Supplementary Information

Random formation of G-quadruplexes in human telomere overhangs leads to a kinetic folding pattern with targetable vacant G-tracts

Jibin Abraham Punnoose,[1] Yue Ma,[2] Mohammed Enamul Hoque,[1] Yunxi Cui,[1] Shogo Sasaki,[2] Athena Huixin Guo,[1] Kazuo Nagasawa[2] and Hanbin Mao[1]*

[1] Department of Chemistry and Biochemistry, Kent State University, Kent, Ohio, 44242, USA

[2] Department of Biotechnology and Life Science Faculty of Technology, Tokyo University of Agriculture and Technology (TUAT), 2-14-16 Naka-cho, Koganeishi, Tokyo 184-8588, Japan

*Corresponding author: HM, hmao@kent.edu (330-672-9380)

Table of contents

18. Matlab code for the statistical calculation of the formation pattern of telomere G-quadruplexes
19. Igor code for the folding simulation of telomere G-quadruplex and intermediates

1. Oligonucleotides Sequences

Name		Sequence
4G		5′- CTAGACGGTGTGAAATACCGCACAGATGCG **(TTAGGG)$_{n=4}$TTA** GCCAGCAAGACGTAGCCCAGCGCGTC-3′
8G		5′- CTAGACGGTGTGAAATACCGCACAGATGCG **(TTAGGG)$_{n=8}$TTA** GCCAGCAAGACGTAGCCCAGCGCGTC-3′
12G	6G1	5′ - CTAGACGGTGTGAAATACCGCACAGATGCG**(TTAGGG)$_{n=5}$TTAGG**-3′
	6G2	5′- **G(TTAGGG)$_{n=6}$TTA**GCCAGCAAGACGTAGCCCAGCGCGTC-3′
16G	12G1	5′- CTAGACGGTGTGAAATACCGCACAGATGCG **(TTAGGG)$_{n=11}$TTAGG**-3′
	4Gfor16G	5′- **G(TTAGGG)$_{n=4}$TTA**GCCAGCAAGACGTAGCCCAGCGCGTC-3′
24G	12G1	5′- CTAGACGGTGTGAAATACCGCACAGATGCG **(TTAGGG)$_{n=11}$TTAGG**-3′
	12G2	5′- **G(TTAGGG)$_{n=12}$TTA**CCAGCAAGACGTAGCCCAGCGCGTC-3
48G	12G1	5′- CTAGACGGTGTGAAATACCGCACAGATGCG **(TTAGGG)$_{n=11}$TTAGG**-3′
	12G2	5′- **GG(TTAGGG)$_{n=12}$T**-3′
	12G3	5′- **TAGGG(TTAGGG)$_{n=10}$ TTAG**-3′
	12G4	5′- **GG(TTAGGG)$_{n=12}$**TTACCAGCAAGACGTAGCCCAGCGCGTC-3
Splint		5′- TAACCCTAACCCTAACCCTAACCC-3′
Splint remover		5′- GGGTTAGGGTTAGGGTTAGGGTTA-3′
8G-Spaced	4GNC1	5′- CTAGACGGTGTGAAATACCGCACAGATGCG **(TTAGGG)$_{n=4}$**TTA(T)$_{n=6}$
	4GNC2	5′- (T)$_{n=15}$ **(TTAGGG)$_{n=4}$**TTAGCCAGCAAGACGTAGCCCAGCGCGTC
12G-Spaced	8GNC1	5′- CTAGACGGTGTGAAATACCGCACAGATGCG **(TTAGGG)**$_{n=4}$TTA(T)$_{n=21}$ **(TTAGGG)**$_{n=4}$ TTA (T)$_{n=6}$
	4GNC2	5′- (T)$_{n=15}$ **(TTAGGG)$_{n=4}$**TTAGCCAGCAAGACGTAGCCCAGCGCGTC
NC Splint- Spaced		5′- GCTACGAAAAAAAAAAAAAAAAAATAACCC
NC Splint Remover-Spaced		5′- GGGTTATTTTTTTTTTTTTTTTTCGTAGC
Chimera 1.0 DNA		/5′Hexynyl/TTTAACCCTAA
Chimera 2.0 DNA		/5′Hexynyl/TTTAACCCTAACCCTAA

Table S1*: Oligonucleotides used in the experiment. The telomere sequences are shown in bold. The shorter telomere sequences used to prepare the longer DNA are categorized under the final sequence name.*

2. Synthesis of the 48 G-tracts Construct

Synthesis of the 48G construct was started with the four longest commercially available oligonucleotides (12G1, 12G2, 12G3 and 12G4) each containing 12G-tracts (Figure S1). The 12G1 and 12G4 have additional non-G-tract sequences in the 5′ and 3′ ends, respectively, for the ligation with the 2028 bp and the 2690 bp handles in the later steps. The 12G1 and 12G2 were phosphorylated and ligated using dynamic splint ligation (1) with a 12G1:12G2:splint DNA ratio of 1:1:6. Similarly, the 12G3 and 12G4 DNA were also phosphorylated and ligated. The reacted species yielded two fragments that contained 24 G-tracts. They were named as 24G1 (ligation of the 12G1 and 12G2 DNA) and 24G2 (ligation of the 12G3 and 12G4 DNA), respectively (Figure S1A). They were purified by an 8% native PAGE gel after removing the splint DNA using a 1:10 ratio of split: splint-removal-DNA (Table S1). These 24G DNA fragments were ligated to the biotinylated 2028 bp and the digoxigenin labelled 2690 bp handles, respectively. The reaction was confirmed through gel shift assays on a 0.8% agarose gel (Figure S1C). The same gel was used to purify the DNA handles ligated with the sequence of interest (24G1 and 24G2). These purified DNA fragments were ligated using dynamic splint ligation with a 24G1:24G2:Splint ratio of 1:1:12. The splint used for the ligation was subsequently removed using a splint:splint-removal-DNA ratio of 1:100. The resulting construct contains 48 G-tracts that can form a maximum of 12 G-quadruplexes in the presence of K^+ ions. The successful synthesis of the 48G construct was confirmed through the gel shift assay (Figure S1C, lane 6).

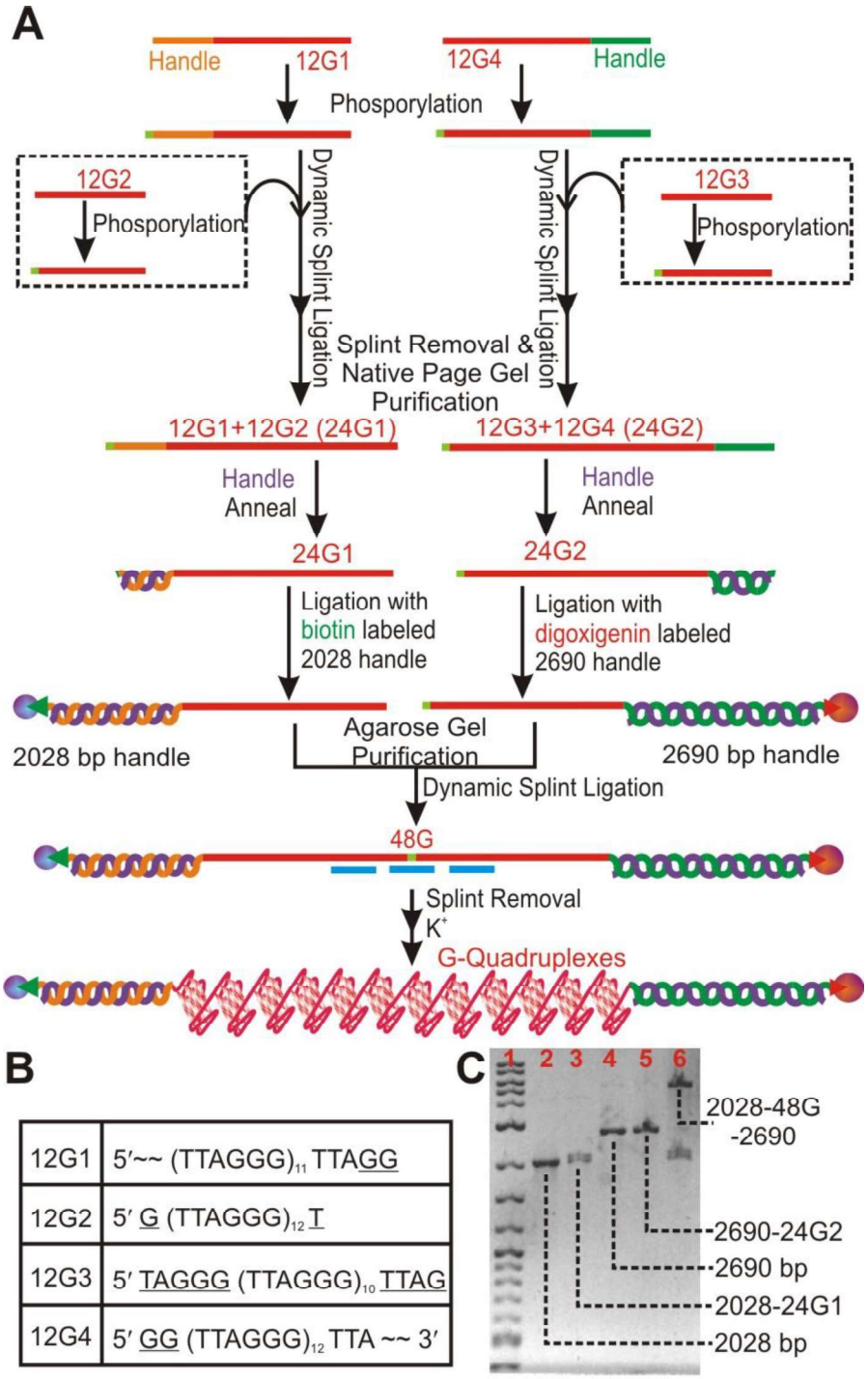
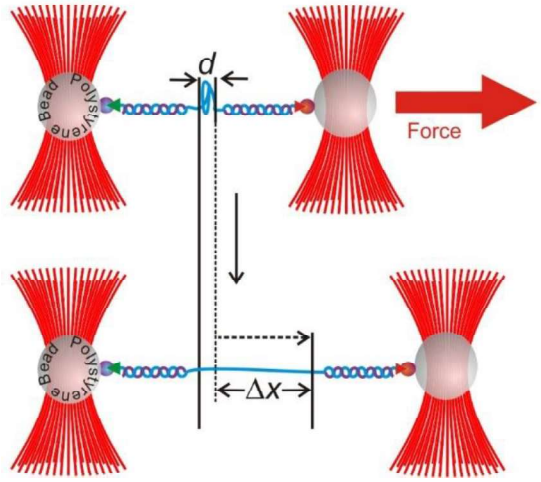


Figure S1: *Synthesis of the 48G construct for single-molecule analysis. (A) Flowchart elaborating the synthesis of the 48G fragment from four commercially available sequences 12G1, 12G2, 12G3, and 12G4. (B) Sequences of the 12G1, 12G2, 12G3, and 12G4. (C) Agarose gel picture confirms the successful synthesis of the 48G construct. Lane 1 depicts the 2-Log DNA Ladder (NEB, 0.1-10.0 kb), lanes 2-5 represent different intermediates leading to the final construct ("2028-48G-2690" in lane 6). The faster migration band in lane 6 represents*

unreacted 2028-24G1 conjugated (see band 3). This compound will not be tethered between the two optically trapped beads. Therefore, it will not influence the single-molecular results.

3. Identification of Structures in Telomere Overhang

A



B

Species	*d* (nm)	Δ*L* (nm)	L_c (nm)	# nts :experimental	# nts :expected
GH	2.3	2.2	4.5	10	9
GT	1.5	5.4	6.9	15	15
GQ	1.1	8.6	9.7	21	21
MF	1.1	11.3	12.4	27	27
HO	2 x 1.1	17.6	19.8	44	45

Figure S2: *(A) Schematic Force-Extension (F-X) experiment depicting the unfolding event and the change in extension. The change-in-extension (Δx) observed during the unfolding of structures is used to retrieve change-in-contour-length (ΔL) using a worm-like-chain model (Materials and Methods, Equation 1), which leads to the estimation of the contour length (L_c=ΔL + d, where d is the end-to-end distance of a folded structure along the pulling direction). The L_c is directly proportional to the number of nucleotides in the structure. (B) Table showing reported values of end-to-end distance (d) (2, 3), calculated ΔL and L_c of folded species, experimentally determined nucleotides involved in the structure, and expected number of nucleotides for different species. Since G-Hairpin (GH), Misfolded G-quadruplex (MF), and Higher-Order structures (HO) have no reported d value, approximations were made. In G-hairpin, d value is determined from the distance between phosphorus atom of the dG^4 and that of the T^{13} backbone from NMR structure of the hybrid-1 G-quadruplex (PDB ID: 2HY9) (3). Misfolded structures are considered as G-quadruplex formed from non-adjacent G-tracts (4, 5), hence the d value is assumed to be similar to that of the hybrid-1 G-quadruplex. The slight change observed in the experimentally calculated number of nucleotides might be originated from the misfolded structures adopting conformations different from the hybrid-1 conformation. Higher-order structures are considered as two stacked G-quadruplexes (6). Therefore, d is about twice the end-to-end distance of the hybrid-1 G-quadruplex.*

4. *ΔL and Unfolding Force Histograms of Secondary Structures Formed in the 4G to 48G Telomere Constructs*

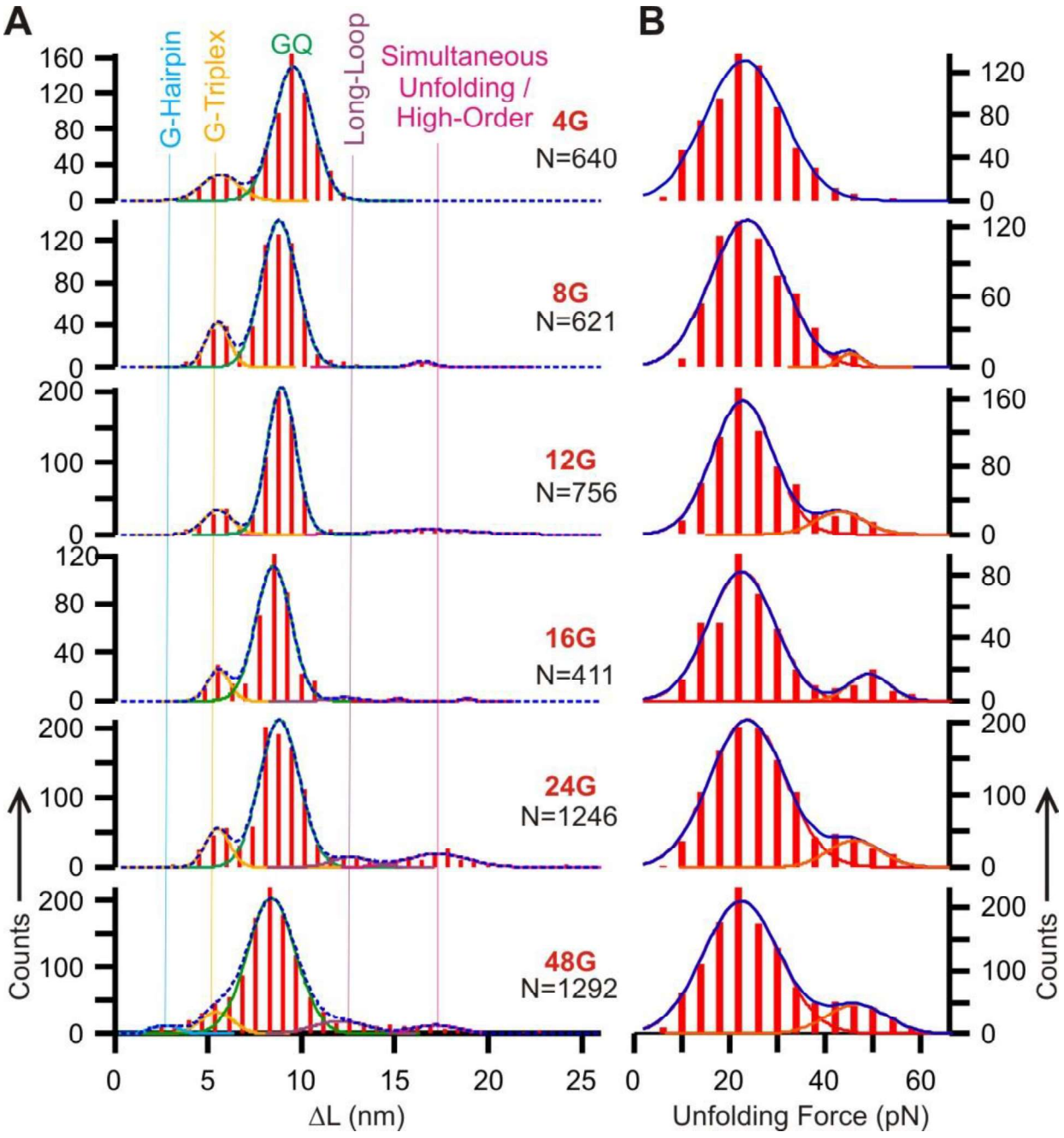


Figure S3: *Characterization of the unfolding events in telomere sequences (4G-48G) in a 10 mM Tris buffer supplemented with 100 mM K^+ at pH 7.4. (A) ΔL histograms of the telomere sequences. Folded species are resolved using multi-peak Gaussian fitting and identified by the length of the folded nucleotides in each structure. See Figure S2 for detailed calculation. (B) Unfolding force histograms of the structures. Curves represent Gaussian fittings.*

5. *Percentage of all species observed in the 4G to 48G constructs*

	GH(%)	GT(%)	GQ(%)	MF(%)	HO(%)
4G	0	15	85	0	0
8G	0	17	83	0	0
12G	0	14	86	0	0
16G	0	14	85	1	0
24G	0	11	82	2	4
48G	2	10	79	7	2

***Table S2*:** *Percentage of all species observed in the 4G to 48G constructs in a 10 mM Tris buffer supplemented with 100 mM K^+ at pH 7.4. GH: G-hairpin; GT: G-triplex; GQ: G-quadruplex; MF: Misfolded Structure (long-loop GQ for example); HO: High-order GQ or simultaneously unfolded GQs.*

6. *Thermodynamic and Kinetic Folding Patterns of the G-quadruplexes in an 8 G-tract Construct*

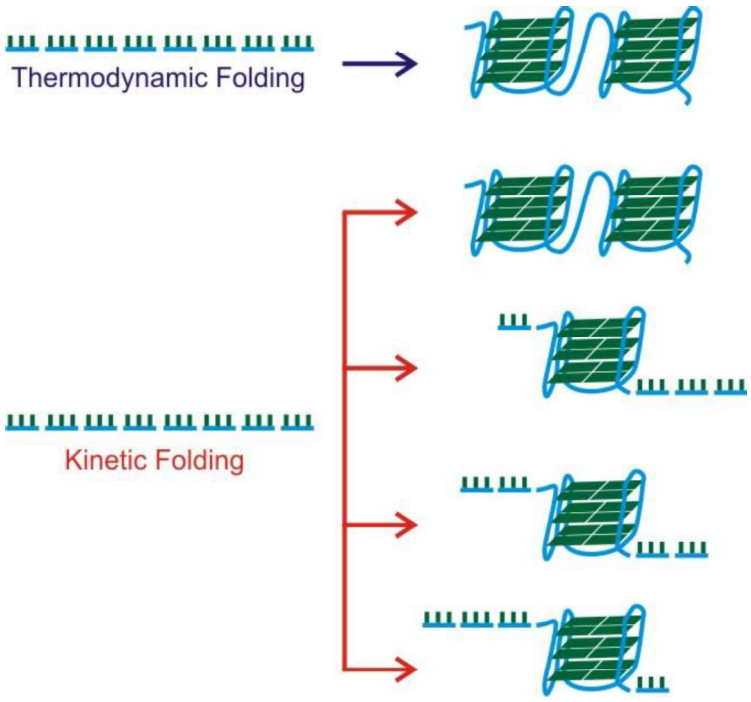


Figure S4: *Thermodynamic and kinetic folding patterns of the G-quadruplexes in the 8G construct. In the thermodynamic folding, two G-quadruplexes are formed, which is the most energetically favored product. In the kinetic folding model, G-quadruplex folds at random locations. As a result, there are three different arrangements in which only one G-quadruplex can form. By contrast, only one arrangement can lead to the most energetically favored product (see the two G-quadruplexes on top).*

7. *All Different Possibilities of G-Quadruplex Formation in the 4G to 48G Telomere Constructs Obtained from Statistical Calculation Using the Matlab Code*

Number of G-Quadruplexes Formed	Telomere Sequence & discrete folding possibilities				
	8G	**12G**	**16G**	**24G**	**48G**
1 GQ	5	9	13	21	45
2 GQ	1	15	45	153	861
3 GQ		1	35	455	9139
4 GQ			1	495	58905
5 GQ				126	237336
6 GQ				1	593775
7 GQ					888030
8 GQ					735471
9 GQ					293930
10 GQ					44913
11 GQ					1365
12 GQ					1
Overall Possibility	6	25	94	1251	2863771

Table S3: *All different possibilities of G-quadruplex formation in the 8G-48G telomere constructs.*

8. *Algorithm Flowchart for Telomere G-quadruplex and Intermediates Folding Simulation*

To evaluate the hypothesis that secondary structures fold randomly in the telomere, we wrote an Igor Pro program (WaveMetrics, Portland, OR) to simulate the folding pattern and calculate the percentage of folded features (See Figure S5 for the algorithm flowchart). First, the number of maximally possible structural formation was determined based on the length of an input template sequence. A loop was then initiated to identify a random location in the template where a structure was randomly selected to form from an ensemble of folded species experimentally observed (Figure S3 and Table S2). The process continued until consecutive G-tracts were no longer available for a particular structure. The total number of folded structures in the template was then recorded. The whole process was iterated to generate a statistical average of features formed per template. The pattern for number of features formed and the average features obtained (calculated using Equation 2) from the simulation were compared to the experimentally observed data. The program was also designed to include spacing between two G-quadruplexes as an input parameter, which was used to identify the minimal spacing between neighboring G-quadruplexes to account for the electrostatic repulsion.

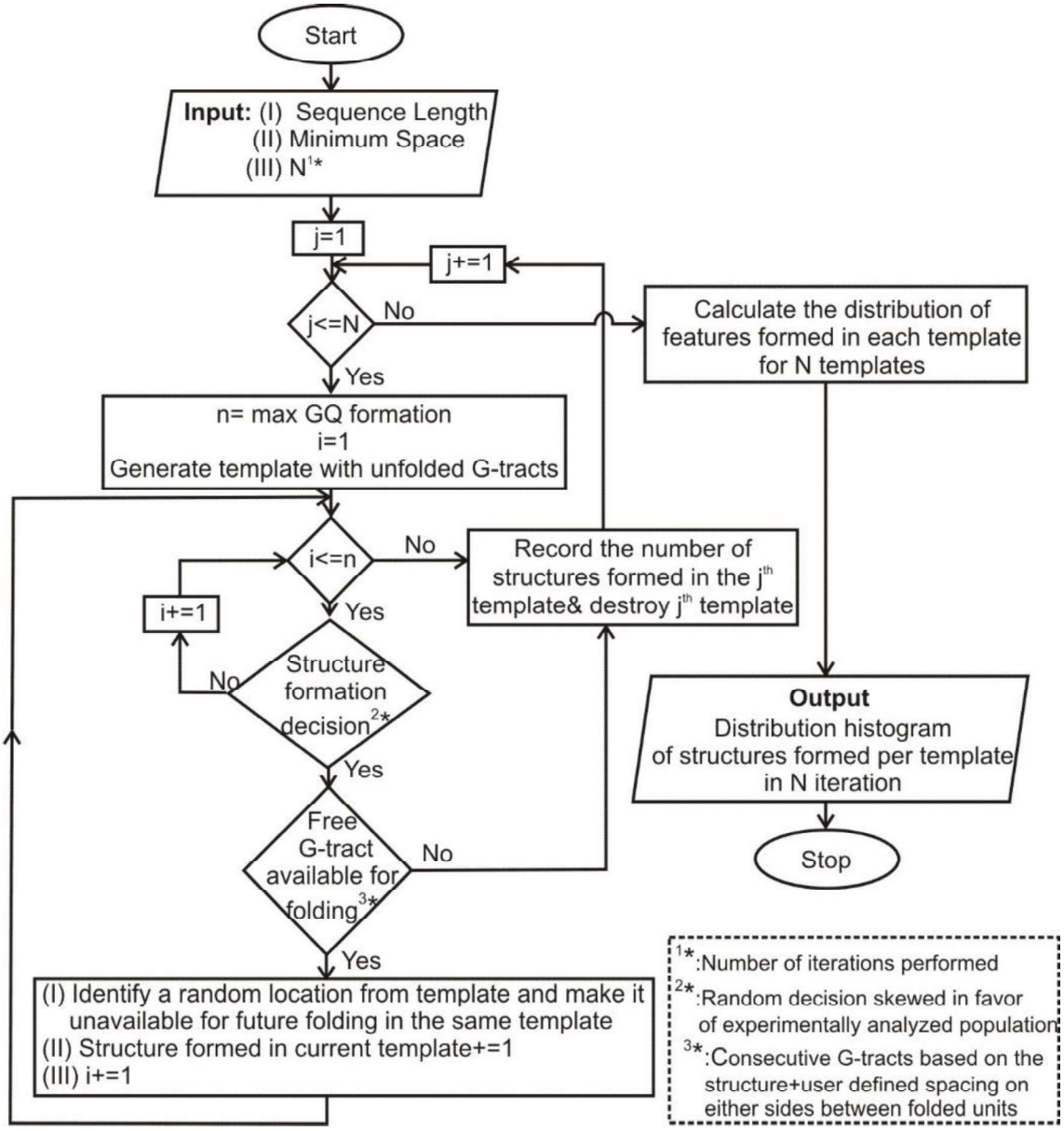


Figure S5*: The algorithm flowchart describing the sequence of action for the folding simulation inside a telomere overhang sequence. The input parameters are the sequence length, minimum space between neighboring G-quadruplex units (in the units of G-tracts) to minimize the repulsion between the G-quadruplexes, and the number of folding iterations to be performed (N). The output gives the distribution histogram of the G-quadruplex formed per template for N iterations performed.*

9. *Folding Patterns Obtained from Kinetic Folding Simulation*

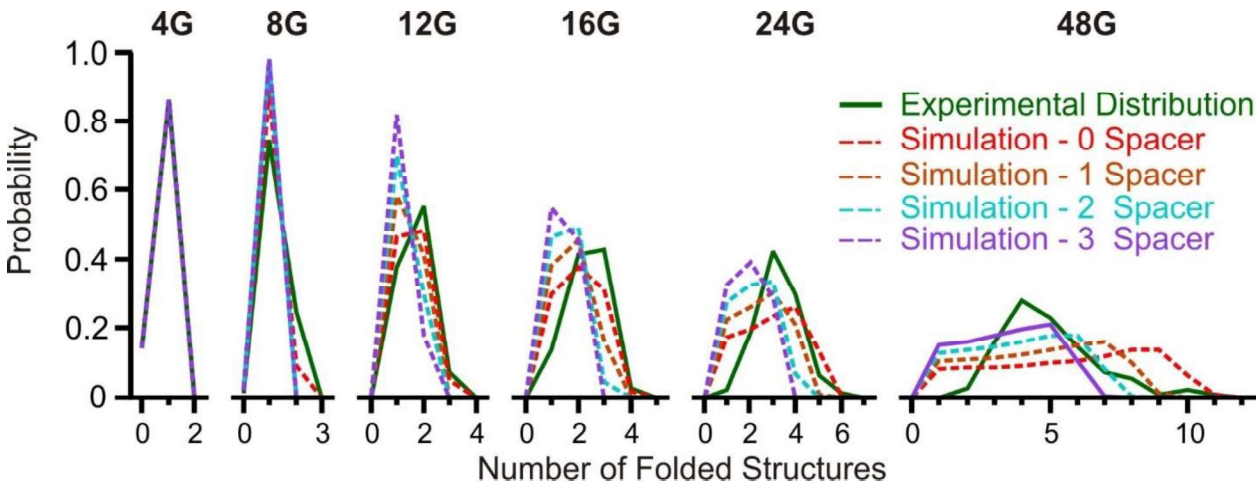


Figure S6: *Population patterns of folded structures in telomere overhangs obtained from the kinetic folding model without and with various minimum spacing between two G-quadruplexes in the 4G to 48G telomere constructs. See Figure 2D for the RMSD deviation of each simulation from the experimentally observed folding pattern.*

10. ***Dudko-Szabo Fitting for G-quadruplexes in the 12G Construct***

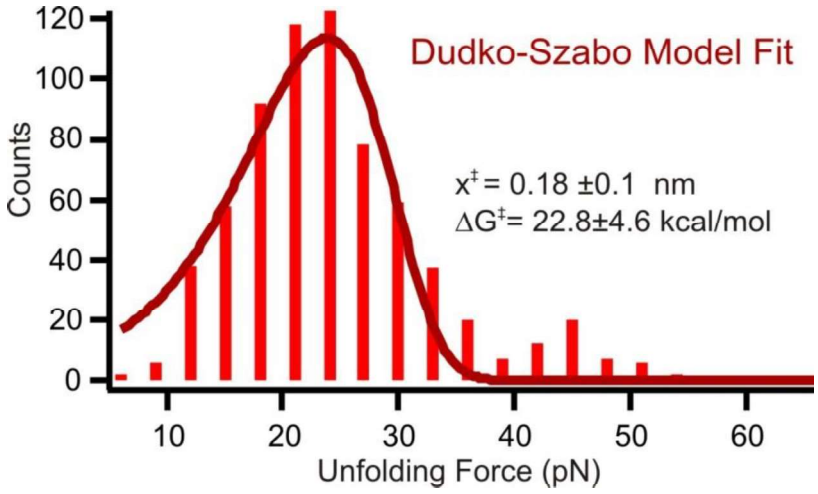


Figure S7: *Dudko-Szabo fitting to determine the kinetic parameters including the activation distance from the unfolded state to the transition state ($x^{\ddagger}$) and the activation energy ($\Delta G^{\ddagger}$).*

11. *G-quadruplex Folding Pathways in the 12G Construct*

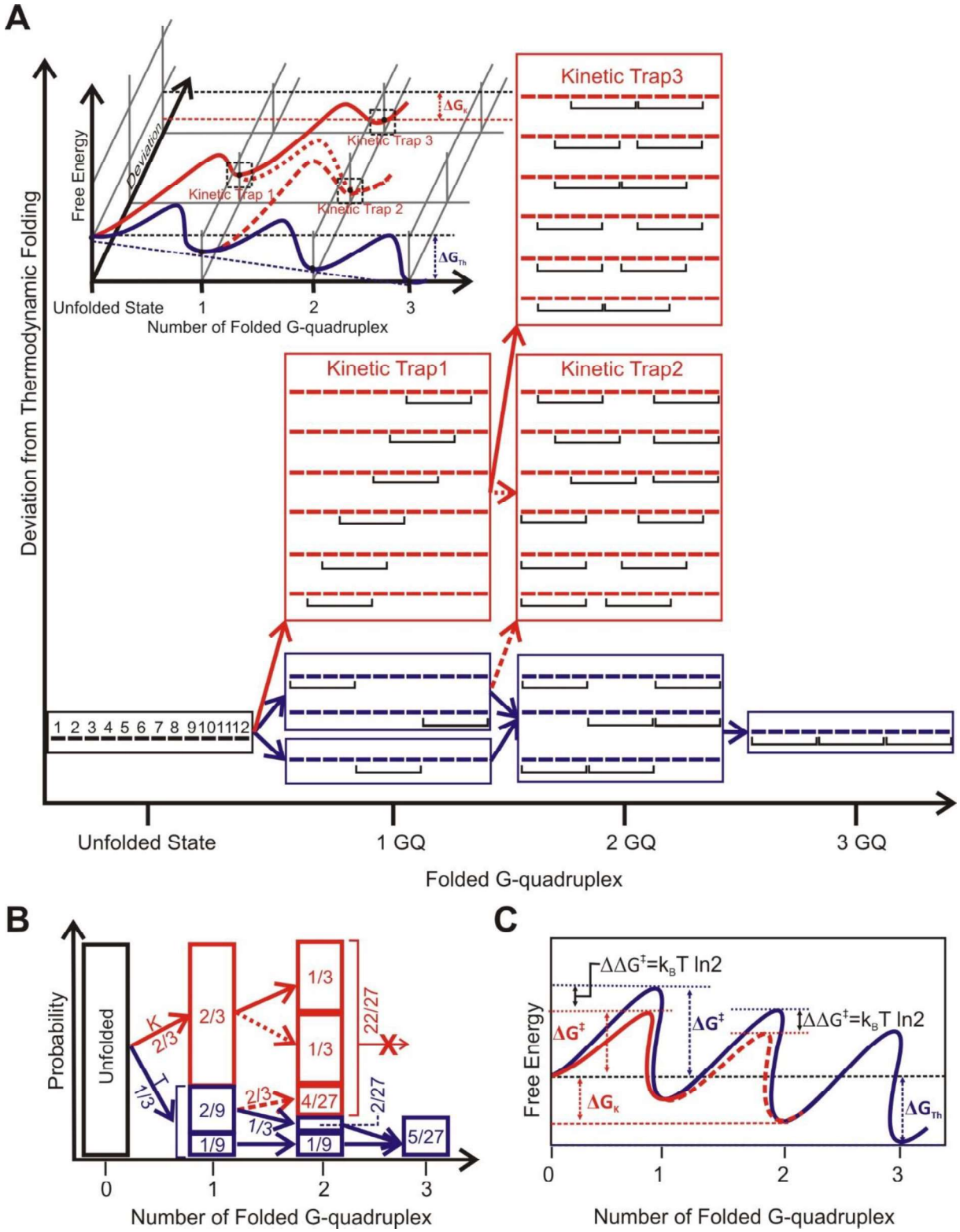


Figure S8*: Folding pathways for the G-quadruplex population pattern in the 12G telomere construct. (A) All folding pathways of the G-quadruplexes (diagrams not drawn to scale). Each 4*

consecutive G-tracts for the formation of a G-quadruplex are indicated by a horizontal bracket. The folding pathway that yields a maximum of 3 G-quadruplexes is considered as the thermodynamically favored folding (blue). In kinetically favored pathways (red), the folding of G-quadruplex occurs at random positions, which leads to kinetically trapped G-quadruplexes. A maximum of two G-quadruplexes is folded in all possible kinetic pathways. The apparent reduction in the activation energy for the kinetic folding pathway is calculated from the Boltzmann distribution. From the unfolded state, there are three thermodynamic possibilities leading to the folding of the first G-quadruplex, whereas six possibilities exist in the kinetic folding pathways leading to the G-quadruplexes in the kinetic trap 1 (see diagram in A). Since there are twice as many possibilities to fold into the first G-quadruplex along the kinetic pathway, the apparent activation energy for the first ensemble kinetic folding to the kinetic trap 1 state is reduced by a factor of k_BTln2 (where k_B is the Boltzmann constant and T is absolute temperature). From the first G-quadruplex, two of the three thermodynamically favored foldings can lead to the formation of the second G-quadruplexes (kinetic trap 2) via a second ensemble kinetic folding pathway. The ratio of the kinetic vs thermodynamic pathway is 2:1, which leads to the apparent reduction of activation energy of the 2^{nd} ensemble kinetic pathway by the same factor of k_BTln2. Notice kinetic trap 1 can lead to either kinetic trap 2 or kinetic trap 3. Due to the complexity, the activation energies for these two pathways are not estimated. Once the 12G construct adopts a particular kinetic folding path, it cannot return to the thermodynamic path without first unfolding the G-quadruplex trapped kinetically. Such an escape is not likely during the experimental time scale given the slow unfolding kinetics of a telomeric G-quadruplex (7). (B) Diagram showing lineage of thermodynamic and kinetic folding probabilities from the unfolded state. Fractions depict probabilities for each state. (C) Diagram depicting the free energy changes of the thermodynamic and the major ensemble kinetic folding pathways.

12. *Calculation of Bound Fraction of G-quadruplexes in the 24G Construct in Presence of L2H2-6OTD Chimera 2.0*

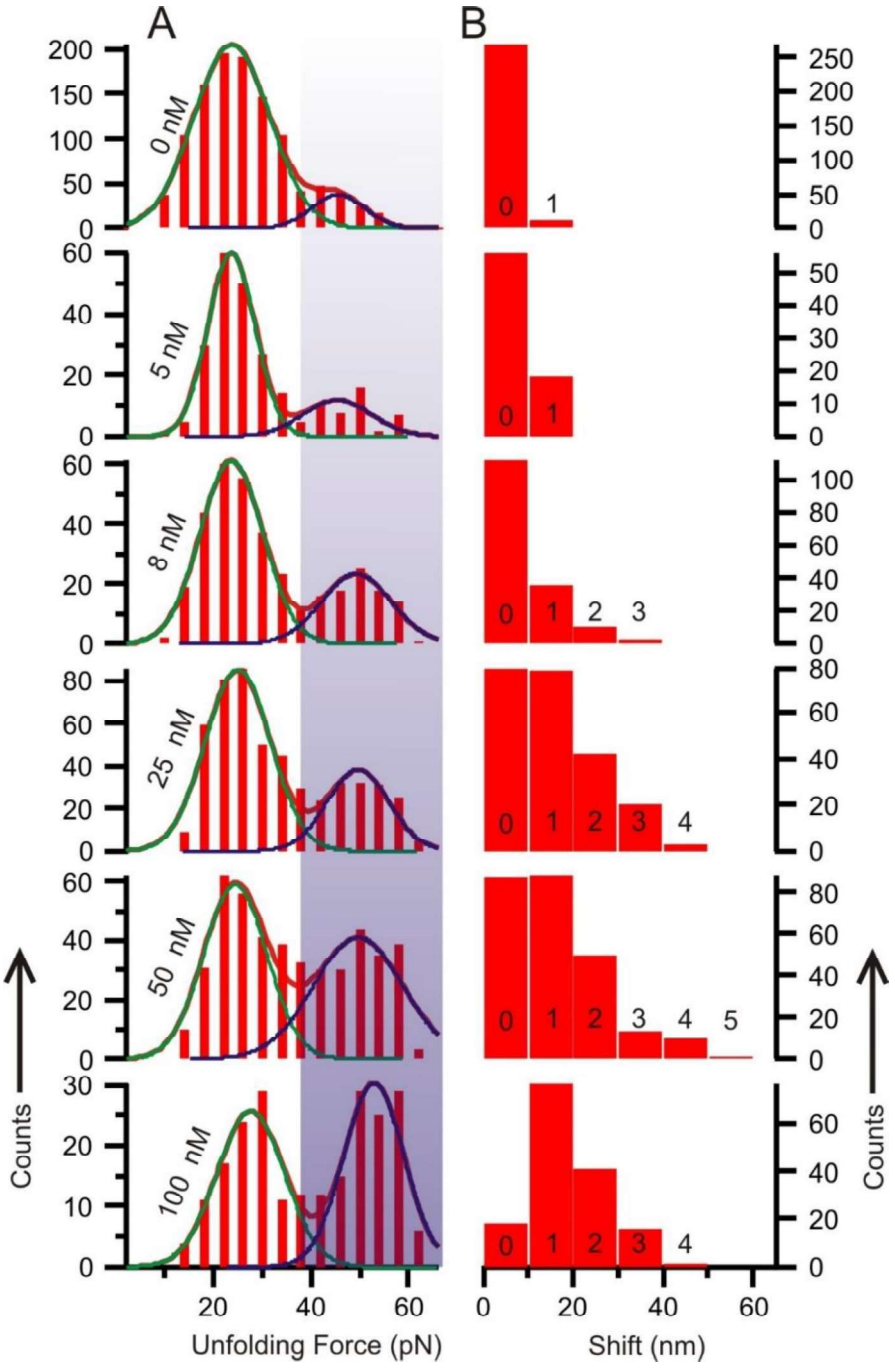


***Figure S9**: (A) Unfolding force histograms of the 24G construct in presence of L2H2-6OTD Chimera 2.0 with various concentrations. (B) Corresponding histograms of non-unfolded G-quadruplexes.*

13. ***Calculation of Bound Fraction of G-quadruplexes in the 24G Construct in the Presence of the L2H2-6OTD Monomer***

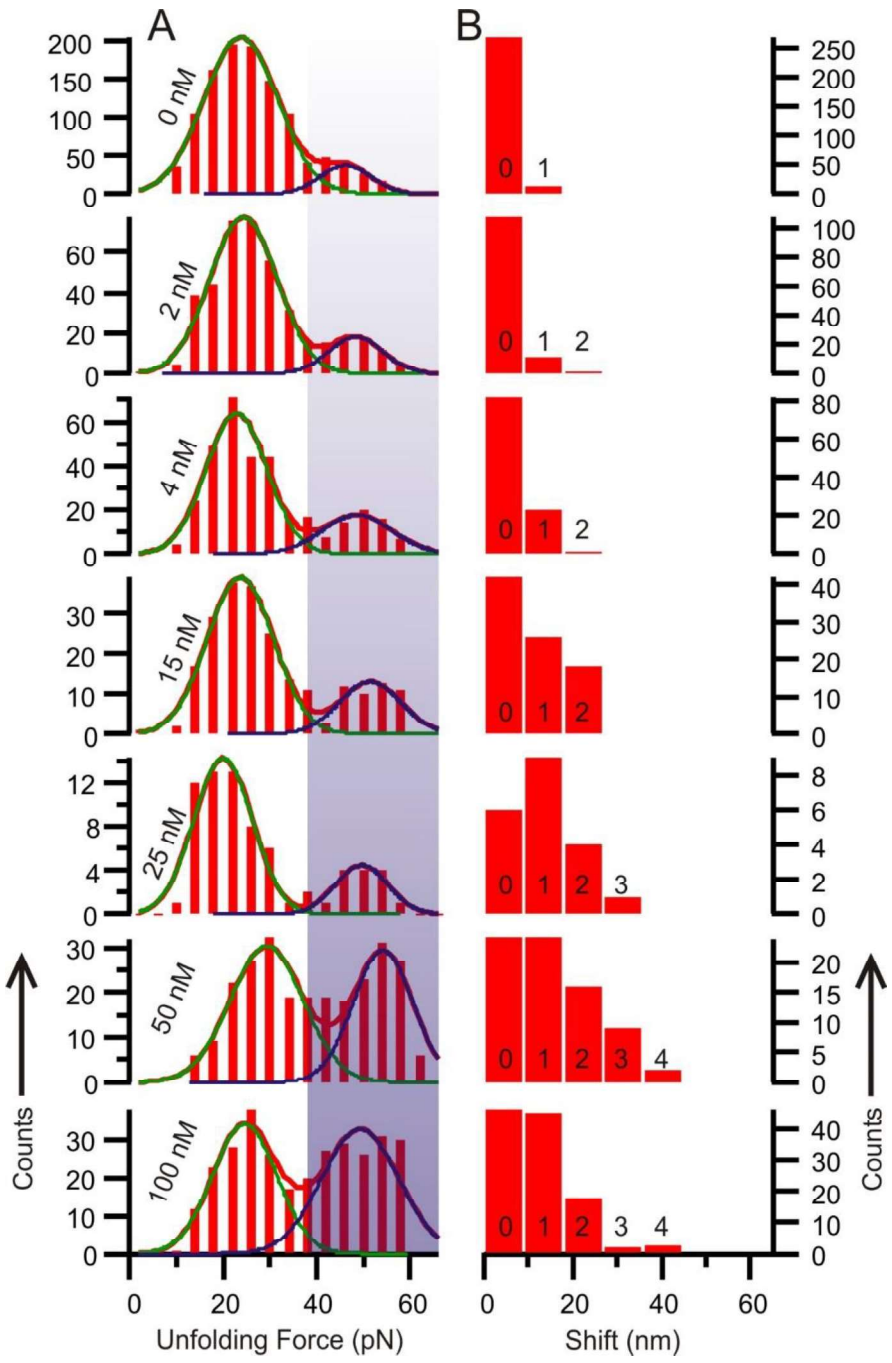


Figure S10: *(A) Unfolding force histograms of the 24G construct in the presence of the L2H2-6OTD monomer with various concentrations. (B) Corresponding histograms of non-unfolded G-quadruplexes.*

14. ***Calculation of Bound Fraction of G-quadruplexes in the 48G Construct in Presence of L2H2-6OTD Chimera 2.0***

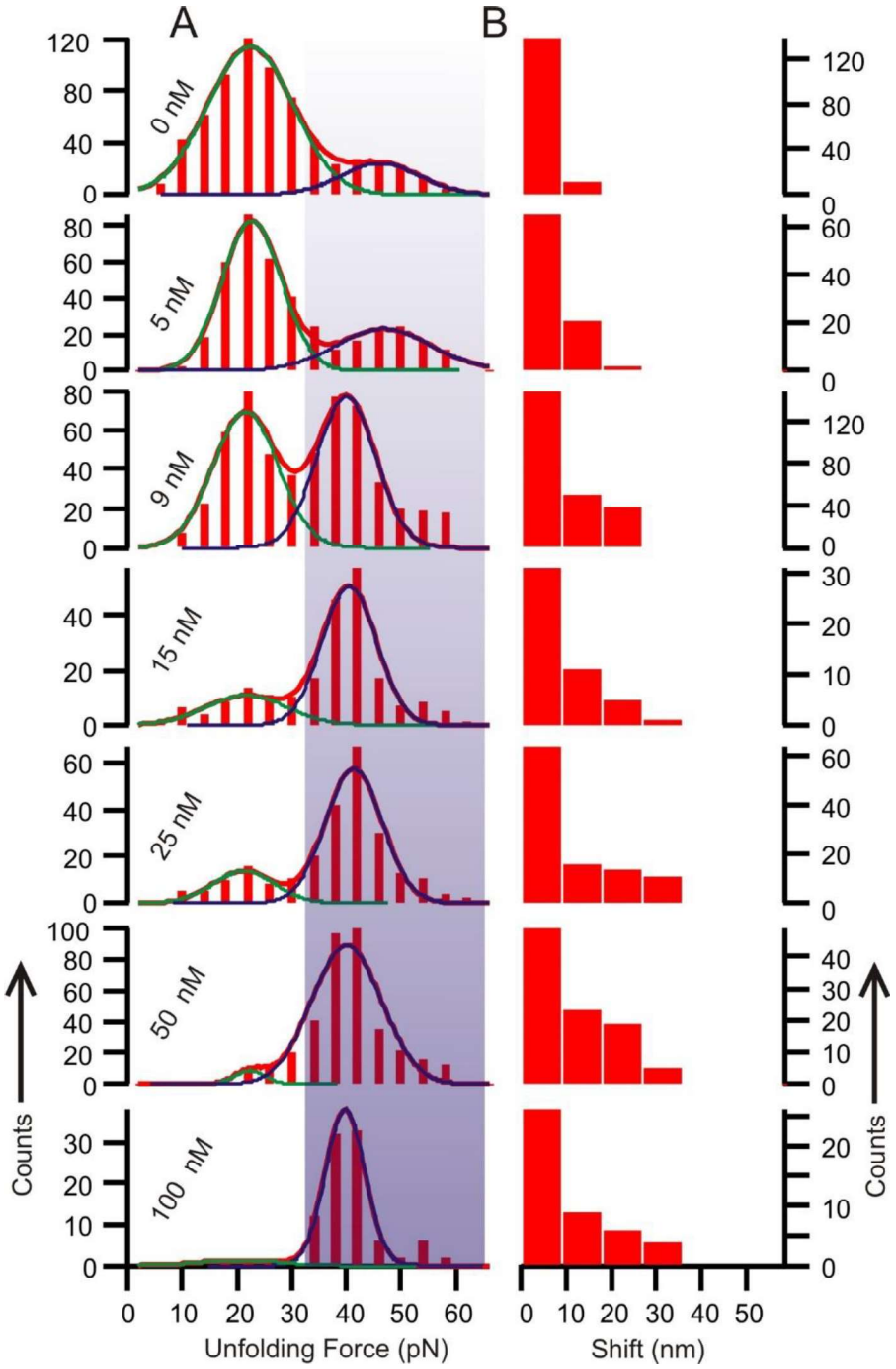


Figure S11: *(A) Unfolding force histograms of the 48G construct in the presence of the L2H2-6OTD Chimera 2.0 with various concentrations. (B) Corresponding histograms of non-unfolded G-quadruplexes.*

15. ***Calculation of Bound Fraction of G-quadruplexes in the 48G Construct in the Presence of the L2H2-6OTD***

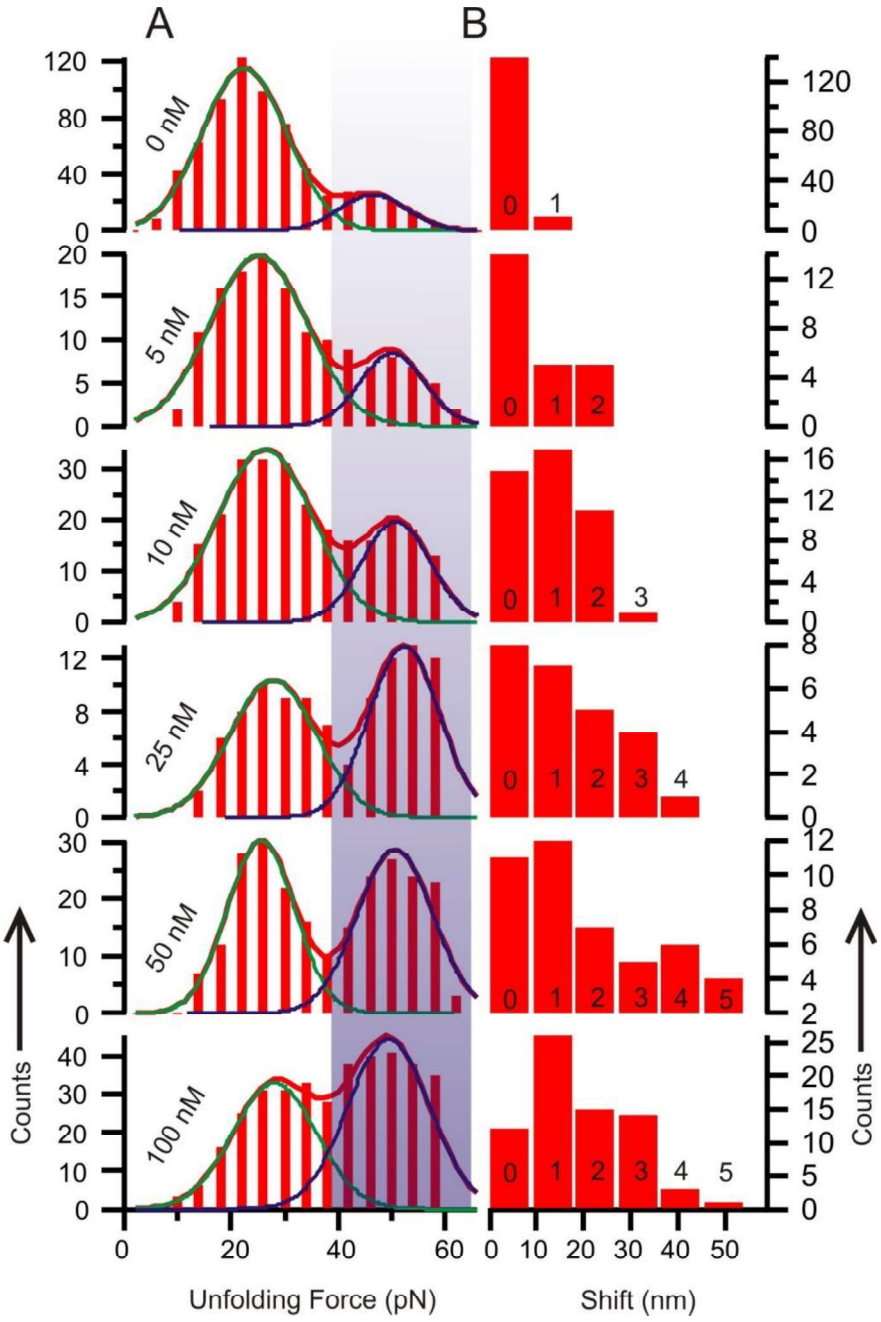


Figure S12: *(A) Unfolding force histograms of the 48G construct in the presence of the L2H2-6OTD monomer with various concentrations. (B) Corresponding histograms of non-unfolded G-quadruplexes.*

16. ***Calculation of Bound Fraction of the G-quadruplexes in the 24G Construct in the Presence of L2H2-6OTD Chimera 1.0***

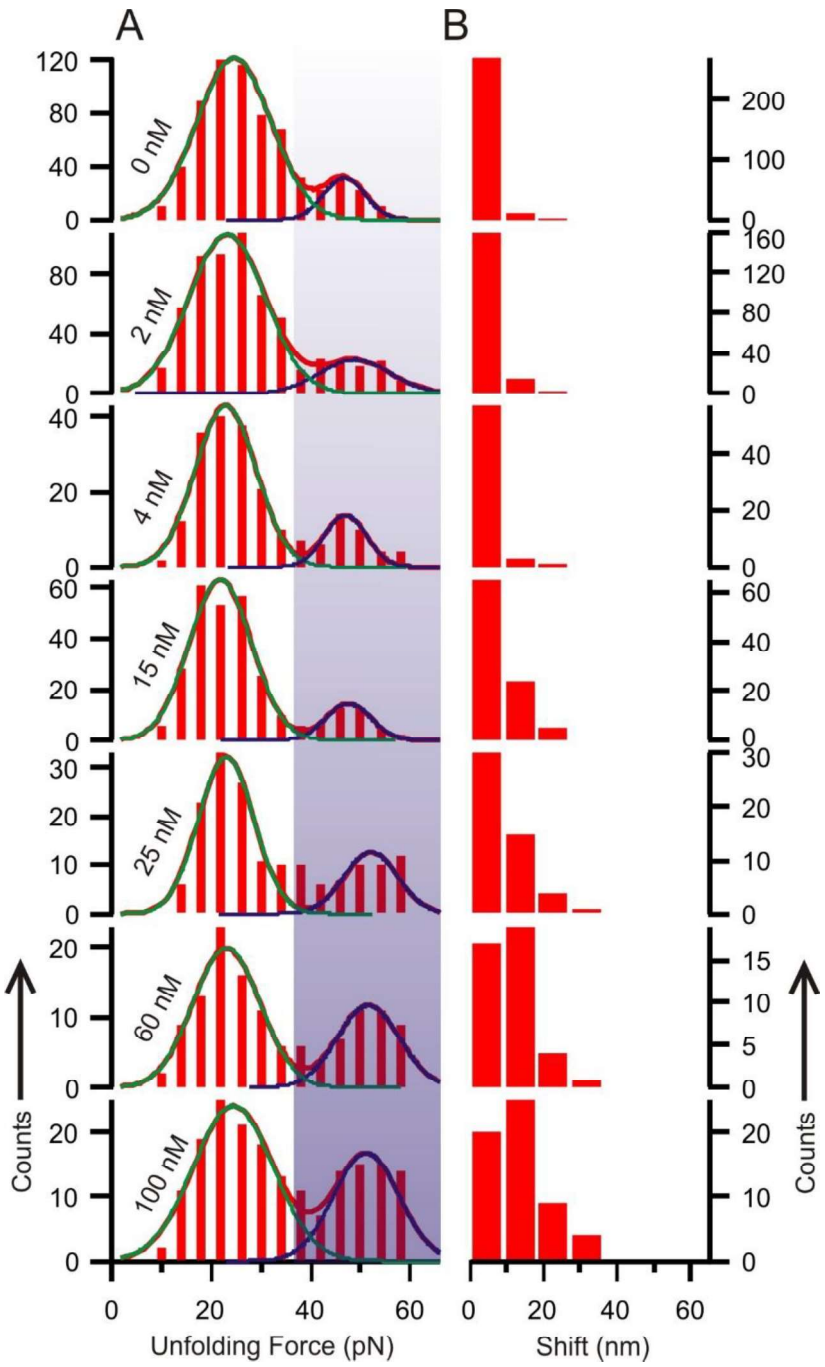


Figure S13: *(A) Unfolding force histograms of the 24G construct in the presence of the L2H2-6OTD Chimera 1.0 with various concentrations. (B) Corresponding histograms of non-unfolded G-quadruplexes.*

17. *L2H2-6OTD Azide Synthesis and Characterization*

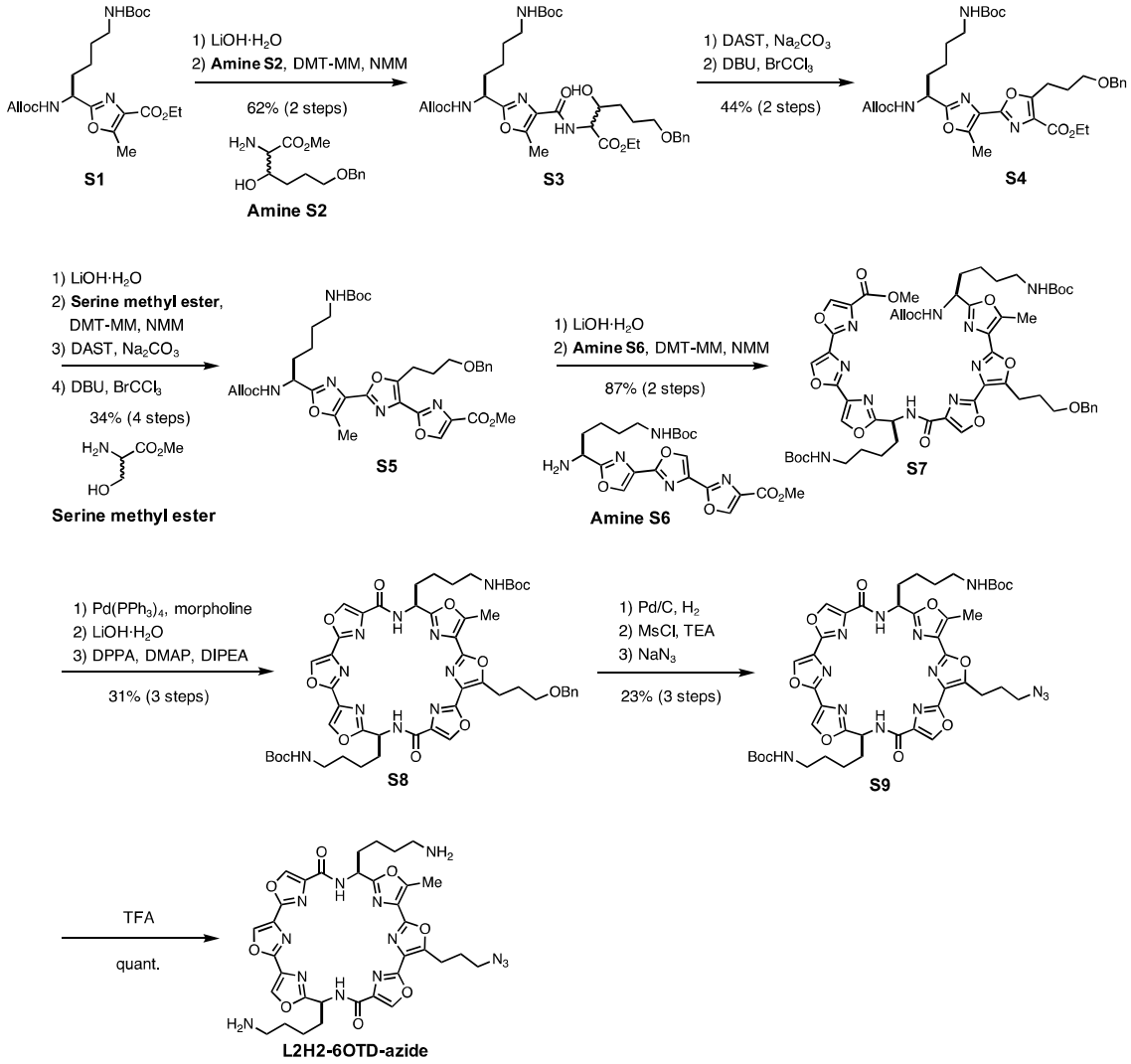
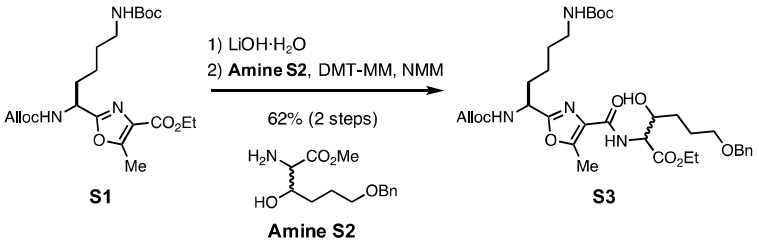
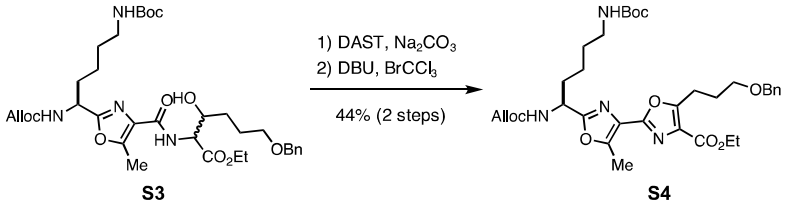


***Figure S14**: Scheme for the synthesis of L2H2-6OTD Azide.*

Flash chromatography was performed on silica gel 60 (spherical, particle size 0.040-0.100 mm; Kanto Co., Inc., Japan). Optical rotations were measured on a JASCO P-2200 polarimeter. ^{1}H and ^{13}C NMR spectra were recorded on JEOL JNM-AL 300, JNM-ECX 400 and JNM-ECA 500. The spectra are referenced internally according to the residual solvent signals of $CDCl_3$ (^{1}H NMR; δ = 7.26 ppm, ^{13}C NMR; δ = 77.0 ppm), DMSO-d_6 (^{1}H NMR; δ = 2.50 ppm, ^{13}C NMR; δ = 39.5 ppm). Data for ^{1}H NMR are recorded as follows; chemical shift (δ, ppm), multiplicity (s, singlet; t, triplet; m, multiplet; br, broad), integration, coupling constant (Hz). Data for ^{13}C NMR are reported in terms of chemical shift (δ, ppm). Mass spectra were recorded on JEOL JMS-T100LC spectrometer with ESI-MS mode using MeOH as solvent.

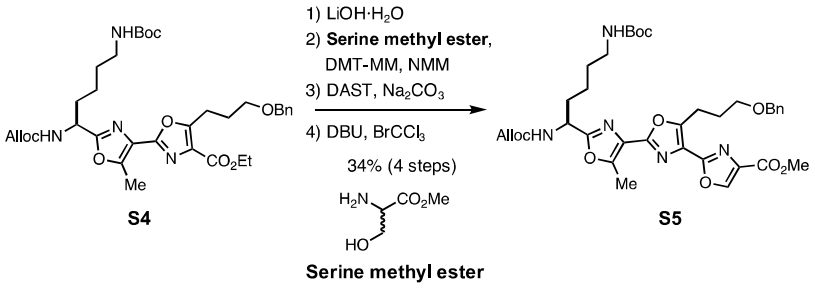


Amide S3: To a solution of **S1** [1] (3.26 g, 7.66 mmol) in THF-H_2O (3:1, 40 mL) was added LiOH·H_2O (643 mg, 15.3 mmol) at room temperature. After stirring for 3.5 h, the mixture was quenched with 3 N HCl to give a solution carboxylic acid, which was used without further purification. To a solution of the crude carboxylic acid was added NMM (2.26 mL, 20.4 mmol), DMT-MM (4.51 g, 15.3 mmol) and amine **S2** [1] (1.44 g, 5.11 mmol) at room temperature. After stirring for 3.5 h, the reaction was quenched with 1.2 N HCl. The organic layer was separated and the aqueous layer was extracted with EtOAc. The combined organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was purified by silica gel column ($CHCl_3$/EtOAc = 4:1) to give amide **S3** (2.15 g, 62%, 2 steps); $[\alpha]_D^{25}$ = −18.0 (*c* 0.71, $CHCl_3$); ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.72-7.42 (m, 1H), 7.31-7.26 (m, 5H), 5.93-5.88 (m, 1H), 5.35-5.20 (m, 2H), 4.88-4.71 (m, 2H), 4.60-4.50 (m, 4H), 4.26-4.20 (m, 2H), 3.51 (t, *J* = 5.5 Hz, 2H), 3.10 (d, *J* = 5.8 Hz, 2H), 2.60 (s, 3H), 2.07-1.60 (m, 6H), 1.51-1.22 (m, 15H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.0, 170.2, 162.2, 156.3, 155.8, 154.0, 138.2, 132.6, 128.8, 128.5, 128.3, 127.8, 127.7, 118.1, 79.3, 73.0, 72.1, 70.3, 66.1, 61.9, 61.8, 56.2, 49.2, 40.1, 33.7, 29.6, 28.5, 26.4, 22.4, 14.3, 11.8; HRMS (ESI, M+Na) calcd for $C_{34}H_{50}N_4O_{10}Na$ 697.3425, found 697.3436.



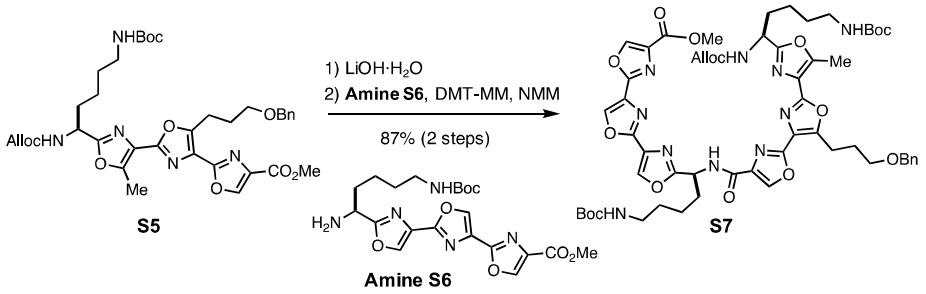
Oxazole S4: To a solution of amide **S3** (2.15 g, 3.19 mmol) in dry CH_2Cl_2 (35 mL) was cooled to 0 °C and was treated with Na_2CO_3 (1.01 g, 9.57 mmol) and DAST (506 μL, 3.83 mmol) under argon atmosphere. After stirring for 10 min, the mixture was quenched with saturated aqueous $NaHCO_3$. The organic layer was separated and the aqueous layer was washed with CH_2Cl_2. The combined organic extracts were dried over $MgSO_4$, filtered, and concentrated *in vacuo*. To a solution of the crude oxazoline

in CH_2Cl_2 (15 mL) was cooled to 0 °C and was treated with DBU (1.43 mL, 9.57 mmol), $BrCCl_3$ (949 μL, 9.57 mmol) under argon atmosphere. After stirring for 13 h, the reaction was quenched with 1.2 N HCl. The organic layer was separated and the aqueous layer was extracted with $CHCl_3$. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was purified by silica gel column (hexane/EtOAc = 2:1) to give oxazole **S4** (928 mg, 44%, 2 steps); $[\alpha]_D^{25} = -12.0$ (*c* 0.90, $CHCl_3$); ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.33-7.25 (m, 5H), 5.96-5.85 (m, 1H), 5.49 (d, *J* = 7.2 Hz, 1H), 5.27 (dd, *J* = 17.2, 10.3 Hz, 2H), 4.93 (dt, *J* = 7.6, 6.2 Hz, 1H), 4.57 (d, *J* = 5.5 Hz, 2H), 4.50 (s, 2H), 4.37 (q, *J* = 7.2 Hz, 2H), 3.54 (t, *J* = 6.2 Hz, 2H), 3.21 (t, *J* = 7.6 Hz, 2H), 3.09 (d, *J* = 5.8 Hz, 2H), 2.65 (s, 3H), 2.10-1.79 (m, 4H), 1.58-1.21 (m, 15H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.8, 162.1, 159.0, 156.1, 155.8, 150.7, 138.3, 132.5, 130.1, 129.5, 128.3, 127.6, 117.9, 79.2, 73.0, 69.2, 66.0, 61.0, 49.1, 40.0, 33.8, 29.5, 28.4, 28.0, 23.2, 22.4, 14.3, 11.8; HRMS (ESI, M+Na) calcd for $C_{34}H_{46}N_4O_9Na$ 677.3163, found 677.3203.

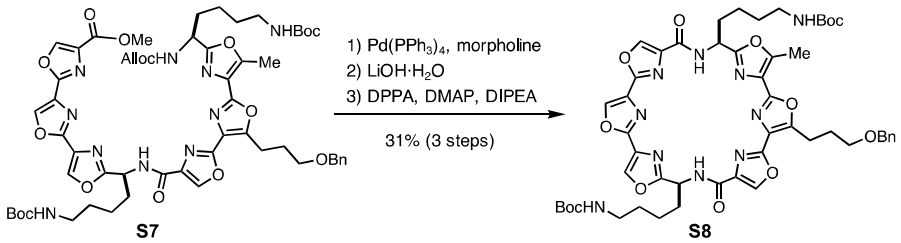


Oxazole S5: To a solution of oxazole **S4** (900 mg, 1.37 mmol) in THF-H_2O (3:1, 20 mL) was added LiOH·H_2O (115 mg, 2.75 mmol) at room temperature. After stirring for 1.5 h, the mixture was quenched with 3 N HCl to give a solution carboxylic acid, which was used without further purification. To a solution of the crude carboxylic acid was added NMM (909 μL, 8.22 mmol), DMT-MM (1.62 g, 5.48 mmol) and serine methyl ester (4.26 mg, 2.74 mmol) at room temperature. After stirring for 1 h, the reaction was quenched with 1.2 N HCl. The organic layer was separated and the aqueous layer was extracted with EtOAc. The combined organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated *in vacuo* to give amide, which was used without further purification. To a solution of crude amide in dry CH_2Cl_2 (30 mL) was cooled to 0 °C and was treated with Na_2CO_3 (277 mg, 2.61 mmol) and DAST (181 μL, 1.37 mmol) under argon atmosphere. After stirring for 10 min, the mixture was quenched with saturated aqueous $NaHCO_3$. The organic layer was separated and the aqueous layer was washed with $CHCl_3$. The combined organic extracts were dried over $MgSO_4$, filtered, and concentrated *in vacuo*. To a solution of the crude oxazoline in CH_2Cl_2 (15 mL) was cooled to 0 °C and was treated with DBU (1.30 mL, 8.70 mmol), $BrCCl_3$ (862 μL, 8.70 mmol) under argon atmosphere. After stirring for 16 h, the reaction was quenched with 1.2 N HCl. The organic layer was separated and the aqueous layer was extracted with $CHCl_3$. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was purified by silica gel column ($CHCl_3$/EtOAc = 5:1) to give oxazole **S5** (300 mg, 34%, 4 steps); $[\alpha]_D^{25} = -12.7$ (*c* 1.10, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.25 (s, 1H), 7.34-7.29 (m, 5H), 5.95-5.88 (m, 1H), 5.47 (d, *J* = 7.8 Hz, 1H), 5.27 (dd, *J* = 16.5, 10.5 Hz, 2H), 4.95 (dt, *J* = 7.8, 6.4 Hz, 1H), 4.59 (d, *J* = 6.0 Hz, 2H), 4.49 (s, 2H), 3.92 (s, 3H), 3.57 (t, *J* = 6,4 Hz, 2H), 3.33 (t, *J* = 7.6 Hz,

2H), 3.1 (d, J = 6.0 Hz, 2H), 2.69 (s, 3H), 2.17-2.08 (m, 2H), 2.04-1.82 (m, 2H), 1.43-1.37 (m, 13H); ^{13}C NMR (100 MHz,$CDCl_3$) δ 162.8, 161.6, 156.6, 156.0, 155.7, 154.2, 151.4, 150.8, 143.4, 143.3, 138.3, 134.2, 132.5, 128.3, 127.6, 127.5, 125.3, 124.6, 117.9, 79.2, 76.6, 72.9, 69.1, 66.0, 52.2, 49.2, 40.0, 33.9, 33.8, 29.5, 28.4, 27.8, 22.7, 22.4, 11.8; HRMS (ESI, M+Na) calcd for $C_{36}H_{45}N_5O_{10}Na$ 730.3064, found 730.3069.

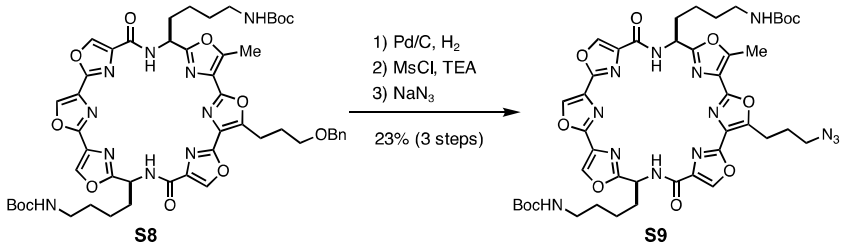


Bistrioxazole S7: To a solution of oxazole S5 (300 mg, 467 μmol) in THF-H_2O (3:1, 8 mL) was added LiOH·H_2O (39.2 mg, 934 μmol) at room temperature. After stirring for 45 min, the mixture was quenched with 1.2 N HCl to give a solution carboxylic acid, which was used without further purification. To a solution of the crude carboxylic acid was added NMM (103 μL, 934 μmol), DMT-MM (206 mg, 700 μmol) and amine S6 [2] (323 mg, 700 μmol) at room temperature. After stirring for 18 h, the reaction was quenched with 1.2 N HCl. The organic layer was separated and the aqueous layer was extracted with $CHCl_3$. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was purified by silica gel column ($CHCl_3$/EtOAc = 8:1) to give bistrioxazole S7 (464 mg, 87%, 2 steps); $[\alpha]_D^{25}$ = −31.5 (*c* 0.32, MeOH); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.40 (s, 1H), 8.31(s, 1H), 8.30 (s, 1H), 8.23 (s, 1H), 7.52 (d, J = 8.7 Hz, 1H), 7.31-7.23 (m, 5H), 5.95-5.86 (m, 1H), 5.48 (m, 2H), 5.27 (dd, J = 16.5, 10.5 Hz, 2H), 4.95 (dt, J = 7.3, 6.4 Hz, 1H), 4.59-4.50 (m, 4H), 4.48 (s, 2H), 3.95 (s, 3H), 3.57 (t, J = 6.0 Hz, 2H), 3.30 (t, J = 7.6 Hz, 2H), 3.13-3.06 (m, 4H), 2.70 (s, 3H), 2.20-1.84 (m, 6H), 1.61-1.41 (m, 26H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.8, 161.3, 160.1, 156.1, 156.0, 155.9, 155.4, 154.1, 150.9, 143.9, 141.2, 139.6, 139.3, 138.2, 136.4, 132.5, 130.8, 129.9, 128.4, 127.6, 127.5, 125.3, 124.6, 117.9, 79.2, 73.0, 69.0, 66.0, 52.3, 49.2, 46.8, 40.2, 40.0, 33.9, 33.4, 29.5, 28.4, 28.0, 22.8, 22.4, 11.2; HRMS (ESI, M+Na) calcd for $C_{56}H_{68}N_{10}O_{16}Na$ 1159.4712, found 1159.4725.



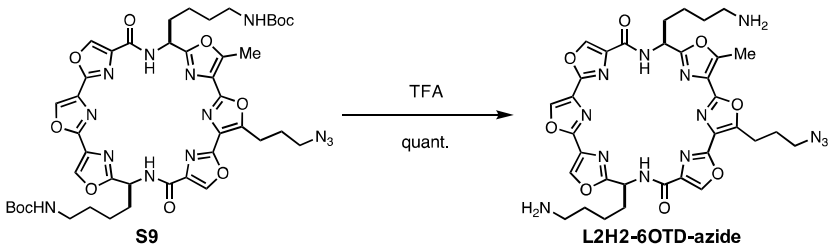
Macrocyclic bisamide S8: To a solution of bistrioxazole S7 (414 mg, 364 μmol) in THF (7 mL) was added morpholine (317 μl, 3.64 mmol) and $Pd(PPh_3)_4$ (126 mg, 109 μmol) at room temperature under argon atmosphere. After stirring for 19 h, the mixture was concentrated *in vacuo* to give amine, which was used without further purification. To a solution of the crude amine in THF-H_2O (3:1, 4 ml), was added LiOH·H_2O (18.0 mg, 429 μmol) at room temperature. After stirring for 2.5 h, the mixture was

quenched with 3 N HCl, and concentrated *in vacuo* to give carboxylic acid, which was used without further purification. To a solution of the crude carboxylic acid in dry DMF-CH_2Cl_2 (1:2, 72 ml) was added $^{i}Pr_2NEt$ (146 μL, 856 μmol), DMAP (52.3 mg, 428 μmol) and DPPA (461 μL, 2.14 mmol). The resulting mixture was refluxed for 17 h under argon atmosphere. The reaction was quenched with 1.2 N HCl and the organic layer was separated and the aqueous layer was extracted with EtOAc. The combined organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was purified by silica gel column (CH_3Cl/MeOH = 50:1) to give macrocyclic bisamide **S8** (68.0 mg, 31%, 3 steps); $[\alpha]_D^{25} = -7.14$ (*c* 0.84, $CHCl_3$); ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.11 (s, 1H), 9.10 (s, 1H) 8.88 (s, 1H), 8.80 (s, 1H), 8.33 (d, *J* = 7.3 Hz, 1H), 8.29 (d, *J* = 7.8 Hz, 1H), 7.38-7.21 (m, 5H), 6.75 (br, 2H), 5.43 (dt, *J* = 7.3, 5.0 Hz, 1H), 5.33 (dt, *J* = 6.9, 5.0 Hz, 1H), 4.40 (s, 2H), 3.51 (t, *J* = 6.0 Hz, 2H), 3.30-3.16 (m, 2H), 2.85-2.75 (m, 4H), 2.65 (s, 3H), 2.07-1.90 (m, 6H), 1.43-1.16 (m, 26H); ^{13}C NMR (100 MHz, DMSO-d_6) δ 164.5, 162.0, 158.8, 158.6, 155.6, 155.5, 154.5, 154.0, 151.3, 142.4, 141.5, 140.9, 138.4, 135.9, 129.7, 128.4, 128.1, 127.2, 124.6, 123.7, 79.1, 77.1, 68.4, 47.3, 47.2, 33.4, 33.3, 29.1, 28.1, 27.3, 22.4, 21.1, 20.9, 11.4; HRMS (ESI, M+Na) calcd for $C_{51}H_{60}N_{10}O_{13}Na$ 1043.4239, found 1043.4219.



S8 **S9**

Azide S9: To a solution of macrocyclic bisamide **S8** (52.9 mg, 51.8 μmol) in MeOH-THF (1:1, 2 mL) was added $Pd(OH)_2$/C (15.7 mg, 30 wt%), and the reaction mixture was stirred at room temperature under hydrogen atmosphere. After 27 h, the mixture was filtrated through a pad of celite and eluted with CH_3Cl-MeOH (9:1). The solution was concentrated *in vacuo* to give alcohol, which was used without further purification. To a solution of the crude alcohol in CH_2Cl_2 (1 ml) was added TEA (29.1 μL, 207 μmol) and MsCl (8.02 μL, 104 μmol) at 0 °C. After stirring for 3 h, the mixture was quenched with H_2O, and the organic layer was separated and the aqueous layer was extracted with CH_2Cl_2. The combined organic layer was dried over $MgSO_4$, filtered, and concentrated *in vacuo* to give Ms protected alcohol, which was used without further purification. To a solution of crude Ms protected alcohol in dry DMF (500 μL) was added NaN_3 (2.32 mg, 11.9 μmol) at room temperature, and the mixture was warmed to 50 °C under argon atmosphere. After 1.5 h, the mixture was allowed to cool to room temperature. The reaction was quenched with H_2O, and the mixture was extracted with EtOAc. The combined organic layer was washed with brine, dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was purified by silica gel column (CH_3Cl/MeOH = 50:1) to give azide **S9** (11.5 mg, 23%, 3 steps); $[\alpha]_D^{25} = -0.25$ (*c* 1.57, $CHCl_3$); ^{1}H NMR (300 MHz, DMSO-d_6) δ 9.12 (s, 1H), 9.02 (s, 1H), 8.91 (s, 1H), 8.89 (s, 1H), 8.33 (d, *J* = 6.5 Hz, 1H), 8.28 (d, *J* = 7.2 Hz, 1H), 6.77 (t, *J* = 5.5 Hz, 2H), 5.38 (dt, *J* = 7.2, 5.5 Hz, 1H), 5.34 (dt, *J* = 7.2, 4.8 Hz, 1H), 3.51-3.44 (m, 2H), 3.28-3.15 (m, 2H), 2.85-2.81 (m, 4H), 2.71 (s, 3H), 2.08-1.86 (m, 6H), 1.42-1.23 (m, 26H); ^{13}C NMR (75 MHz, DMSO-d_6) δ 164.5, 162.1, 158.8, 158.7, 155.6, 155.5, 155.4, 154.6, 154.5, 153.2, 151.5, 142.5, 142.3, 141.7, 141.0, 136.1, 136.0, 129.8, 128.4, 124.8, 123.7, 79.2, 77.2, 49.8, 47.3, 33.3, 29.1, 28.1, 26.5, 22.7, 21.1, 21.0, 11.5; HRMS (ESI, M+Na) calcd for $C_{44}H_{53}N_{13}O_{12}Na$ 978.3834, found 978.3834.



S9 **L2H2-6OTD-azide**

L2H2-6OTD-azide: To a solution of azide S9 (6.50 mg, 6.80 μmol) was added CH_2Cl_2-TFA (1:1, 1 mL) at 0 °C. The reaction mixture was stirred for 10 min, concentrated *in vacuo* to give L2H2-6OTD-azide (6.2 mg, quant.); $[\alpha]_D^{25}$ = +67.7 (*c* 0.62, MeOH); ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.14 (s, 1H), 9.12 (s, 1H), 8.93 (s, 1H), 8.90 (s, 1H), 8.31 (d, *J* = 7.5 Hz, 1H), 8.28 (d, 7.5 Hz, 1H) 7.69 (br, 6H), 5.45 (dt, *J* = 7.4, 5.7 Hz, 1H), 5.35 (dt, *J* = 6.9, 5.7 Hz, 1H), 3.49 (t, *J* = 6.6 Hz, 2H), 3.23 (t, *J* = 5.7 Hz, 2H), 2.77-2.72 (m, 7H), 2.11-1.87 (m, 6H), 1.56-1.38 (m, 6H), 1.23-1.17 (m, 2H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 164.4, 162.0, 158.9, 158.7, 158.0, 157,8, 155.6, 155.4, 154.6, 154.5, 153.3, 151.6, 142.5, 142.4, 141.8, 141.0, 136.0, 135.8, 129.7, 128.5, 124.8, 123.7, 49.8, 47.2, 47.0, 33.4, 26.7, 26.7, 26.4, 22.7, 20.9, 20.9, 11.5; HRMS (ESI, M+Na) calcd for $C_{34}H_{37}N_{13}O_8Na$ 778.2786, found 778.2780.

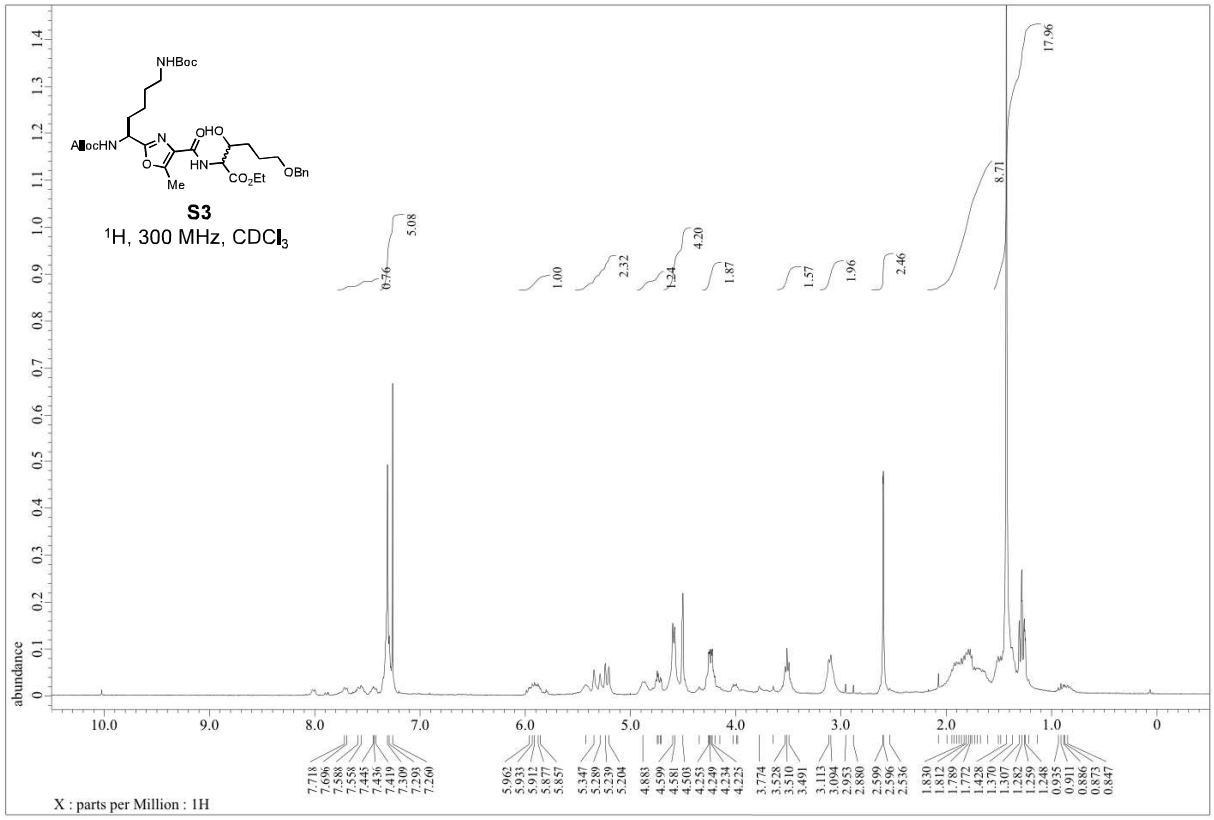


Figure S15: ^{1}H NMR spectrum for compound **S3**.

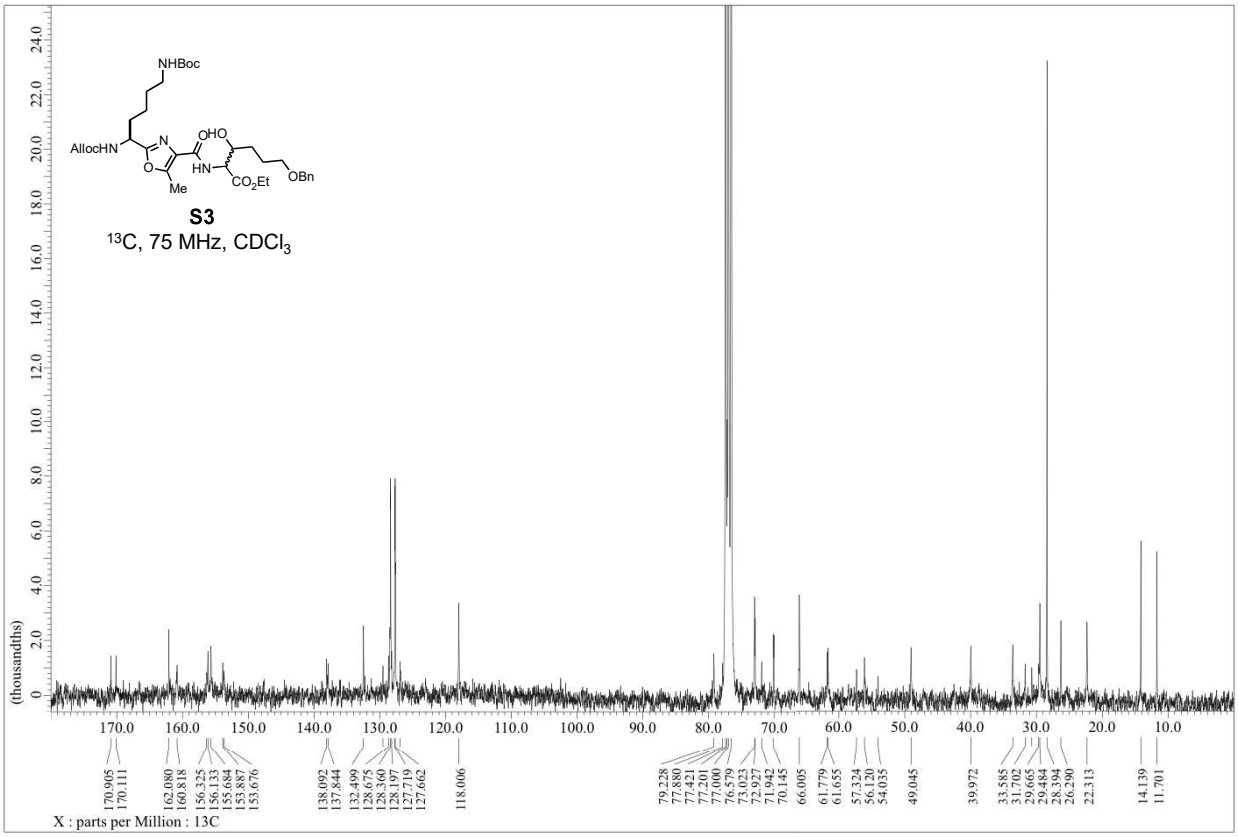


Figure S16: ^{13}C NMR spectrum for compound **S3**.

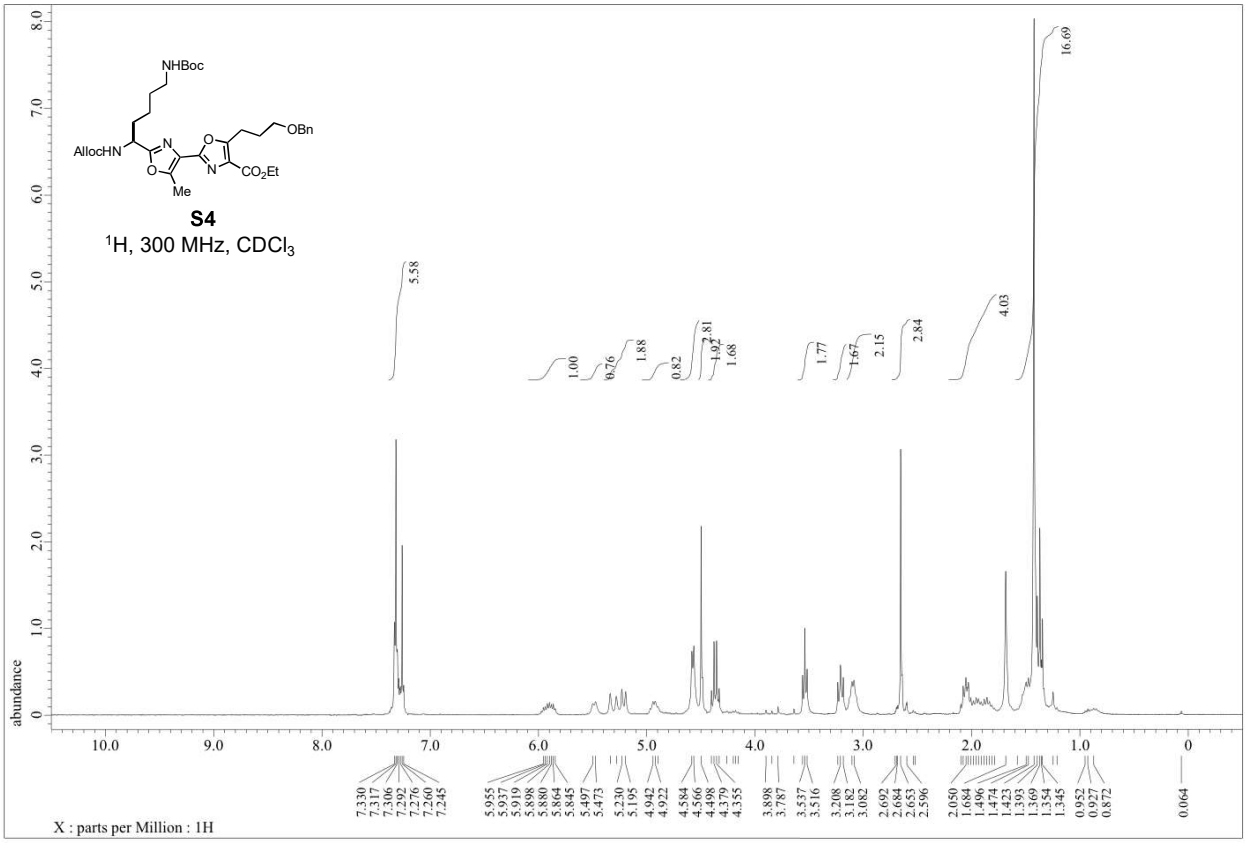


Figure S17: ^{1}H NMR spectrum for compound **S4**.

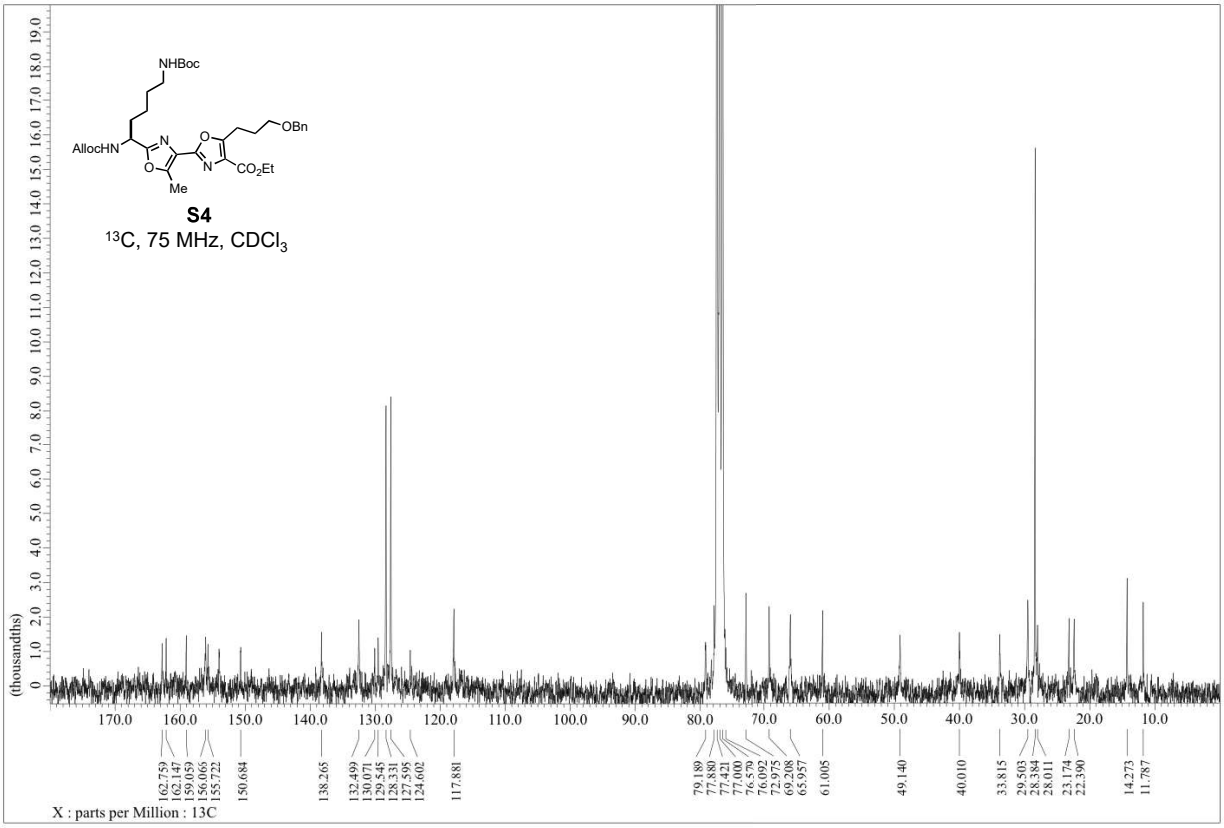


Figure S18: ^{13}C NMR spectrum for compound **S4**.

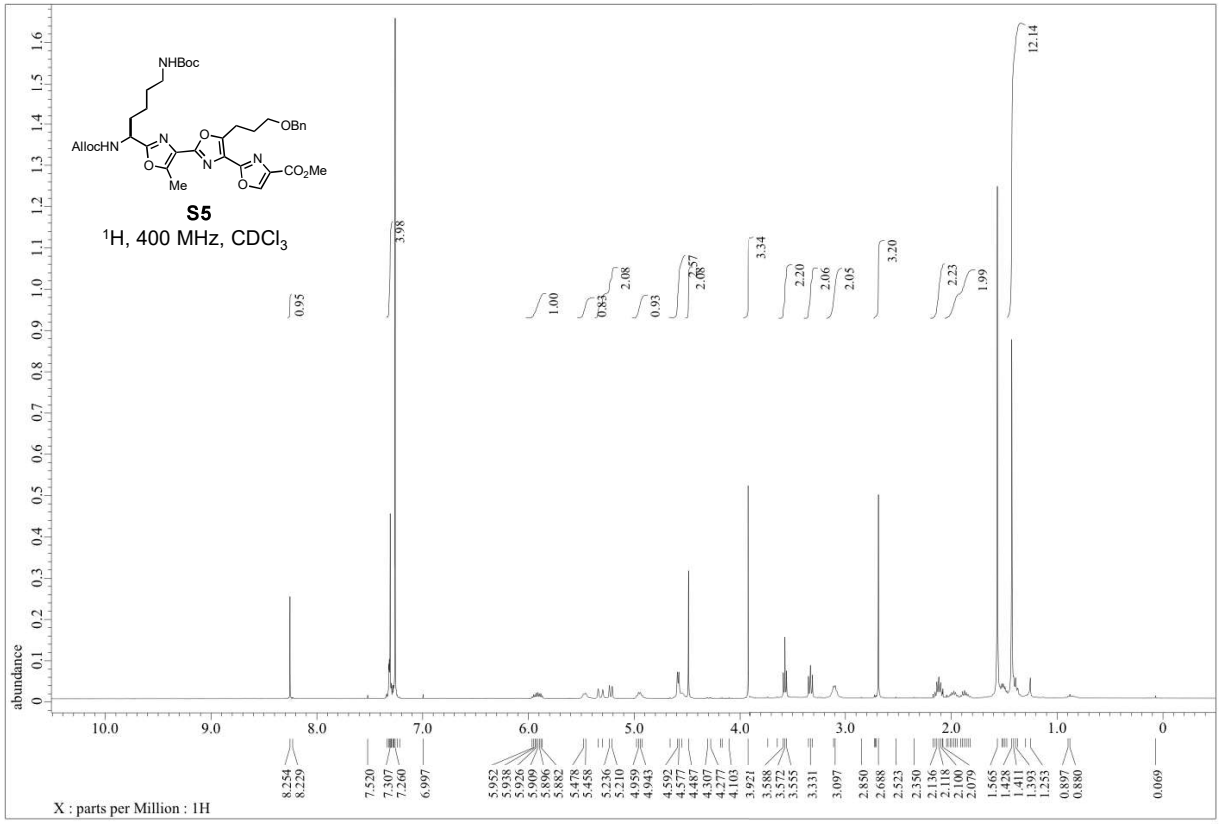


Figure S19: ^{1}H NMR spectrum for compound **S5**.

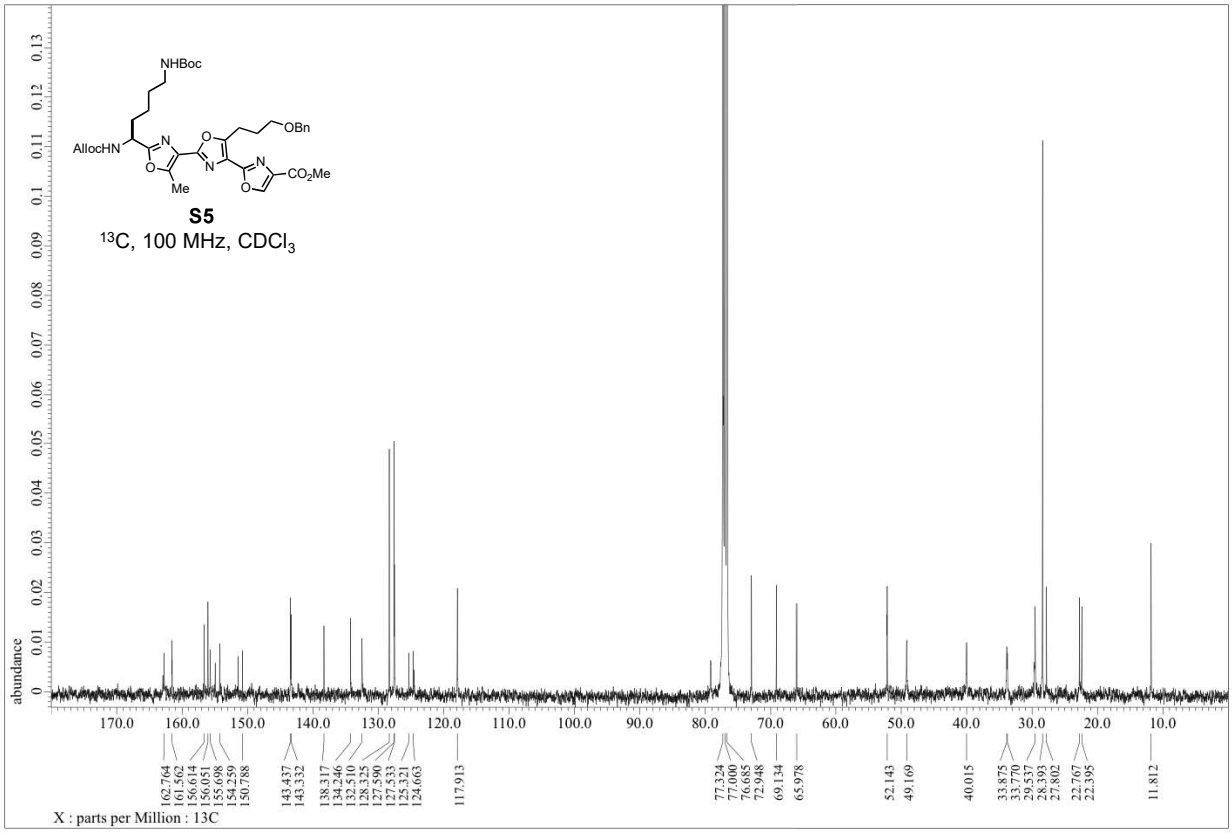


Figure S20: ^{13}C NMR spectrum for compound **S5**.

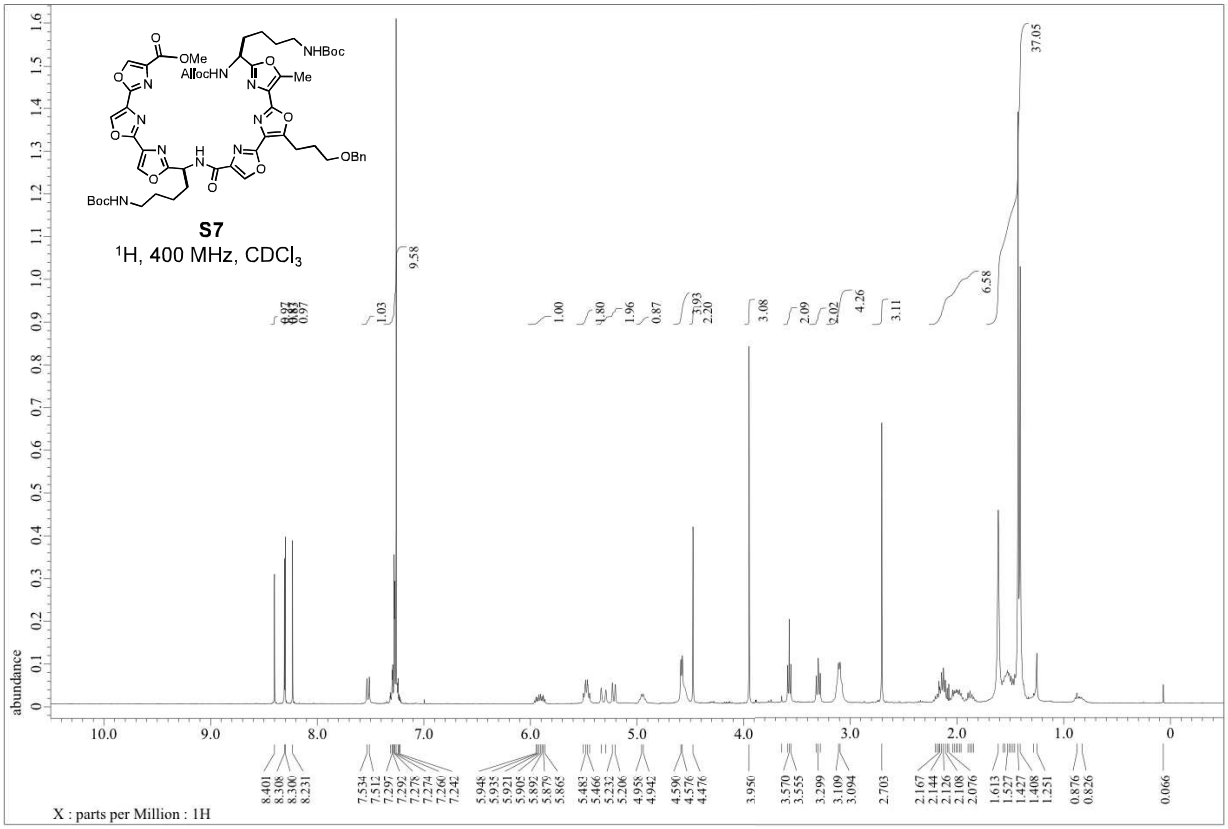


Figure S21: ^{1}H NMR spectrum for compound **S7**.

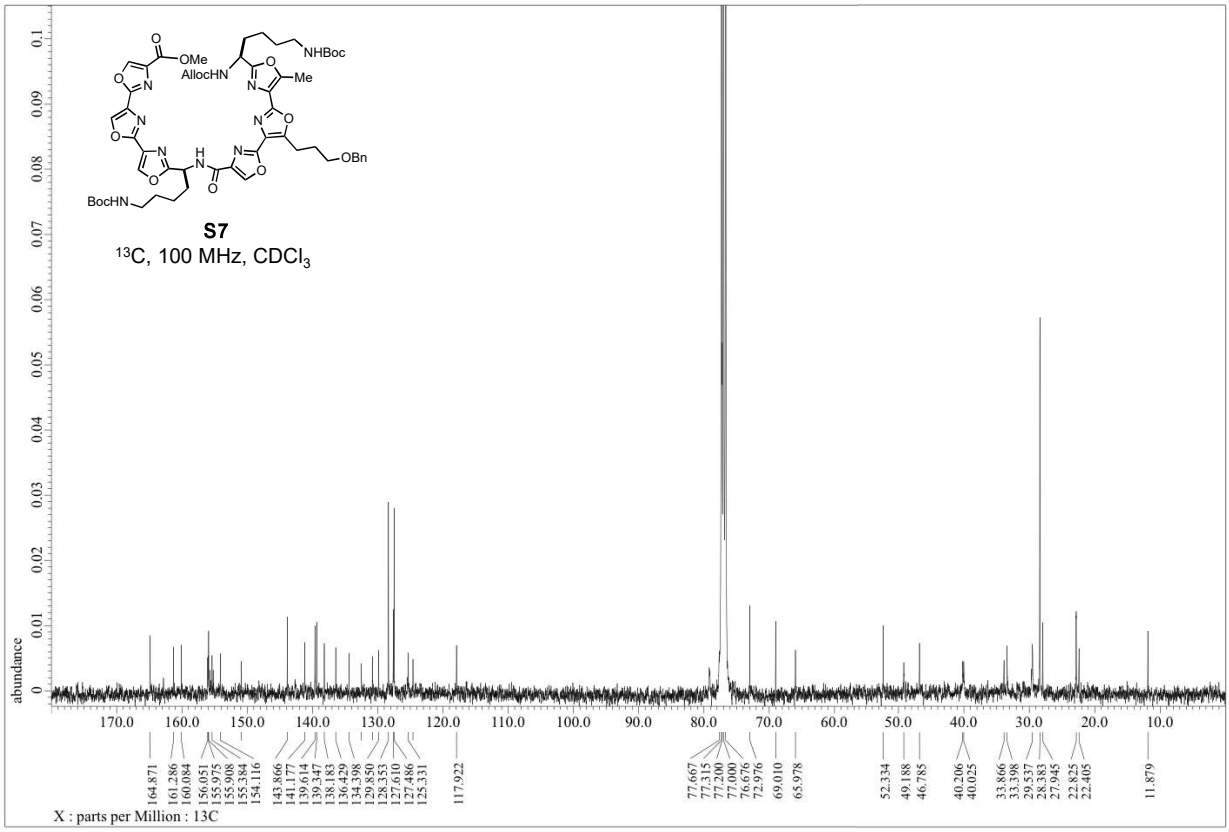


Figure S22: ^{13}C NMR spectrum for compound **S7**.

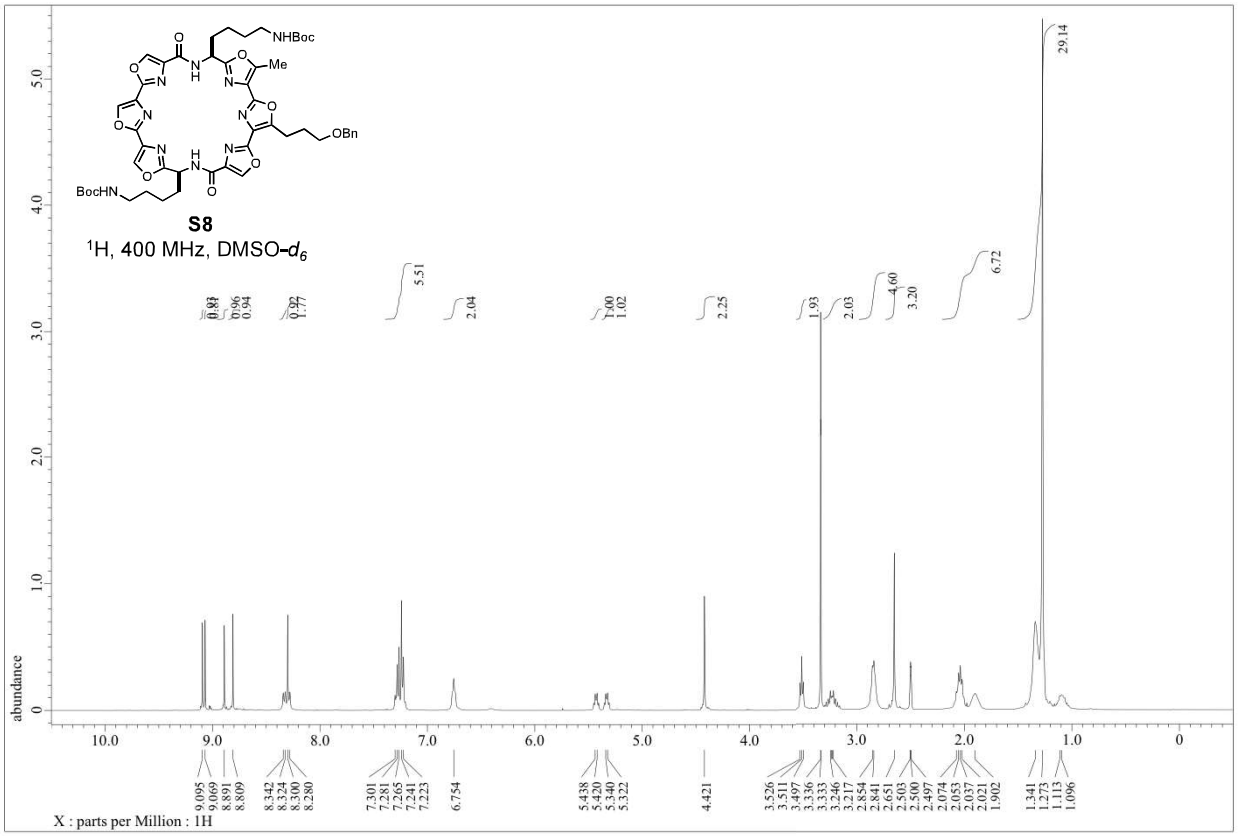


Figure S23: ^{1}H NMR spectrum for compound **S8**.

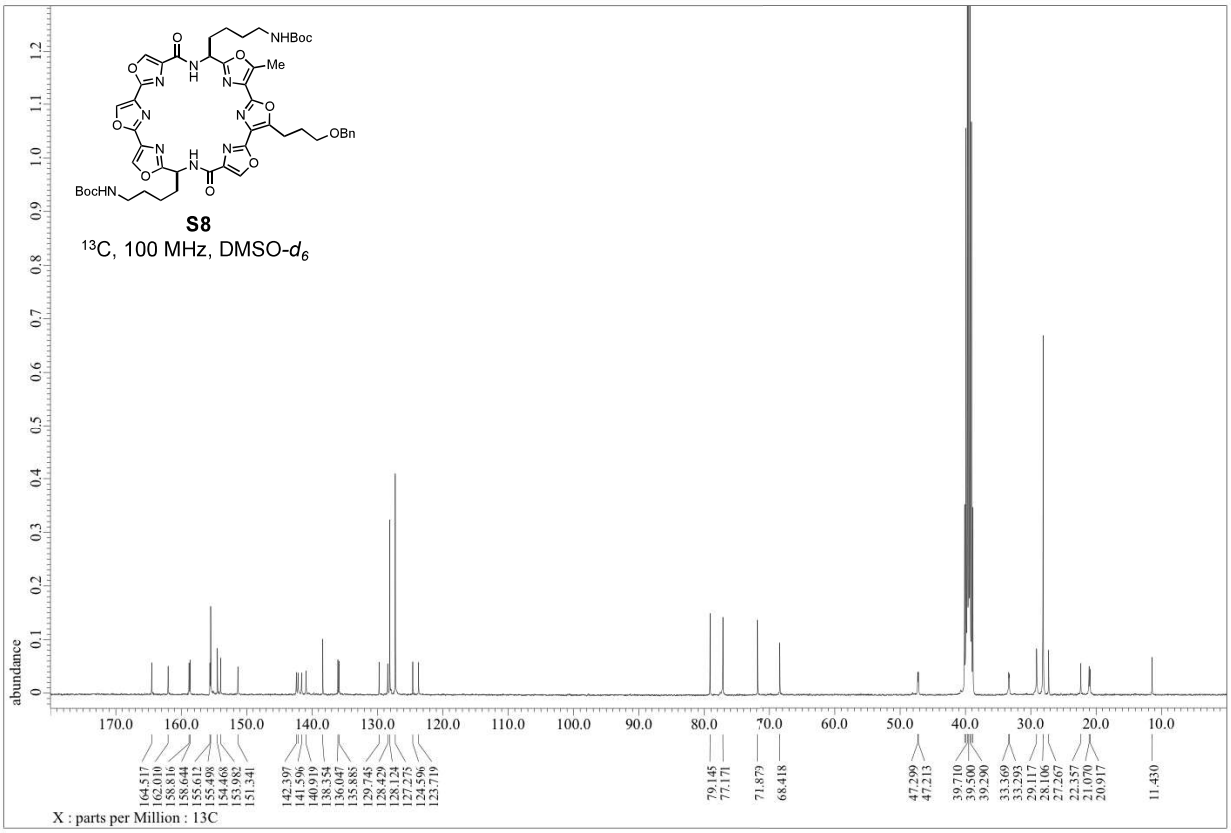


Figure S24: ^{13}C NMR spectrum for compound **S8**.

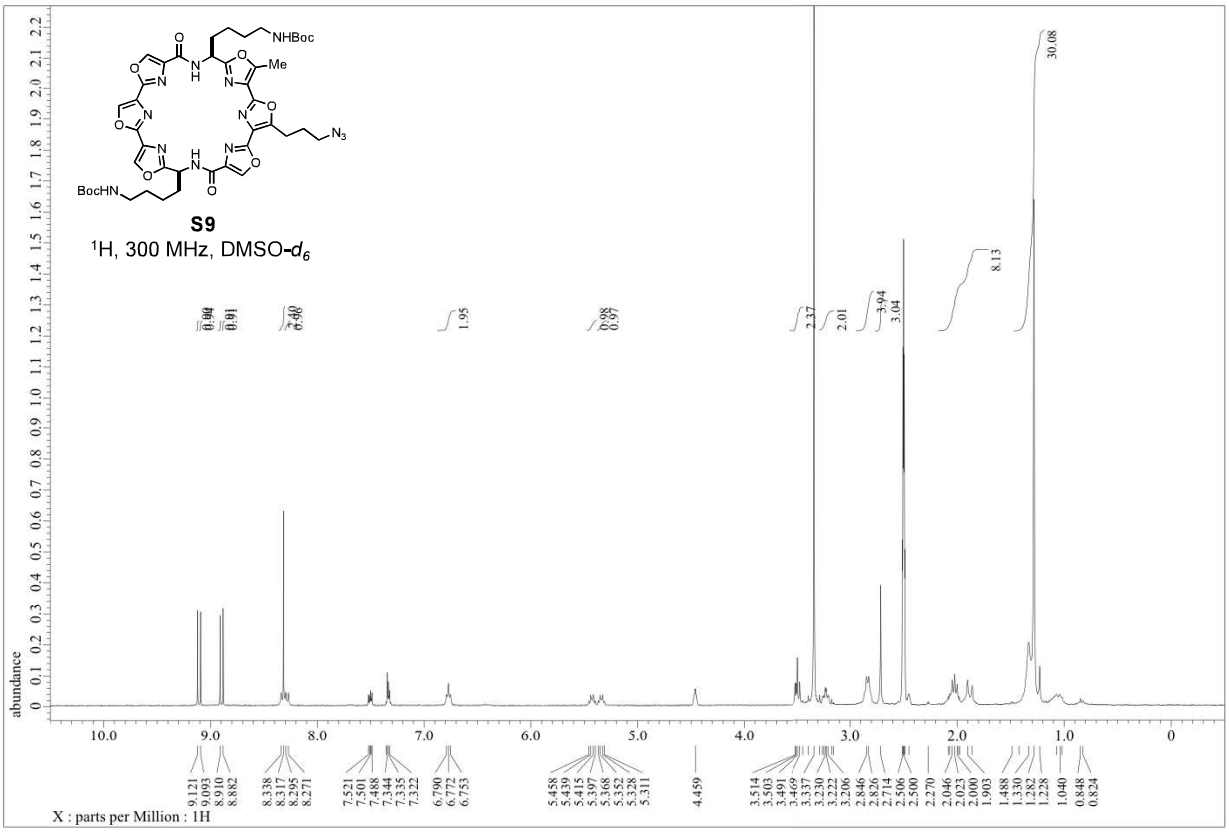


Figure S25: ^{1}H NMR spectrum for compound **S9**.

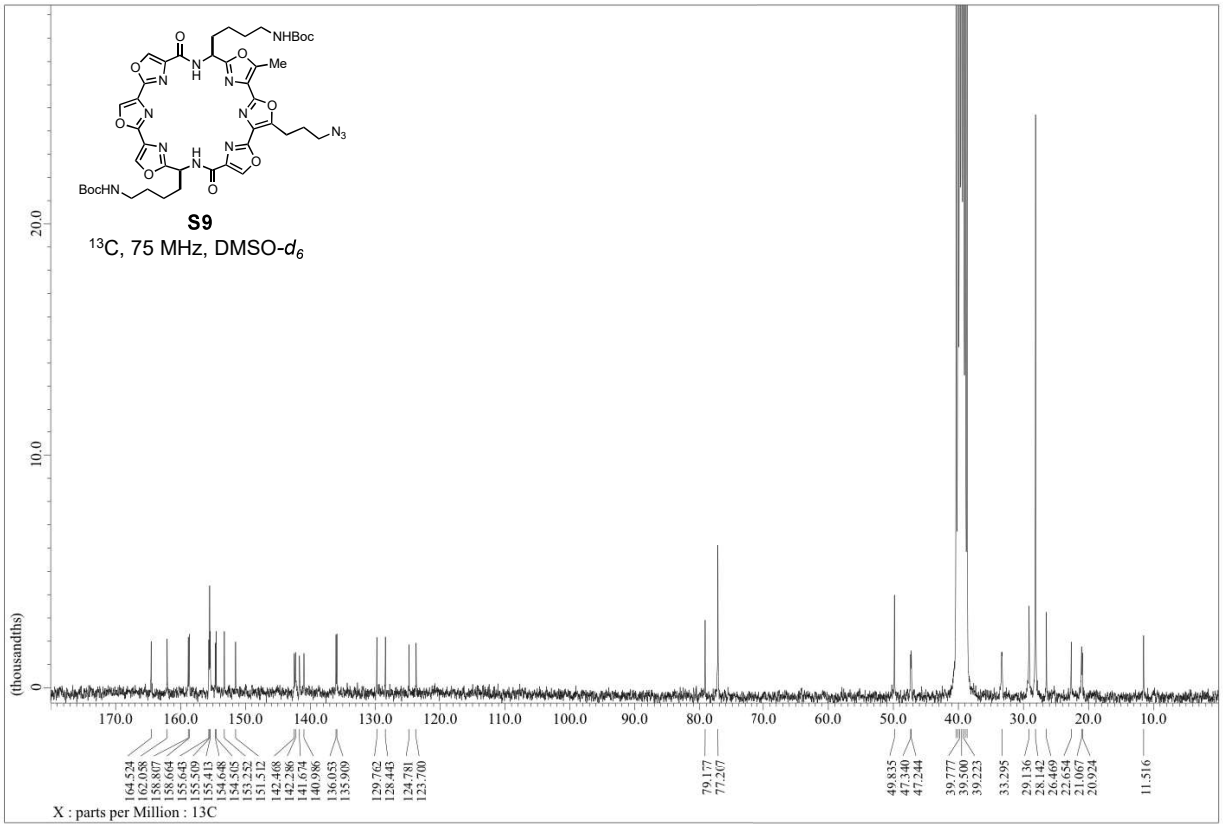


Figure S26: ^{13}C NMR spectrum for compound **S9**.

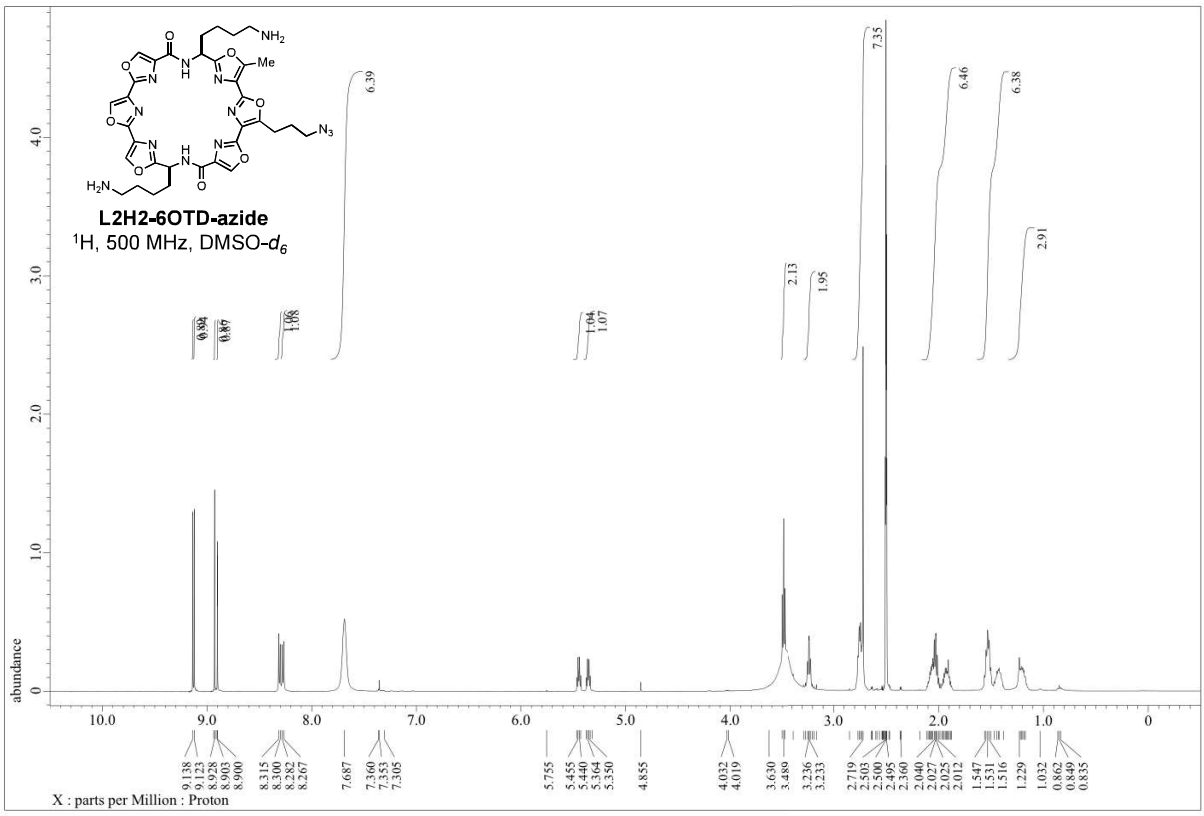


Figure S27: ^{1}H NMR spectrum for compound **L2H2-6OTD-azide**.

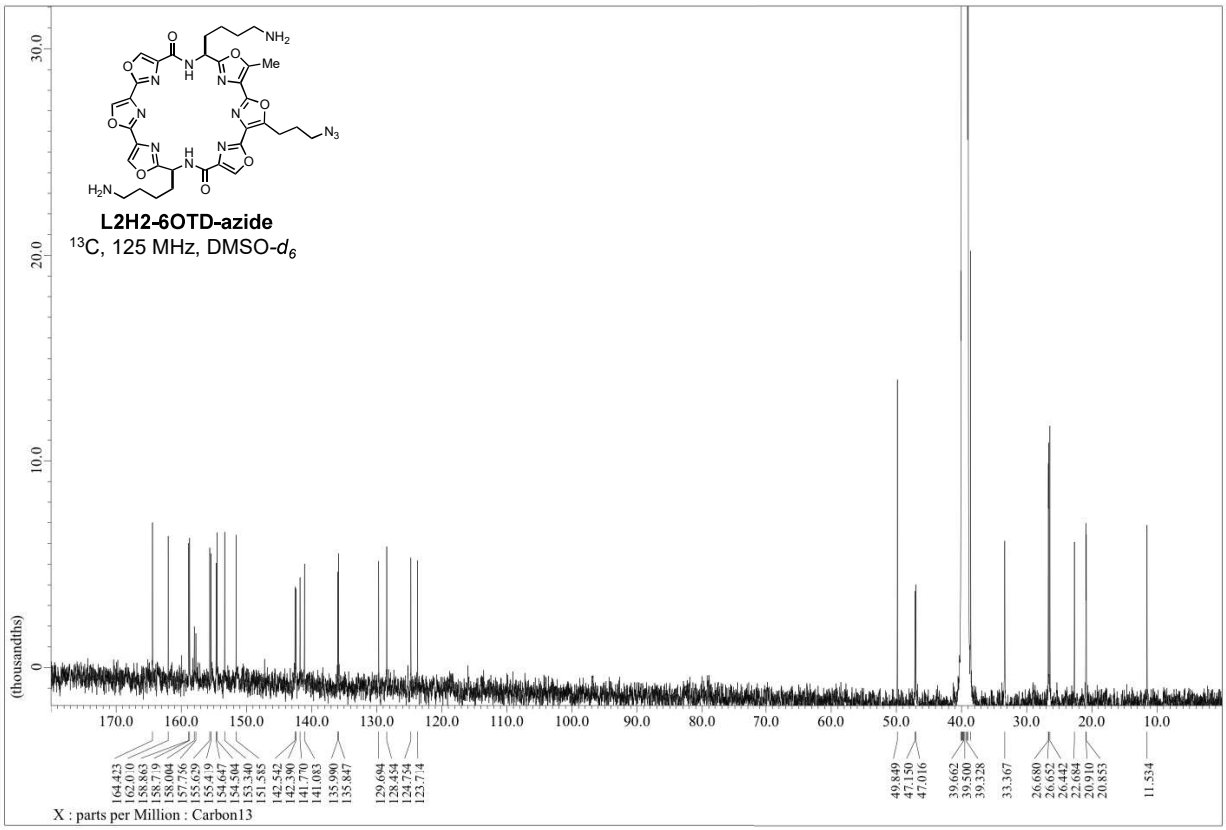


Figure S28: ^{13}C NMR spectrum for compound **L2H2-6OTD-azide**.

18. *Matlab Code for the Statistical Calculation of the Formation Pattern of Telomere G-quadruplexes*

```
clear all
lengtht=24;                 // this equals to number of G-tracts
ng=4;
nouse=mod(lengtht,ng);
n=(lengtht-nouse)/ng;
for i=1:n
sumn=lengtht-i*ng;
tn=i+1;
index=linspace(0,sumn,sumn+1);
l=length(index);
id=nchoosek(1:l+tn-1,tn)-repmat(0:tn-1,nchoosek(l+tn-1,tn),1);
temp=index(id);
r1=temp(sum(temp,2)==sumn,:);
eval(['j',num2str(i),'=[];']);
p=size(r1);
p=p(1);
for k=1:p
qr1=r1(k,:);
rr=perms(qr1);
eval(['j',num2str(i),'=[j',num2str(i),';rr];']);
end
eval(['j',num2str(i),'=sortrows(j',num2str(i),');']);
eval(['trj=j',num2str(i),';']);
rrj=unique(trj,'rows');
tsrrj=size(rrj);
srrj(i)=tsrrj(1);
rsrrj(i)=srrj(i)*i;
eval(['zrj',num2str(i),'=rrj;']);
end
z1=sum(srrj)*n;
z2=sum(rsrrj);
zeff=z2/z1;
```

19. *Igor Code for Folding Simulation of Telomere G-quadruplex and Intermediates*

```
function TeloFoldingSimulator()
  Variable GTract, N, Spacer, GHP, GTP, GQP, MFP, HOP
   Prompt GTract "G-Tracts"
   Prompt N "Iterations"
   Prompt Spacer "Spacer Length(G-tracts)"
   Prompt GHP "G-Hairpin %"
   Prompt GTP "G-Triplex %"
   Prompt GQP "G-Quadruplx %"
   Prompt MFP "Misfolded %"
   Prompt HOP "Higher-order %"
   doprompt "Enter the number of G-Tracts & Iterations ", GTract, N, Spacer, GHP, GTP, GQP, MFP, HOP

 //Following calculate the absolute percentage formation of each structure from experimental data to a
range later used by program to guide the formation in the simulated template

   Variable TotalPercentage=GHP+GTP+GQP+MFP+HOP
   GHP=GHP/ TotalPercentage*86.4
   GTP=GTP/TotalPercentage*86.4
   GQP=GQP/TotalPercentage*86.4
   MFP=MFP/TotalPercentage*86.4
   HOP=HOP/TotalPercentage*86.4

   GTP=GTP+GHP
   GQP=GQP+GTP
   MFP=MFP+GQP
   HOP=MFP+HOP


  //Following code will calculate the maximum GQ formation in the Sequence entered
    Variable maxform
    maxform=  trunc (GTract/4)

    Make/N= (maxform+1) /D   Distribution
    Make/N=  (maxform+1) /D DistributionLim
    Make/N=(N) /D codekeeper
    Make/N=1 /D x1=0
    Make/N=1 /D TotalFeatures=0
    Make/N=1 /D TotalGQ=0
    Make/N=1 /D TotalGT=0
    Make/N=1 /D TotalGH=0
    Make/N=1 /D TotalMF=0
    Make/N=1 /D TotalHO=0
```

```
  //following code will Generate random numbers which will be used for folding from a random place in the template

 variable j,k,m, GQn, o, p, q, r, s, t, u,usp,uspLL, uspUL, v=0,v1=0, w, y, z=0, counter=0, counter1=0, npoints, nucleationGQ, nucleationGT, nucleationGH, nucleationMF, nucleationHO, featurespertrace
     Appendtotable Distribution, DistributionLim, TotalFeatures, TotalGH, TotalGQ, TotalGT, TotalMF, TotalHO, Codekeeper , x1
     ModifyTable sigDigits(Distribution)=20
     ModifyTable sigDigits(DistributionLim)=20
     ModifyTable sigDigits(Totalfeatures)=20
     ModifyTable sigDigits(TotalGT)=20
     ModifyTable sigDigits(TotalGH)=20
     ModifyTable sigDigits(TotalGQ)=20
     ModifyTable sigDigits(TotalMF)=20
     ModifyTable sigDigits(TotalHO)=20
     ModifyTable sigDigits(Codekeeper)=20
     ModifyTable sigDigits(x1)=20
     delayupdate

     nucleationGH=Gtract-2
     nucleationGT=Gtract-3
     nucleationGQ=Gtract-4
     nucleationMF=Gtract-5
     nucleationHO=Gtract-8


     DistributionLim=0

    for(k=0; k<N; k+=1)


      Make/N=(GTract) Template
      Make/N=(GTract) Template2
      Template=0
      featurespertrace=0

      for(GQn=0; GQn< maxform; GQn=GQn+1)
         j = abs (enoise(100 , 1))
         if (j<=GHP)
            j=1
          elseif(j>GHP && j<=GTP)
            j=2
            elseif(j>GTP && j<=GQP)
              j=3
```

```
            elseif(j>GQP && j<=MFP)
             j=4
              elseif(j>MFP && j<=HOP)
               j=5
                elseif(j>HOP)
                 j=6
        endif

//G-Hairpin Formation
        if(j==1)
          for(u=0, v=0; u<Gtract-1; u+=1)
             if((u-spacer)<0)
              uspLL=0
              else
               uspLL=u-spacer
             endif
             if((u+1+spacer)>=Gtract)
               uspUL=Gtract
              else
                uspUL=u+Spacer+1
              endif

            for(usp=uspLL; usp<=uspUL; usp+=1)
             v=v+Template[usp]
             endfor
             if (v==0)
               break
             endif
           endfor
         endif

//G-Triplex Formation

        if(j==2)
          for(u=0,v=0; u<Gtract-2; u+=1)
             if((u-spacer)<0)
                uspLL=0
              else
                uspLL=u-spacer
             endif
             if((u+2+spacer)>=Gtract)
                uspUL=Gtract
              else
                uspUL=u+Spacer+2
             endif
```

```
            for(usp=uspLL; usp<=uspUL; usp+=1)
               v=v+Template[usp]
             endfor

              if (v==0)
                break
              endif
            endfor

        endif

//G-Quadruplex formation


          if(j==3)
            for(u=0, v=0; u<Gtract-3; u+=1)
               if((u-spacer)<0)
                 uspLL=0
                 else
                   uspLL=u-spacer
                endif
                if((u+3+spacer)>=Gtract)
                  uspUL=Gtract
                  else
                    uspUL=u+Spacer+3
                  endif

              for(usp=uspLL; usp<=uspUL; usp+=1)
               v=v+Template[usp]
               endfor
               if (v==0)
                 break
               endif
             endfor
          endif

 //Misfolded Formation

  if(j==4)
          for(u=0, v=0; u<Gtract-4; u+=1)
             if((u-spacer)<0)
               uspLL=0
               else
                 uspLL=u-spacer
```

```
            endif
            if((u+4+spacer)>=Gtract)
              uspUL=Gtract
             else
               uspUL=u+Spacer+4
             endif

          for(usp=uspLL; usp<=uspUL; usp+=1)
           v=v+Template[usp]
           endfor
           if (v==0)
            break
           endif
         endfor
       endif

//Higher-order Formation

   if(j==5)
        for(u=0, v=0; u<Gtract-7; u+=1)
          if((u-spacer)<0)
           uspLL=0
           else
            uspLL=u-spacer
           endif
           if((u+7+spacer)>=Gtract)
             uspUL=Gtract
            else
              uspUL=u+Spacer+7
             endif

          for(usp=uspLL; usp<=uspUL; usp+=1)
           v=v+Template[usp]
           endfor
           if (v==0)
            break
           endif
         endfor
       endif


     if(j==1)                          //for G-hairpin
       do
           m= abs (round (enoise(nucleationGH , 1)))
```

```
        if((m-spacer)<0)
         uspLL=m
        else
          uspLL=m-spacer
        endif
        if((m+1+Spacer)>=Gtract)
          uspUL=Gtract
         else
           uspUL=m+Spacer+1
         endif


        for(u=uspLL,v1=0; u<=uspUL; u+=1)
          v1=v1+Template[u]
        endfor
    while ( v==0 && v1!=0 )

  elseif(j==2)                    //for G-triplex
      do
      m= abs (round (enoise(nucleationGT , 1)))

        if((m-spacer)<0)
         uspLL=m
        else
          uspLL=m-spacer
        endif
        if((m+2+Spacer)>=Gtract)
          uspUL=Gtract
         else
           uspUL=m+Spacer+2
         endif


        for(u=uspLL,v1=0; u<=uspUL; u+=1)
          v1=v1+Template[u]
        endfor

      while ( v==0 && v1!=0 )



elseif(j==3)                    //for G-Quadruplex
      do
      m= abs (round (enoise(nucleationGQ , 1)))
```

```
        if((m-spacer)<0)
         uspLL=m
        else
          uspLL=m-spacer
        endif
        if((m+3+Spacer)>=Gtract)
          uspUL=Gtract
         else
           uspUL=m+Spacer+3
         endif


        for(u=uspLL,v1=0; u<=uspUL; u+=1)
          v1=v1+Template[u]
        endfor

      while ( v==0 && v1!=0 )


elseif(j==4)                    //for Mis-folded
      do
      m= abs (round (enoise(nucleationMF , 1)))

        if((m-spacer)<0)
         uspLL=m
        else
          uspLL=m-spacer
        endif
        if((m+4+Spacer)>=Gtract)
          uspUL=Gtract
         else
           uspUL=m+Spacer+4
         endif

        for(u=uspLL,v1=0; u<=uspUL; u+=1)
          v1=v1+Template[u]
        endfor

      while ( v==0 && v1!=0 )

   elseif(j==5)                    //for higher order
      do
      m= abs (round (enoise(nucleationHO , 1)))

        if((m-spacer)<0)
```

```
        uspLL=m
       else
         uspLL=m-spacer
       endif
       if((m+7+Spacer)>=Gtract)
         uspUL=Gtract
        else
          uspUL=m+Spacer+7
        endif


       for(u=uspLL,v1=0; u<=uspUL; u+=1)
         v1=v1+Template[u]
       endfor

    while ( v==0 && v1!=0 )


endif


  if(j==1 && v==0)
    Template[m]=1
    Template[m+1]=1

    featurespertrace=featurespertrace+1
    Totalfeatures[0]=Totalfeatures[0]+1
    TotalGH[0]=TotalGH[0]+1

   elseif (j==2 && v==0)
    Template[m]=1
    Template[m+1]=1
    Template[m+2]=1

    featurespertrace=featurespertrace+1
    Totalfeatures[0]=Totalfeatures[0]+1
    TotalGT[0]=TotalGT[0]+1

   elseif (j==3 && v==0)
    Template[m]=1
    Template[m+1]=1
    Template[m+2]=1
    Template[m+3]=1

    featurespertrace=featurespertrace+1
```

```
            Totalfeatures[0]=Totalfeatures[0]+1
            TotalGQ[0]=TotalGQ[0]+1

            elseif (j==4 && v==0)
            Template[m]=1
            Template[m+1]=1
            Template[m+2]=1
            Template[m+3]=0
            Template[m+4]=1

            featurespertrace=featurespertrace+1
            Totalfeatures[0]=Totalfeatures[0]+1
            TotalMF[0]=TotalMF[0]+1

            elseif (j==5 && v==0)
            Template[m]=1
            Template[m+1]=1
            Template[m+2]=1
            Template[m+3]=1
            Template[m+4]=1
            Template[m+5]=1
            Template[m+6]=1
            Template[m+7]=1

            featurespertrace=featurespertrace+1
            Totalfeatures[0]=Totalfeatures[0]+1
            TotalHO[0]=TotalHO[0]+1

          endif
      endfor

        Distribution[featurespertrace]=Distribution[featurespertrace]+1

  Print"---------------------------Summary---------------------------"
  Print "Number of G-Tracts:",  GTract
  Print "Maximum GQ possible:",  maxform
  Print "Over all Distribution of Features:", Distribution
  Print "Non Repeated Pattern of Features:", DistributionLim
  Print "Avarage Spacing (User Defined):", Spacer
  Print "Total Features:", Totalfeatures
  Print "Total GQ:", TotalGQ
  Print "Total GT:", TotalGT

end
```

Supplementary References

1. Punnoose, J. A., Cui, Y., Koirala, D., Yangyuoru, P. M., Ghimire, C., Shrestha, P., and Mao, H. (2014) Interaction of G-Quadruplexes in the Full-Length 3′ Human Telomeric Overhang, *J Am Chem Soc 136*, 18062-18069.
2. Koirala, D., Mashimo, T., Sannohe, Y., Yu, Z., Mao, H., and Sugiyama, H. (2012) Intramolecular folding in three tandem guanine repeats of human telomeric DNA, *Chem. Commun. 48*, 2006-2008.
3. Dai, J., Punchihewa, C., Ambrus, A., Chen, D., Jones, R. A., and Yang, D. (2007) Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation, *Nucleic Acids Res. 35*, 2440-2450.
4. Koirala, D., Ghimire, C., Bohrer, C., Sannohe, Y., Sugiyama, H., and Mao, H. (2013) Long-Loop G-Quadruplexes Are Misfolded Population Minorities with Fast Transition Kinetics in Human Telomeric Sequences, *J. Am. Chem. Soc. 135*, 2235-2241.
5. Yue, D. J. E., Lim, K. W., and Phan, A. T. (2011) Formation of (3+1) G-Quadruplexes with a Long Loop by Human Telomeric DNA Spanning Five or More Repeats, *J. Am. Chem. Soc. 133*, 11462-11465.
6. Schonhoft, J. D., Bajracharya, R., Dhakal, S., Yu, Z., Mao, H., and Basu, S. (2009) Direct experimental evidence for quadruplex-quadruplex interaction within the human ILPR, *Nucleic Acids Res. 37*, 3310-3320.
7. Yu, Z., Koirala, D., Cui, Y., Easterling, L. F., Zhao, Y., and Mao, H. (2012) Click Chemistry Assisted Single-Molecule Fingerprinting Reveals a 3D Biomolecular Folding Funnel, *J. Am. Chem. Soc. 134*, 12338-12341.