SUPPLEMENTARY INFORMATION

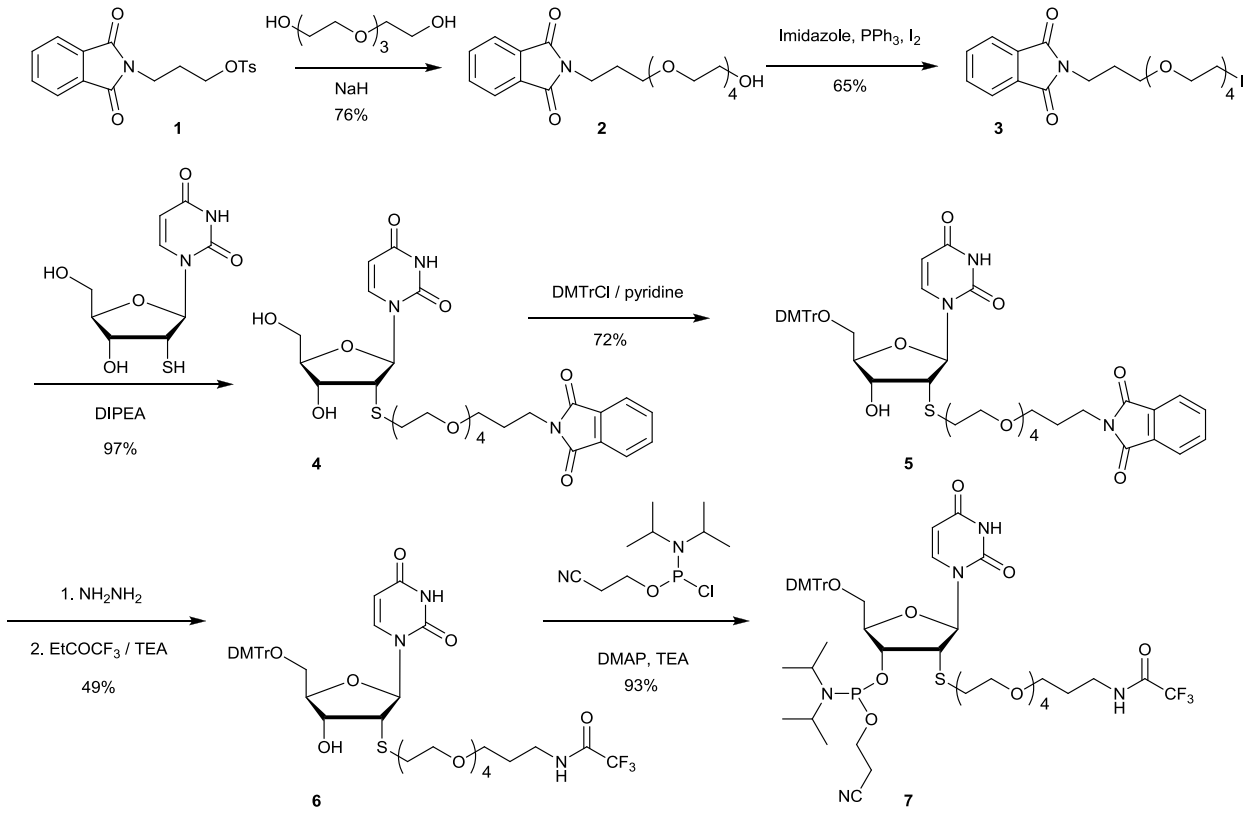
Site-specific inter-strand cross-links of DNA duplexes

Miao Ye, Johan Guillaume, Yu Liu, Ruojie Sha, Risheng Wang, Nadrian C. Seeman*, James W. Canary*

Department of Chemistry, New York University, New York, NY 10003, USA

ned.seeman@nyu.edu; james.canary@nyu.edu

1. Synthesis of uridine phosphoramidite reagent.



Scheme S1. Synthesis of the uridine phosphoramidite reagent (i.e., for U_N^{16}) with a pendent amine group.

N-(3-hydroxy-1-propyl)-phthalimide tosylate (compound 1)

Compound **1** was synthesized previously using different reagents and conditions.[i] In our preparation, 2-(3-hydroxypropyl)isoindoline-1,3-dione, (4.1 g, 20 mmol) was mixed at -10 °C with p-toluene sulfonyl chloride (12.4 g, 65 mmol) in the presence of anhydrous chloroform (20 mL) obtained by passing chloroform through an alumina column and stored with molecular sieves. Dry pyridine (13 mL) was added dropwise and the reaction was stirred vigorously for 2 h. The reaction was then cooled to -20 °C. The solvent was removed and the residue was vacuum-

dried. The reaction was dissolved in cold water (40 mL) and the product was extracted with CH_2Cl_2 (3 x 40 mL). The organic extracts were combined and extracted successively with 2 M HCl (2 x 40 mL), 5% $NaHCO_3$ (40 mL), and water (40 mL). The organic layer was dried over anhydrous $MgSO_4$. The mixture was filtered, vacuum dried, and purified by silica column chromatography using ethyl acetate/hexane as eluant (10:1 to 10:3.5). The yield was 77%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.9 (m, 6H), 7.3 (d, 2H, J = 5 Hz), 4.15 (t, 2H), 3.75 (t, 2H), 2.45 (s, 3H), 2.11 (m, 2H). MS (ESI) calculated for $C_{18}H_{17}NO_5S$: 359.08 found: ($M+H^+$) 360.2. Lit: ^{1}H NMR ($CDCl_3$, ppm) 2.0 (2H, t, J = 6 Hz), 2.4 (3H. s), 3.7 (2H, 1, J = 6 Hz), 4.6 (2H, 1, J = 6 Hz), 7.2 (2H, d, J = 8Hz), 7.7 (6H, m).

Compound 2

All reagents for the reaction yielding compound **2** were used fresh and kept dry in a glove box. Tetraethylene glycol (1.68 g/mL, 3.2 mL, 28 mmol) was dissolved in dry DMF (130 mL), and then was mixed with 60% NaH (0.60 g) for 30 min at 40°C. Compound **1** was dissolved in DMF (60 mL) and was added slowly in portions to the reaction vessel. The reaction was maintained at 0°C during the addition of each portion. The final reaction mixture was left for 40 h, when it was quenched by adding fifteen drops of methanol. The solvent was removed under high vacuum and the residue was partitioned between CH_2Cl_2 and water (200 mL each). The organic layers were combined and washed with saturated $NaHCO_3$ (300 mL), then brine (200 mL). The organic layer was dried over anhydrous Na_2SO_4. The product was purified by silica chromatography with CH_2Cl_2/MeOH (100:0 to 100:1 using an extremely slow gradient). The yield was 76%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.83 (m, 4H), 3.75 (t, 2H, J= 7 Hz), 3.6 (m, 18H), 1.95 (p, 2H, J= 5 Hz). ^{13}C NMR (400 MHz, $CDCl_3$, ppm): 28.6, 35.8, 61.6, 68.8, 70.2, 70.3, 70.4, 70.5, 70.5, 70.6, 72.5, 123.1, 132.2, 133.8, 168.3. MS (ESI) calculated for $C_{19}H_{27}NO_7$: 381.18 found: 382.1 ($M+H^+$), 404.1 ($M+ Na^+$).

Compound 3

The alcohol **2** was co-evaporated with dry acetonitrile three times to remove water. The alcohol was further dried under vacuum overnight prior to reaction. The alcohol **2** (0.115 g, 0.303 mmol) was mixed with 2 equivalents each of molecular iodine (0.113 g, 0.606 mmol), imidazole (0.043 g, 0.61 mmol), and triphenyl phosphine (0.159 g, 0.606). Tetrahydrofuran (2.5 mL) was added and the solution was stirred at room temperature for 1.5 h. A white precipitate was often, but not always, observed upon reaction. The solution was partitioned between water (50 mL) and CH_2Cl_2 (50 mL). The aqueous layer was extracted twice with CH_2Cl_2 (30 mL) and the organic layers were combined, washed with water (30 mL), and dried over anhydrous Na_2SO_4. The solvent was removed under vacuum and the residue was dried and purified by silica chromatography, starting with pure CH_2Cl_2 then very slowly changing to a gradient of CH_2Cl_2/ethyl acetate (1:1). Some side products have very similar retention factor in the column as the desired product, **3** (Rf = 0.76, ethyl acetate). The pure product was oily, but appears solid in the presence of residual triphenyl phosphine oxide side product, which does not have any apparent effect on the next step of the reaction. The yield was estimated 65%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.75 (m, 4H), 3.75 (m, 4H), 3.5 (m, PEG H), 3.15 (t, 2H, J = 4 Hz), 1.92 (m, 2H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm): 2.95, 28.7, 34.2, 35.6, 68.8, 70.2, 70.3, 70.5, 70.6, 70.6, 72.0, 123.2, 132.2, 133.8, 134.1, 168.4. MS (ESI) calculated for $C_{19}H_{26}INO_6$: 491.08 found: ($M+H^+$) 492.1.

Compound 4

This compound **4** was synthesized as soon as halide **3** was prepared. Mercaptouridine (35 mg, 0.13 mmol) was dissolved in CH_3CN, the iodinated product **3** was added (60 mg, 0.12 mmol) at 0 °C. DIEA was added dropwise under argon and the reaction was stirred for 20 h at room temperature. The solvent was removed under vacuum. The solution was dissolved in dichloromethane (45 mL) and was washed with brine (2 x 15 mL) and saturated sodium bicarbonate solution (15 mL). The organic layer was dried with anhydrous Na_2SO_4 and the product was purified using silica column chromatography with MeOH in dichloromethane as eluant (start with pure dichloromethane to 1:30 solution). The yield was 97%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 9 (s, 1H, NH), 7.82 (d, 1H, J = 8.5 Hz), 7.65-7.79 (m, 4H, ArH), 5.78 (d, 1H, J = 9 Hz), 5.69 (d, 1H, J = 9 Hz), 4.3 (d, 2H, J = 6 Hz), 3.9 (dd, 1H, J = 7 Hz), 3.4-3.8 (m, PEG H), 2.8 (m, 2H), 1.91 (m, 2H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm): 28.6, 31.6, 35.5, 54.5, 62.8, 68.8, 70.1, 70.4, 70.4, 70.5, 70.6, 71.9, 72.3, 88.4, 88.9, 102.7, 123.2, 132.2, 134.0, 142.0, 150.7, 163.3, 168.5. MS (ESI) calculated for $C_{28}H_{37}N_3O_{11}S$: 623.21 found: ($M+H^+$) 624.3, ($M+Na^+$) 646.3.

Compound 5

Compound **4** (500 mg, 0.8 mmol) was dissolved in anhydrous pyridine, and then a first portion of dimethoxytrityl chloride (200 mg, 0.6 mmol) was added to the mixture. The solution was stirred at room temp for 1 h when a second portion of dimethoxytrityl chloride was added (200 mg, 0.6 mmol). The reaction was stirred for an additional 1.5 h and was then quenched with methanol (1 mL). The solution was stirred for at least 5 min after stirring before the solvent was removed under vacuum. The solution was kept away from heat, including from the rotary evaporator. The product was dissolved in ethyl acetate (100 mL) and was washed with sodium bicarbonate saturated solution (2 x 33 mL) and brine (33 mL). The organic layer was dried with sodium sulfate. The solvent was removed under vacuum and the product was purified using silica column chromatography with CH_2Cl_2 and ethyl acetate (1:0 slowly to 2:1). 5% TEA was added to the CH_2Cl_2 to stabilize the product. The yield was 72%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.6-7.8 (m, 5H), 7.1-7.4 (m, 9H), 6.85 (d, 4H), 6.0 (d, 1H, J = 9 Hz), 5.2 (d, 5H, J = 9 Hz), 4.44 (dd, 1H), 4.15 (s, 1H), 4.1 (q, 1H), 4.05 (s, 1H), 3.4-3.75 (m, PEG H), 2.75 (m, 2H), 1.95 (s, 1H), 1.85 (p, 2H), 1.4 (t, 4H), 1.25, (t, 4H), 0.75 (m, 2H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm): 8.1, 14.2, 21.8, 28.6, 32.4, 35.5, 53.0, 55.3, 56.4, 60.4, 63.4, 63.5, 78.8, 70.2, 70.3, 70.4, 70.5, 70.6, 71.6, 84.8, 87.0, 87.4, 102.6, 113.3, 123.2, 127.2, 128.0, 128.2, 130.1, 130.2, 132.2, 133.9, 135.0, 135.2, 140.0, 144.3, 150.5, 158.7, 163.0, 168.4. MS (ESI) calculated for $C_{49}H_{55}N_3O_{13}S$: 925.35 found: ($M+Na^+$) 948.6.

Compound 6

Compound **6** was obtained via a two-step process, involving the cleavage of the phthalate group, followed by the addition of a trifluoroacetyl protecting group. Compound **5** (1.7 g, 1.8 mmol) was dissolved in methanol. Hydrazine was added to the solution, which was heated to reflux for 3 h. The solution was cooled and the solvent was removed under vacuum. The residue was redissolved in CH_3CN (2 x 15 mL), co-evaporated to remove remaining hydrazine, and vacuum dried for 1 h. The reaction was dissolved in anhydrous MeOH (12 mL). TEA (0.57mL) and ethyl trifluoroacetate (1.31 mL) were added with stirring at room temperature for 16 h. The product was vacuum dried and was then dissolved in CH_2Cl_2 (200 mL), 5% sodium bicarbonate (3 x 65 mL). The organic layer was dried over sodium sulfate. The product was purified by column chromatography with MeOH/CH_2Cl_2 (0/100 to 2/100). The yield was 49%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.6-7.8 (m, 5h), 7.1-7.4 (m, 9H), 6.80 (d, 4H, J= 6 Hz), 6.0 (d, 1H, 9 Hz), 5.2 (d, 5H, J = 10 Hz), 4.45 (t, 1H,), 4.15 (s, 1H), 4.05 (s, 1H), 3.4-3.75 (m, PEG H), 2.65-2.85 (m, 5H), 1.85 (p, 2H, J = 6 Hz), 1.3 (t, 1H, J = 8 Hz). ^{13}C NMR (400 MHz, $CDCl_3$, ppm): 14.2, 28.0, 32.5, 39.0, 55.3, 56.6, 60.4, 63.4, 70.2, 70.2, 70.3, 70.4, 70.40, 70.44, 70.57, 71.60, 78.8, 70.2, 70.3, 70.4, 70.5, 70.6, 71.64, 84.7, 87.23, 87.28, 102.6, 113.3, 127.2, 128.0, 128.2, 130.1, 130.2, 135.0, 135.2, 140.0, 144.2, 150.4, 158.77, 158.78, 163.0. MS (ESI) calculated for $C_{43}H_{52}F_3N_3O_{12}S$: 891.94 found: ($M+Na^+$) 914.4.

Compound 7

Compound **7** was synthesized from **6**. Dried compound **6** (0.595 g, 0.668 mmol) was added to TEA (0.44 mL), and DMAP (5.36 mg). N,N-diisopropyl-O-cyanoethyl-phosphoramidite chloride (0.352 g, 1.33 mmol) was warmed to room temperature and added dropwise. The solution was stirred at ambient temperature for 40 min. The solvent was removed and the product was vacuum-dried for 10 min. The mixture was dissolved in ethyl acetate (75 mL) and the pH of the solution was kept above 7. The solution was washed with 4.6% sodium bicarbonate solution (2 x 45 mL), then with brine (45 mL). The solvent was removed after drying with anhydrous Na_2SO_4. The product was purified by silica chromatography with the eluant CH_2Cl_2 (0.4% TEA) /MeOH (200:1 to 100:1). The product was co-evaporated with CH_3CN three times and was further dried under vacuum overnight. The yield was 93%. ^{1}H NMR (400 MHz, $CDCl_3$, ppm): 7.7-7.85 (m, 2H), 7.15-7.45 (m, 8H), 6.85 (q, 4H), 6.2 (q, 1H), 5.35 (q, 5H), 4.65 (t, 1H), 4.15 (m, 1H), 4.05 (m, 1H), 3.4-3.75 (m, PEG H), 2.65-2.85 (m, 3H), 1.85 (p, 2H), 1.0-1.4 (m, 18H). ^{31}P NMR (400 MHz, $CDCl_3$, ppm): 150 ppm (d). MS (ESI) calculated for $C_{52}H_{69}F_3N_5O_{12}PS$: 1075.4479 found: ($M^+$) 1075.7. HRMS found: 1075.4490.

2. Time course studies of the N-1 to N-6 and N+1 to N+4 coupling reactions.

Reaction rates of time course studies were estimated based on relative band intensity under ethidium bromide staining. A crude calibrating assay was performed to estimate relative band intensities of all starting materials and products using 100 pmole of each.

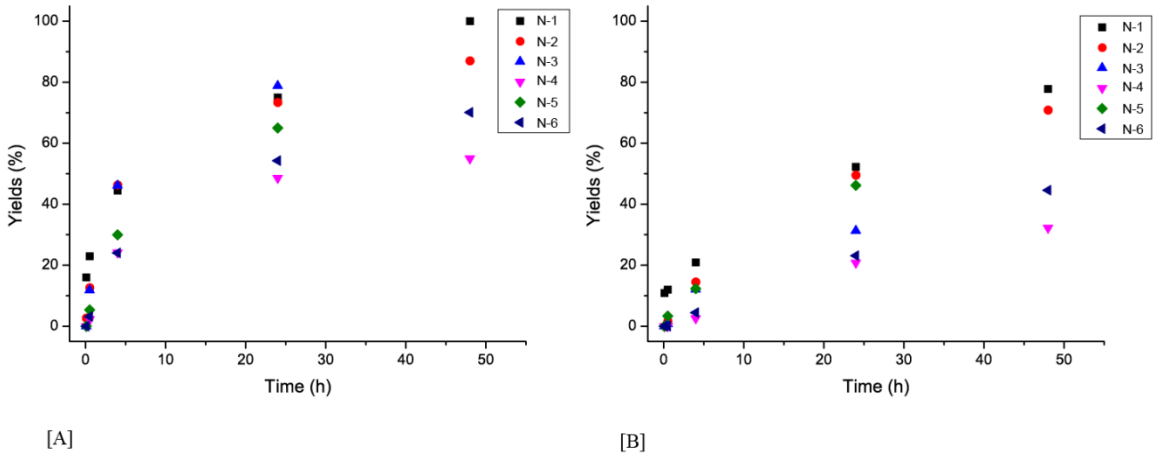


Figure S1 Time course studies of the N-1 to N-6 coupling reactions. (A) Reaction at 20°C. (B) Reaction at 4 °C.

Table S1 Time course study of the N-i series reactions.

(a) N-i series reactions at 20°C.

Reaction time	Products yields (%)					
	N-1	N-2	N-3	N-4	N-5	N-6
5 min	16.0	2.7	0	0	0	0
30 min	23.0	12.6	11.8	2.2	5.3	3.2
4 h	44.5	46.3	46.1	24.2	29.9	24.0
24 h	75.0	73.4	78.9	48.6	65.0	54.3
48 h	100	87	[a]	55.0	[a]	70.1

(b) N-i series reactions at 4 °C.

Reaction time	Products yields (%)					
	N-1	N-2	N-3	N-4	N-5	N-6
5 min	10.9	0	0	0	0	0
30 min	18.8	1.6	0.7	0	3.3	0
4 h	20.9	14.5	12.0	6.0	12.3	4.5
24 h	52.2	49.5	31.3	20.8	46.2	23.1
48 h	77.8	70.9	[a]	32.2	[a]	44.6

[a] Reaction yields not collected

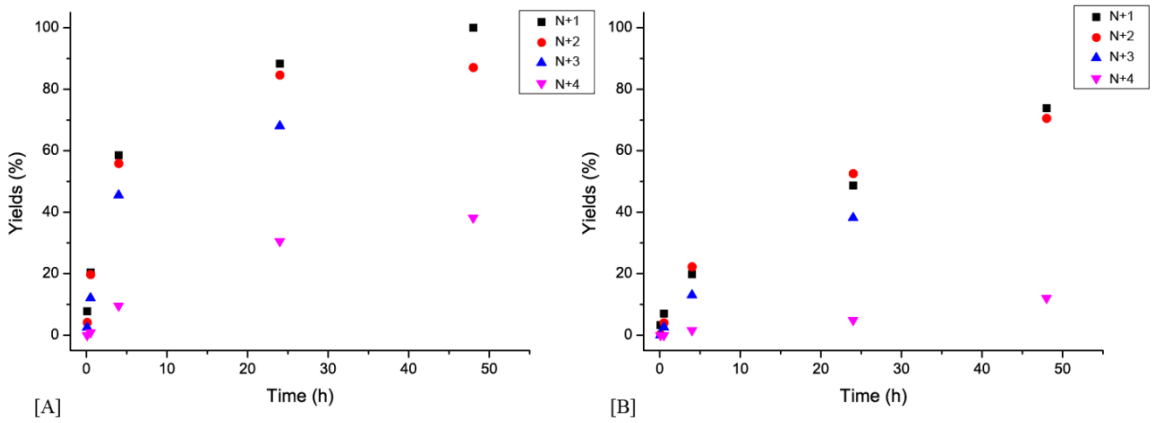


Figure S2 Time course studies of N+1 to N+4 reactions. (A) Reaction at 20°C. (B) Reaction at 4 °C.

Table S2 Time course study of the N+i series reactions.

(a) N+i series reactions at 20°C.

Reaction time	Products yields (%)			
	N+1	N+2	N+3	N+4
5 min	7.8	4.2	2.5	0
30 min	20.4	19.8	12.0	0.9
4h	58.5	55.9	45.5	9.5
24 h	88.3	84.6	68.1	30.6
48 h	100	87.1	[a]	38.2

(b) N+i series reactions at 4°C.

Reaction time	Products yields (%)			
	N+1	N+2	N+3	N+4
5 min	3.3	0.3	0	0
30 min	7.0	4.0	2.6	0
4h	19.8	22.3	13.1	1.6
24 h	48.7	52.6	38.2	4.9
48 h	73.9	70.6	[a]	12.1

[a] Reaction yields not collected

3. MALDI-TOF mass spectra of coupled products of 10-atom linkers.

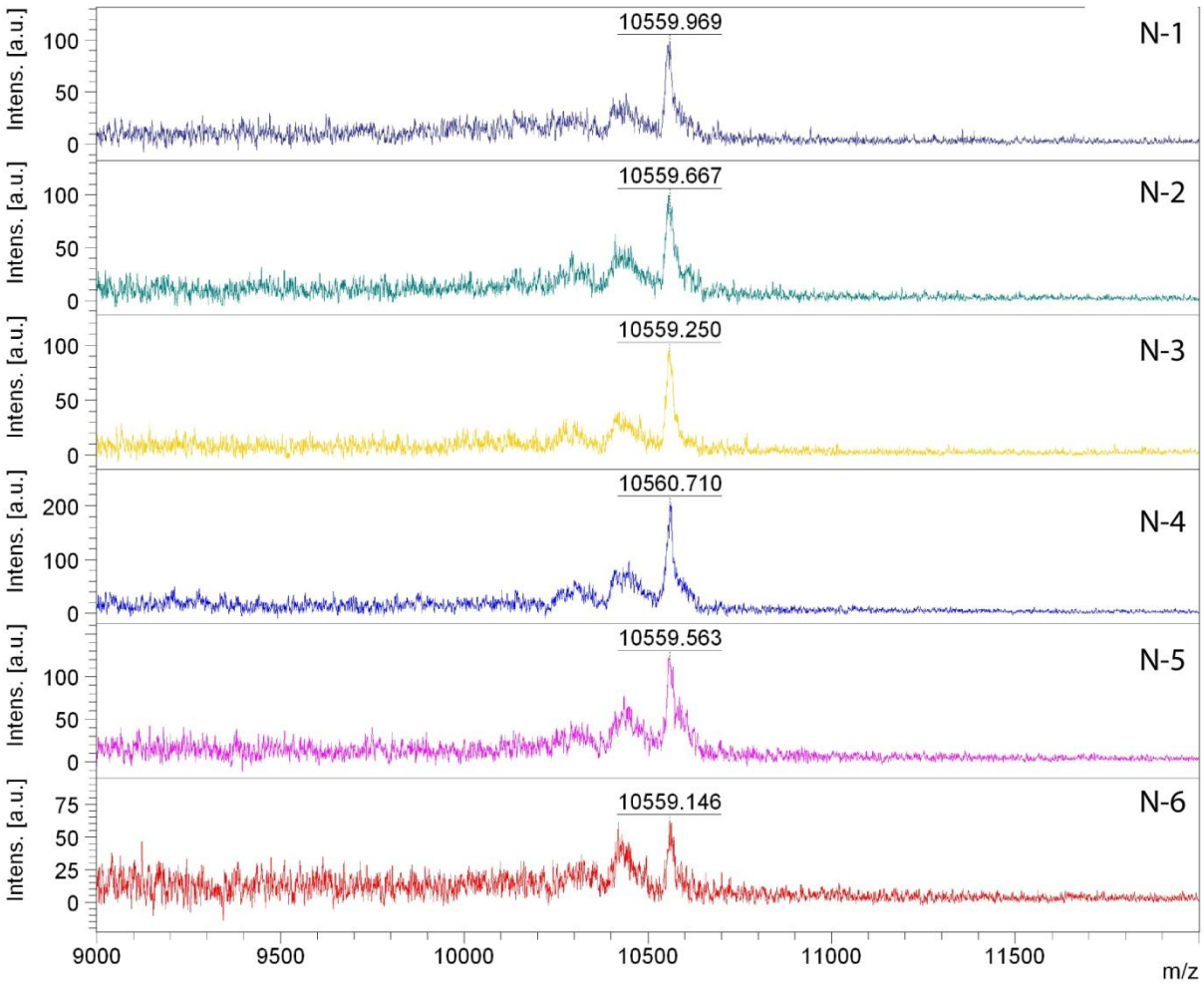


Figure S3. Mass spectra of the N-i series products.

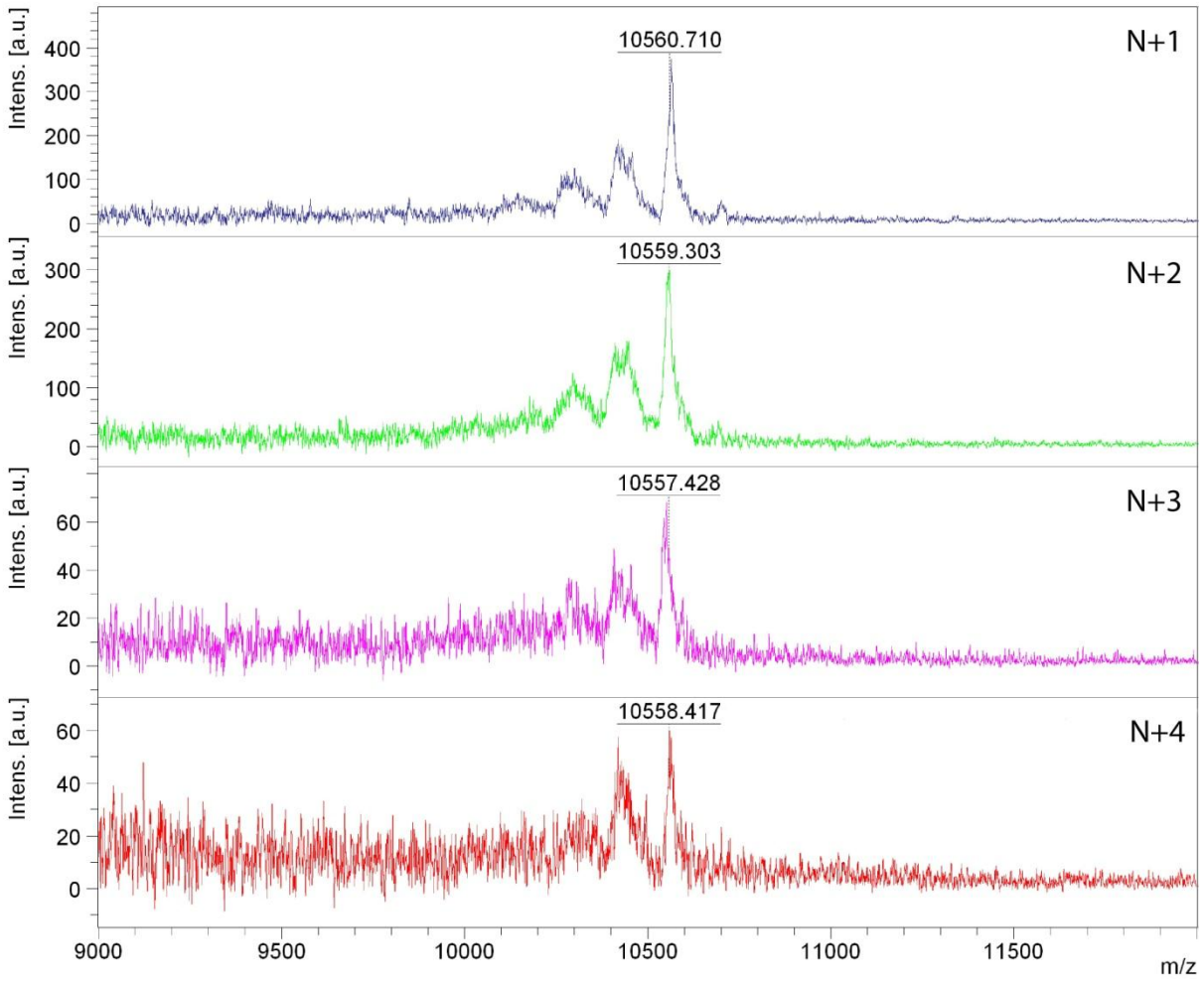


Figure S4. Mass spectra of the N+i series products.

4. MALDI-TOF data for coupled products of 32-atom linkers.

Table S3. MALDI-TOF data for the coupled products of 32-atom linkers.

Coupled products	Mass	
	Calculated	Found
N-6(L)	10879	10884
N-7(L)	10879	10882
N-8(L)	10879	10885

5. Melting curves of the uncoupled N-1 duplex, single modified duplex with U_C or U_N modifed nucleotides).

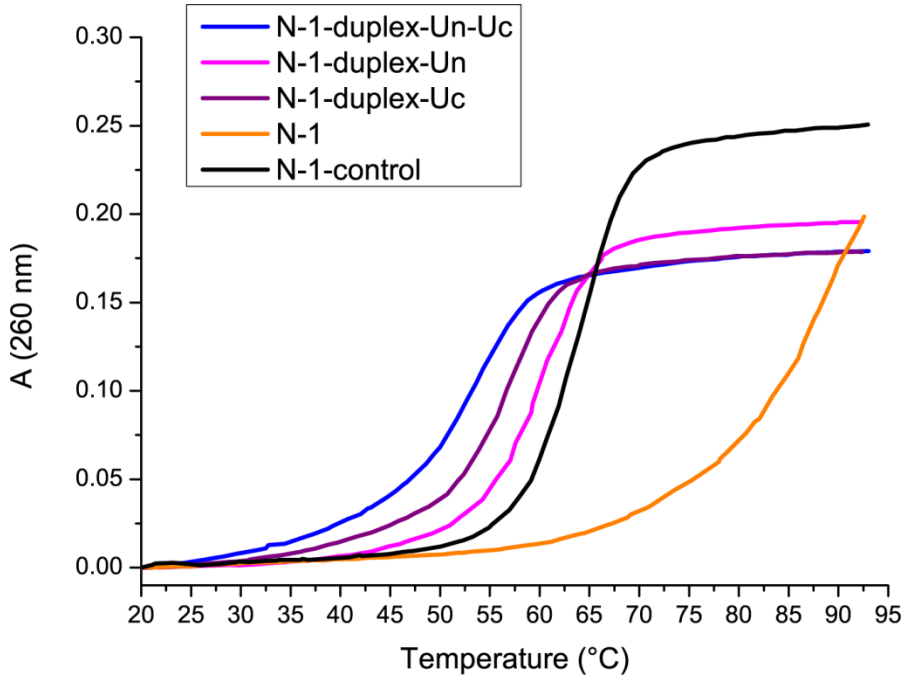


Figure S5. Melting curves of uncoupled the N-1 duplex, duplex with U_C modified nucleotide and U_N modified nucleotides.

Reference

[i]K. J. Martinkus, C. H. Tann, and S. J. Gould, *Tetrahedron* 1983, **39**, 3493–3505