixture was stirred for 30 min. The reaction mixture was concentrated *in vacuo* to give **4** (6.6 mg, 24%, 2 steps); $[a]^{25}_D$ = +3.4 (c 0.5 MeOH); ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.03 (d, J = 8.5 Hz, 1H), 8.99 (s, 1H), 8.83 (s, 1H), 8.56 (d, J = 9.0 Hz, 1H), 7.96 (t, J = 6.0 Hz, 1H), 7.84-7.94 (m, 3H), 5.22 (dt, J = 8.0, 15.5 Hz, 1 H), 5.07 (dt, J = 5.0, 14.0 Hz, 1H), 3.86 (s, 3H), 3.64-3.53 (m, 4H), 3.47-3.40 (m, 2H), 3.15-3.10 (m, 4H), 2.71 (s, 3H), 2.69 (s, 3H) 2.15-2.08 (m, 2H), 2.03-1.82 (m, 2H), 1.60-1.35 (m, 8H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 163.2, 161.9, 161.4, 160.3, 157.6, 157.5, 157.4, 156.1, 155.5, 155.4, 155.1, 151.9, 151.4, 151.3, 145.8, 142.8, 136.8, 133.7, 126.6, 126.3, 124.4, 124.2, 52.6, 49.4, 47.2, 32.7, 31.7, 28.5, 28.3, 23.1, 22.5, 12.1, 11.9; HRMS (ESI, $[M+Na]^+$) calcd for C42H59N22O9 1015.4835, found 1015.4788.

HPLC Data for 2-4

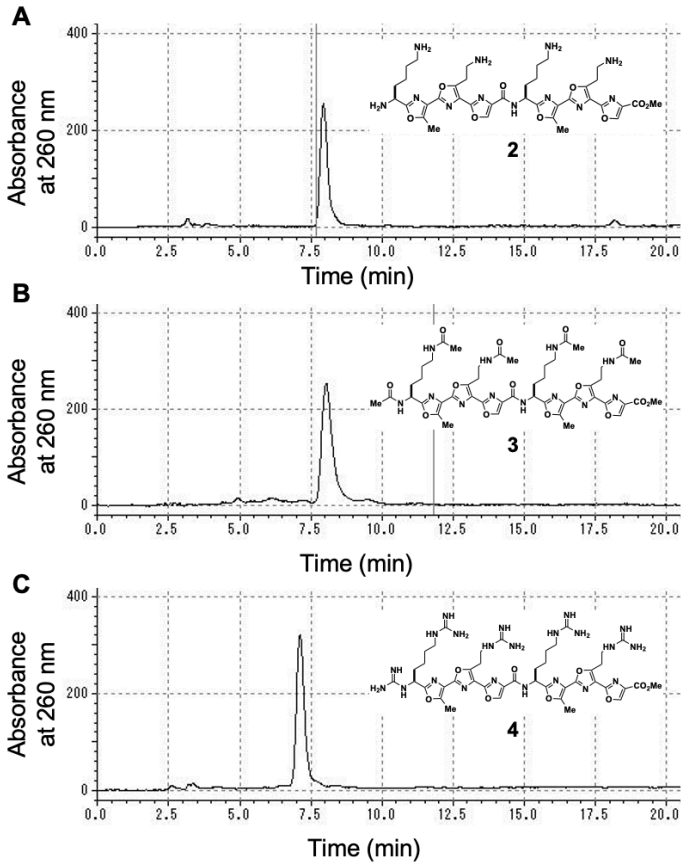
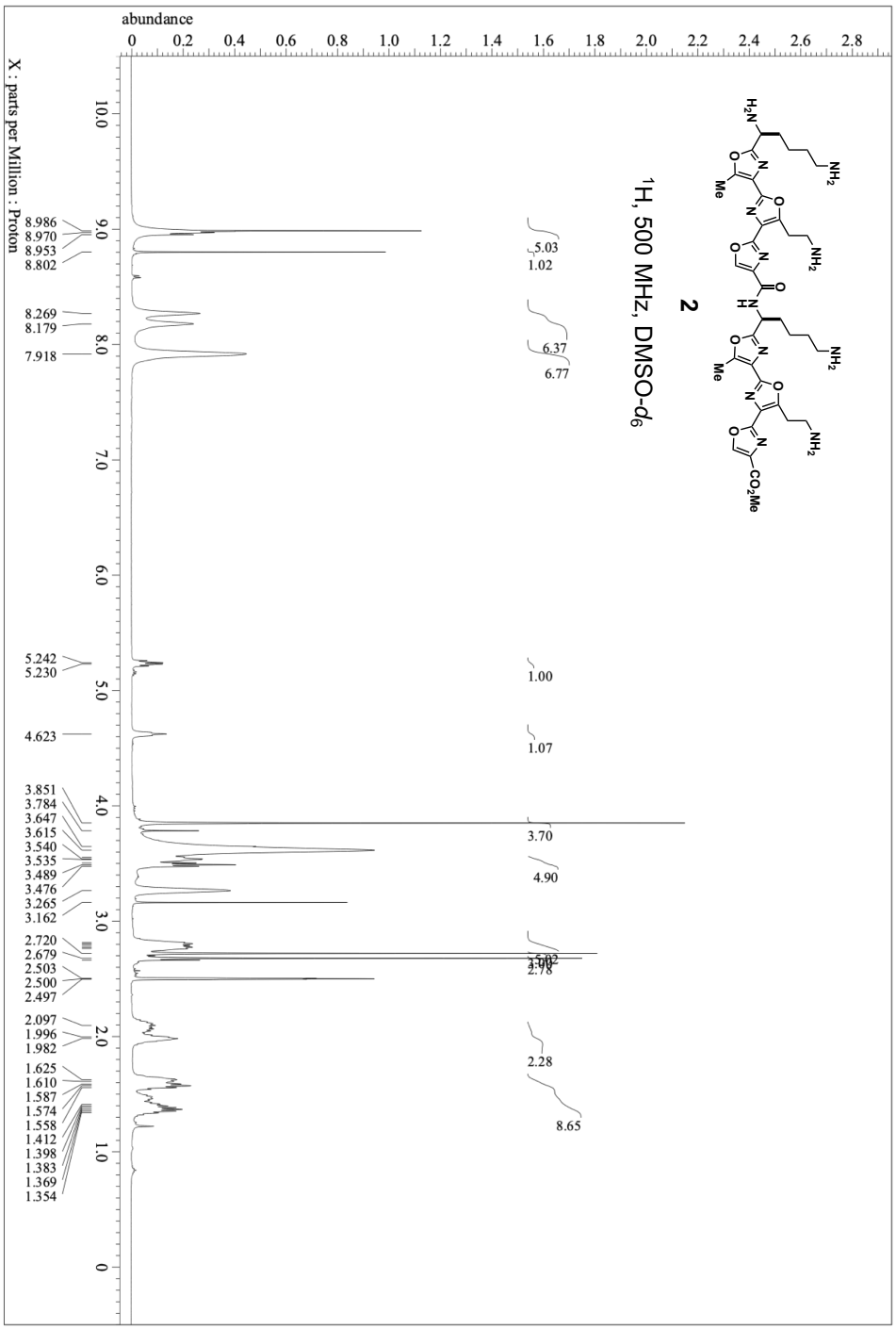
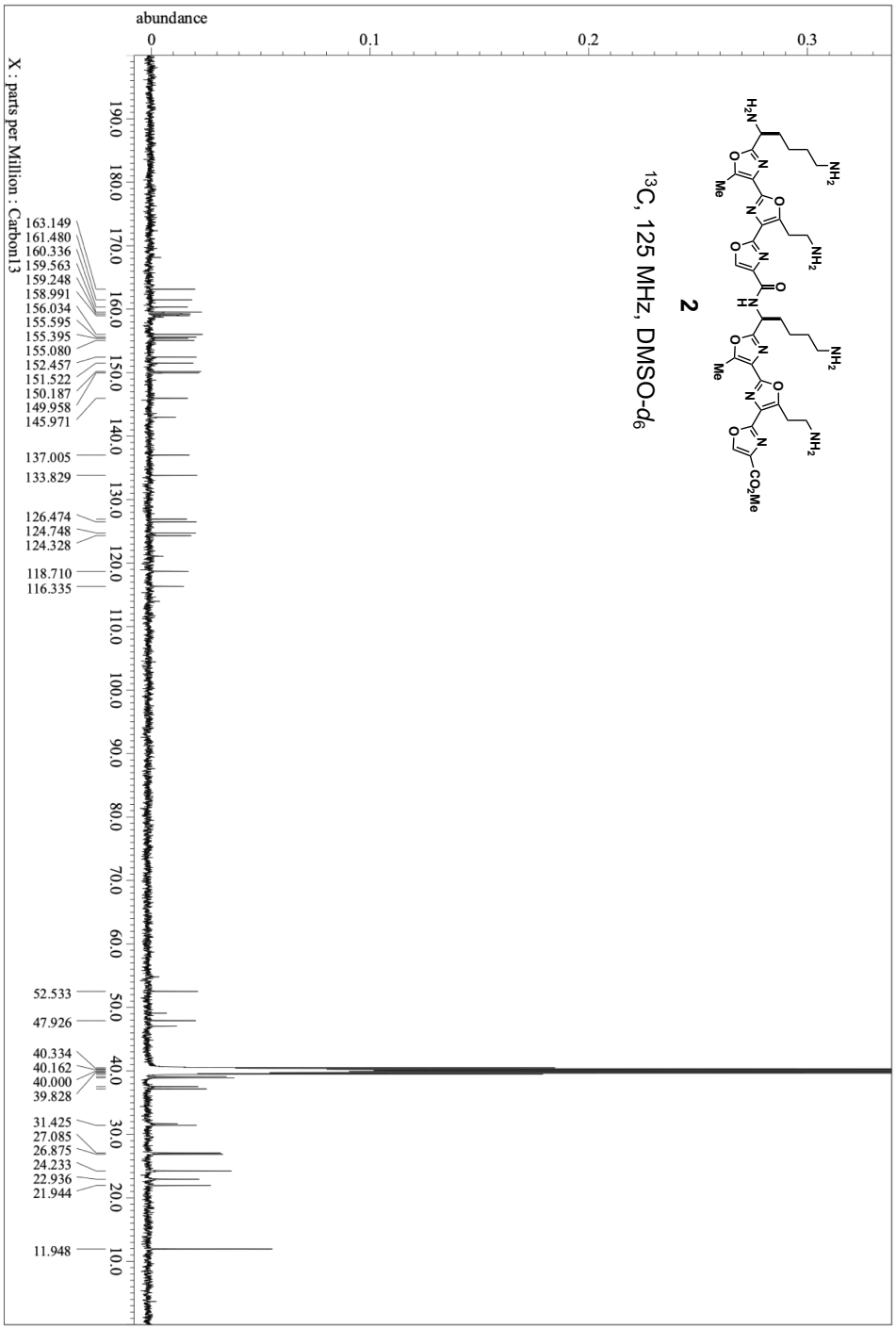
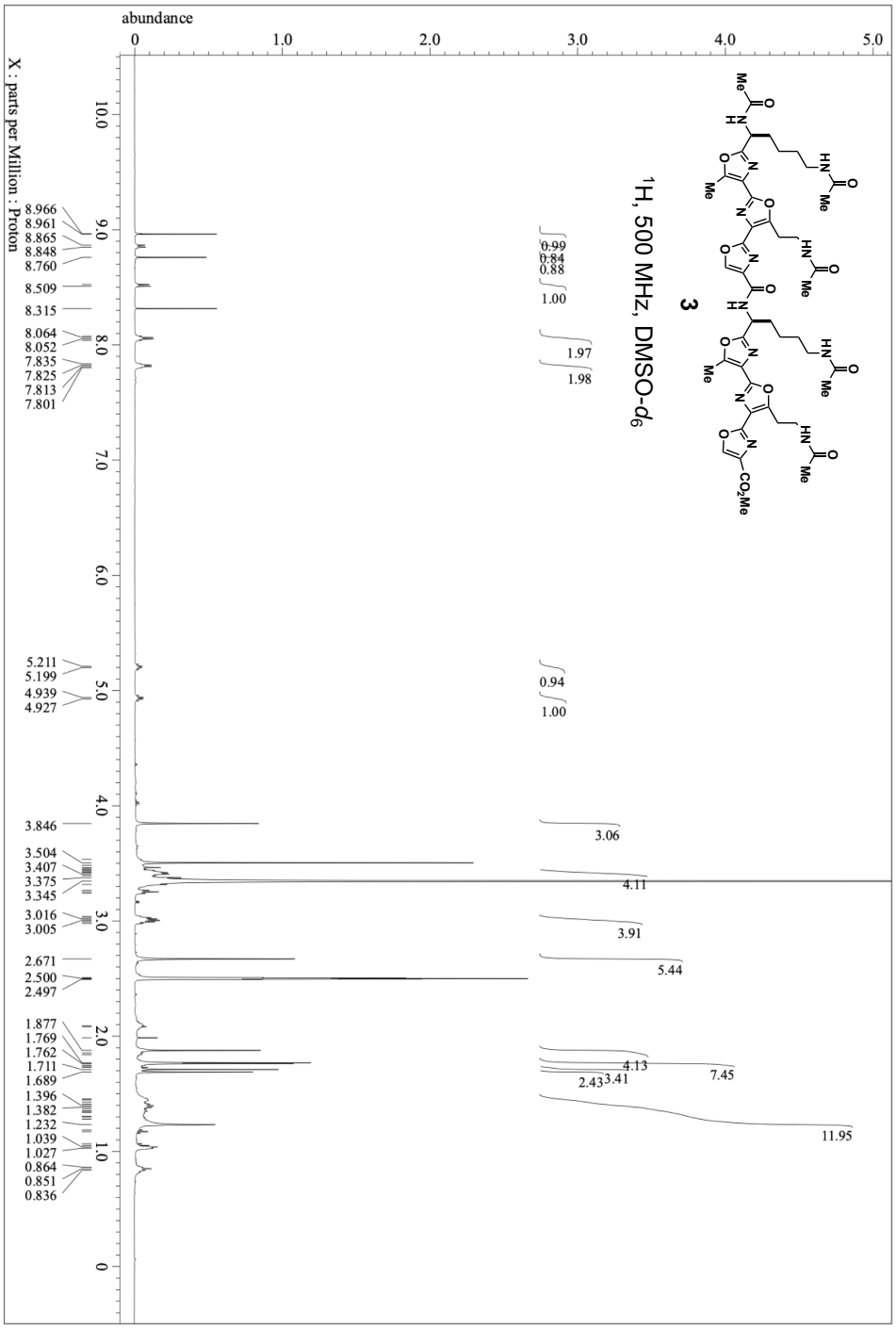


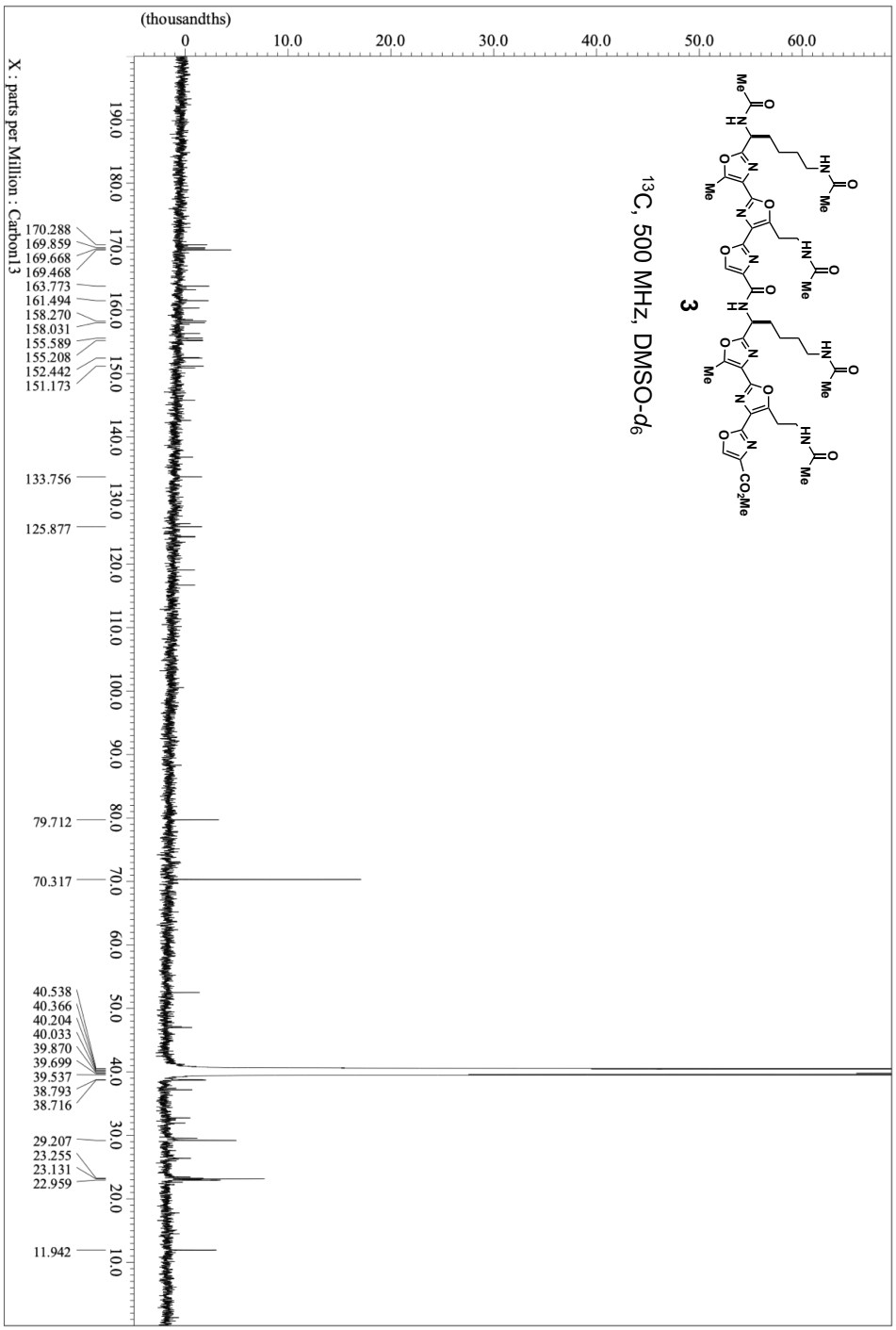
Fig. S6 Reverse phase HPLC Data of Ligands (**2-4**) was checked at 260 nm absorbance by using Shiseido Cancel Pak AQ C18 column. The mobile phase conditions is (A) H_2O (in 0.1% TFA) : MeOH (in 0.1% TFA) = 70 : 30 to 50 : 50 for 20 min; (B) H_2O (in 0.1% TFA) : MeOH (in 0.1% TFA) = 40 : 60 to 20 : 80 for 20 min; (C) H_2O (in 0.1% TFA) : MeOH (in 0.1% TFA) = 60 : 40 to 40 : 60 for 20 min.

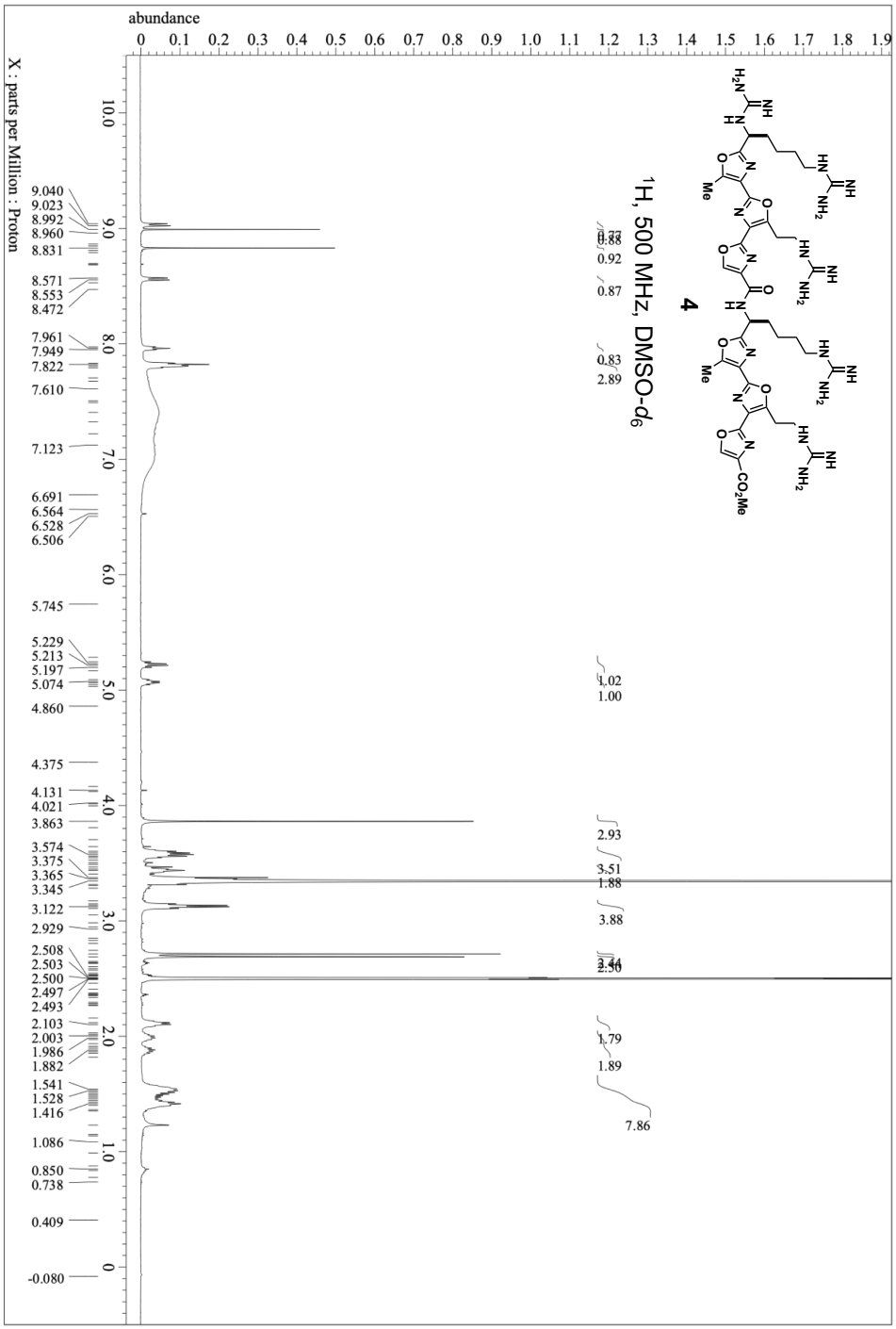
^{1}H and ^{13}C NMR spectra for **2, 3, 4**

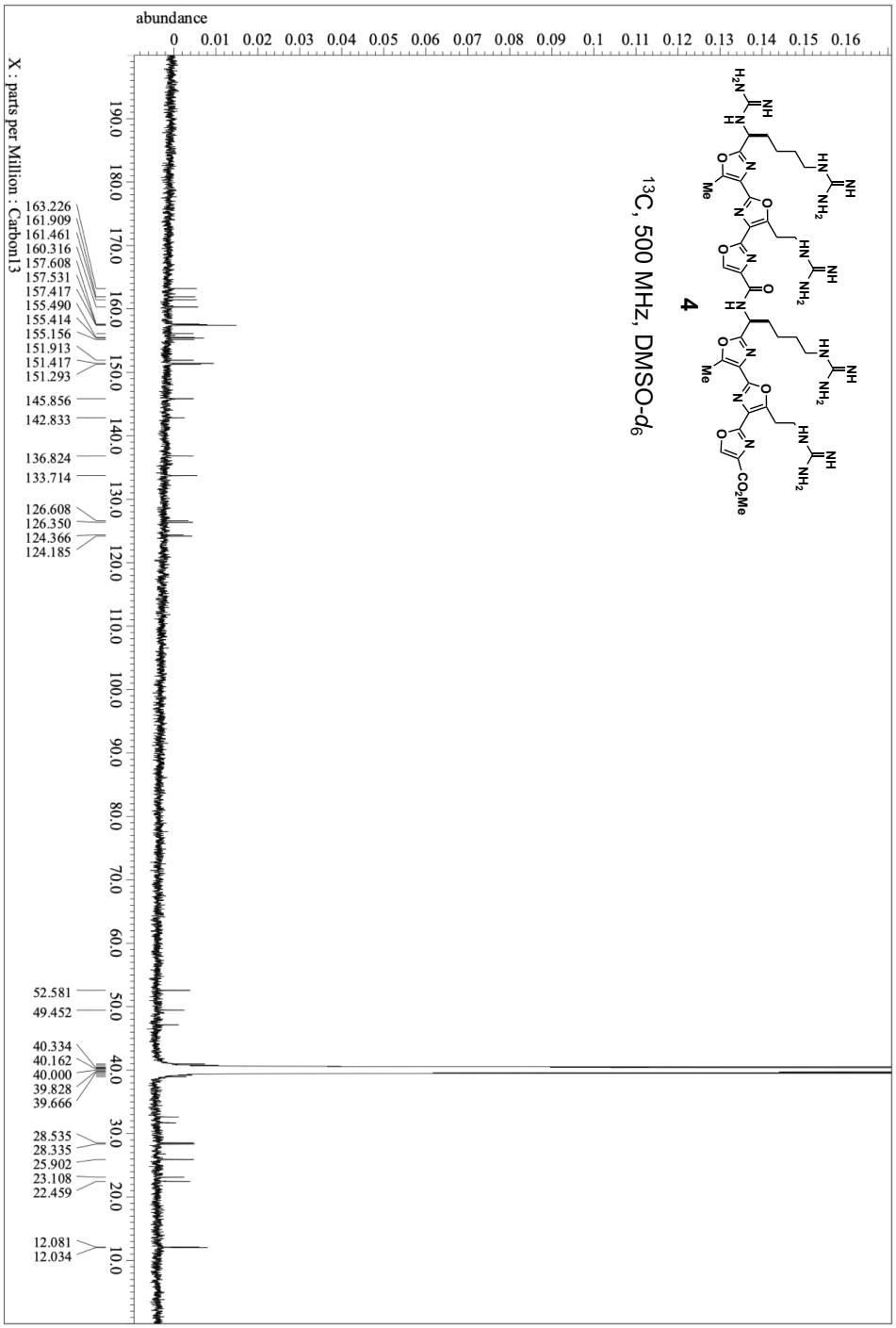



abundance
0
0.1
0.2
0.3
X : parts per Million : Carbon13
190.0
180.0
170.0
160.0
150.0
140.0
130.0
120.0
110.0
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
163.149
161.480
160.336
159.563
159.248
158.991
156.034
155.595
155.395
155.080
152.457
151.522
150.187
149.958
145.971
137.005
133.829
126.474
124.748
124.328
118.710
116.335
52.533
47.926
40.334
40.162
40.000
39.828
31.425
27.085
26.875
24.233
22.936
21.944
11.948
2
^{13}C, 125 MHz, DMSO-d_6
Me
Me
NH2
NH2
NH2
NH2
CO2Me


abundance
X : parts per Million : Proton
3
^{1}H, 500 MHz, DMSO-d_6


(thousandths)
0
10.0
20.0
30.0
40.0
50.0
60.0
X : parts per Million : Carbon13
190.0
180.0
170.0
160.0
150.0
140.0
130.0
120.0
110.0
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
170.288
169.859
169.668
169.468
163.773
161.494
158.270
158.031
155.589
155.208
152.442
151.173
133.756
125.877
79.712
70.317
40.538
40.366
40.204
40.033
39.870
39.699
39.537
38.793
38.716
29.207
23.255
23.131
22.959
11.942
^{13}C, 500 MHz, DMSO-d_6
3


abundance
X : parts per Million : Proton
^{1}H, 500 MHz, DMSO-d_6
4
Me
Me
CO$_2$Me


abundance
0 0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.1 0.11 0.12 0.13 0.14 0.15 0.16
X : parts per Million : Carbon13
190.0 180.0 170.0 160.0 150.0 140.0 130.0 120.0 110.0 100.0 90.0 80.0 70.0 60.0 50.0 40.0 30.0 20.0 10.0
163.226
161.909
161.461
160.316
157.608
157.531
157.417
155.490
155.414
155.156
151.913
151.417
151.293
145.856
142.833
136.824
133.714
126.608
126.350
124.366
124.185
52.581
49.452
40.334
40.162
40.000
39.828
39.666
28.535
28.335
25.902
23.108
22.459
12.081
12.034
4
13C, 500 MHz, DMSO-d6

Supplementary Reference

[1] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739-1749.

[2] Y. Ma, Y. Tsushima, M. Sakuma, S. Sasaki, K. Iida, S. Okabe, H. Seimiya, T. Hirokawad, and K. Nagasawa, *Org. Biomol. Chem*., 2018, **16**, 7375-7382