Electronic Supplementary Material (ESI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2023

Experimental Supporting Information For:

A Modular Aldol Approach for Internal Fluorescent Molecular Rotor Chalcone Surrogates for DNA Biosensing Applications

Ryan E. Johnson,[a] Makay T. Murray,[b] Lucas J. Bycraft,[a] Stacey D. Wetmore,[b,*] and Richard A. Manderville[a,*]

[a]Departments of Chemistry and Toxicology, University of Guelph, Guelph, Ontario, Canada, N1G2WI

[b]Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge, Alberta, Canada, T1K 3M4

Table of Contents:

Experimental.

Materials and Methods.

Unmodified oligonucleotides, enamine catalysts, 6-hydroxy-indanone (6HI), (*R*)-(+)-glycidol, DNA synthesis materials, solvents, and all other reagents were purchased from commercial sources and used as received excluding a few select aldehydes, which were synthesized via known procedures (7-diethylaminocoumarin-3-aldehyde (Cou), 3-(5-(Dimethylamino)thiophen-2-yl)acrylaldehyde (*E*th), and 3-dimethylaminonapthalene-1-carbaldehyde(Nap)).[1–3] Buffers were prepared from their respective salts using water from a filtration system (18.2MΩ) with neutralization to the desired pH using 5M aq. HCl/NaOH. Modified oligonucleotides were synthesized on a 1μmol scale with DMT-ON protection using an ABI394 DNA/RNA synthesizer with trityl monitor. Purification of oligos after solid-phase synthesis was performed using Glen-Pak DNA purification cartridges or RP-HPLC on an Agilent 1200 series HPLC system equipped with a diode array detector, fluorescence detector, autosampler, and fraction collector. Ultraviolet−visible (UV− Vis) spectra were obtained on a Cary 300-Bio UV−Vis spectrophotometer equipped with a 6 × 6 multicell block Peltier temperature control unit. Fluorescence measurements were acquired on either a Cary Eclipse Fluorescence spectrophotometer or an Edinburgh spectrofluorometer FS5 at ambient temperature. MN4 quinine titrations were analyzed using a BioTek Synergy H1 plate reader with 96 well plates. NMR spectra were recorded on Bruker Avance 300 or 400MHz spectrometers at room temperature. Low-resolution mass spectra were acquired on a LTQ XL Ion Trap mass spectrometer using an electrospray ionization source ESI. High-resolution mass spectra were recorded on an Agilent LC-UHD Q-Tof instrument with LC bypass.

Solid Phase DNA synthesis.

Solid-phase DNA synthesis of 6HI modified oligos utilized standard and 6HI modified phosphoramidites with standard synthesis reagents. The 6HI-phosphoramidite was inserted into various sequences utilizing standard coupling times for all modified and unmodified phosphoramidites. Post-synthesis, oligos containing the 5′ DMT-ON protection was deprotected in 1mL of 30% ammonium hydroxide for 24 hours at room temperature or in 2 hours at 65°C. Crude oligos were then purified using Glen-Pak DNA purification cartridges. Cartridge eluent purity was verified via RP-HPLC, and the solvent mixture was concentrated to ~1-1.5mM using a ThermoSavant DNA 120 SpeedVac at a medium drying rate with quantification via UV-Vis measurements at 260nm. Prepared

oligo mixtures were then stored at 4°C and used as is in downstream aldol condensation reactions.

On-Strand Aldol Reactions.

Method 1 (Enamine Catalyst)

All aldehydes were prepared to a concentration of 1 M in DMSO (except for Cou (0.2M) and Ju (0.75M), which were dissolved in DMAC). 20μL of the DMSO or DMAC stock aldehyde solution was then added to 35μL of the 6HI modified oligo (~1-1.5mM) which was dissolved in MΩ H2O to give an approximately 30:70 DMSO or DMAC:H2O mixture. Reaction mixture at this point typically contained some quantity of precipitated aldehyde. At this point, 3μL of enamine catalyst was then added to the reaction mixture (piperidine in all cases, except for Cou, which utilized morpholine) followed by subsequent heating at 75°C with mixing via vortex every 1-2 hours. Aldehyde underwent solvation readily upon heating. After 3-5 hours, the reaction mixture was cooled to room temperature and 5% (V/V) 5M NaCl was added along with 3× the volume of absolute ethanol. The mixture was stored in the freezer at −20°C overnight. After centrifugation for 1 hour at 13,400 RPM in ambient temperature, the supernatant was decanted, and the DNA pellet was air-dried. The pellet was redissolved in purified 18.2MΩ water and the yield was examined via HPLC. In cases where yields were ~90%, purification was carried out using a Glen Gel-Pak (Glen Research) desalting column with MΩ water as eluent to remove excess aldehyde. For yields under 90%, HPLC purification was carried out at 70 °C using a 5μm reversed phase semipreparative C18 column (100 × 10mm^2) with a flow rate of 3.3mL/min or a 8μm PLRP-S column (7.5 × 150mm^2) with a flow rate of 2.7mL/min in various gradients of buffer A in buffer B (buffer A = 30:70 aqueous 50mM TEAA, pH 7.2/acetonitrile; buffer B = 95:5 aqueous 50mM TEAA, pH 7.2/acetonitrile). Peaks showing high absorbance at both 254nm (DNA) and 430-600nm (chalcone modification) were collected. Yields were estimated using the relative integrals of product and reactant DNA peaks. Following purification, the collected samples were lyophilized to dryness and desalted using a Glen Gel-Pak (Glen Research) desalting column with elution in purified water. Samples were subsequently quantified using UV−Vis measurements and analyzed by ESI-MS.

Method 2 (Sodium Hydroxide Catalyst)

Aldehydes were prepared to a stock concentration of 1M in DMSO. 35μL of the 6HI modified oligo (~1-1.5mM) was then combined with 20μL of stock aldehyde solution followed by the addition of 5μL aq. NaOH (1M) and subsequent heating at 75°C with mixing via vortex every 0.5 hours. Aldehyde underwent solvation readily upon heating.

After 1-2 hours, the reaction mixture was cooled to room temperature and 5% (V/V) 5M NaCl was added along with 3× the volume of absolute ethanol. Downfield purification from this point mirrored Method 1 (see above). Typical purified yields for this reaction gave 350μL of product oligo in a concentration of 100-140μM.

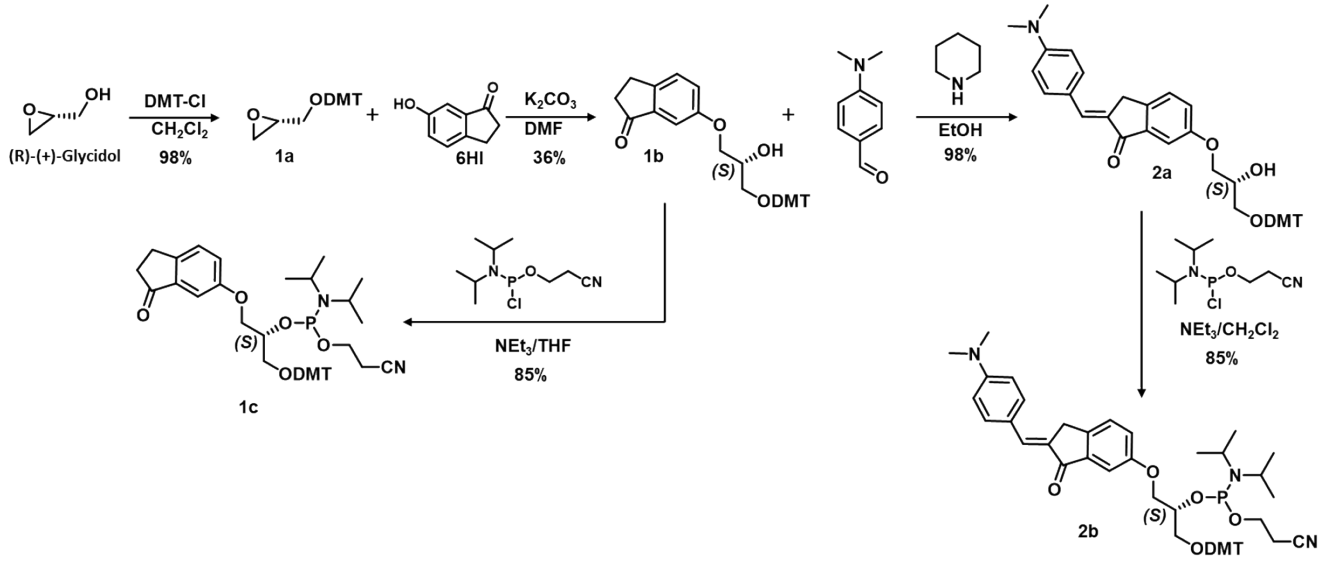
Thermal Melting and Spectroscopic Measurements.

All fluorescent/UV−Vis spectra and thermal melting measurements (T_m) were carried out using a 10 mm light path quartz glass cells with a baseline correction. For thermal melting measurements with oligo substrates, DNA stock solutions were diluted to 5μM in binding buffer (50mM sodium phosphate with 100mM sodium chloride (pH 7)) to which 1 equivalent of complementary strand with either cytosine or a THF modification opposite the dye was added. UV absorbance was monitored at 260nm as a function of temperature with five alternating ramps from 10−90 and 90− 10 °C at a heating/cooling rate of 0.5 °C/min. Average T_m values were calculated using hyperchromicity calculations performed with the Varian Thermal melting software. Relative fluorescent quantum yields (Φ_{fl}) and molar absorption coefficients (ε_{max}) were measured at three different concentrations via titration of the modified oligo into excess complementary strand using either fluorescein (Φ_s=0.95)[4] in 0.1M NaOH or Rhodamine 101 (Φ_s=1.0)[5] in absolute ethanol as fluorescent standards.

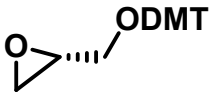
MN4 Quinine Titration Platform.

Measurements were carried out in triplicate on a BioTek Synergy H1 plate reader with 96 well plates. To perform the titration, a stock buffered oligo solution (2.0mL, 18.75mM sodium phosphate, 37.5mM NaCl, pH 7) was prepared, which contained both the 11mer, labelled with An6HI at position G9 (6.25μM) and the native 22mer (7.5μM). 160μL of the stock solution was dispensed into 12 wells after which varying volumes of quinine/ MΩ water were added to give a final volume of 200μL. Wells were then shaken and allowed to equilibrate for 10 minutes followed by subsequent fluorescent measurements.

Phosphoramidite Synthesis.

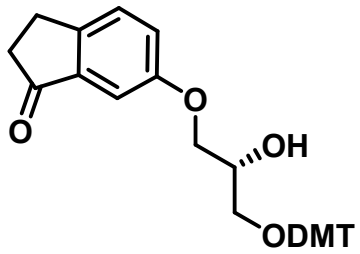


Scheme S1. Synthesis of (*S*)-6-hydroxy-indanone phosphoramidite (**1c**) and (*S*)-An6HI phosphoramidite (**2b**).

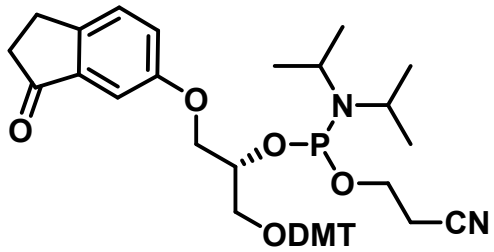


Glycidol-DMT (1a): Compound **1a** was synthesized according to literature procedure.[6] (*R*)-glycidol (1.00mL, 15.1mmol, 1.00 eq.) was added to an oven-dried round bottom flask followed by stirring under argon in 5mL dry DCM and the addition of freshly distilled triethylamine (8.00mL, 57.4mmol, 3.82 eq.). DMT-Chloride solution (20mL, 0.88M, 17.67 mmol, 1.17 eq.) in DCM was added dropwise at 0°C followed by an additional 10mL DCM to promote stirring of the solution. Reaction mixture was stirred at room temperature for 16 hours under argon followed by washing with half saturated aq. $NaHCO_3$. Organic extraction was carried out using DCM (1 x 60mL) followed by drying over $MgSO_4$ and removal of solvent under reduced pressure to give a dark red residue. Residue was then purified using silica gel chromatography (99:1, dichloromethane: triethylamine) to provide **1a** as a colourless viscous oil (5.5838g, 98% yield). ^1H NMR ($CDCl_3$, 400 MHz): δ 7.51-

7.47 (m, 2H), 7.41-7.35 (m, 4H), 7.35-7.28 (m, 2H), 7.27-7.20 (m, 1H), 6.89-6.84 (m, 4H), 3.82 (s, 6H), 3.50 (dd, 1H, *J* = 9.9 Hz, 2.3 Hz), 3.20-3.10 (m, 2H), 2.82-2.78 (m, 1H), 2.68-2.63 (m, 1H). $^{13}C\{^{1}H\}$ NMR ($CDCl_3$, 100 Hz) δ 158.5, 144.8, 136.0, 130.0, 128.2, 127.9, 126.8, 113.1, 86.1, 64.6, 55.2, 51.1, 44.7. MS (ESI) m/z: $[M+Na]^+$ calculated for $C_{24}H_{24}O_4$ = 399.16, found 398.94.

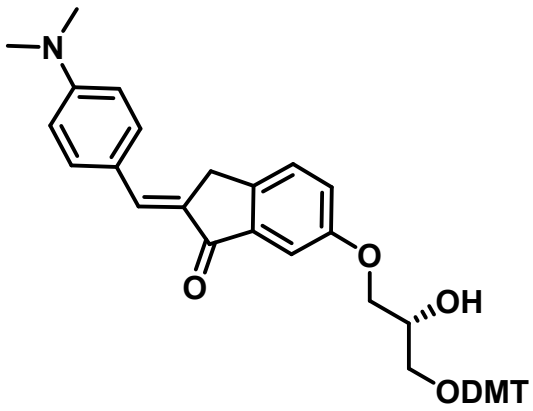


DMT protected (*S*)-6-hydroxy-indanone glycerol nucleoside (1b): 6-hydroxy-indanone (1.2372g, 8.35mmol, 1.20 eq.), **1a** (2.6089g, 6.95mmol, 1.00 eq.), KI (140mg, 0.84mmol, 0.12 eq.) and K_2CO_3 (2.8234g, 20.42mmol, 2.44 eq.) were combined in 20mL DMF and heated at 110°C overnight. Reaction mixture was then cooled to room temperature and diluted with 75 mL EtOAc followed by washing of the organic layer with deionized water (3 X 125mL) and 5% LiCl (1 x 100mL). Organic layer was then filtered through celite and dried over Na_2SO_4. Solvent was then removed under reduced pressure to give an oily residue which was purified via silica gel chromatography (8:91:1 ethyl acetate: DCM: triethylamine) to give **1b** as a foamy white solid (1.2870g, 36%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.48-7.44 (m, 2H), 7.40 (d, 1H, *J* = 8.5 Hz), 7.37-7.34 (m, 4H), 7.33-7.29 (m, 2H), 7.27-7.24 (m, 1H), 7.23-7.19 (m, 2H), 6.88-6.84 (m, 4H), 4.18 (q, 1H, *J* = 4.6 Hz), 4.14-4.07 (m, 2H), 3.83 (s, 6H), 3.41-3.35 (m, 2H), 3.11 (t, 2H, *J* = 5.6 Hz), 2.77-2.74 (m, 2H). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 Hz) δ 206.9, 158.6, 158.4, 148.3, 144.7, 138.3, 135.8, 130.0, 128.1, 127.9, 127.4, 126.9, 124.2, 113.2, 106.0, 86.3, 69.6, 69.5, 63.9, 55.2, 37.0, 25.2. MS (QTOF) m/z: $[M+Na]^+$ calculated for $C_{33}H_{32}O_6$ = 547.2091, found 547.2064.

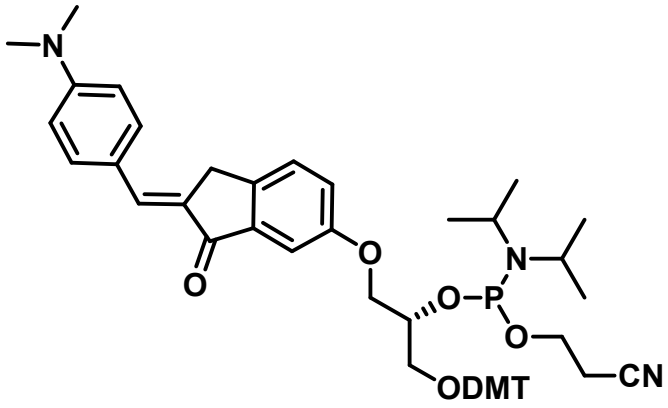


(*S*)-6-hydroxy-indanone phosphoramidite (1c): 1b (0.2844g, 0.54mmol, 1.00 eq.) was placed in an oven dried round bottom flask and co-evaporated with dry toluene (3 x 8mL) followed by solvation in dry THF (8mL) and the addition of freshly distilled TEA (0.32mL, 2.29mmol, 4.25 eq.). Reaction mixture was then stirred under argon for 10 minutes followed by the dropwise addition of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.20mL, 0.89mmol, 1.65 eq.) and stirring for a further 2 hours. Solvent was then removed under reduced pressure and organic layer was extracted via the addition of ethyl acetate

containing 3% triethylamine. Organic layer was then washed with half sat. $NaHCO_3$ (2 x 50 mL) and brine (1 x 30 mL). Solvent was removed under reduced pressure and the residue was purified via silica gel chromatography (39:59:2 ethyl acetate: hexanes: triethylamine) to afford product **1c** (330mg, 85%) which was characterized via ^{31}P NMR and QTOF-HRMS, followed by storage under argon at -20°C. ^{31}P NMR (162 MHz, CDCl3): δ 149.90, 149.72. MS (QTOF) m/z: $[M+H]^+$ calculated for $C_{42}H_{49}N_2O_7P$= 725.3350, found 725.3355.



DMT protected (*S*)-An6HI glycerol nucleoside (2a): 1b (213mg, 0.40mmol, 1.00 eq.) was combined with 4-(Dimethylamino)benzaldehyde (208mg, 1.39mmol, 3.48 eq.) in 5 mL EtOH followed by the addition of 3 drops piperidine and reflux overnight. Solvent was then removed under reduced pressure and the residue was purified using silica gel chromatography (24:75:1 acetone: hexanes: triethylamine) to afford **2a** as a bright orange crystalline solid (256mg, 98%). ^1H NMR ($CDCl_3$, 400 MHz): 7.67-7.58 (m, 3H), 7.48-7.42 (m, 3H), 7.37-7.29 (m, 6H), 7.26-7.15 (m, 2H), 6.87-6.82 (m, 4H), 6.79-6.74 (m, 2H), 4.16-4.08 (m, 2H), 3.92 (d, 2H, *J* = 1.3 Hz), 3.80 (s, 6H), 3.40-3.34 (m, 2H), 3.07 (s, 6H), 2.52-2.47 (m,1H). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 Hz) δ 194.1, 158.5, 158.4, 151.2, 144.7, 142.4, 140.0, 135.9, 135.8, 134.9, 132.8, 130.7, 130.1, 128.1, 127.9, 126.9, 126.7, 123.2, 123.1, 113.2, 111.9, 106.7, 86.3, 69.6, 69.5, 63.9, 55.2, 46.3, 40.1, 32.0. MS (QTOF) m/z: $[M+H]^+$ calculated for $C_{42}H_{41}NO_6$ = 656.3007, found 656.2998.

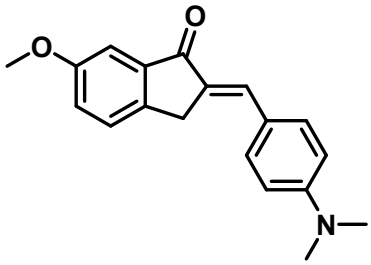


(*S*)-An6HI phosphoramidite (2b): 2a (254mg, 0.39mmol, 1.00 eq.) was placed in a flame dried round bottom flask and dissolved in 6mL of dry DCM, followed by the addition of dry TEA (0.4mL, 2.86mmol, 7.35 eq.) and stirring under argon for a further 10 minutes. 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.11mL, 0.49mmol, 1.26 eq.) was then added and the reaction mixture was stirred under argon for 2 hours with monitoring via TLC. Upon reaction completion, the mixture was quenched with 15 mL of half sat. $NaHCO_3$ and the organic layer was washed with another 15mL of half sat. $NaHCO_3$ followed by drying over Na_2SO_4 and removal of solvent under reduced pressure. Resulting residue was then purified using silica gel chromatography (49:50:1 ethyl acetate: hexanes: triethylamine) to afford **2b** as a foamy bright orange solid (278mg, 85%) which was characterized via ^{31}P NMR and QTOF-HRMS, followed by storage under argon at -20°C. ^{31}P NMR (162 MHz, CDCl3): δ 149.87, 149.73. MS (QTOF) m/z: $[M+H]^+$ calculated for $C_{51}H_{58}N_3O_7P$ = 856.4085, found 856.4061.

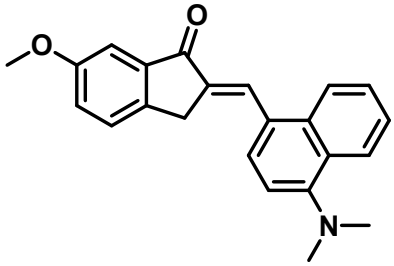
Chalcone Free Dye Synthesis.

All chalcone free dyes were synthesized using 6-methoxyindanone in the following procedure:

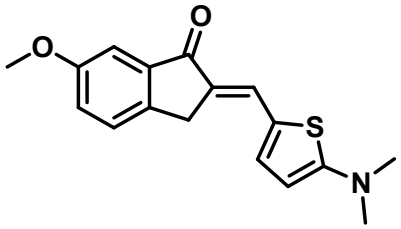
To a stirring solution of 6-methoxyindanone (0.4mmol) and aldehyde (0.4mmol) in anhydrous ethanol (3mL) was added several drops of 10 N NaOH until a colour change was observed. After stirring for ~2 hours, the formed precipitate was filtered and washed with ice cold EtOH (3 x 10mL) yielding pure chalcone free dyes. In all cases, only the *E*-isomer was obtained.



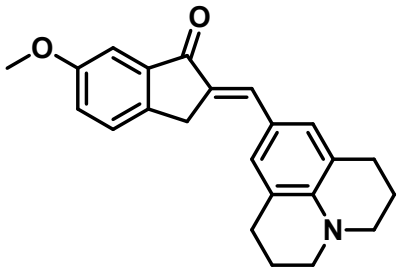
An6MI: Yellow powder (150mg, 93%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.65 (t, 1H, *J* = 1.5 Hz), 7.62 (d, 2H, *J* = 8.9 Hz), 7.46 (d, 1H, *J* = 8.2 Hz), 7.38 (d, 1H, *J* = 2.5 Hz), 7.20 (dd, 1H, *J* = 2.6, 8.40 Hz), 6.77 (d, 2H, *J* = 8.9 Hz), 3.95 (d, 2H, *J* = 1.5 Hz), 3.89 (s, 3H), 3.08 (s, 6H). ^{13}C{1H} NMR ($CDCl_3$, 100 Hz) δ 194.3, 159.5, 151.2, 142.1, 140.0, 134.9, 132.8, 130.8, 126.7, 123.3, 123.0, 112.0, 105.6, 55.7, 40.1, 32.1. MS (ESI) m/z: $[M+H]^+$ calculated for $C_{19}H_{19}NO_2$ = 294.15, found 294.25.



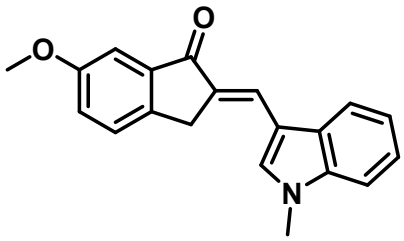
Nap6MI: Yellow powder (55mg, 67%). ^{1}H NMR ($CDCl_3$, 400 MHz): 8.46 (t, 1H, *J* = 1.8 Hz), 8.33-8.27 (m, 2H), 7.81 (d, 1H, *J* = 8.0 Hz), 7.63-7.53 (m, 2H), 7.46-7.41 (m, 2H), 7.22 (dd, 1H, *J* = 2.8, 8.4 Hz), 7.13 (d, 1H, *J* = 8.0 Hz), 3.97 (d, 2H, *J* = 1.8 Hz), 3.92 (s, 3H), 2.99 (s, 6H). ^{13}C{1H} NMR ($CDCl_3$, 100 Hz) δ 194.0, 159.6, 152.8, 142.8, 139.7, 135.9, 133.9, 130.8, 128.5, 127.9, 126.9, 126.7, 126.2, 125.3, 125.0, 124.4, 123.6, 113.1, 105.8, 55.7, 45.0, 31.8. MS (ESI) m/z: $[M+H]^+$ calculated for $C_{23}H_{21}NO_2$ = 344.16, found 344.25.



Th6MI: Red powder (65mg, 64%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.77 (dt, 1H, *J* = 1.4, 0.4 Hz), 7.43 (dd, 1H, *J* = 8.3, 0.4 Hz), 7.35 (d, 1H, *J* = 2.6 Hz), 7.23 (d, 1H, *J* = 4.2 Hz), 7.17 (dd, 1H, *J* = 8.3, 2.6 Hz), 5.95 (d, 1H, *J* = 4.2 Hz), 3.88 (s, 3H), 3.80 (d, 1H, *J* = 1.4 Hz), 3.08 (s, 6H). ^{13}C{1H} NMR ($CDCl_3$, 100 Hz) δ 193.1, 165.0, 159.4, 141.4, 140.9, 136.7, 128.7, 126.7, 126.6, 124.0, 122.4, 105.5, 103.3, 55.6, 42.3, 31.5. MS (ESI) m/z: $[M+H]^+$ calculated for $C_{17}H_{17}NO_2S$ = 300.10, found 300.17.



Ju6MI: Orange powder (55mg, 44%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.56 (t, 1H, *J* = 1.8 Hz), 7.45 (d, 1H, *J* = 8.4 Hz), 7.37 (d, 1H, *J* = 2.5 Hz), 7.19-7.16 (m, 3H), 3.92 (d, 2H, *J* = 1.4 Hz), 3.88 (s, 3H), 3.28 (t, 4H, *J* = 6.0 Hz), 2.81 (t, 4H, *J* = 6.0 Hz), 2.00 (p, 4H, *J* = 6.0 Hz), 3.83 (s, 6H), 3.41-3.35 (m, 2H), 3.11 (t, 2H, *J* = 5.6 Hz), 2.77-2.74 (m, 2H). ^{13}C{1H} NMR ($CDCl_3$, 100 Hz) δ 194.1, 159.4, 144.5, 142.1, 140.2, 135.4, 130.6, 129.8, 126.6, 122.7, 122.3, 121.1, 105.6, 55.6, 50.0, 32.2, 27.8, 21.6. MS (ESI) m/z: $[M+H]^+$ calculated for $C_{23}H_{23}NO_2$ = 346.18, found 346.33.



Ind6MI: Yellow powder (73mg, 67%). ^{1}H NMR ($CDCl_3$, 400 MHz): 8.09 (t, 1H, *J* = 1.6 Hz), 7.97 (d, 1H, *J* = 7.4 Hz), 7.50 (s, 1H), 7.44 (d, 1H, *J* = 8.0 Hz), 7.40-7.28 (m, 4H), 7.18 (dd, 1H, *J* = 8.3, 2.6 Hz), 3.90 (s, 3H), 3.90 (s, 3H), 3.79 (d, 2H, *J* = 1.6 Hz). ^{13}C{1H} NMR ($CDCl_3$, 100 Hz) δ 193.7, 159.5, 141.3, 140.5, 136.9, 130.8, 130.7, 128.6, 126.7, 125.3, 123.2, 122.9, 121.2, 119.2, 112.4, 109.7, 105.7, 55.7, 33.5, 32.6. MS (ESI) m/z: $[M+H]^+$ calculated for $C_{20}H_{17}NO_2$ = 304.13, found 304.25.

Table S1. Photophysical Properties of Free Chalcone Probes in Various Solvents.

Donor	λ_{Ex}[a]	λ_{Em}[a]	ε_{Max}	λ_{Ex}[b]	λ_{Em}[b]	I_{rel}[c]	λ_{Ex}[d]	λ_{Em}[d]	φ_{fl}[e]	I_{rel}[f]	*B* ($cm^{-1}M^{-1}$)
Th6Ml	460	556	50,000	505	569	10.3	450	516	0.04	47	1,800
Ju6Ml	465	579	30,300	488	607	17	438	524	0.16	166	4,787
An6Ml	430	547	33,000	449	574	10.3	413	499	0.13	57	4,422
Ind6Ml	415	510	33,000	450	520	5.6	400	460	0.03	4	825
Nap6Ml	415	579	20,000	432	611	10.9	400	529	0.16	31	3,140

[a] Photophysical data recorded in 100% MeOH. [b] Photophysical data recorded in 75% Glycerol/ 25% MeOH. [c] Relative emission intensity (I_{rel}) of the free chalcone probes in 75% Glycerol/ 25% MeOH versus 100% MeOH. [d] Photophysical data recorded in 1,4-dioxane. [e] Φ_{fl} of free chalcone probes in 1,4-dioxane. [f] Relative emission intensity (I_{rel}) of the free chalcone probes in 1,4-dioxane versus 100% MeOH.

NMR Spectra for Phosphoramidite Synthesis.

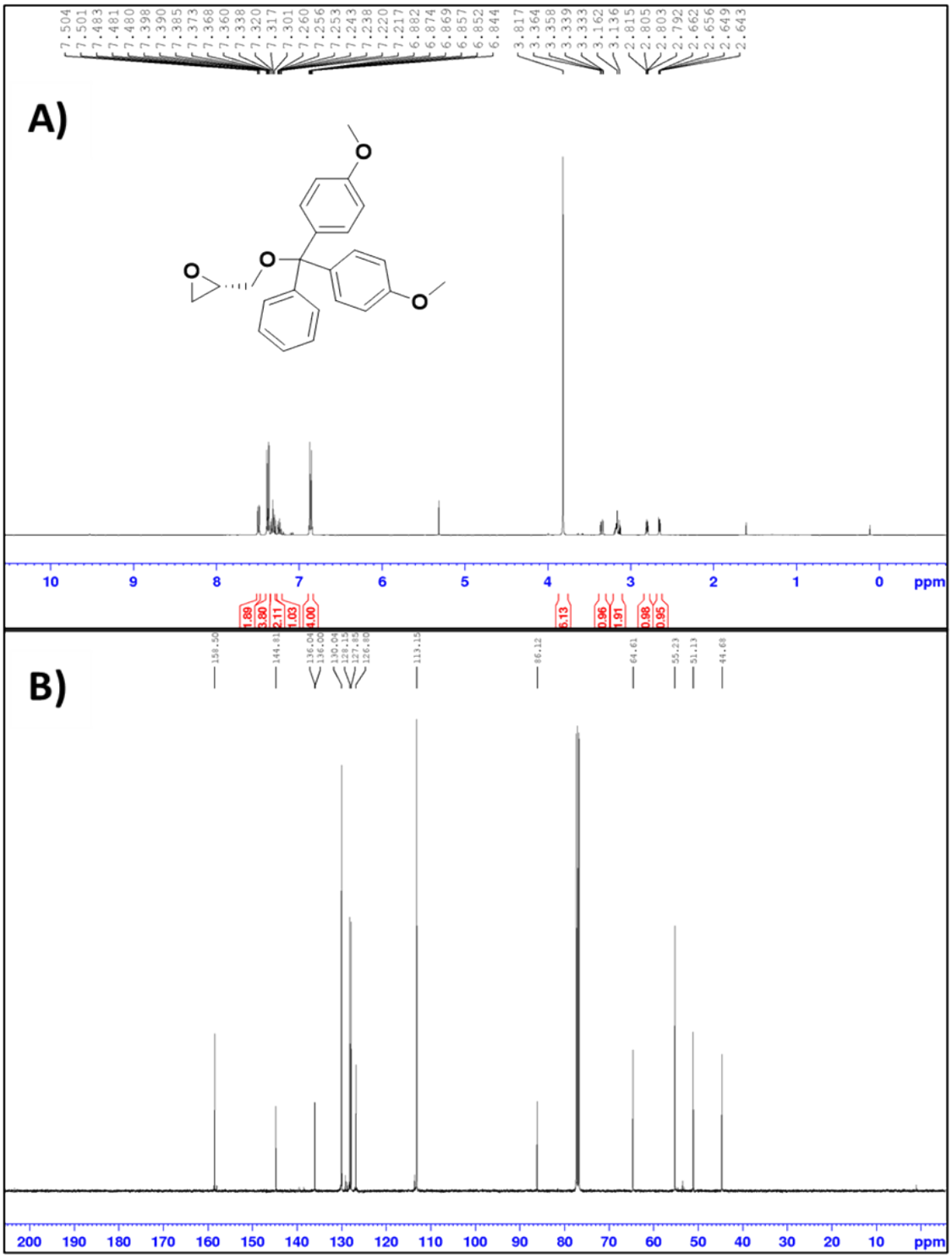


Figure S1. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

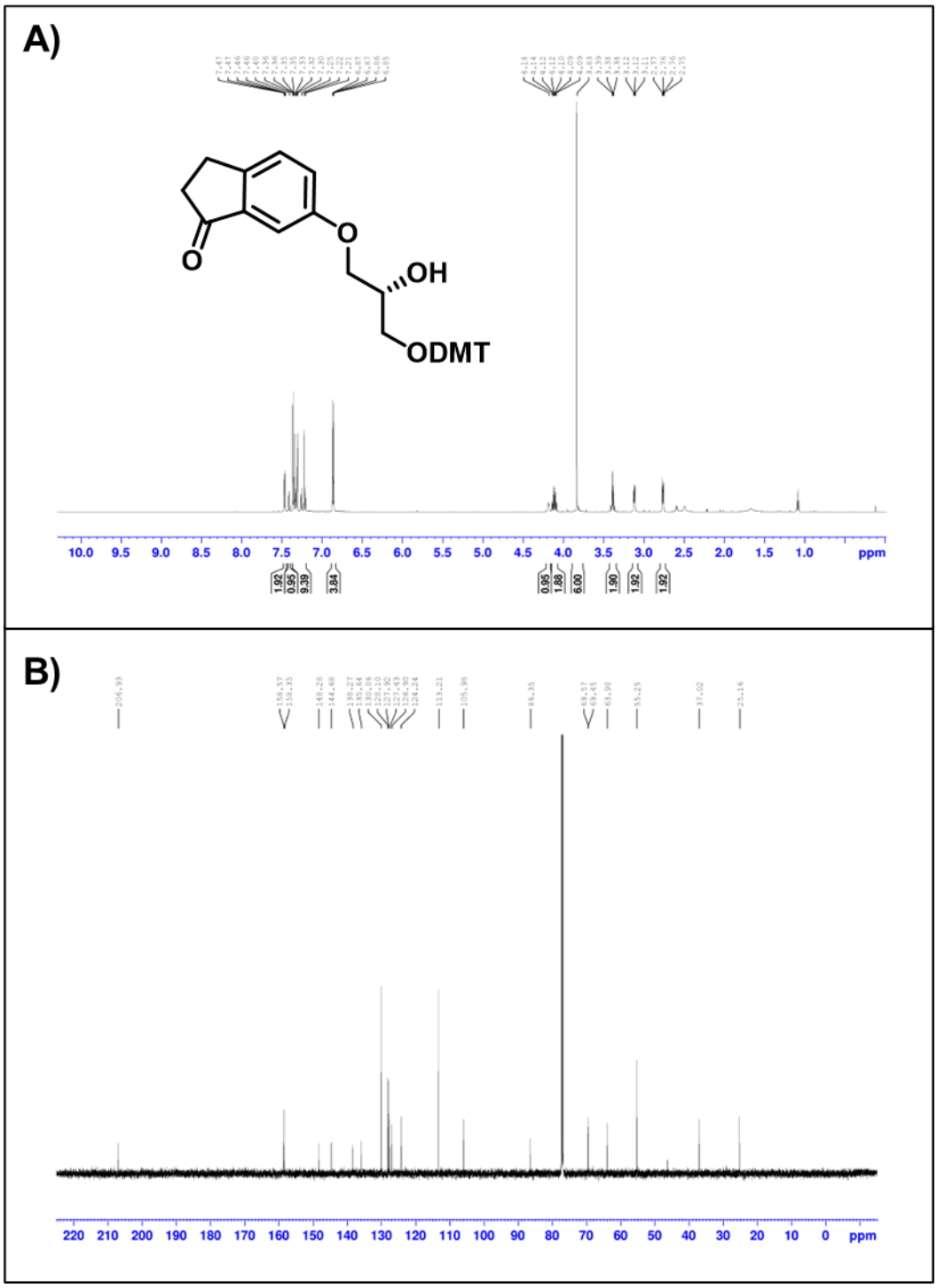


Figure S2. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

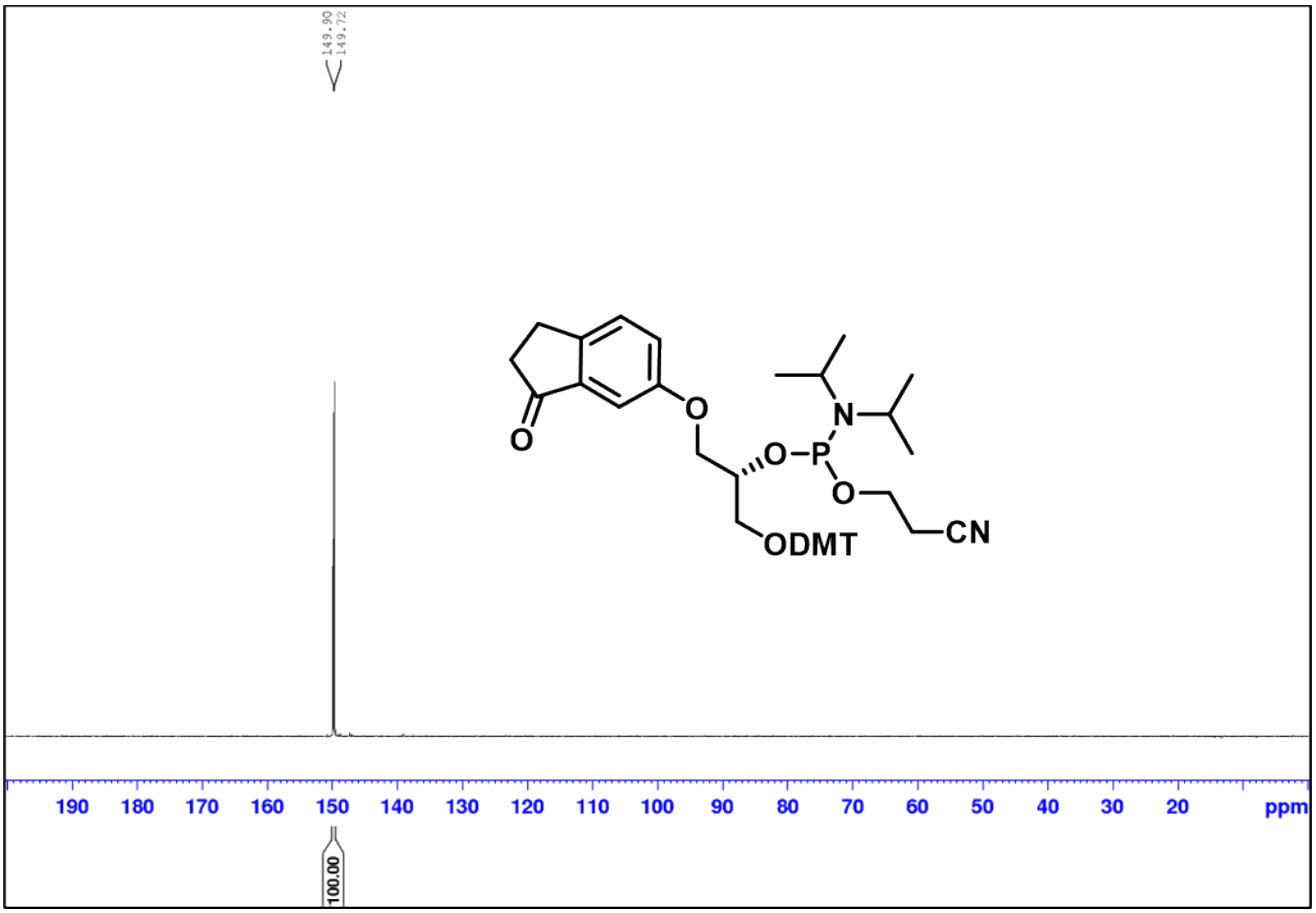


Figure S3. $^{31}P\{1H\}$ NMR spectrum in $CDCl_3$.

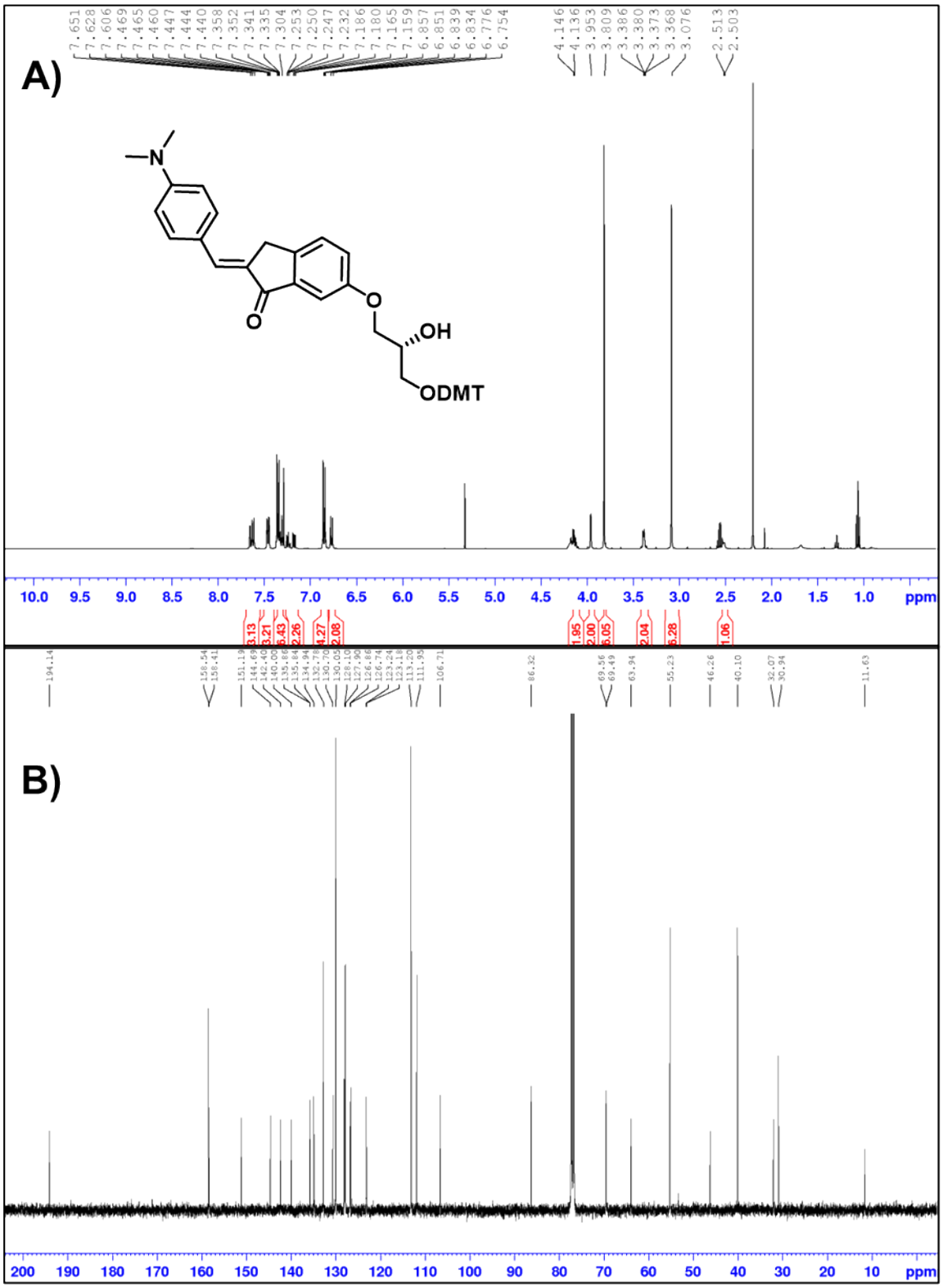


Figure S4. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

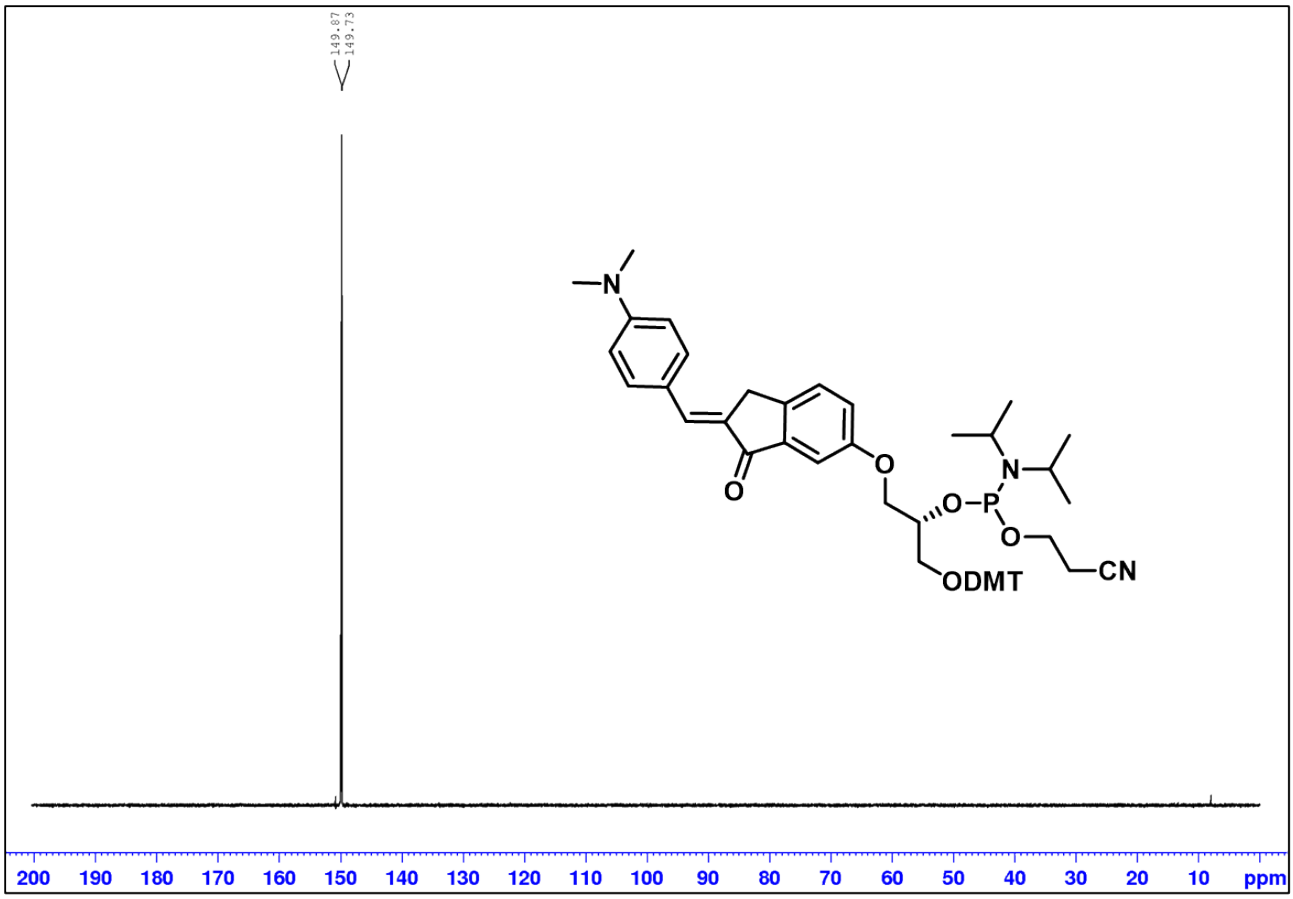


Figure S5. $^{31}P\{1H\}$ NMR spectrum in $CDCl_3$.

HRMS for Phosphoramidite Synthesis.

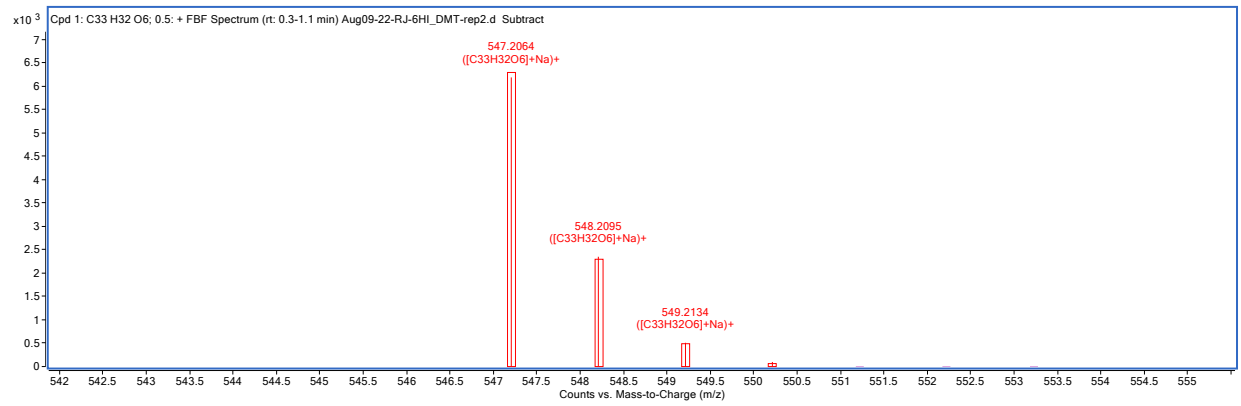


Figure S6. High-resolution QTOF-MS in positive ion mode for compound **1b.**

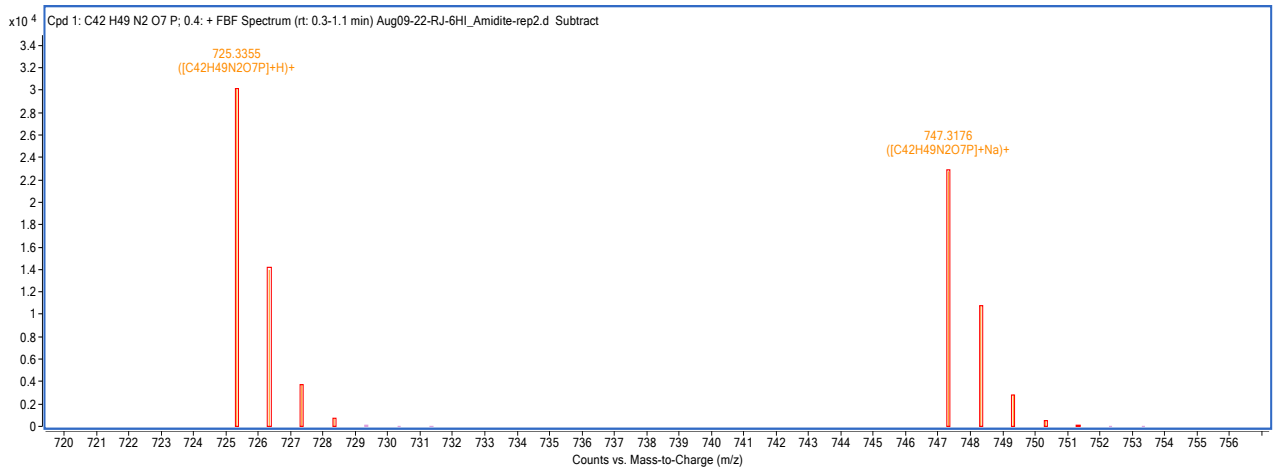


Figure S7. High-resolution QTOF-MS in positive ion mode for compound **1c.**

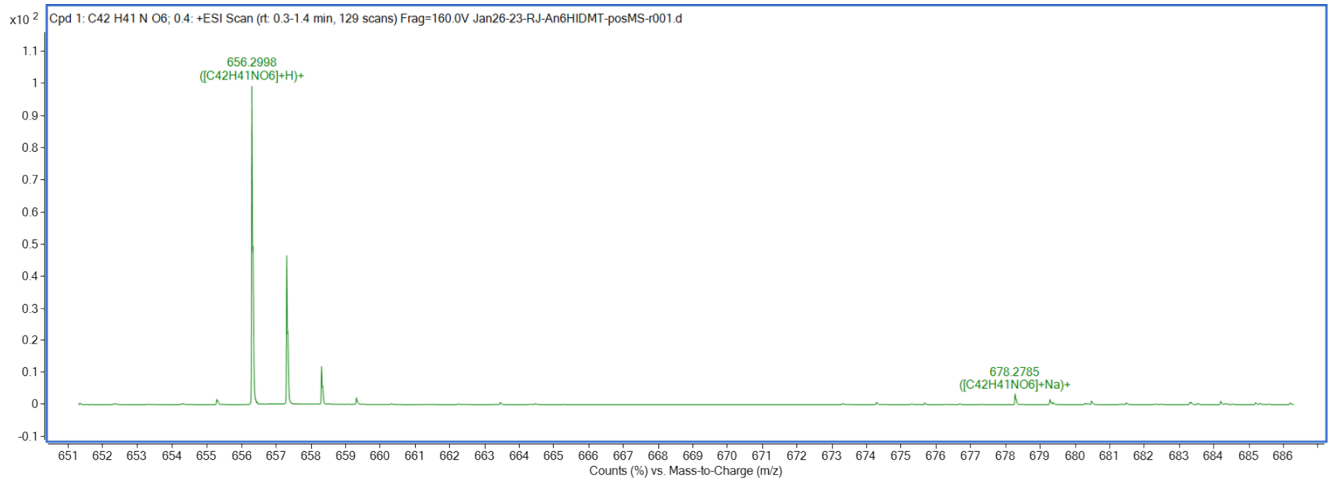


Figure S8. High-resolution QTOF-MS in positive ion mode for compound **2a.**

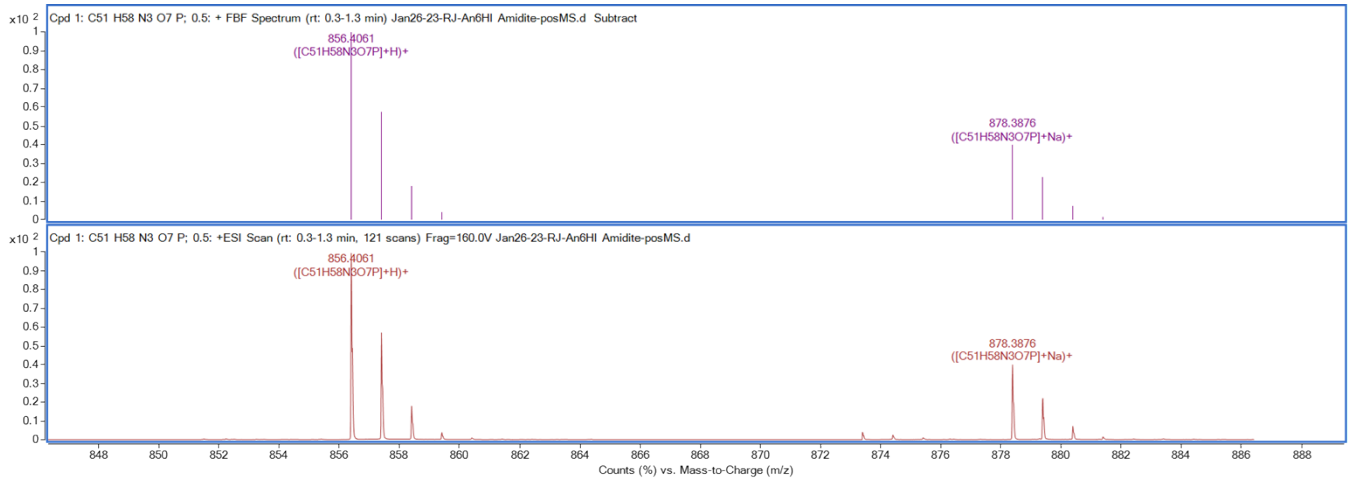


Figure S9. High-resolution QTOF-MS in positive ion mode for compound **2b.**

NMR Spectra for Free Chalcone Synthesis.

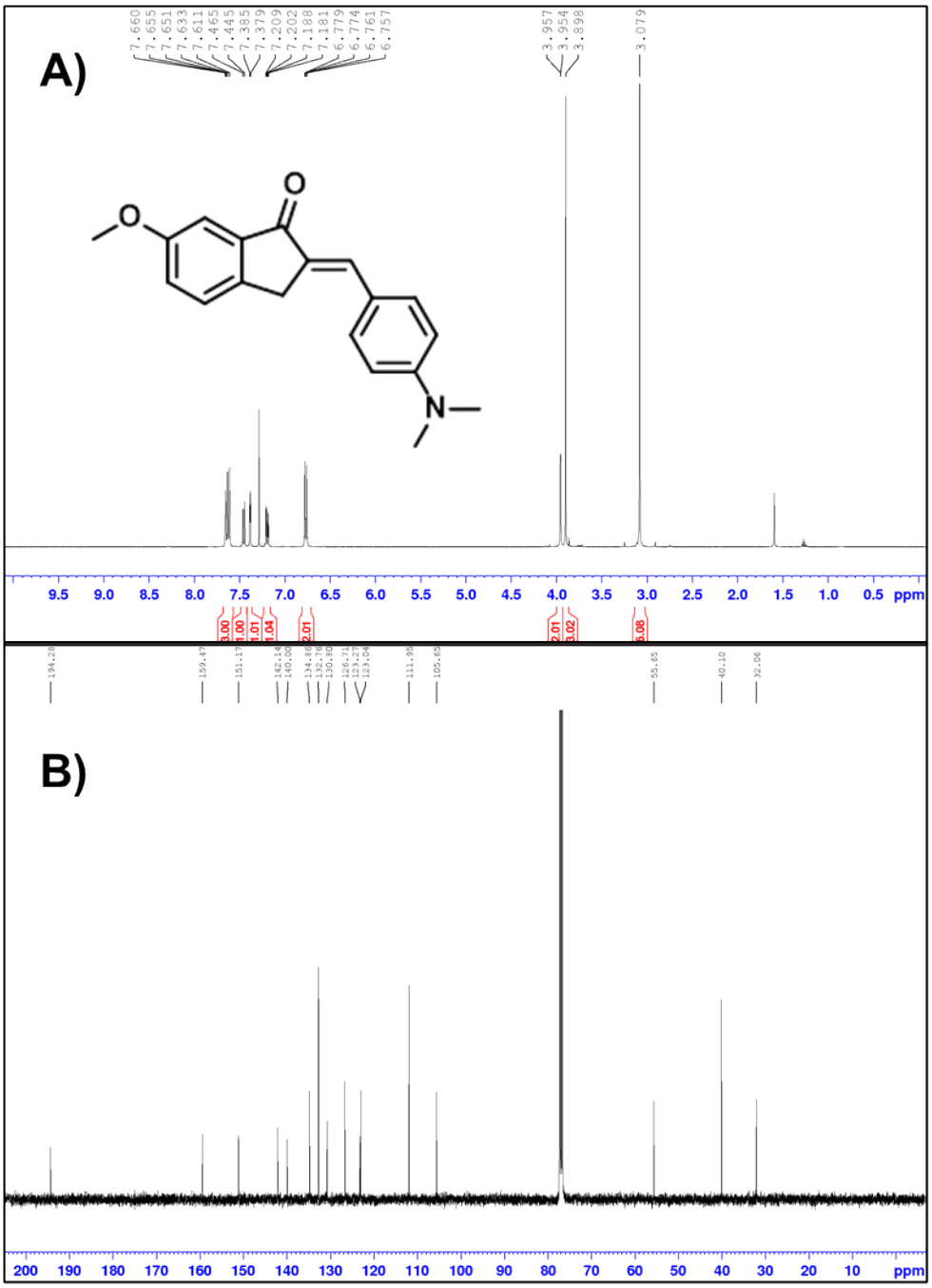


Figure S10. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

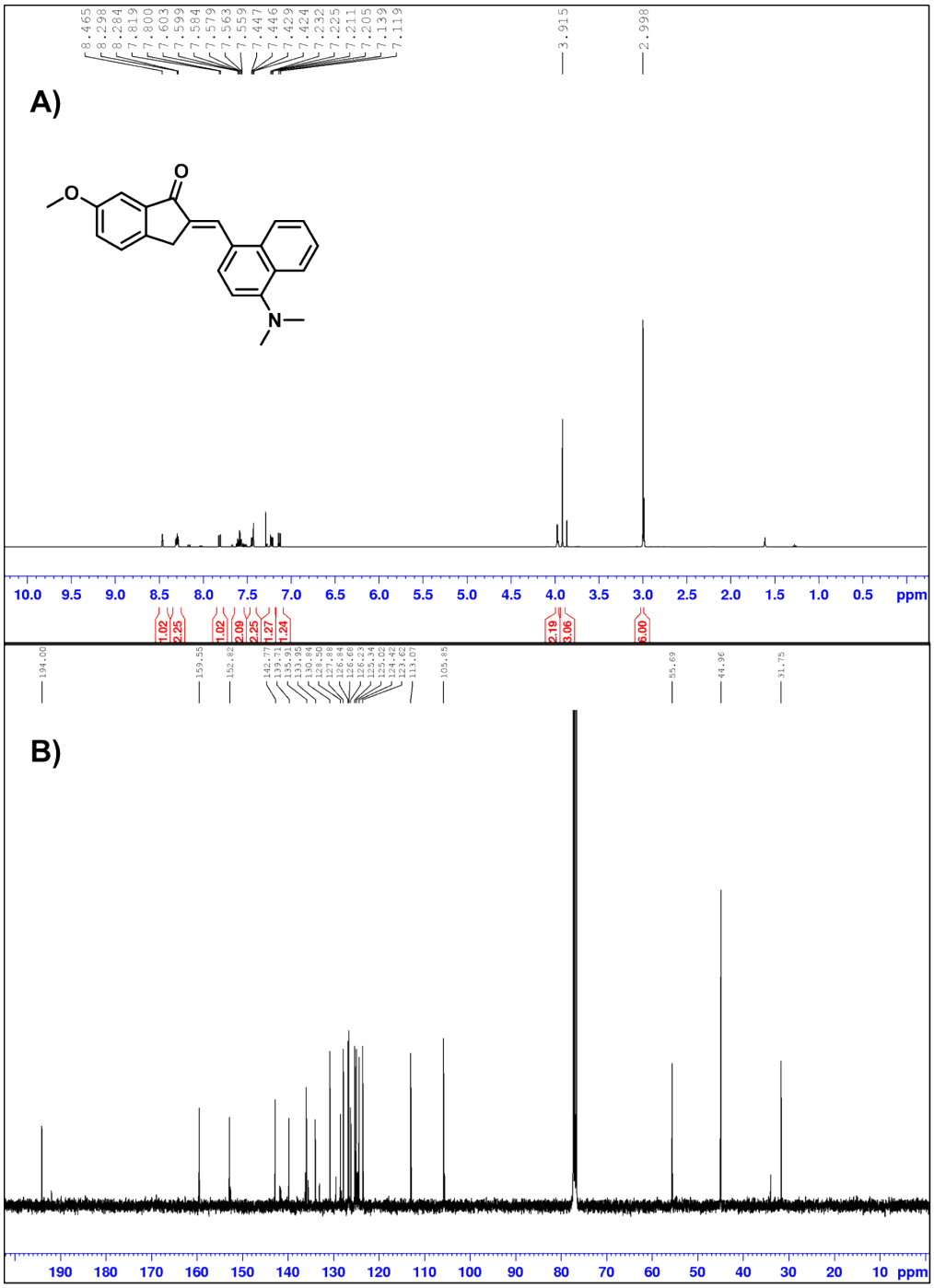


Figure S11. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

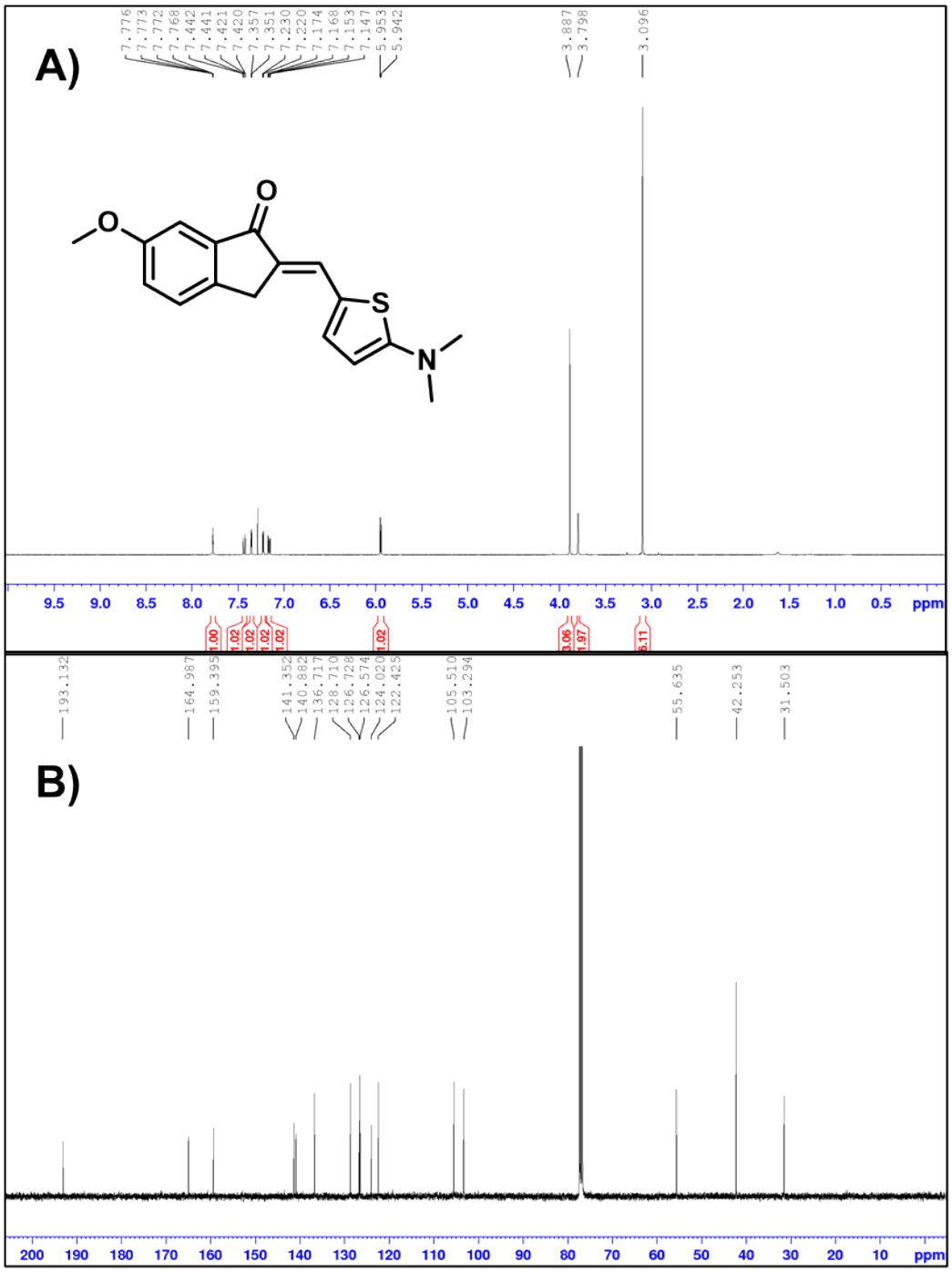


Figure S12. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

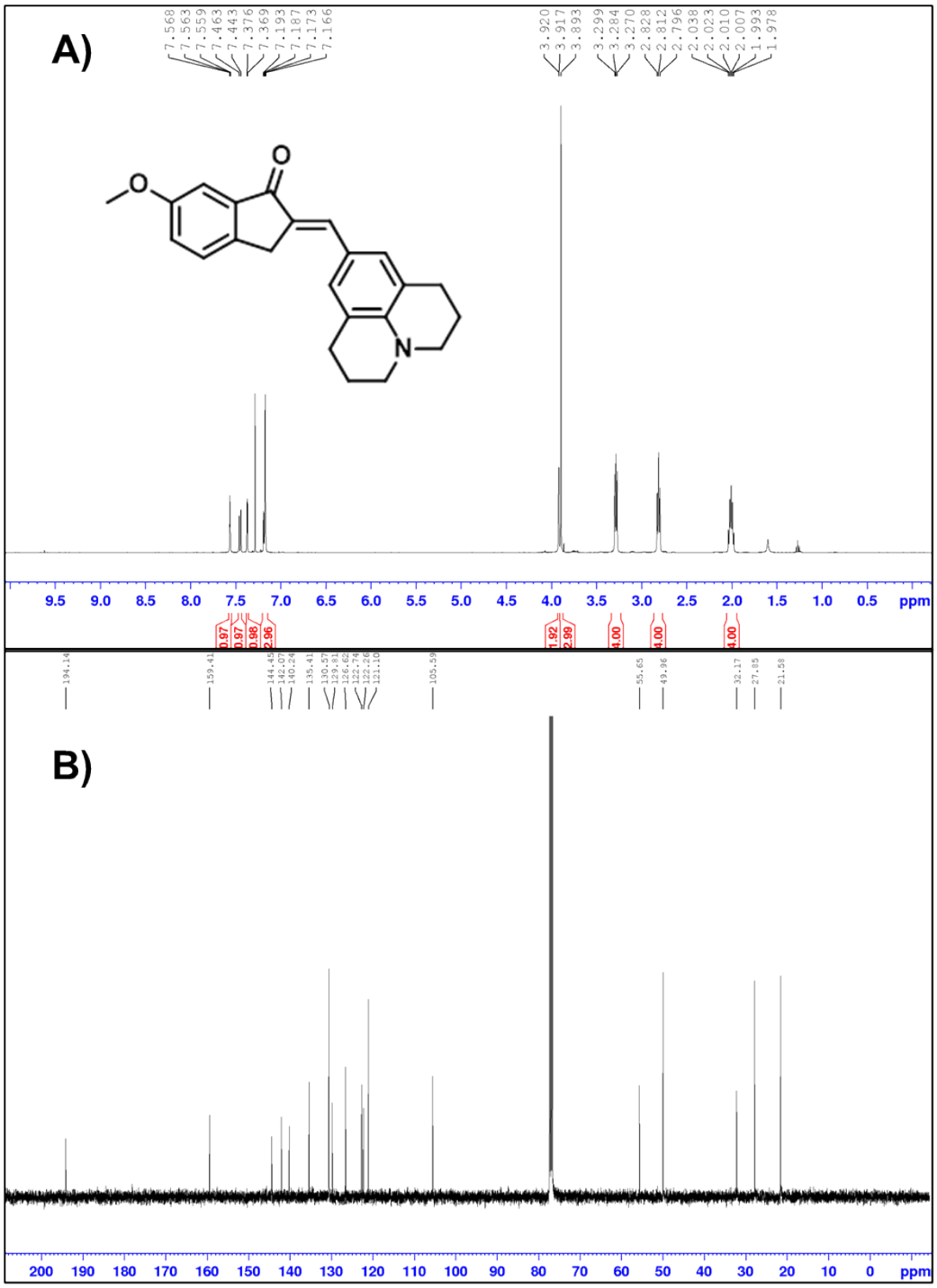


Figure S13. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

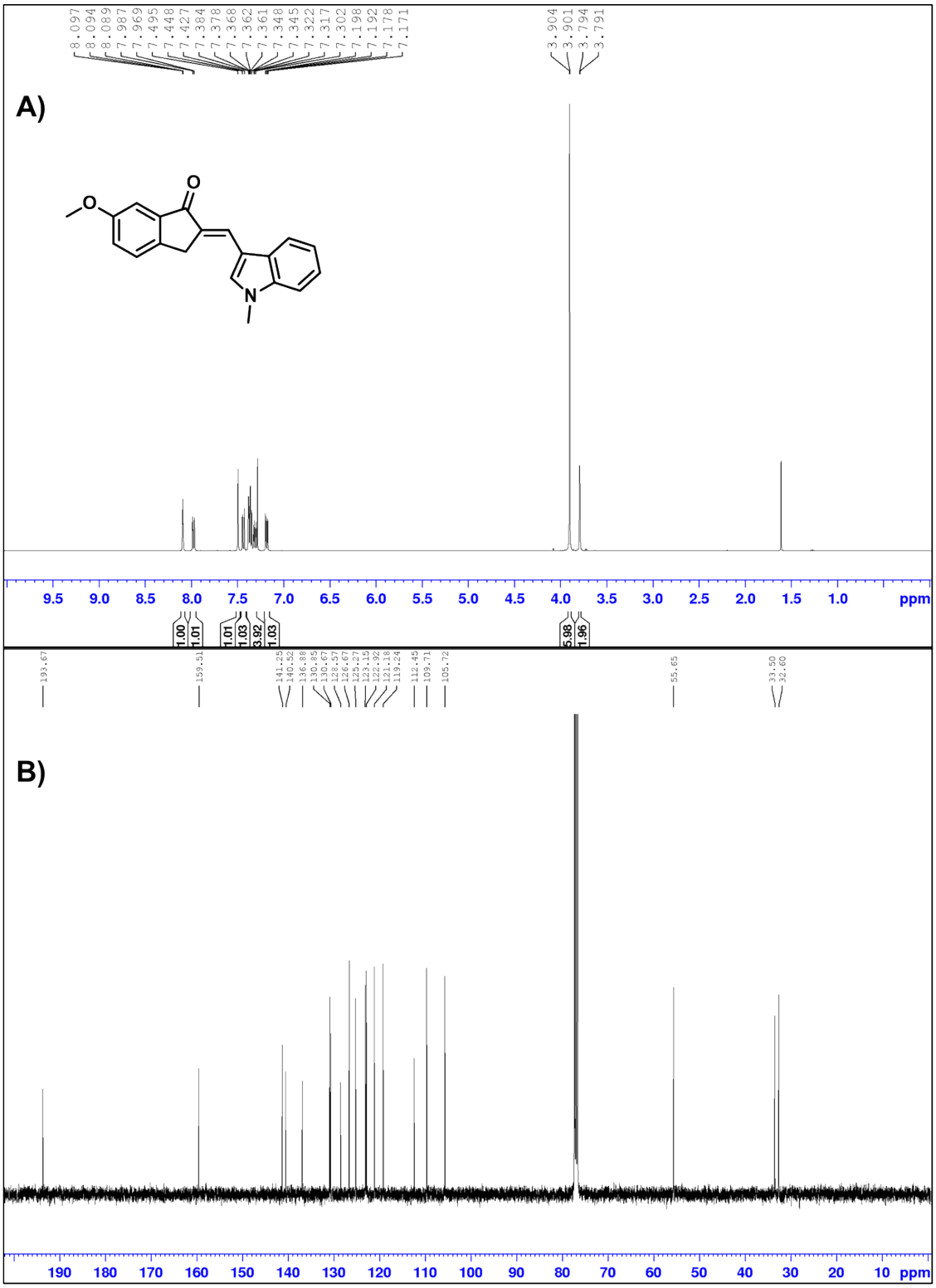


Figure S14. A) ^{1}H NMR spectrum in $CDCl_3$. B) $^{13}C\{^{1}H\}$ NMR spectrum in $CDCl_3$.

Table S2. MS Analysis of Modified *Nar*I Oligonucleotides with modifications.

***Nar*I**[a]	**Formula**	**Calc. Mass**[b]	**Exptl (*m/z*) (ESI$^-$)**[c]	**Exptl Mass**
6HI	$C_{116}H_{148}N_{37}O_{71}P_{11}$	3536.63	$[M-6H]^{6-}$= 588.55	3537.30
An6HI	$C_{125}H_{157}N_{38}O_{71}P_{11}$	3667.70	$[M-6H]^{6-}$= 610.30	3667.80
Th6HI	$C_{123}H_{155}N_{38}O_{71}P_{11}S$	3673.66	$[M-4H]^{4-}$= 917.56	3674.24
Ind6HI	$C_{126}H_{155}N_{38}O_{71}P_{11}$	3677.68	$[M-6H]^{6-}$= 612.20	3679.20
Ju6HI	$C_{129}H_{161}N_{38}O_{71}P_{11}$	3719.73	$[M-6H]^{6-}$= 619.15	3720.90
Nap6HI	$C_{129}H_{159}N_{38}O_{71}P_{11}$	3717.71	$[M-6H]^{6-}$= 618.80	3718.80
Cou6HI	$C_{130}H_{161}N_{38}O_{73}P_{11}$	3763.72	$[M-3H]^{3-}$=1253.40	3763.20
ETh6HI	$C_{125}H_{157}N_{38}O_{71}P_{11}S$	3699.67	$[M-4H]^{4-}$=923.99	3699.96

[a]Modification at G7 (X) of the 12mer *Nar*I oligonucleotide (5'-CTC-GGC-X-CCA-TC-3'). [b]Monoisotopic mass of most abundant isotopologue containing one ^{13}C isotope. [c]Measured *m/z* from mass spectrum.

Table S3. Yields and MS Analysis of Modified *Nar*I Oligonucleotides with modifications

Position[a]	**Yield**[b]	**Formula**	**Calc. Mass**[c]	**Exptl (*m/z*) (ESI$^-$)**[d]	**Exptl Mass**
An6HI	~90	$C_{127}H_{159}N_{36}O_{73}P_{11}$	3697.70	$[M-5H]^{5-}$=738.67	3698.35
Th6HI	~90	$C_{125}H_{157}N_{36}O_{73}P_{11}S$	3703.65	$[M-7H+Na]^{6-}$= 620.08	3704.48
Ind6HI	~90	$C_{128}H_{157}N_{36}O_{73}P_{11}$	3707.68	$[M-7H+Na]^{6-}$= 620.75	3708.50
Nap6HI	~90	$C_{131}H_{161}N_{36}O_{73}P_{11}$	3747.71	$[M-7H+Na]^{6-}$= 627.42	3748.52

flanked by thymine. Reactions performed using Method 2.

[a]Modification at G7 (X) of the 12mer *Nar*I oligonucleotide (5'-CTC-GGT-X-TCA-TC-3'). [b]Percent yield from integration of HPLC trace assuming the same extinction coefficients for the 6HI labelled *Nar*I precursor and chalcone product, *Nar*I (ε_{260} = 106,900 $M^{-1}cm^{-1}$). [c]Monoisotopic mass of most abundant isotopologue containing one ^{13}C isotope. [d]Measured *m/z* from mass spectrum.

Table S4. Yields and MS Analysis of Modified *Nar*I Oligonucleotides with modifications

Position[a]	Yield[b]	Formula	Calc. Mass[c]	Exptl (*m/z*) (ESI$^-$)[d]	Exptl Mass
An6HI	~90	$C_{127}H_{157}N_{42}O_{69}P_{11}$	3715.72	$[M-6H]^{6-}$= 618.33	3715.98
Th6HI	~90	$C_{125}H_{155}N_{42}O_{69}P_{11}S$	3721.68	$[M-7H+2Na]^{5-}$= 752.33	3722.65
Ind6HI	~90	$C_{128}H_{155}N_{42}O_{69}P_{11}$	3725.71	$[M-7H+Na]^{6-}$= 623.50	3725.00
Nap6HI	~90	$C_{131}H_{159}N_{42}O_{69}P_{11}$	3765.74	$[M-7H+Na]^{6-}$= 630.17	3765.02

flanked by adenine. Reactions performed using Method 2.

[a]Modification at G7 (X) of the 12mer *Nar*I oligonucleotide (5'-CTC-GGA-X-ACA-TC-3'). [b]Percent yield from integration of HPLC trace assuming the same extinction coefficients for the 6HI labelled *Nar*I precursor and chalcone product, *Nar*I (ε_{260} = 114,900 $M^{-1}cm^{-1}$). [c]Monoisotopic mass of most abundant isotopologue containing one ^{13}C isotope. [d]Measured *m/z* from mass spectrum.

Table S5. Yields and MS Analysis of Napthyl Modified 22mer *Nar*I Oligonucleotides.

Position[a]	Yield[b]	Formula	Calc. Mass[c]	Exptl (*m/z*) (ESI$^-$)[d]	Exptl Mass
G7	90	$C_{226}H_{283}N_{74}O_{131}P_{21}$	6767.22	$[M-9H+2Na]^{7-}$= 972.11	6767.77
C13	90	$C_{226}H_{281}N_{76}O_{131}P_{21}$	6807.22	$[M-11H+2Na]^{9-}$= 760.32	6807.88
A17	90	$C_{225}H_{281}N_{74}O_{132}P_{21}$	6783.21	$[M-7H+Na]^{6-}$= 1133.24	6783.44

Reactions performed using Method 2.

[a]Modification positions in 22mer *Nar*I oligonucleotide (5'-CTC-GGC-**G7**-CCA-TC-**C13**-TT-**A17**-CGA-GC-3'). [b]Percent yield from integration of HPLC trace assuming the same extinction coefficients for the 6HI labelled *Nar*I 22mer precursor and chalcone product, *Nar*I (ε_{260} = 192,900 $M^{-1}cm^{-1}$). [c]Monoisotopic mass of most abundant isotopologue containing one ^{13}C isotope. [d]Measured *m/z* from mass spectrum.

Table S6. MS Analysis of 11mer MN4 Split aptamer. An6HI incorporated using authentic phosphoramidite standard (**2b**).

Position[a]	Formula	Calc. Mass[c]	Exptl (*m/z*) (ESI$^-$)[d]	Exptl Mass
G9	$C_{120}H_{145}N_{47}O_{60}P_{10}$	3499.70	$[M-6H]^{6-}$= 582.42	3500.52

[a]Modification at G9 (X) of the 11mer *MN4* split aptamer (SA) oligonucleotide (5'-GGC-GAC-AAX-GA-3'). [c]Monoisotopic mass of most abundant isotopologue containing one ^{13}C isotope. [d]Measured *m/z* from mass spectrum.

Supplementary Fluorescent Hybridization Studies.

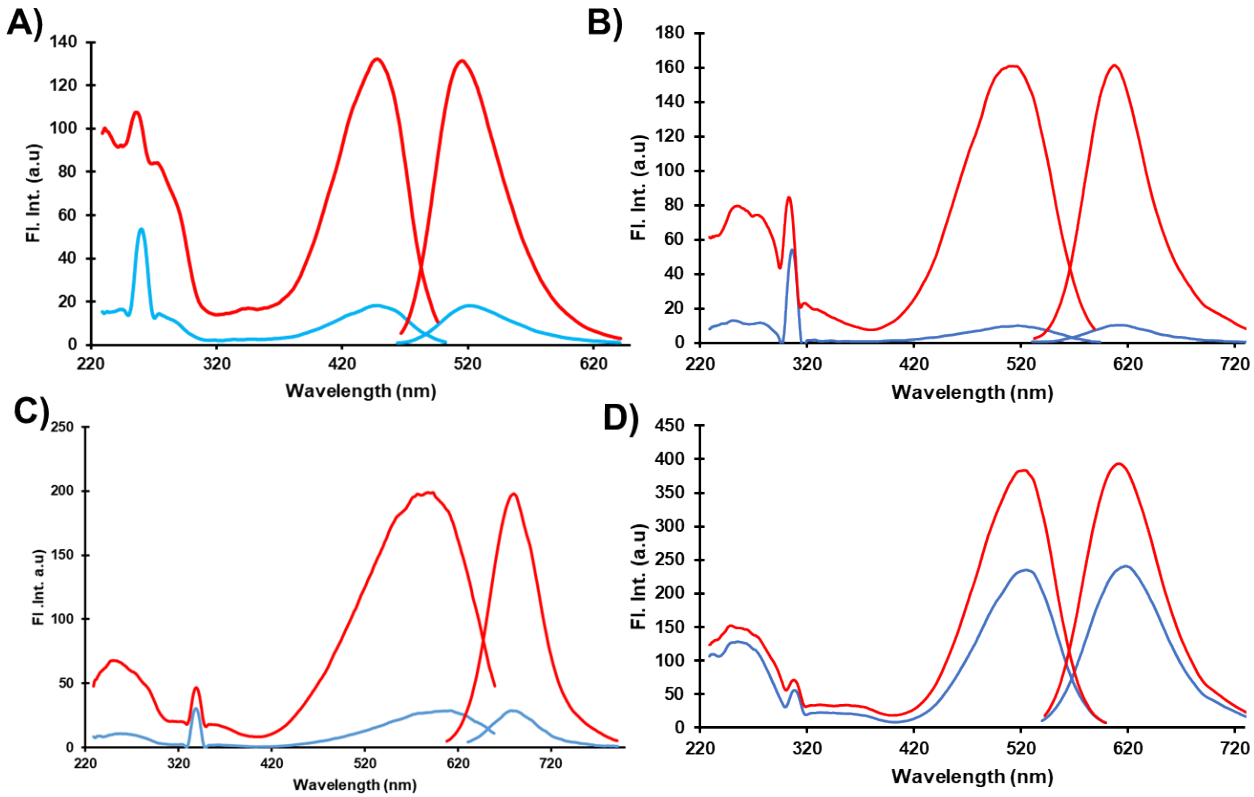


Figure S15. Fluorescence spectra of *Nar*I modified oligos in the absence (solid blue) and presence (solid red) of complementary DNA (across from abasic site or cytosine). Single strand and duplex concentrations at 5μM in 50mM sodium phosphate, pH 7, containing 100mM sodium chloride. A) Ind6HI (abasic). B) Ju6HI (cytosine). C) Eth6HI (cytosine). D) Cou6HI (Cytosine). Excitation Spectra on left, emission spectra on right.

Supplementary UV-Vis Hybridization Studies.

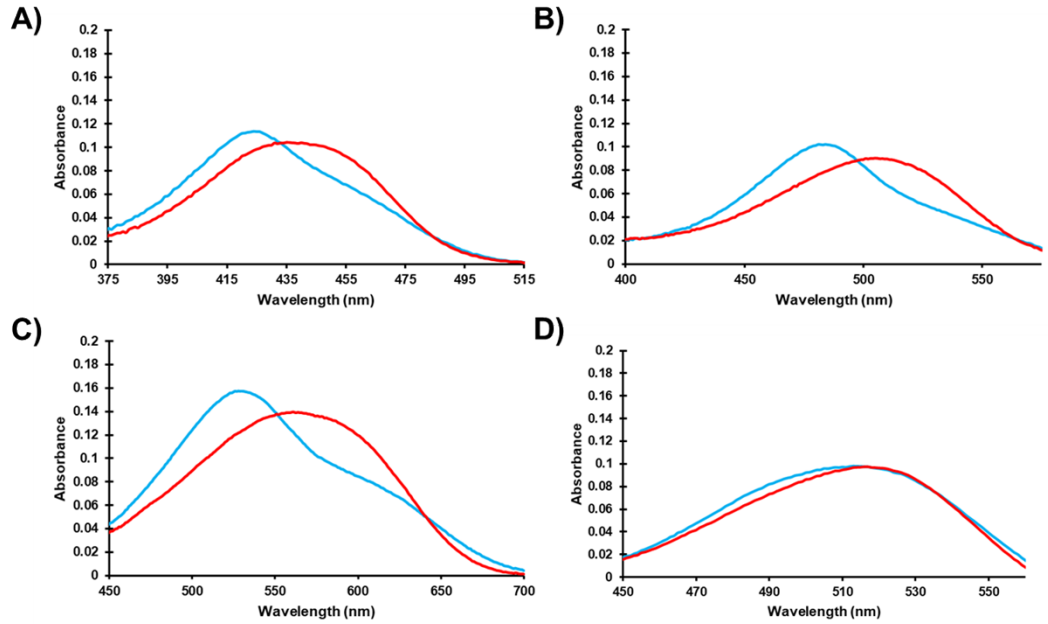


Figure S16. UV-Vis spectra of *Nar*I modified oligos in the absence (solid blue) and presence (solid red) of complementary DNA (across from abasic site or cytosine). Single strand and duplex concentrations at 5μM in 50mM sodium phosphate, pH 7, containing 100mM sodium chloride. A) Ind6HI (abasic). B) Ju6HI (cytosine). C) Eth6HI (cytosine). D) Cou6HI (cytosine).

Supplementary Circular Dichroism (CD) Studies.

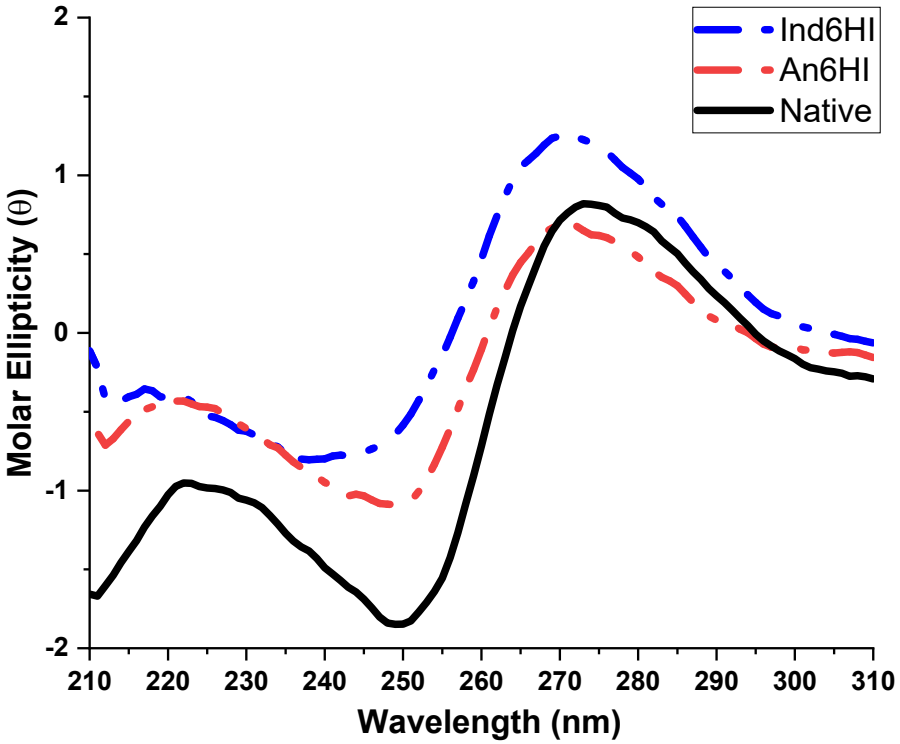


Figure S17. CD spectra of native *Nar*I duplex (black solid line) and modified *Nar*I duplexes containing An6HI (red dash-dot line) and Ind6HI (blue dash-dot line) across from cytosine. Duplex concentrations at 5μM in 50mM sodium phosphate, pH 7, containing 100mM sodium chloride.

Fluorescent Thermal Melting for Nap6HI Modified Oligos.

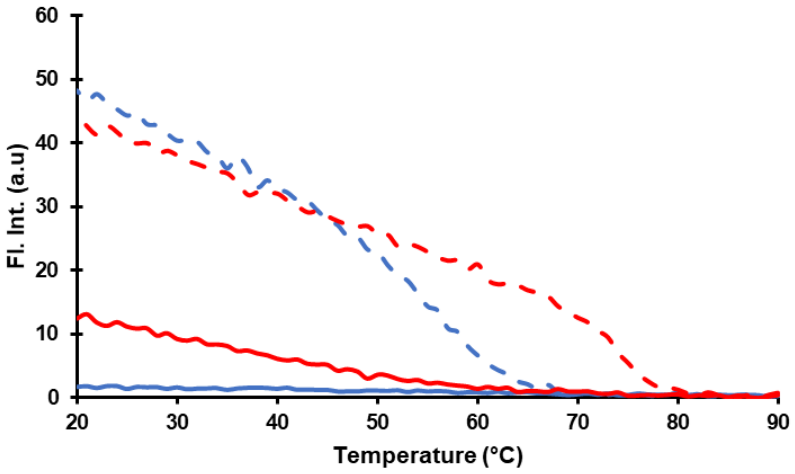


Figure S18. Fluorescent thermal melting curve for Nap6HI Modified *Nar*I oligonucleotides (12mer (blue lines) and 22mer (red lines), at the same position (G7). Solid lines indicate thermal melting in absence of full-length complement, dashed lines indicate thermal melting in presence of full-length complement. Experiments carried out at a cooling rate of 1°C/min at a concentration of 1μM (single strand and duplex), in 50mM sodium phosphate, pH 7, containing 100mM sodium chloride.

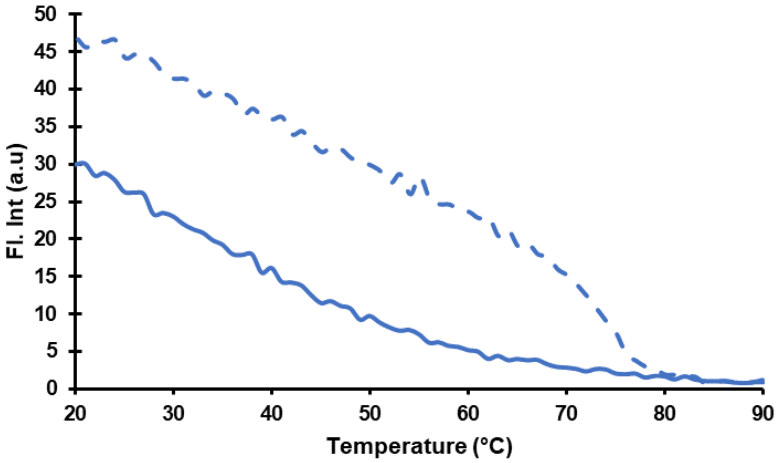


Figure S19. Fluorescent thermal melting curves for 22mer*Nar*I modified at position C13 with Nap6HI. Solid line indicates thermal melting in absence of full-length complement, dashed line indicates thermal melting in presence of full-length complement. Experiments carried out at a cooling rate of 1°C/min at a concentration of 1μM (single strand and duplex), in 50mM sodium phosphate, pH 7, containing 100mM sodium chloride.

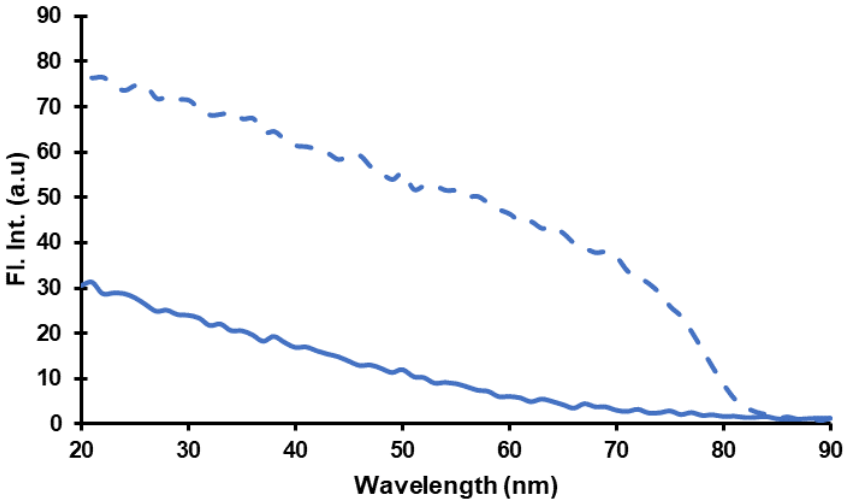


Figure S20. Fluorescent thermal melting curves for 22mer*Nar*I modified at position A17 with Nap6HI. Solid line indicates thermal melting in absence of full-length complement, dashed line indicates thermal melting in presence of full-length complement. Experiments carried out at a cooling rate of 1°C/min at a concentration of 1μM (single strand and duplex), in 50mM sodium phosphate, pH 7, containing 100mM sodium chloride.

Ratiometric Fluorescent Detection of Hybridization by ETh6HI.

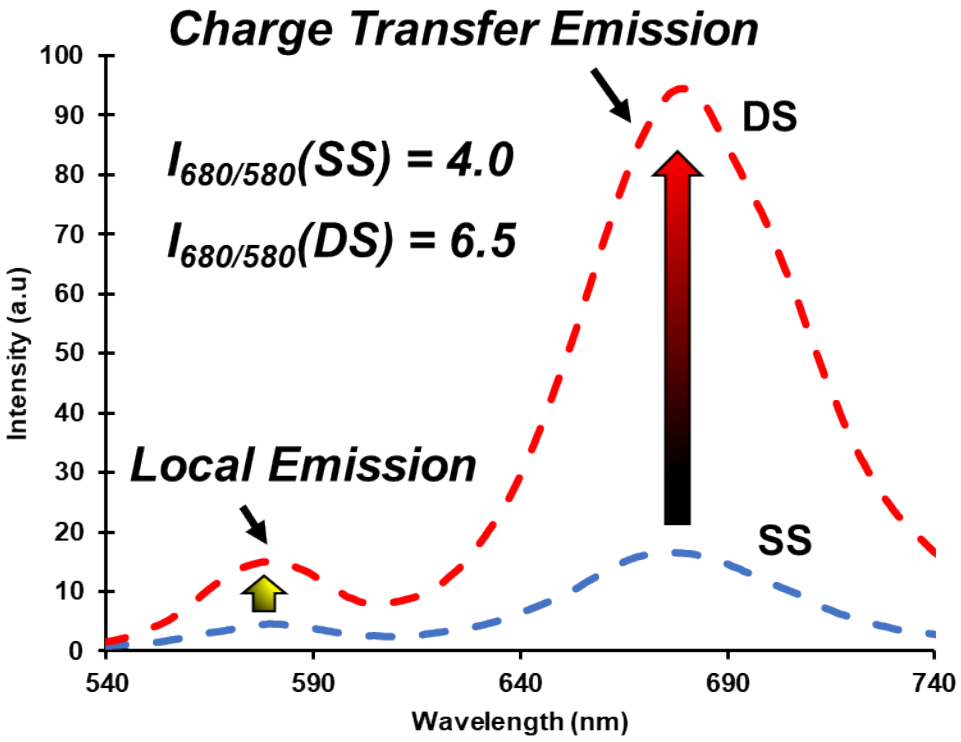


Figure S21. Fluorescence emission spectra (λ_{ex}=520nm) of *Nar*I functionalized with *E*Th6HI in the absence (solid blue) and presence (solid red) of complementary DNA (across from cytosine). Single strand and duplex concentrations at 5μM in 50mM sodium phosphate, pH 7, containing 100mM sodium chloride.

ESI$^-$-MS Spectra of Modified *NarI* Oligonucleotides with Flanking Cytosines.

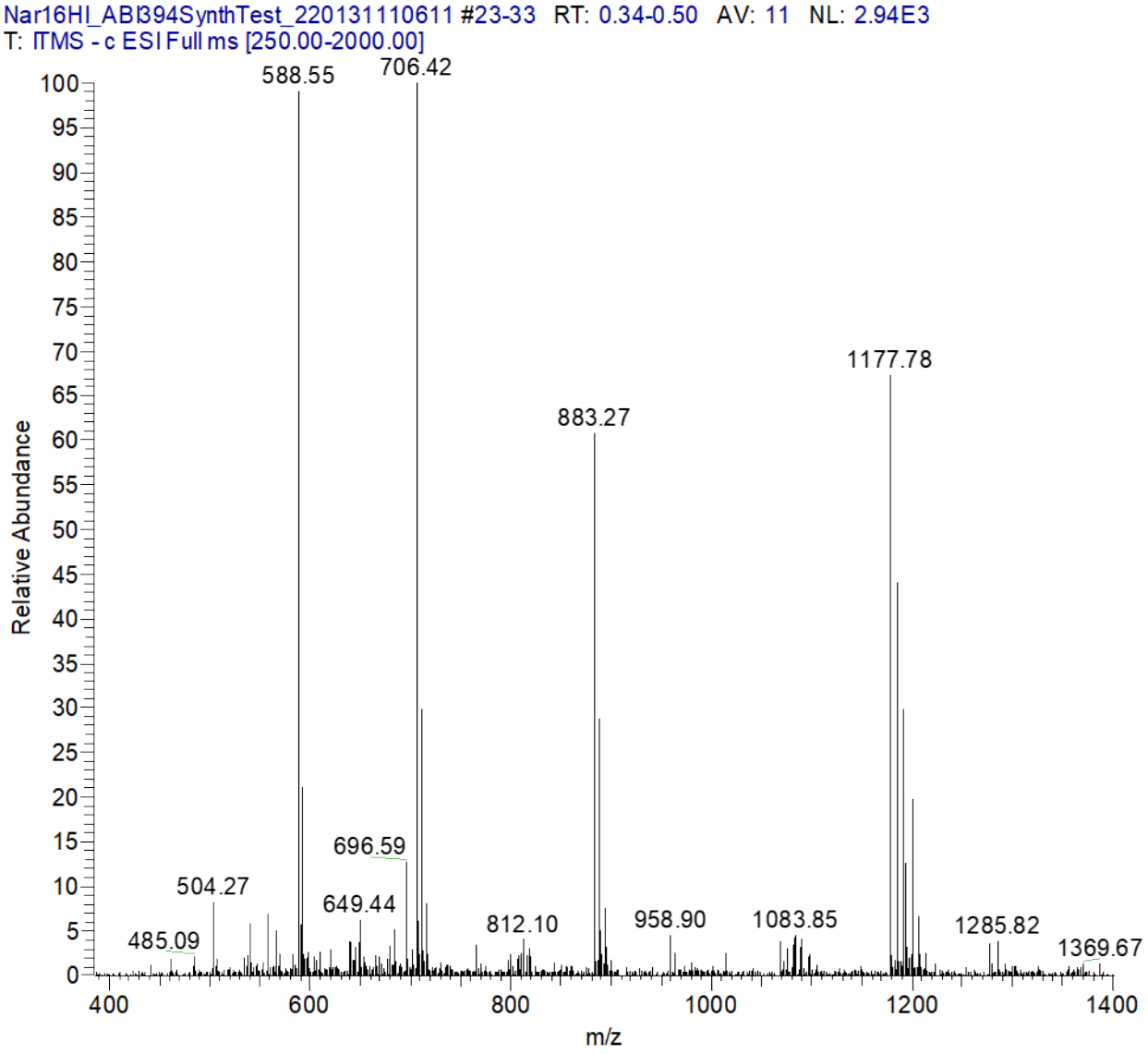


Figure S22. ESI$^-$ MS Spectrum for *NarI* [***FC***-6HI] 12mer sequence.

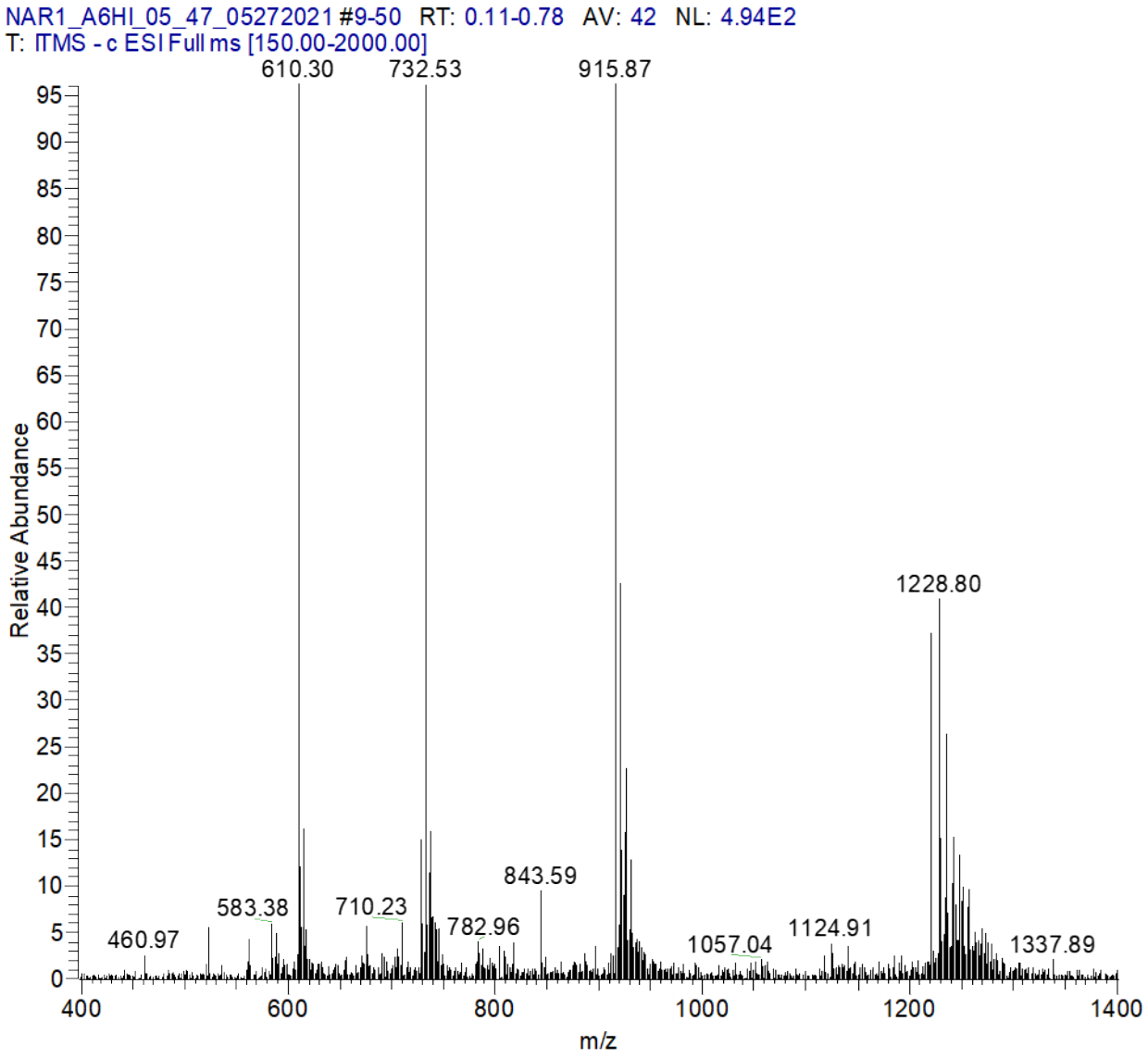


Figure S23. ESI$^-$ MS Spectrum for *Nar*I [***FC***-An6HI] 12mer sequence.

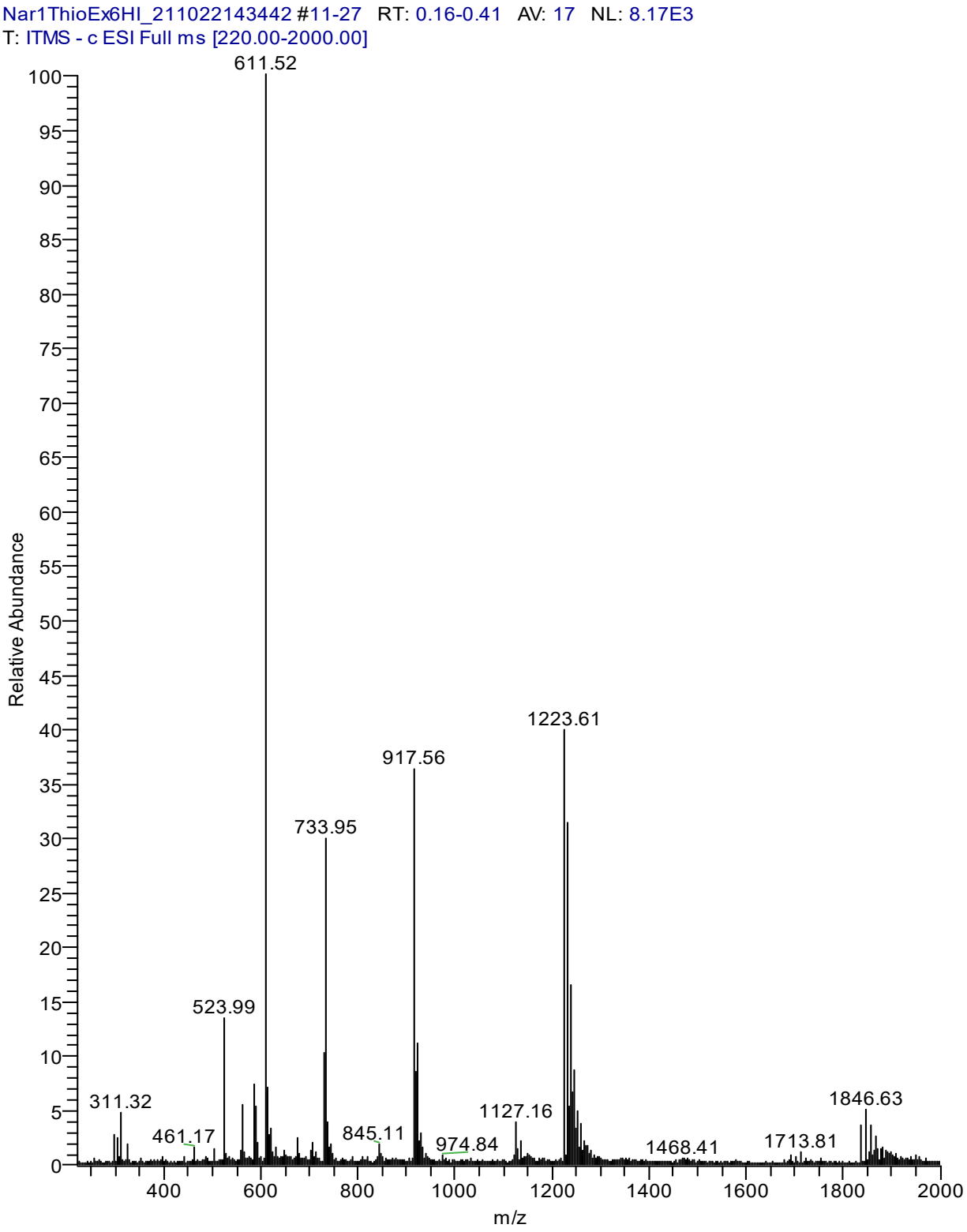


Figure S24. ESI$^-$ MS Spectrum for *Nar*I [***FC***-Th6HI] 12mer sequence.

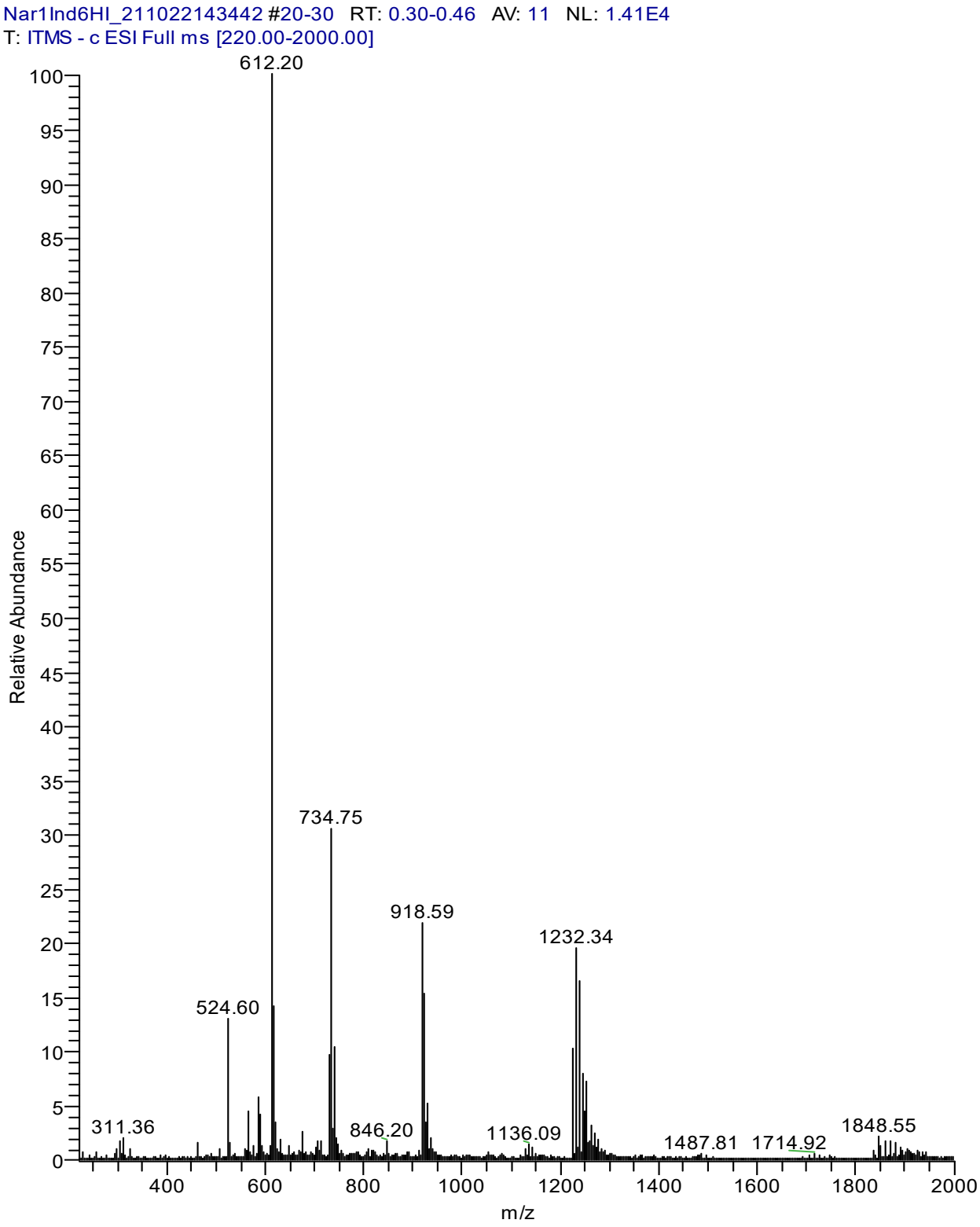


Figure S25. ESI$^-$ MS Spectrum for *Nar*I [***FC***-Ind6HI] 12mer sequence.

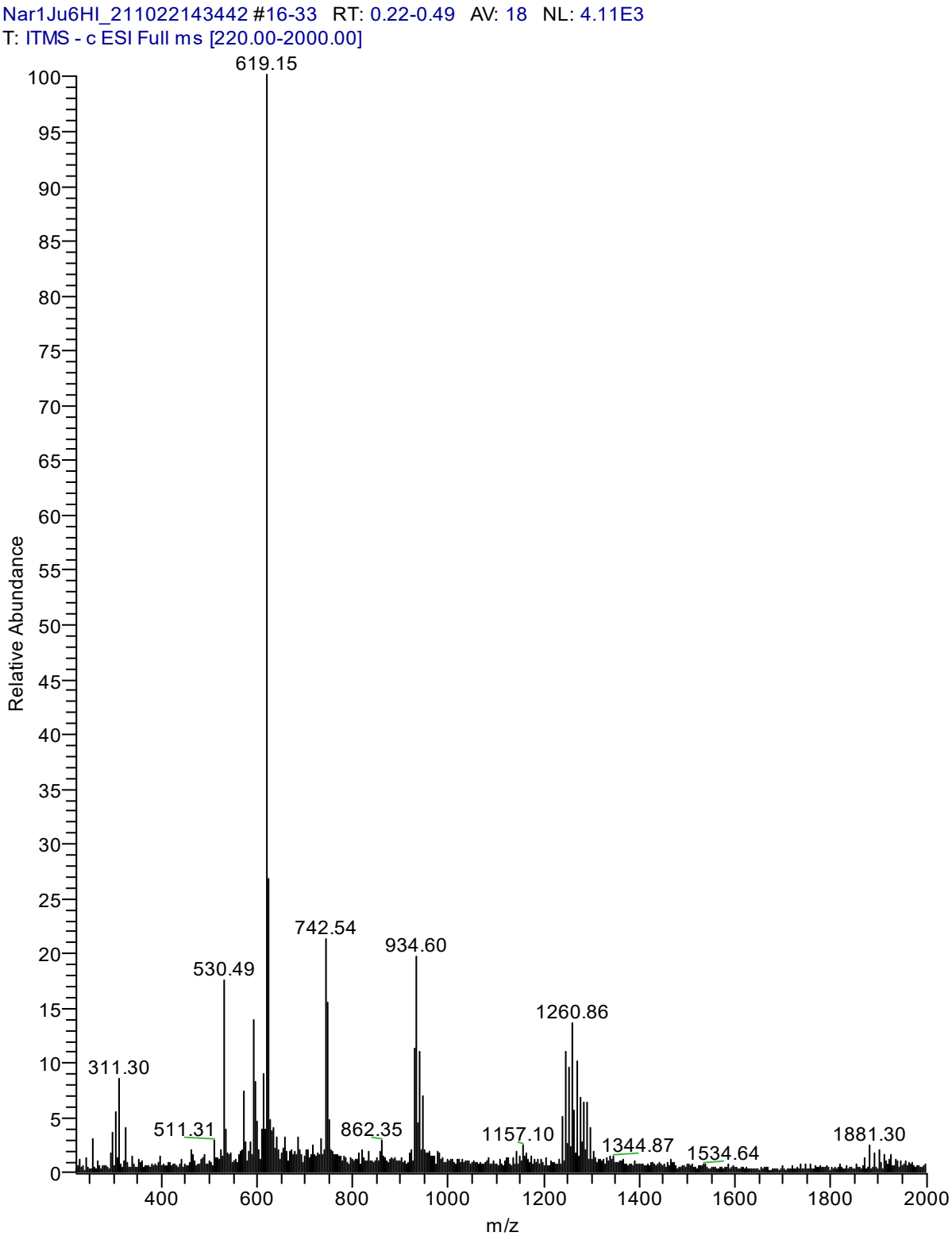


Figure S26. ESI$^-$ MS Spectrum for *Nar*I [***FC***-Ju6HI] 12mer sequence.

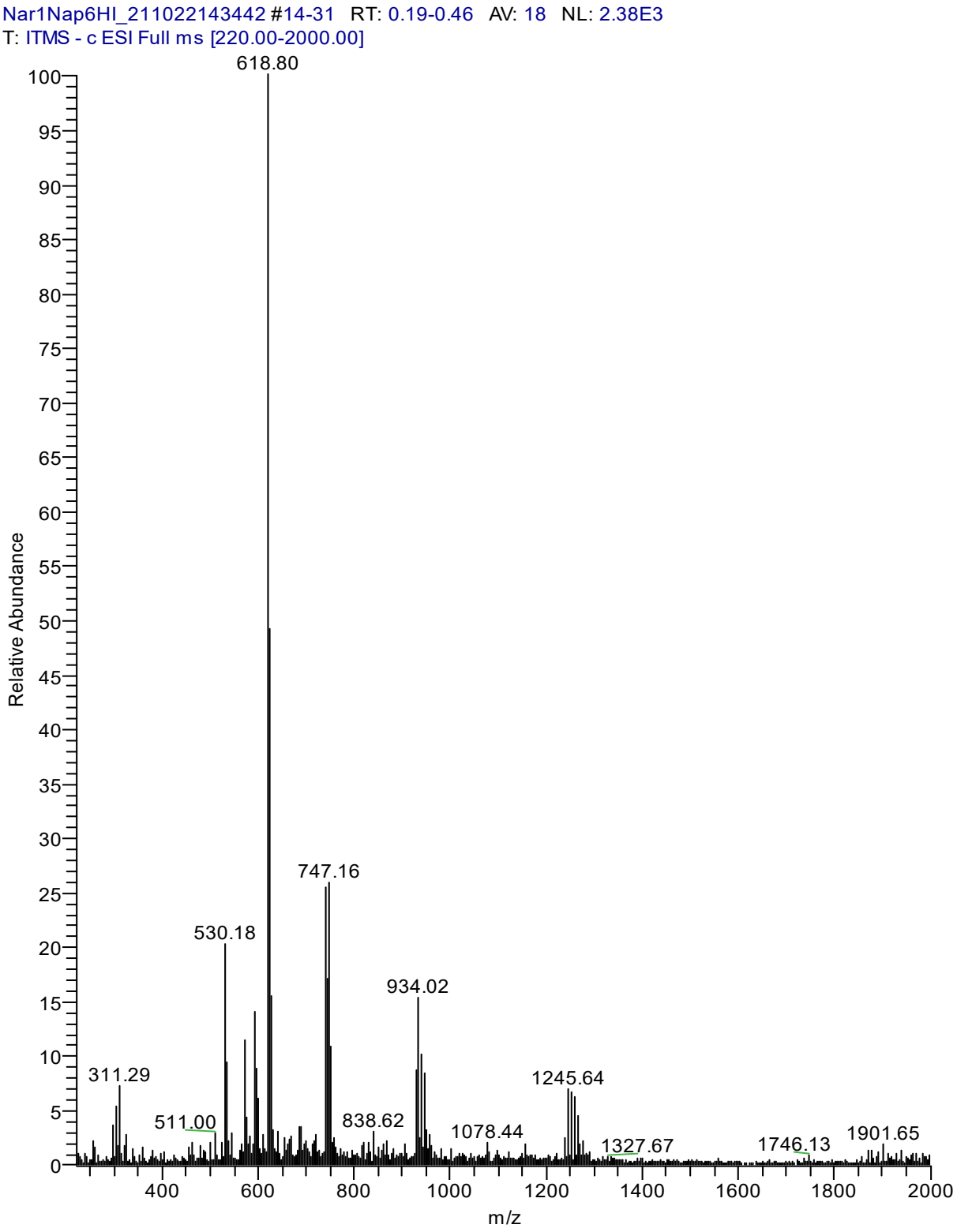


Figure S27. ESI$^-$ MS Spectrum for *Nar*I [***FC***-Nap6HI] 12mer sequence.

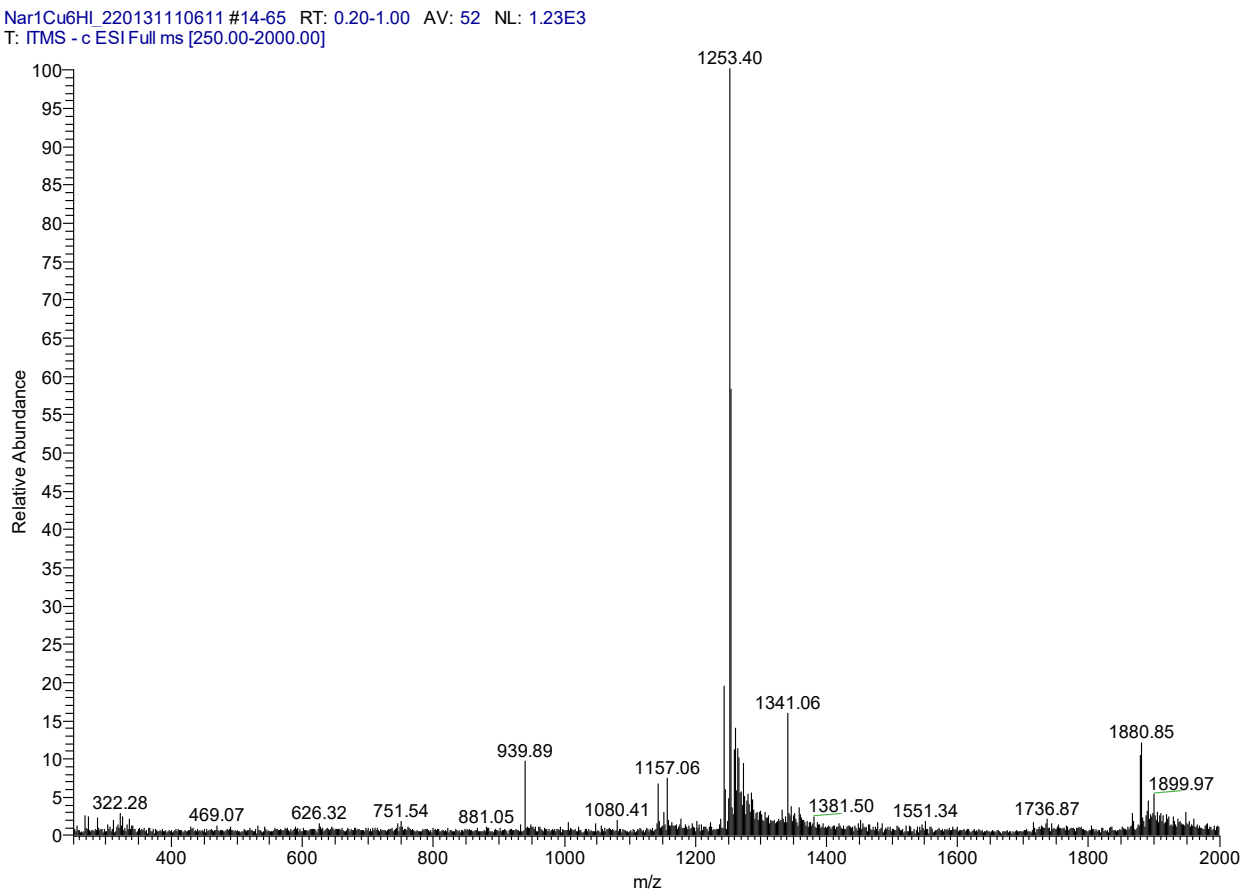


Figure S28. ESI$^-$ MS Spectrum for *Nar*I [***FC***-Cou6HI] 12mer sequence.

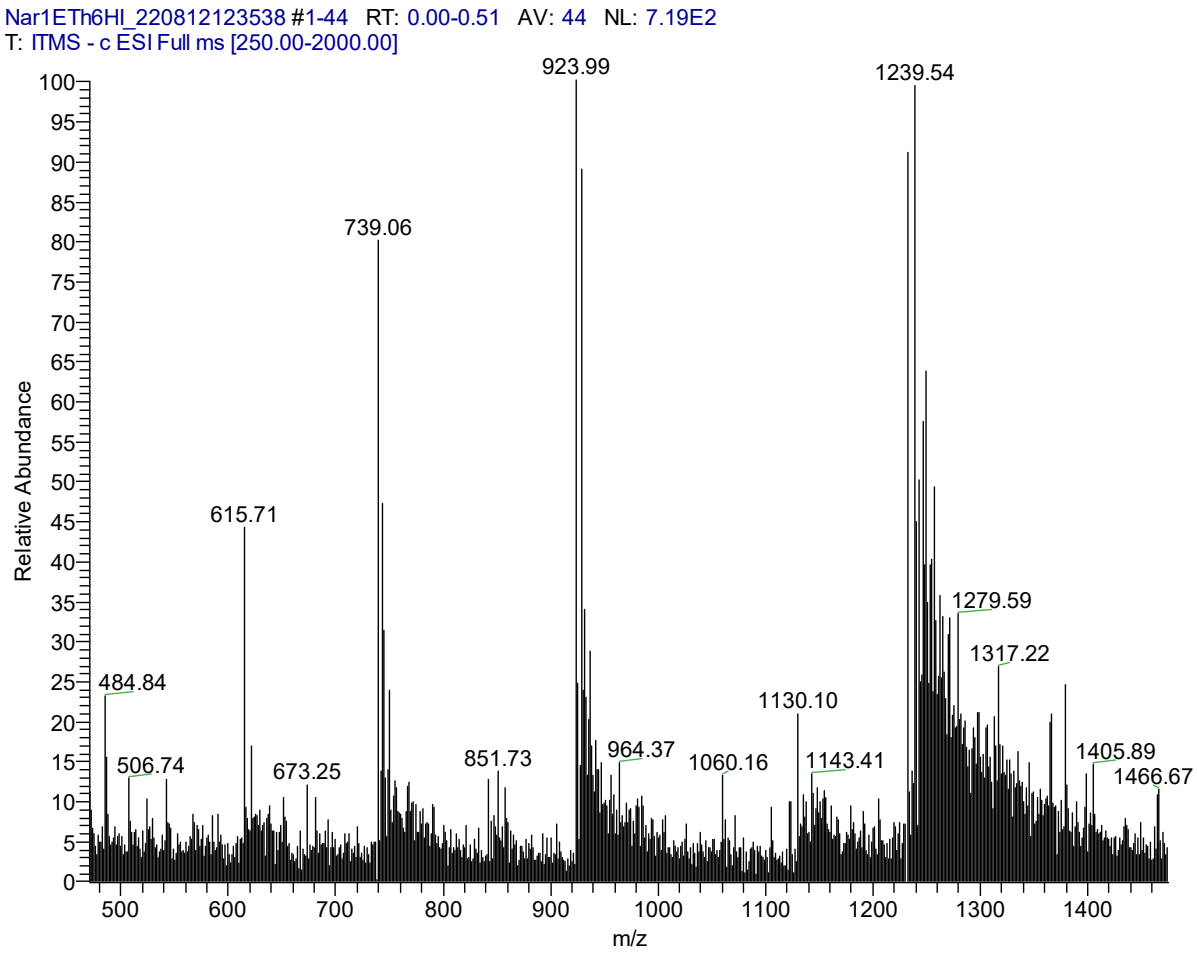


Figure 29. ESI$^-$ MS Spectrum for *Nar*I [***FC***-ETh6HI] 12mer sequence.

ESI$^-$-MS Spectra of Modified *Nar*I Oligonucleotides with Flanking Thymidines.

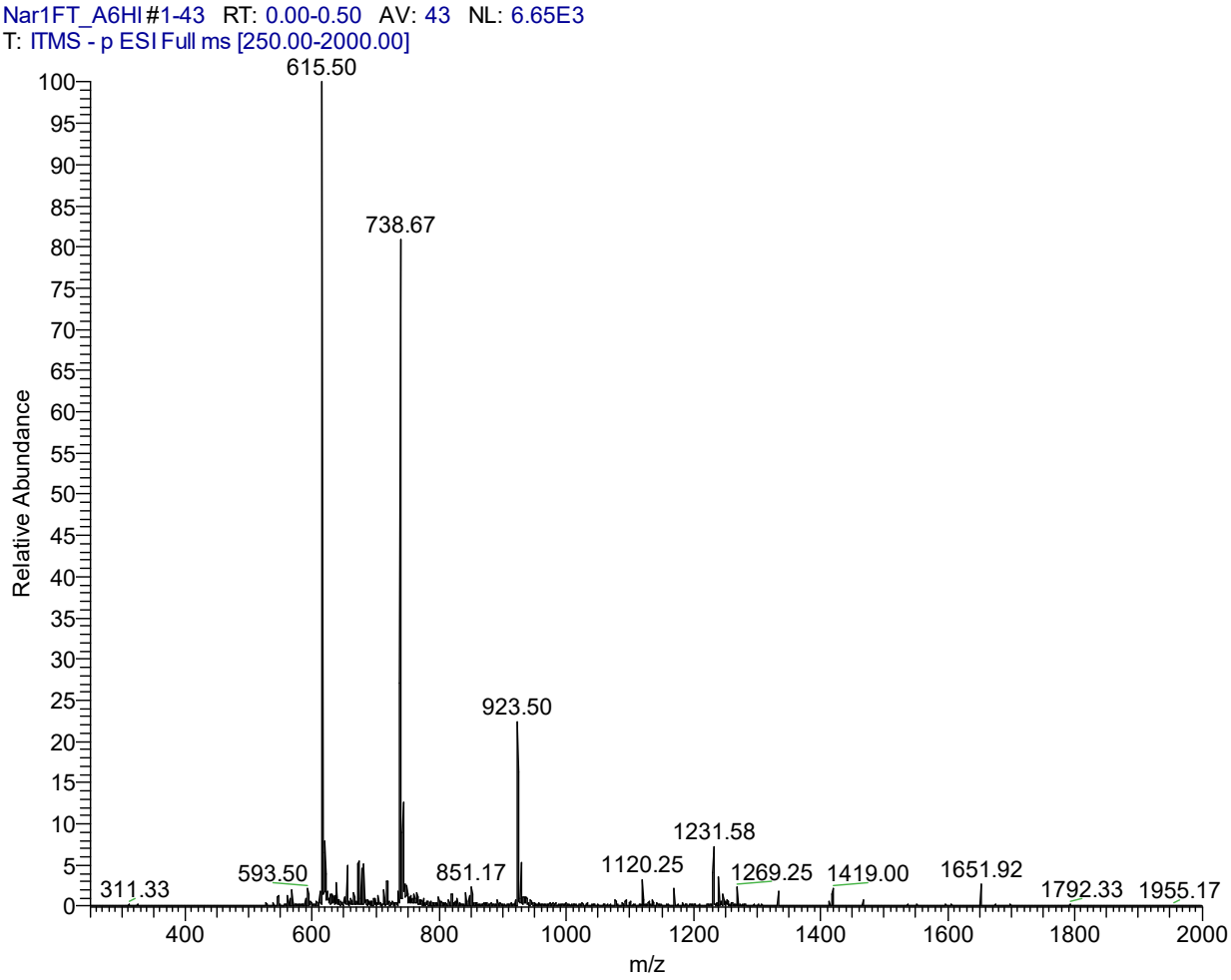


Figure S30. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FT***-An6HI] sequence.

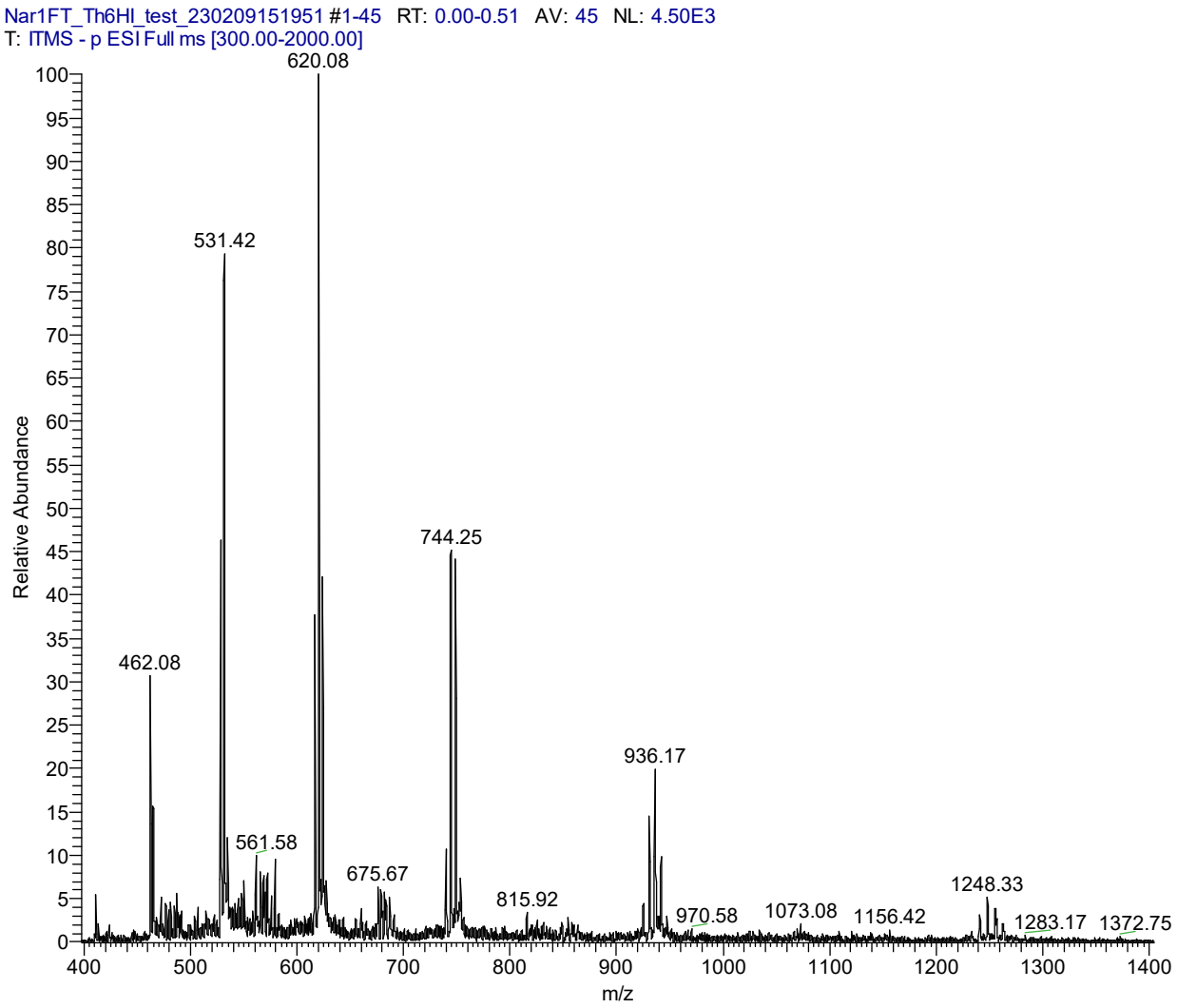


Figure S31. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FT***-Th6HI] sequence.

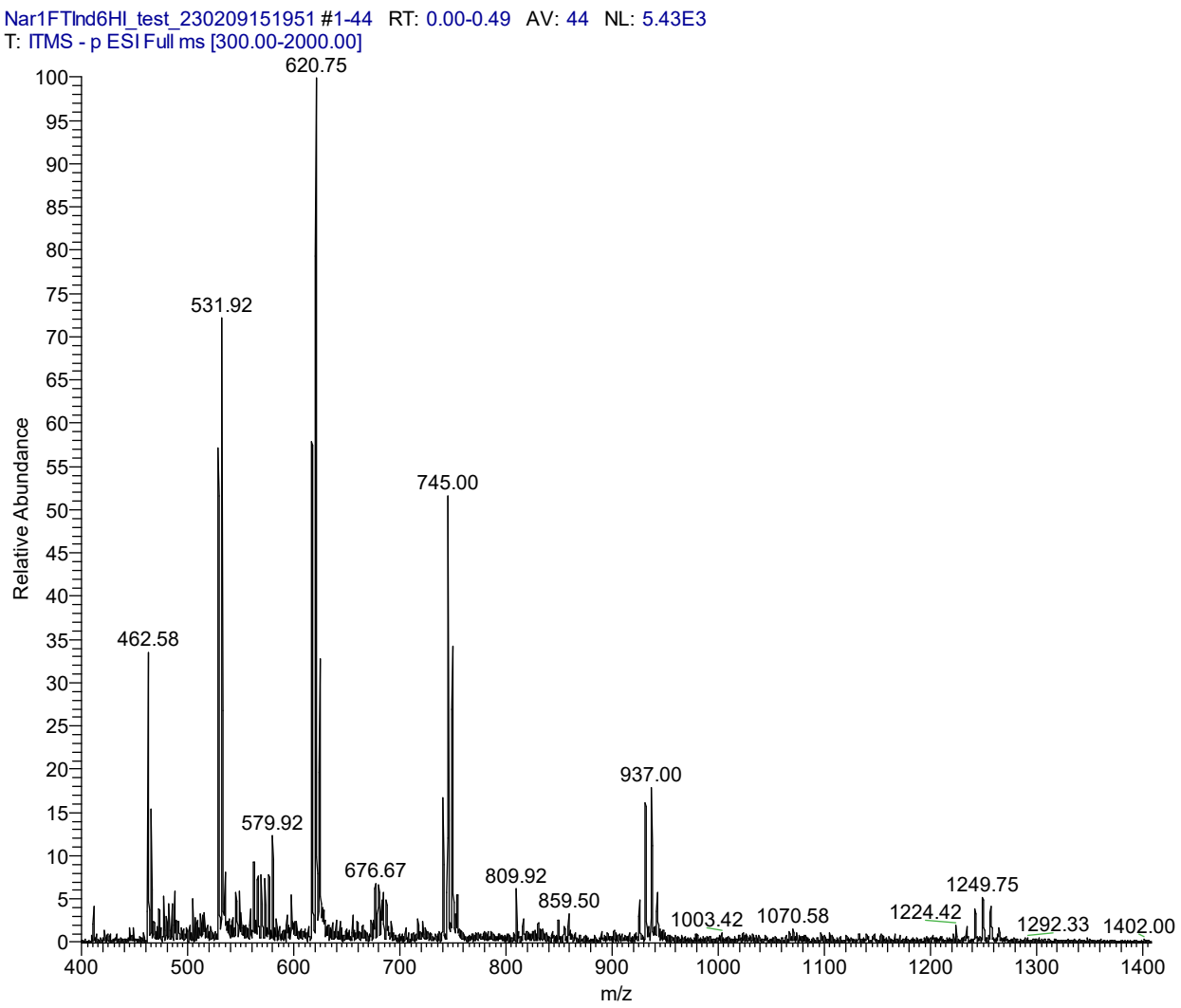


Figure S32. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FT***-Ind6HI] sequence.

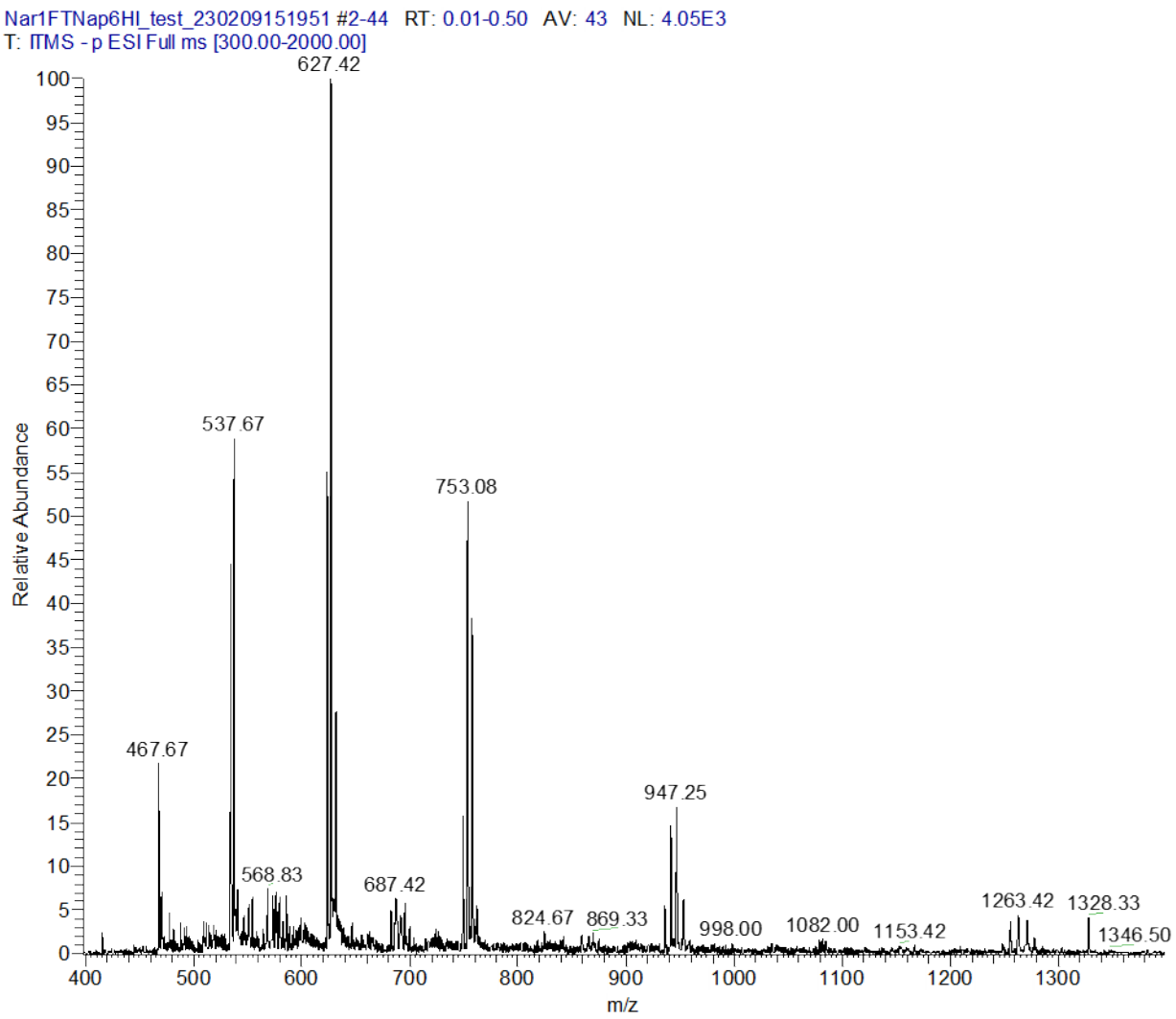


Figure S33. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FT***-Nap6HI] sequence.

ESI$^-$-MS Spectra of Modified *Nar*I Oligonucleotides with Flanking Adenines.

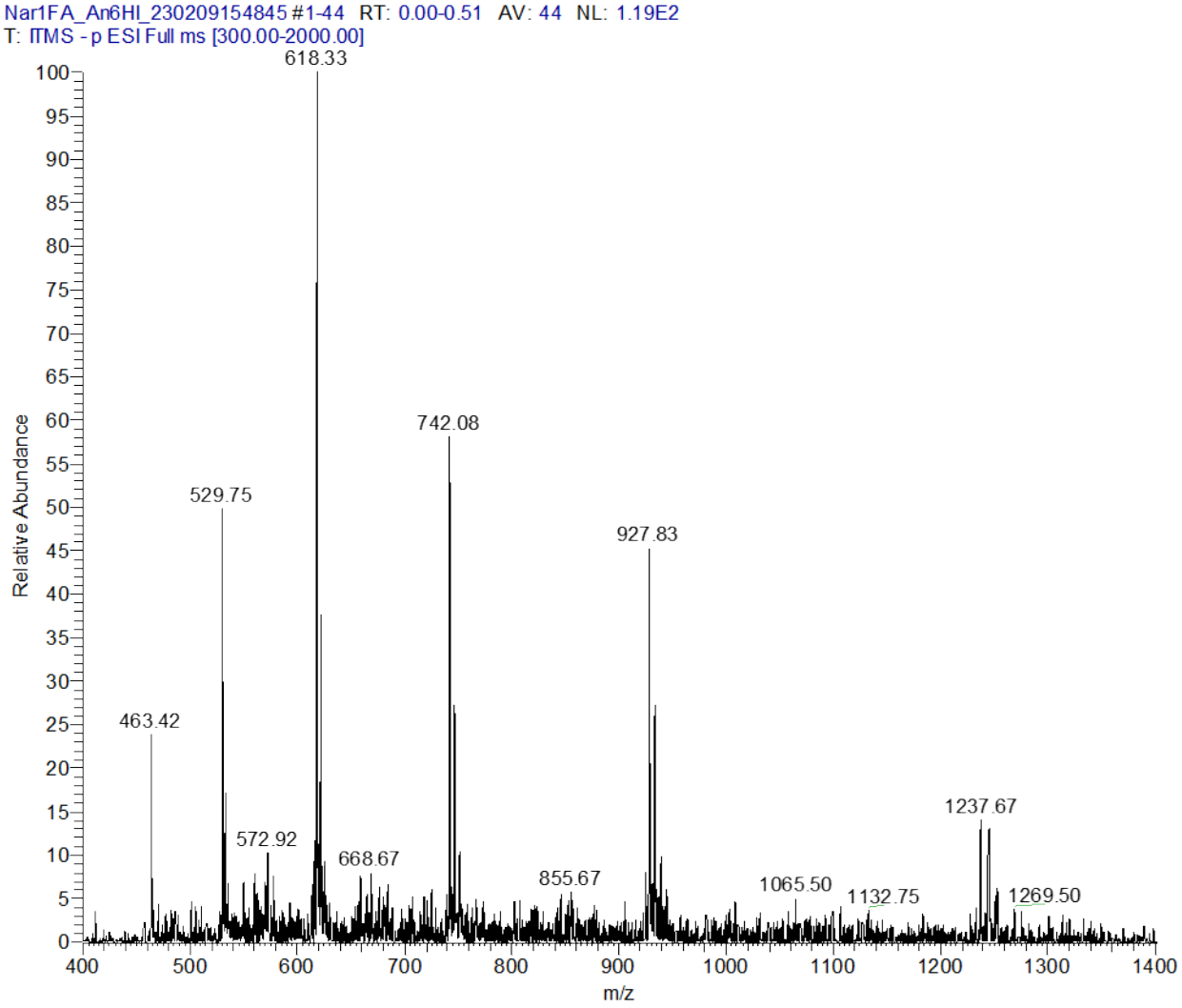


Figure S34. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FA***-An6HI] sequence.

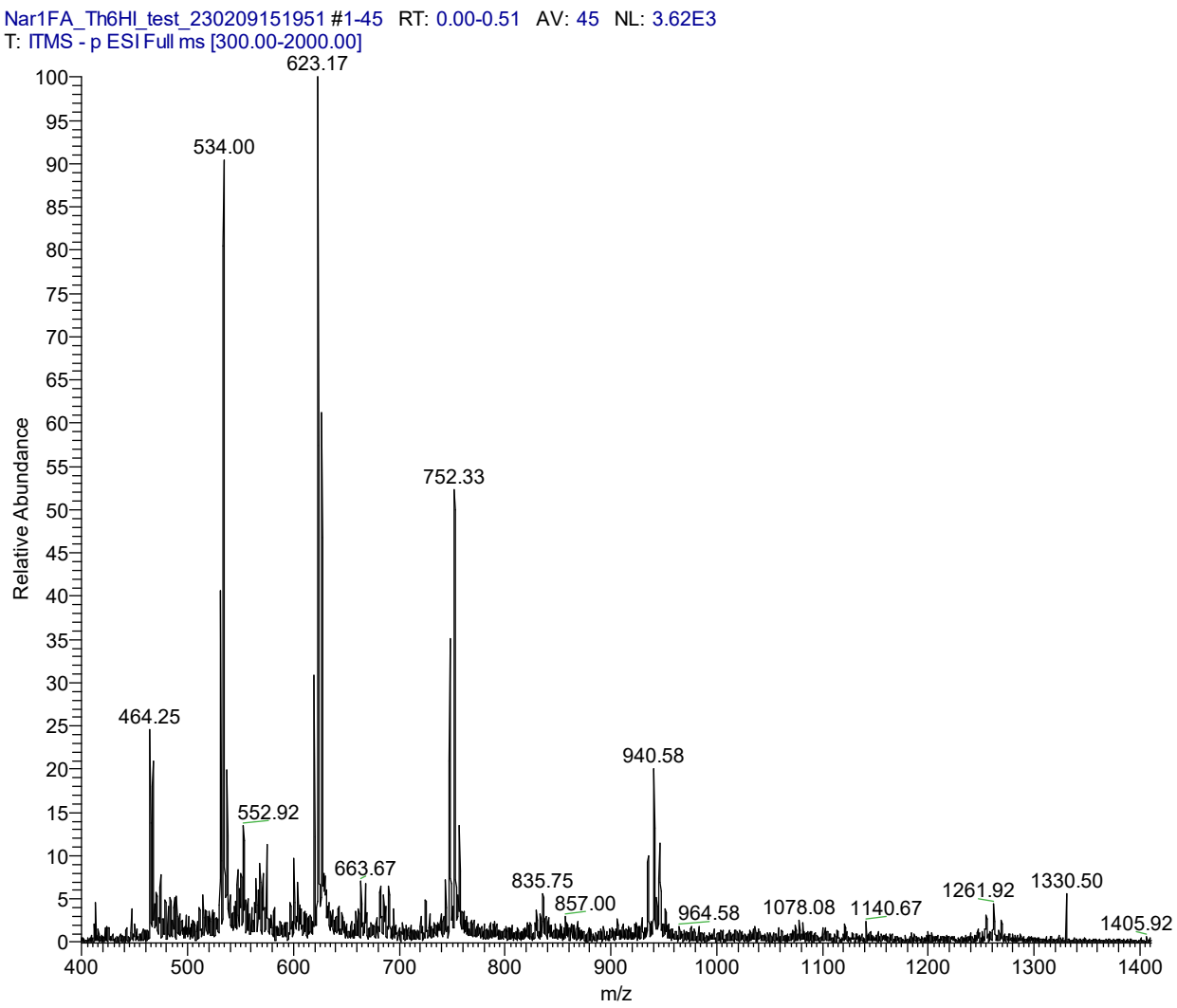


Figure S35. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FA***-Th6HI] sequence.

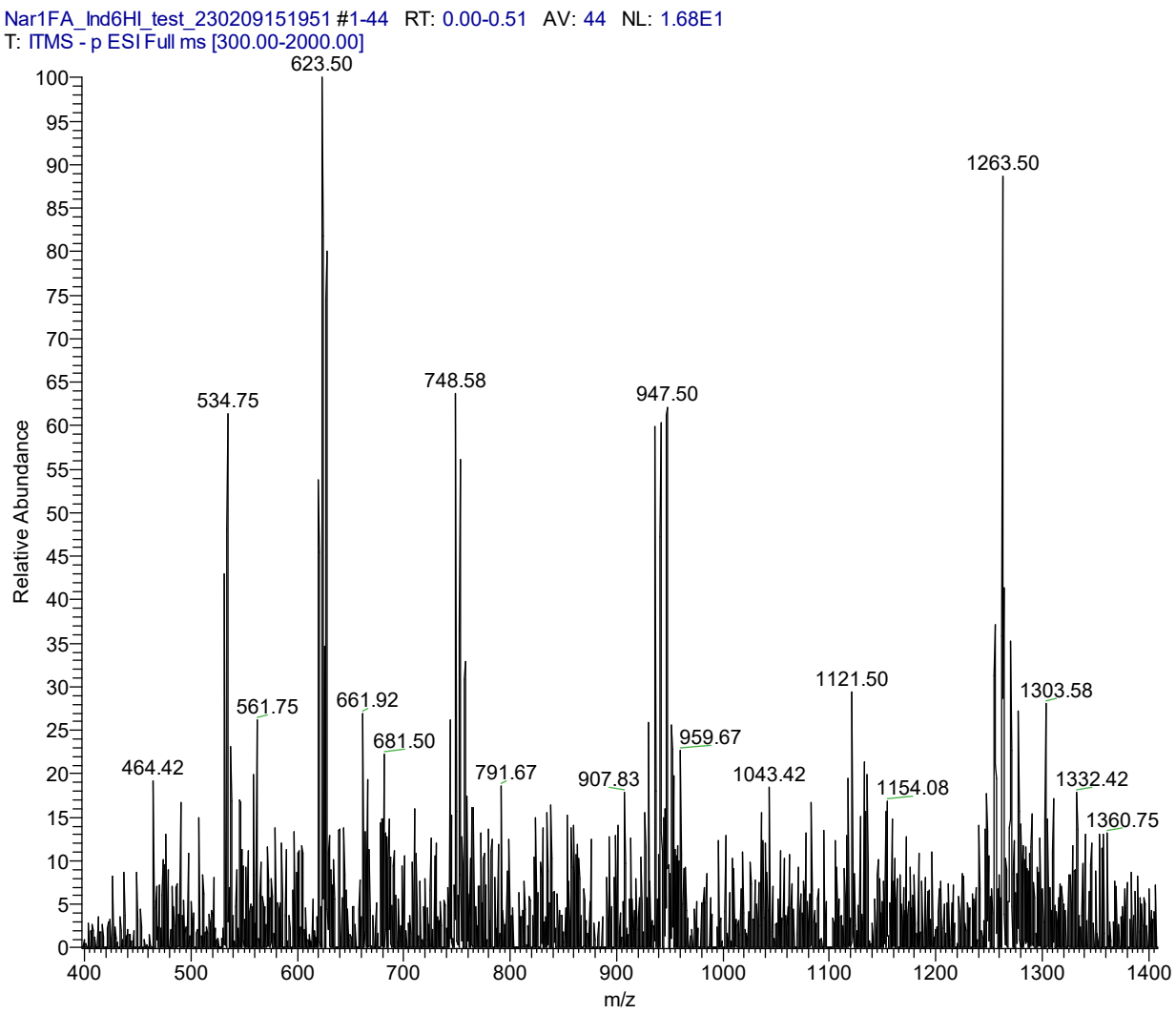


Figure S36. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FA***-Ind6HI] sequence.

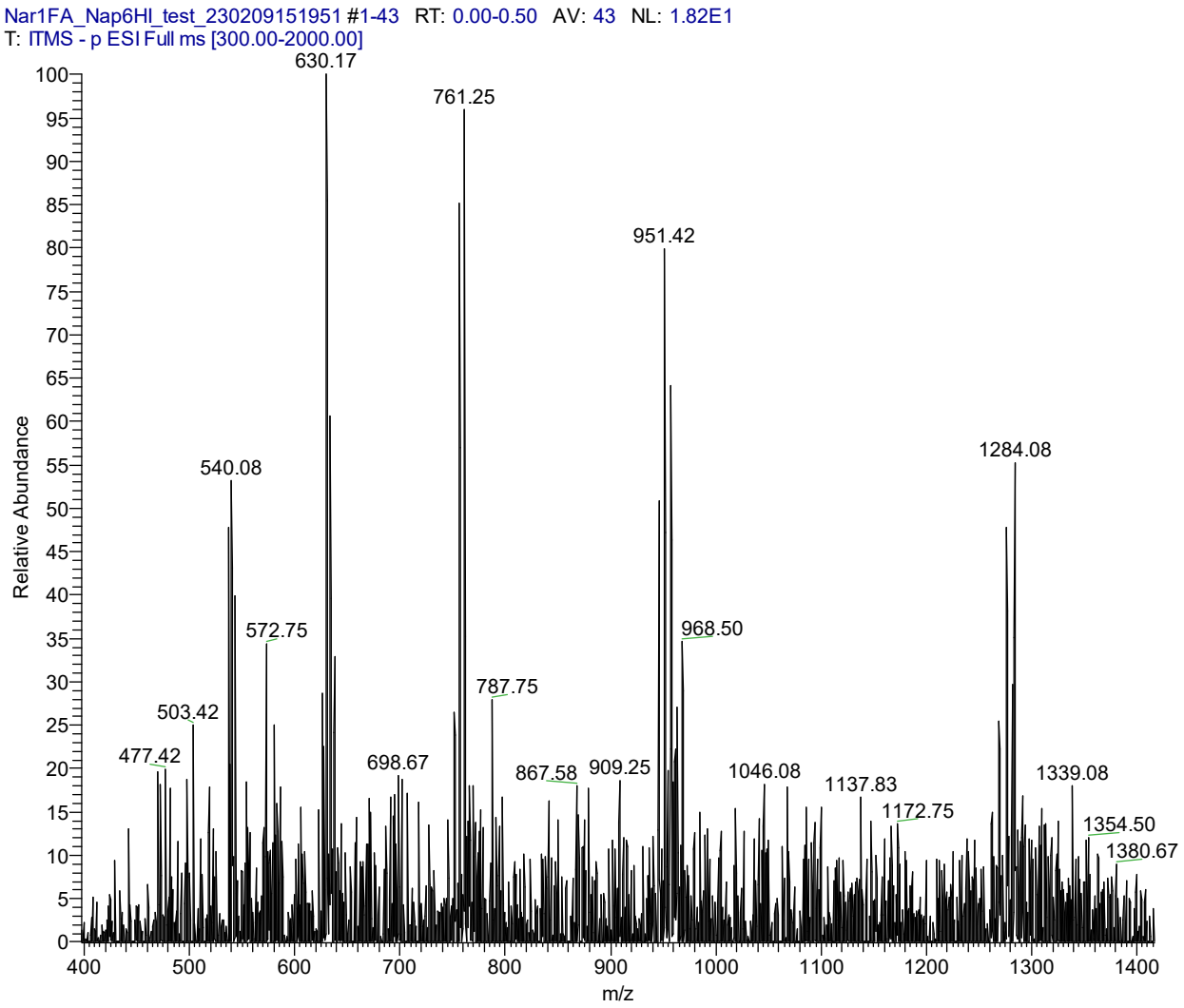


Figure S37. ESI$^-$ MS Spectrum for 12mer*Nar*I [***FA***-Nap6HI] sequence.

ESI$^-$-MS Spectra of Nap6HI Modified 22mer*Nar*I Oligonucleotides.

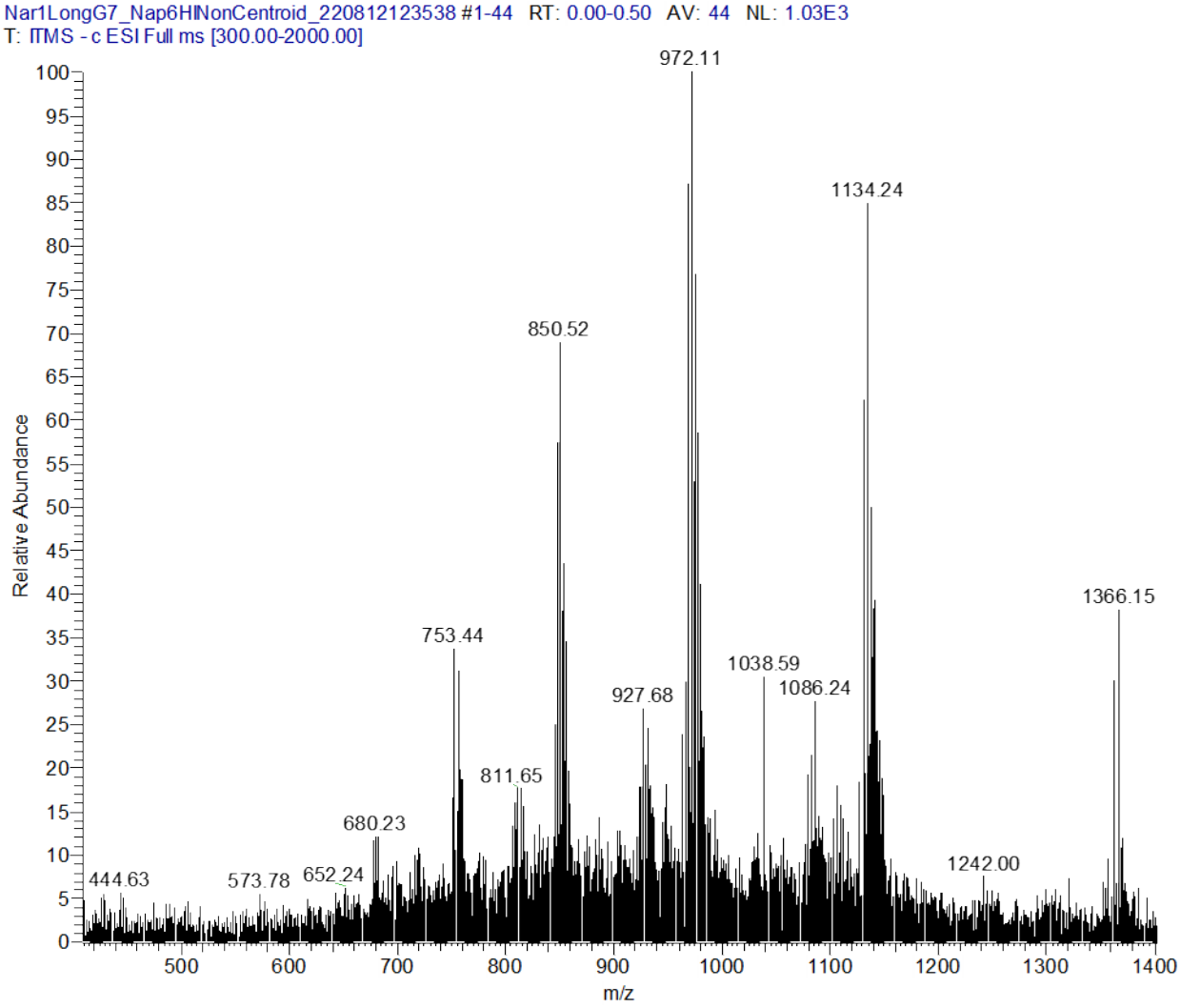


Figure S38. ESI$^-$ MS Spectrum for 22mer*Nar*I[***G7***-Nap6HI] sequence.

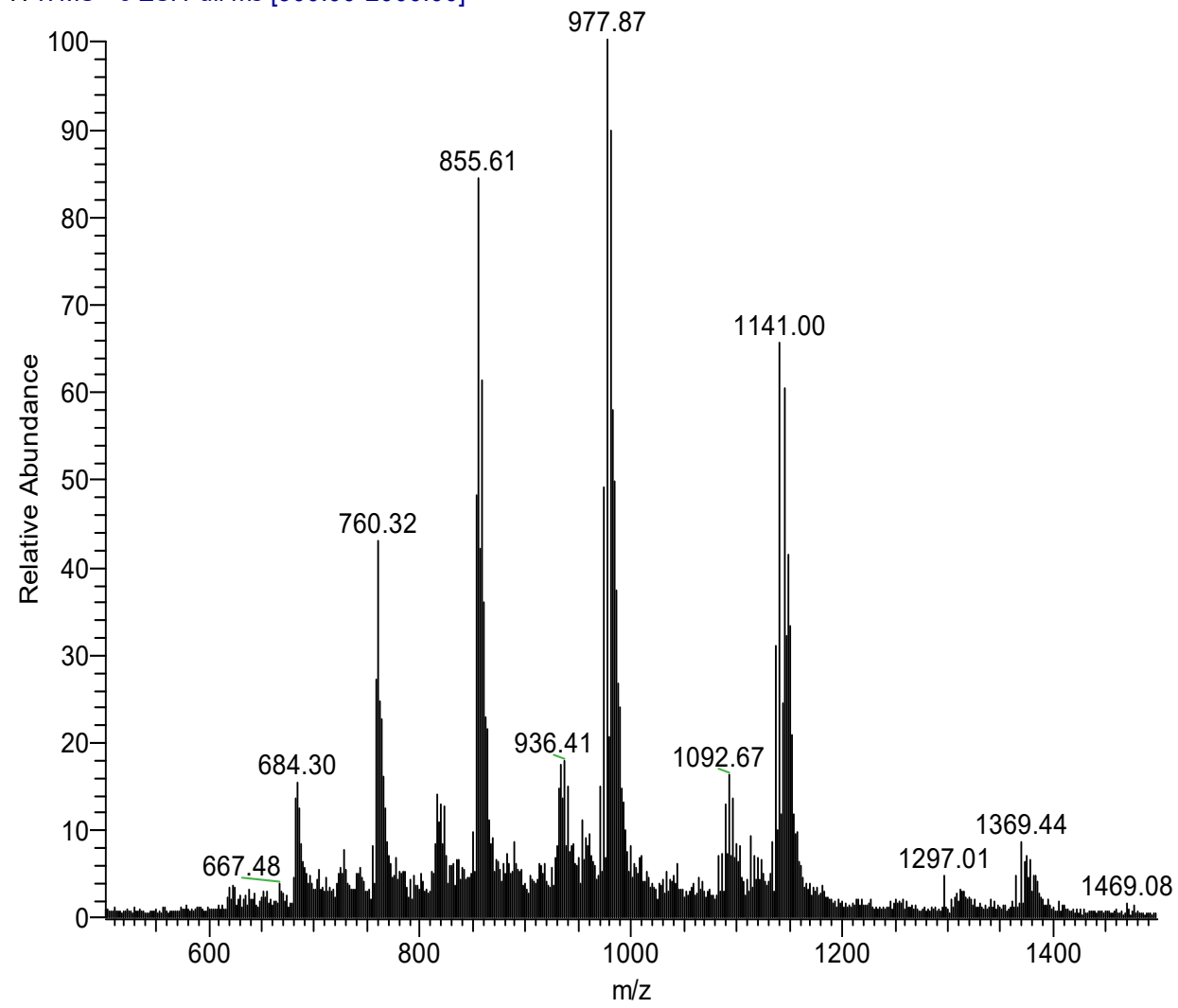


Figure S39. ESI$^-$ MS Spectrum for 22mer*Nar*I[***C13***-Nap6HI] sequence.

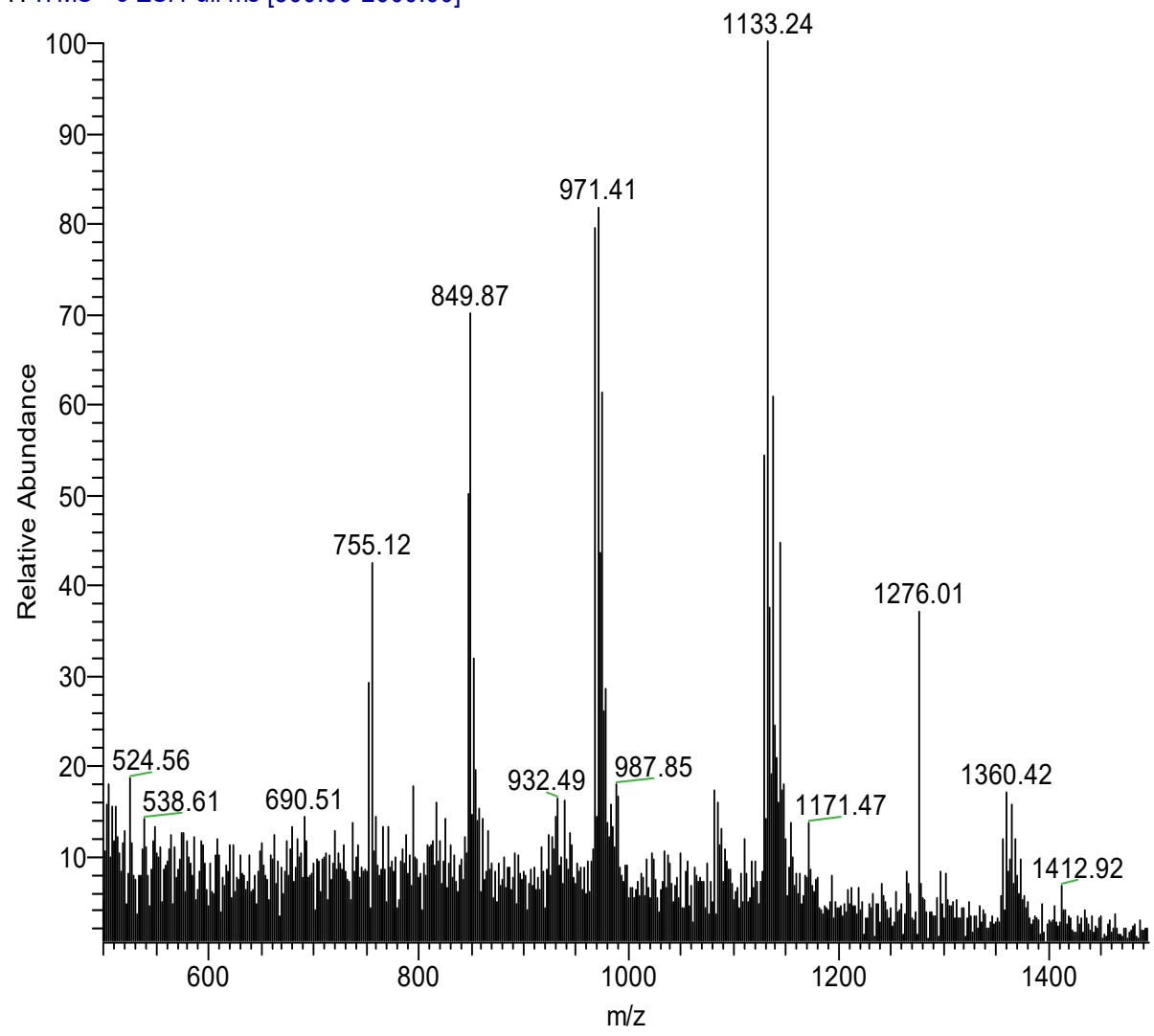


Figure S40. ESI$^-$ MS Spectrum for 22mer*Nar*I[***A17***-Nap6HI] sequence.

ESI$^-$-MS Spectra of Modified 11mer MN4 Split Aptamer labelled with An6HI.

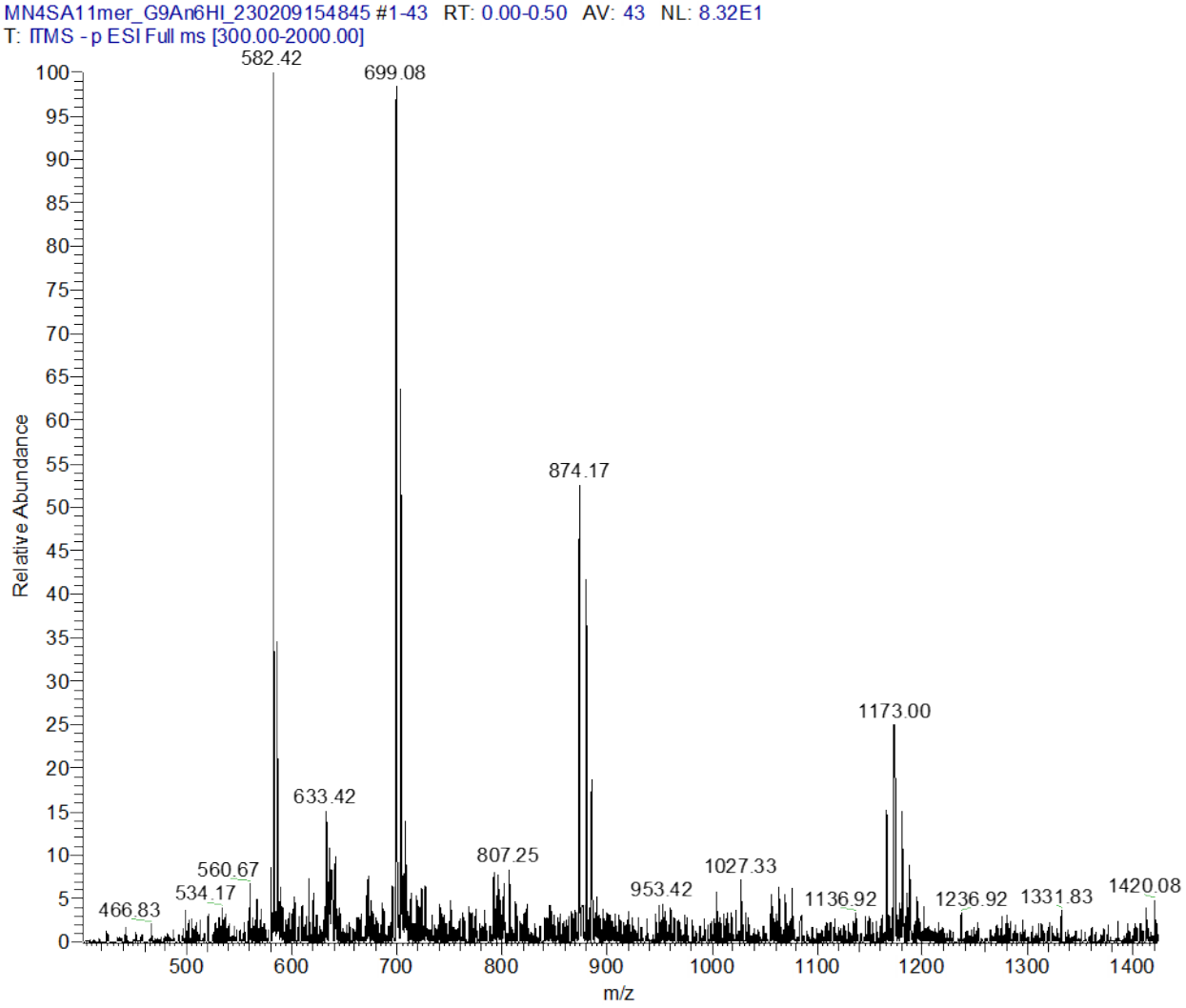


Figure S41. ESI$^-$ MS Spectrum for 11mer MN4 split aptamer [***G9***-An6HI] sequence.

Representative HPLC Chromatograms of Modified *Nar*I Oligonucleotide Reactions Carried Out According to Method 1 (Enamine Catalyst):

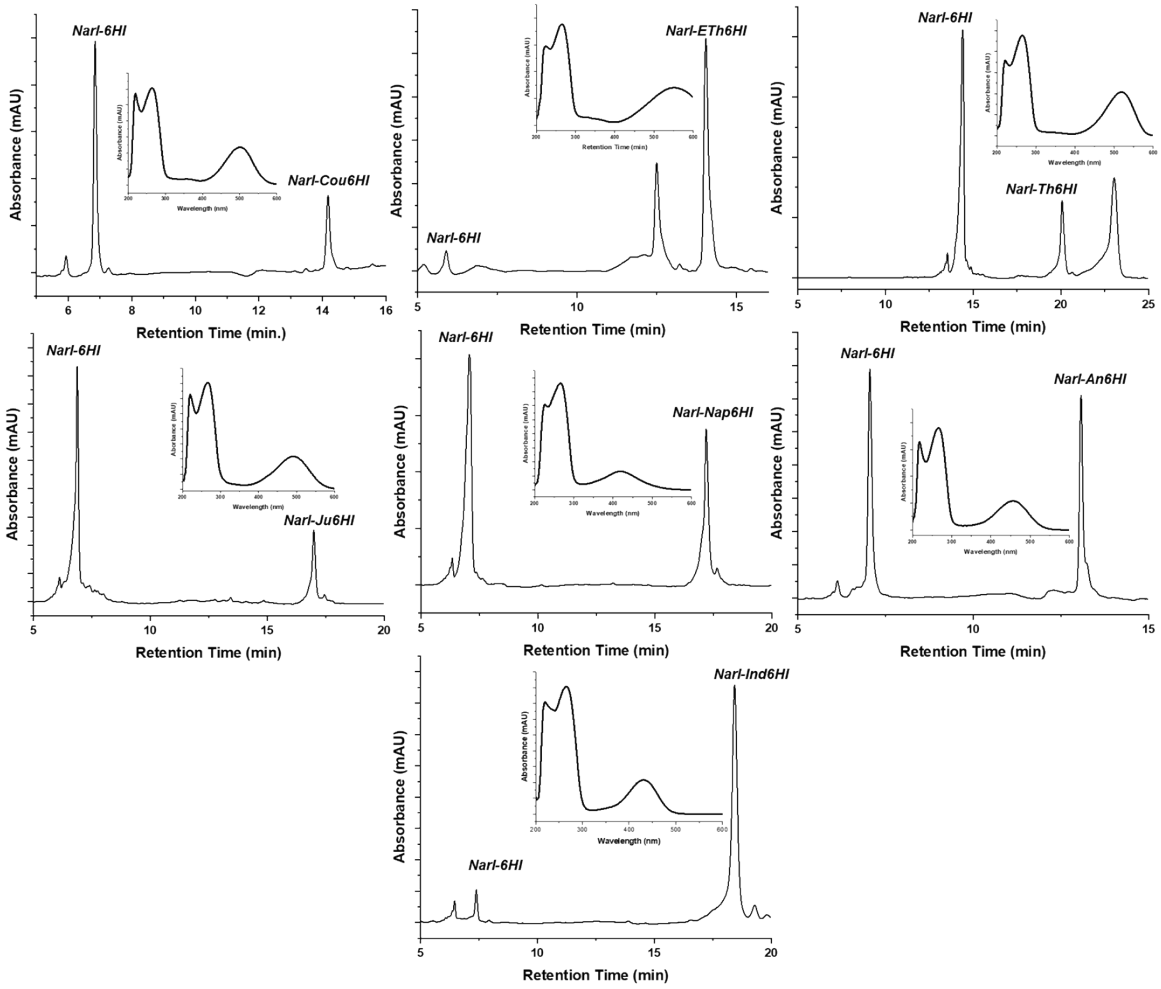


Figure S42. HPLC chromatograms for *Nar*I-6HI reaction mixtures post ethanol precipitation.

Representative HPLC Chromatograms of Modified *Nar*I Oligonucleotide Reactions Carried Out According to Method 2 (Sodium Hydroxide):

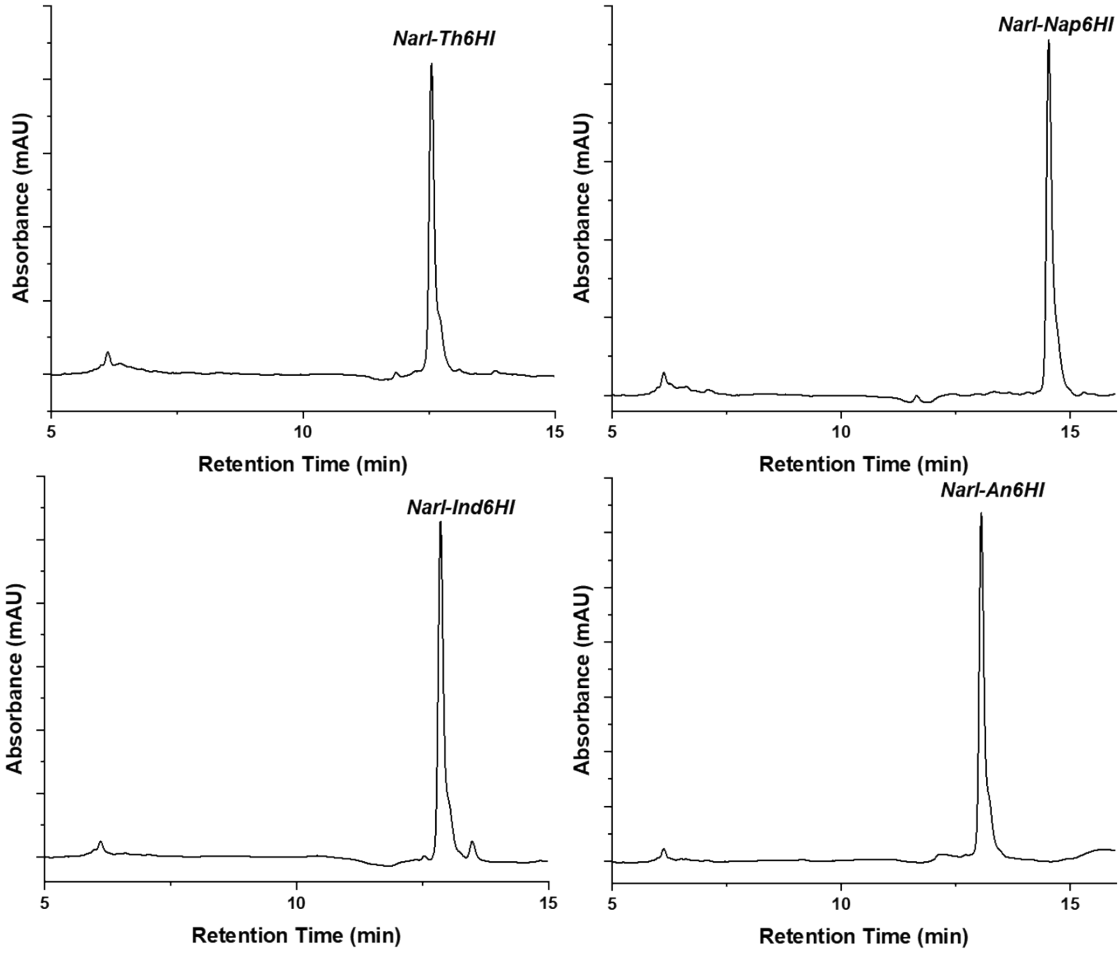


Figure S43. HPLC chromatograms for *Nar*I-6HI reaction mixtures post ethanol precipitation.

Computational Methodology

Model Building.

The canonical *Narl*12 DS (5′–CTCGGC-**G7**-CCATC–3′) starting structure was created using the proto-Nucleic Acid Builder (pNAB)[7] according to B-DNA structural parameters. The THF and the S isomeric forms of the dyes (An6HI and Ind6HI) were built using Gaussview 6.0.16.[8] The *Narl*12 DS containing THF opposite G7 was built by aligning THF with the existing sugar moiety of the 2′-deoxycytidine opposite G7 using the `align` tool in Pymol 2.8[9] followed by deleting the cytosine base from the model. *Narl*12 DS sequences containing dyes opposing C or THF were then built from the corresponding canonical sequence. The glycol adduct of each dye was aligned with the backbone atoms of the G7 nucleotide using the `align` tool of Pymol and G7 was then deleted. Steric clashes between the dye and the flanking base pairs were eliminated by minimally adjusting the dye conformation about the glycol–dye ester bond using '3-button editing mode' in Pymol. In the case of an opposing C, the dyes were minimally displaced toward the major groove to eliminate steric clashes. In total, 6 *Narl*12 DS models were generated, including canonical DNA, the Ind and An chalcone surrogates opposite C, THF complementary to G, and each dye opposite THF.

System Preparation.

Canonical residues were described using the AMBER force field (OL15).[10] To obtain parameters for the modified residues, the THF moiety and all fluorescent dyes with the S stereoisomer of the glycol backbone were first built using Gassview 6.0.16.[8] An initial unguided and exhaustive conformational search was performed on the isolated molecules using RDKit with the UFF force field.[11] The 50 lowest energy conformers were further refined with B3LYP-D3(BJ)/6-31G(d) optimizations performed using Gaussian 16 (B.01).[12] Two structurally distinct minima (conformations) were used for parametrization of THF that correspond to the 3′-endo and 2′-endo 'sugar pucker' conformations, while three conformations of the dyes were used that capture different low energy backbone conformations. The partial charges were calculated using the PyRED online server,[13] while AMBER atom types were supplemented by GAFF[14] parameters.

Each model was prepared for simulation using the `tleap` module of AMBER 22.[15] Specifically, each system was solvated in a truncated octahedral box of TIP4P-EW water,[16] with the nearest box edge being a minimum distance of 10 Å from the solute. This added ~5100 water molecules for each *Narl*12 DS model. Each system was neutralized with Na^+ (20 Na^+ ions). Finally, the system was brought to a NaCl concentration of 150 mM, using the SLTCAP[17] calculator to determine additional ion counts.

Molecular Dynamics Simulation Protocol.

Each system was minimized in four stages: 1) solvent and ions were minimized while holding the solute rigid; 2) the hydrogens on the solute were minimized, while restraining the rest of the model; 3) all atoms of the solute were minimized, while holding the solvent and ions rigid; and 4) the entire system was minimized. Each minimization stage involved 2,000 cycles of steepest descent followed by 2,000 steps of conjugate gradient minimization. At each sub-step, a force constant of 100.0 kcal·mol^{-1}Å^{-2} was used for restraints. Heating was next carried out to bring the systems from 0 K to 310 K in 50 K increments, starting at 10 K and using 20,000 steps of heating per segment. The solute was restrained using a 25 kcal·mol^{-1}·Å^{-2} force constant during heating. Lastly, five equilibration steps were performed, with decreasing restraint on the solute from 20, 15, 10, 5, and finally 1.5 kcal·mol^{-1}·Å^{-2}. Each of the five rounds of equilibration were run for 10,000 steps. The time step for all minimization, heating, and equilibration was 1 fs. During production, the SHAKE[18] algorithm was used, allowing for a time step of 2 fs. The resultant .rst7 files were then used as kernels to start production.

Each of the 6 *Narl*12 DS models were simulated in triplicate for 1 μs, which resulted in 3 μs of total simulation time for every system (18 μs total across the 6 *Narl*12 DS models). For each system containing a dye, additional equilibration in the form of a 300 ns MD simulation was performed to allow the systems to relax. Each simulation was run as an NPT ensemble at 310 K using the Langevin thermostat[19], at 1 bar of pressure, a water density of 0.997 g/mL, and the periodic boundary condition enabled. To balance storage space and sampling frequency, frames were saved to disk every 20 ps for analysis. For a total of 20,000 frames per simulation replica.

Molecular Dynamics Simulation Analysis Protocol.

Simulations were analyzed using `cpptraj` from the AMBER suite and a collection of in-house scripts. Analysis was conducted over the last 1 μs of each replica for each system and reported as averages with standard deviations. Visualizations were created with Pymol 2.8, ChemDraw 22.0, GraphPad Prism 9.5.0, Adobe illustrator 26.5, and Adobe Photoshop 23.5.1. MPED-V AVC (MP4) videos were created by exporting every 20th frame of the trajectory using Pymol. Videos were visually inspected to better understand the structural dynamics prior to applying analysis scripts.

Trajectory analysis included evaluation of the root-mean-square deviations (RMSDs) of all heavy atoms, each residue, and nucleotide backbone atoms (P, O5′, C5′, C4′, C3′, O3′) and per-residue root-mean-square fluctuation (RMSF). Backbone dihedral angles were calculated, including α = ∠($O3'_{5'\text{-base}}$, P, O5′, C5′), β = ∠(P, O5′, C5′, C4′), γ = ∠(O5′, C5′, C4’, C3’), and δ = ∠(C5′, C4′, C3′, O3′). Hydrogen-bonding occupancies were calculated based on a distance cut-off of <3.2 Å and an angle cut-off of >130°. Stacking occupancies were calculated based on the center of mass between the aromatic rings being closer than 5 Å and having an inter-planar angle being <30° or >150°. For each system, hierarchical agglomerative clustering was carried out with respect to the RMSD of the probe, and heavy atoms of the opposing nucleotide/THF moiety and flanking nucleotides. Due to difficulties identifying modified nucleotides with standard tools, helical parameters

were calculated using a pseudo-step between the 5′ and 3′-base pairs flanking the probe using the `vector` and `vectormath` functions in `cpptraj`.

Discussion of MD Simulation Data.

***Narl*12 DS systems**.

Throughout all MD simulation replicas, canonical *Narl*12 DS (5′–CTCGGC-**G7**-CCATC–3′) (Figure S49 A) exhibits B-DNA character consistent with a GC rich DNA sequence (Figure S44). All-atom RMSDs relative to the first production frame are consistently below 2.0 Å throughout each simulation (Figure S45). The central sextet of nucleotides (C6, G7, C8, and their base pair partners) are structurally stable, with average RMSFs ranging from 6.1±0.1 Å to 8.9±0.0 Å, while the terminal ends are considerably more dynamic (average RMSFs of up to 19.1±0.4 Å, Table S10). Regardless, the Watson-Crick hydrogen bonds are long lasting in all base pairs (Table S7). The helical twist in the region including the five base pair steps around G7 is consistent with B-DNA (Table S9). These simulations served as a reference point for the behavior of modified *Narl*12 DS.

In the modified duplexes containing the Ind6HI and An6HI dyes opposite cytosine, the FMRs fully intercalate into the helix and force the opposing C into an extrahelical position. Regardless of incorporation of the dye, the structure of the modified *Narl*12 DS resembles the canonical sequence, including hydrogen bonding in the base pairs flanking both dyes (Table S7). These structural features support the high reported T_m values (Table 1). In the case of Ind6HI, the opposing C is predominantly located in the major groove, forming a long-lasting hydrogen bond between N4 of C and the phosphate of G17 (Figure S44F and S47D). As a result, Ind6HI persistently stacked with the flanking base pairs (Figure S44F S47D). An6HI opposite C also intercalates into the duplex and forces the opposing C into the extrahelical space. However, in contrast to In6HI, An6HI is more flexible owing to steric contacts between the terminal methyl groups the aniline moiety and the deoxyribose of the opposing nucleotide. In fact, the steric bulk of An6HI prevents C from maintaining the same hydrogen bond to G17 as discussed for Ind6HI, resulting in the opposing C adopting both major and minor groove extrahelical positions. Nevertheless, the overall structural dynamics of An6HI is drastically reduced in the *Narl*22 DS relative to the solvent exposed dye due to long-lasting stacking interactions with the flanking bases (Figure 5 and S47). In addition, there is a reduction in the canonical helical twist angle between C6–C8 base pairs by 23.4° and 17.3° for Ind6HI and An6HI, respectively (Table S9). This enhances the stacking interactions with all four flanking bases. For both dyes, the reduced dynamics compared to the free dye and long-lasting stacking of the dyes in *Narl*12 is conducive to fluorescence, providing a structural explanation for the turn-on fluorescence when opposite C.

Upon incorporation of a single THF opposite G7, the furan ring of THF rotates about the backbone axis to direct the hydrophilic O4′ into solvent at the start of all production replicates and this solvent facing orientation is maintained throughout the remainder of

all simulations. This rotation stretches the backbone, increasing the pseudo-base pair step distance between the base pairs flanking G7 (C6–C8) from 7.9±0.4 Å in the canonical duplex to 8.2±0.7 Å. The introduction of THF creates an intrahelical void opposite G7 that was formerly occupied by cytosine, which eliminates stacking interactions with the base pairs flanking the G7 site, as well as three interstrand hydrogen bonds (Table S7). Furthermore, the RMSFs of the nucleotides in the core region are increased slightly (Table S10) and the backbone is periodically distorted from the native B-DNA structure (Figure S48). This local distortion permits two short lived hydrogen bonds, (G7)N2–H···O2P(THF) 7.8% occupancy and (G7)N1–H···O2P(THF) 6.7% occupancy (Figure S48). To accommodate this periodic contraction, the 3′-C base is forced into the major groove and breaks hydrogen bonding with the complementary G17 (Table S7). This local distortion at the G7 site causes a bend in the DNA duplex (Figure S48 B). This structural distortion suggests a single THF moiety strongly disrupts the stability of canonical DNA, which is consistent with a previously reported significant decrease in the T_m upon incorporation of THF into duplexes, albeit with a different sequence context than considered in the present work.[20]

Upon mutation of G7 opposite THF to either Ind6HI or An6HI, the overall helical shape is consistent with B-DNA (Figure S44 C and D) despite the opposing THF maintaining an extrahelical position. As a result, the dye is fully intercalated into the helix and spans the entire interstrand distance (Figure S44) without steric conflicts. Upon incorporation of either probe opposite THF, the dynamics (RMSD and RMSF) of the *Narl*12 DS is restored to that of natural DNA (Figure S45 and Table S10). As a result, unlike the duplex containing THF opposite canonical G7, the hydrogen bonds in the flanking base pairs are maintained (occupancies >90%) despite interstrand hydrogen bonds being absent at the modified G7 position (Table S7). Nevertheless, there is a reduction in the canonical helical twist angle (specifically, the pseudo-twist calculated between C6–C8 base pairs by 22.2° and 19.0° for Ind6HI and An6HI, respectively). This reduction in helical twist enhances the stacking between the intercalated probe and the flanking base pairs (Figure 5). Despite the anticipation of the probes being highly dynamic in solution, the persistent intrahelical stacking interactions restrict the conformational freedom of the probe (Figure S46). As a result, the dihedral angle about the bridge between the donor and acceptor rings reflects a near planar probe, being –7.2°±17.5° for Ind6HI and –15.4°±18.5° for An6HI (Figure S46). The enforced inter-ring planarity coupled with the persistent π–π stacking interactions rationalizes the turn-on capabilities of the chalcone surrogates containing Ind or An donors.

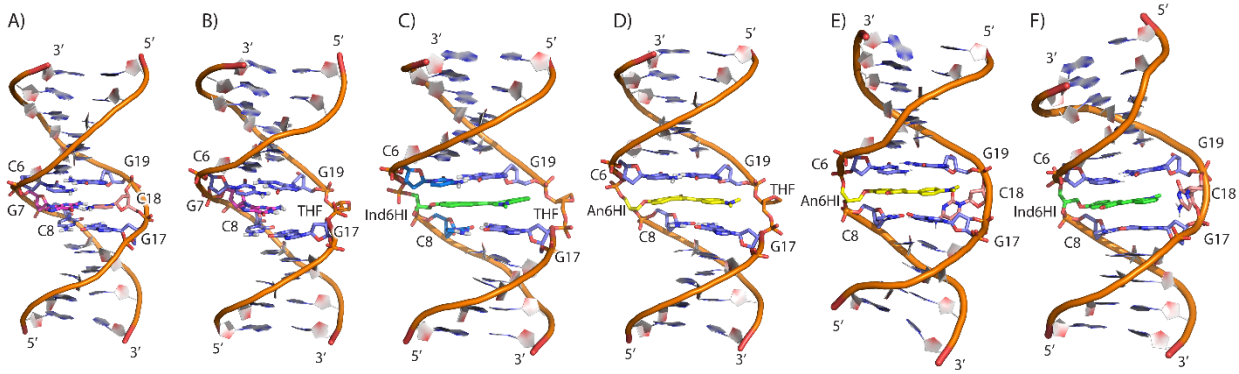


Figure S44. Representative MD structures of *Nar*I12 containing A) canonical nucleotides [G7 (magenta) opposite C18 (tan)], B) THF (orange) opposite G7, C) Ind6HI (green) at G7 opposite THF, and D) An6HI (yellow) at G7 opposite THF. E) An6HI (yellow) at G7 opposite canonical C18. F) Ind6HI (green) at G7 opposite canonical C18. Sequences shown as cartoon for residues distal to the modification site and stick representation for the G7:C18 and flanking base pairs. Water, ions, and non-polar hydrogens hidden for clarity.

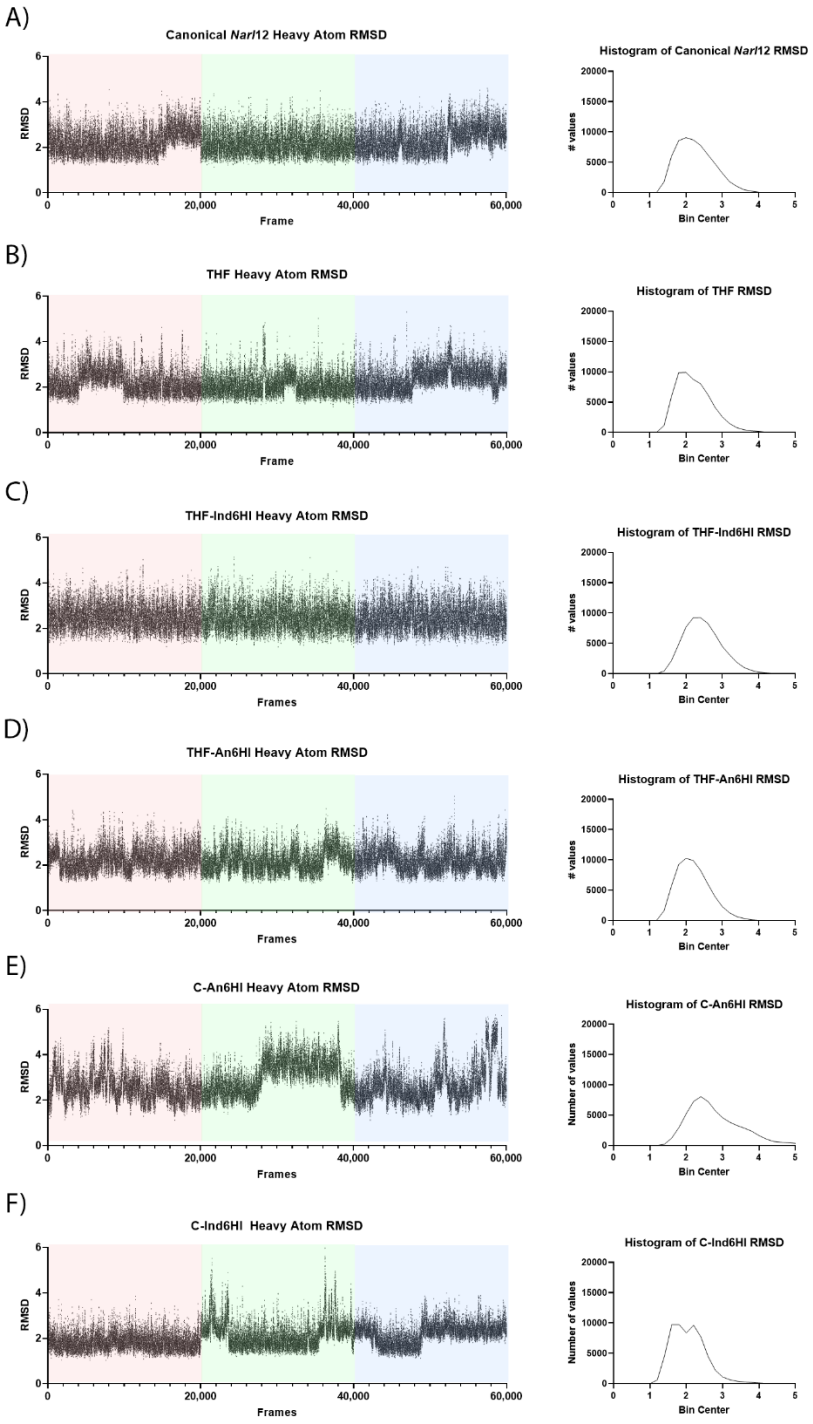


Figure S45. All heavy atom RMSDs across the three MD production simulation replicas. RMSD vs time shown left, colored regions represent replica 1 (pink), replica 2 (green), and replica 3 (blue). Histograms (right) for *Narl*12 DS containing A) canonical nucleotides, B) THF opposite G7, C) Ind6HI opposite THF, D) An6HI opposite THF. E) Cytosine opposite An6HI and F) Cytosine opposite Ind6HI.

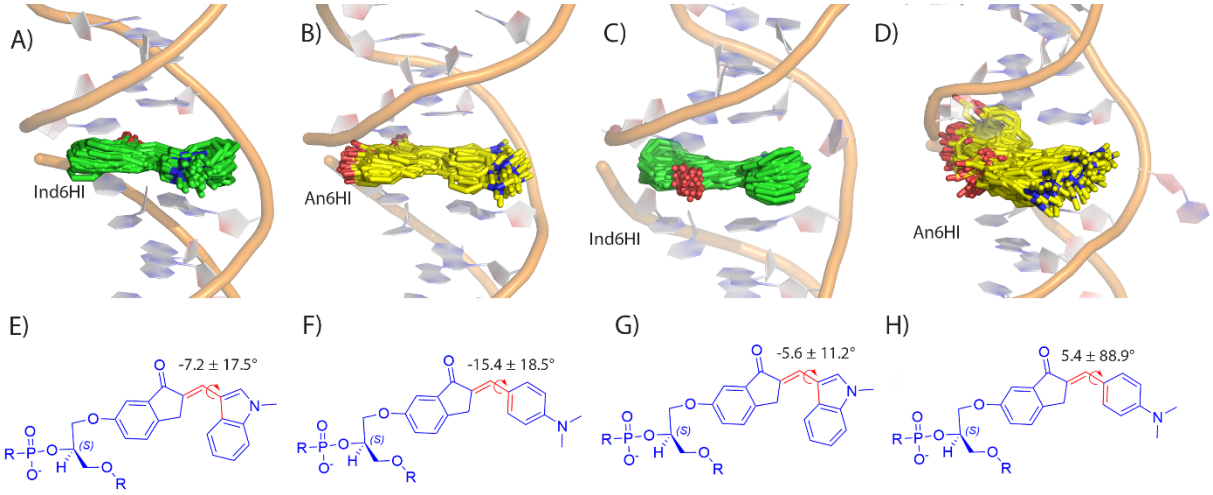


Figure S46. 100 frames overlays from all replicas, taken every 30 ns. Models of *Narl*12 DS containing A) Ind6HI (green) opposite THF. B) An6HI (yellow) at G7 opposite THF. C) Ind6HI opposite C, D) An6HI opposite C. Overlays were aligned based on the two base pairs flanking the probe. A single representative structure of the full *Narl*12 sequence in cartoon representation was used for context. Water, ions, and non-polar hydrogens hidden for clarity. Average dihedral angle (and standard deviation) across MD production replicas about the donor–acceptor bridge of E) Ind6HI opposite THF. F) An6HI opposite THF. G) Ind6HI opposite C. H) An6HI opposite C.

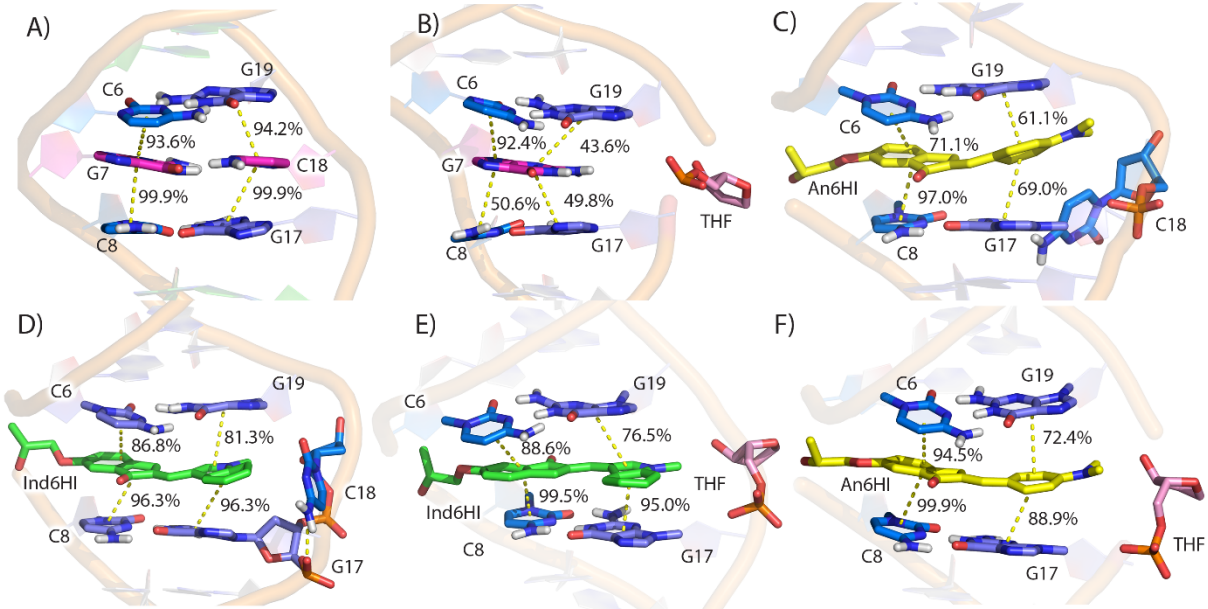


Figure S47. Representative MD structures of *Narl*12 DS containing A) Canonical G opposite C (magenta), B) canonical G opposite THF (pink), C) An6HI (yellow) opposite C (light blue), D) Ind6HI (green) opposite canonical C (light blue), E) Ind6HI (green) opposite THF (pink), and F) An6HI (yellow) opposite THF (pink). Base–base stacking

occupancies are provided. All images are side views from the major groove. Water, ions and nonpolar hydrogen atoms are hidden for clarity.

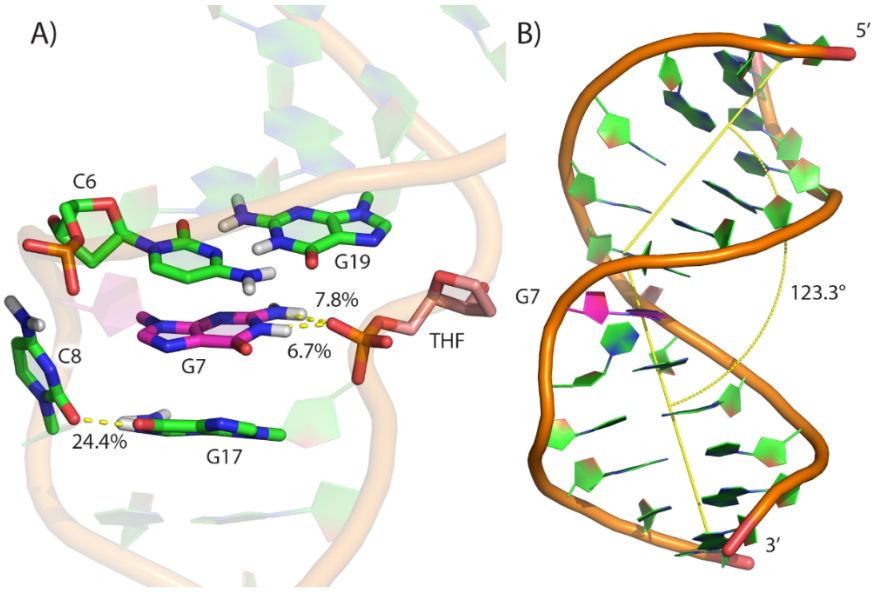


Figure S48. Representative MD structure of *Narl*12 DS containing THF opposite G7, highlighting A) periodic hydrogen bonding between G7 and the non-bridging phosphate oxygen of THF and B) the bend in the modified duplex.

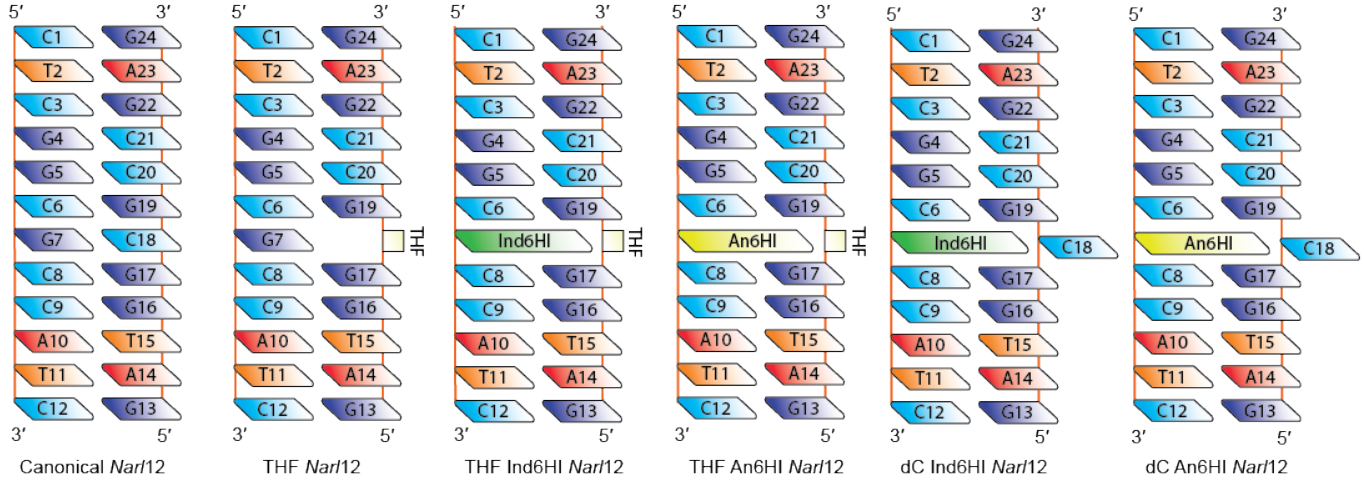


Figure S49. Schematics of the *Narl*12 DS systems considered in the present study, including residue numbering

Table S7. Watson-Crick hydrogen-bonding occupancies (%), average distance (Å) and average angle (deg.) in *Narl*12 DS containing canonical nucleotides, THF opposite G7, Ind6HI at G7 or An6HI at G7.[a,b]

		Canonical			THF			THF-Ind6HI			THF-An6HI			C-Ind6HI			C-An6HI		
Nucleobases	**Interactions**	**Occ**	**Dist**	**Ang**	**Occ**	**Dist**	**Ang**	**Occ**	**Dist**	**Ang**	**Occ**	**Dist**	**Ang**	**Occ**	**Dist**	**Ang**	**Occ**	**Dist**	**Ang**
G5 – C20	N2–H···O2	98 %	2.88	162.4	98 %	2.88	161.9	98 %	2.87	162.4	97 %	2.88	162.5	99 %	2.90	162.4	98 %	2.90	162.4
	N1–H···N3	99 %	2.95	165.6	99 %	2.95	165.6	99 %	2.95	165.6	99 %	2.95	165.6	99 %	2.95	165.7	99 %	2.97	165.5
	O6···H–N4	93 %	2.92	161.6	93 %	2.91	161.8	93 %	2.92	162.1	94 %	2.91	161.4	95 %	2.93	160.6	96 %	2.93	161.3
C6 – G19	O2···H–N2	99 %	2.86	163.6	98 %	2.87	162.6	98 %	2.87	163.9	98 %	2.87	163.5	99 %	2.90	163.6	99 %	2.88	163.5
	N3···H–N1	99 %	2.94	165.0	99 %	2.94	164.3	99 %	2.96	165.3	99 %	2.95	164.9	99 %	2.96	164.6	99 %	2.95	164.3
	N4–H···O6	95 %	2.91	162.5	95 %	2.91	163.7	95 %	2.91	163.4	94 %	2.92	163.4	96 %	2.93	160.9	96 %	2.92	162.3
G7 – C18	N2–H···O2	99 %	2.86	163.6	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	N1–H···N3	99 %	2.94	165.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
	O6···H–N4	94 %	2.91	162.3	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
C8 – G17	O2···H–N2	98 %	2.88	162.4	75 %	2.87	160.6	98 %	2.88	162.3	98 %	2.87	163.0	98 %	2.90	160.6	98 %	2.89	160.9
	N3···H–N1	99 %	2.95	165.5	70 %	2.95	164.0	98 %	2.96	164.9	98 %	2.96	164.2	99 %	2.97	164.9	99 %	2.96	164.9
	N4–H···O6	93 %	2.92	161.6	65 %	2.92	162.1	90 %	2.93	162.1	91 %	2.93	162.2	96 %	2.93	161.2	96 %	2.93	162.4
C9 – G16	O2···H–N2	99 %	2.86	163.5	98 %	2.86	163.3	99 %	2.86	163.4	99 %	2.86	163.4	99 %	2.86	163.1	99 %	2.86	163.4

	N3···H–N1	99 %	2.95	165.2	99 %	2.95	164.9	99 %	2.95	164.9	99 %	2.94	165.2	99 %	2.95	164.9	99 %	2.95	165.0
	N4–H···O6	93 %	2.92	162.3	94 %	2.92	162.6	93 %	2.92	162.8	93 %	2.92	162.5	97 %	2.93	162.6	97 %	2.93	162.5

[a] Hydrogen-bond occupancy was evaluated over all MD production replicas based on a geometric cutoff for donor–acceptor heavy atom distance less than 3.2 Å and a donor–donor-hydrogen–acceptor angle greater than 130°. [b]En-dash represents the lack of a particular interactions (i.e., due to G7 mutation)

Table S8. Average (standard deviation in parentheses, Å) nucleobase–nucleobase vertical separation in *Narl*12 DS.[a]

	Nucleobase–Nucleobase Pair					
	Canonical	**THF**	**THF-Ind6Hl**	**THF-An6Hl**	**C-Ind6Hl**	**C-An6Hl**
G5–C6	3.7(0.2)	3.7(0.2)	3.7(0.2)	3.8(0.3)	3.7(0.2)	3.8(0.3)
C6–C8[b]	7.9(0.4)	8.2(0.7)	7.6(0.4)	7.7(0.4)	7.6(0.5)	7.9(0.6)
C8–C9	3.9(0.3)	4.8(1.2)	4.1(0.3)	4.0(0.3)	4.1(0.3)	4.0(0.4)
C20–G19	3.8(0.3)	3.7(0.2)	3.9(0.3)	3.8(0.3)	3.9(0.3)	3.9(0.3)
G19–G17[c]	7.5(0.3)	8.0(0.8)	7.7(0.4)	7.8(0.5)	7.6(0.5)	7.9(0.7)
G17–G16	4.2(0.3)	3.9(0.3)	4.0(0.3)	4.1(0.3)	4.0(0.3)	4.1(0.3)

[a] Pseudo base step rise evaluated over all MD production replicas. Base pair rise defined as the distance between the center of masses of stacked bases. [b] These nucleobases span three residues centered on the probe. [c] These nucleobases spans three residues centered on THF.

Table S9. Average (standard deviation in parentheses, deg.) helical twist in *Narl*12 DS calculated using pseudo base pair steps.[a,b]

Base Span	Canonical	THF	THF-Ind6HI	THF-An6HI	C-Ind6HI	C-An6HI
5′-1 to 5′[b]	34.2(5.9)	34.6(4.4)	33.9(4.3)	35.2(4.6)	36.4(4.5)	34.6(6.8)
5′ to 3′[c]	71.3(6.8)	65.1(13.6)	49.1(7.0)	52.3(12.4)	47.9(8.7)	54.0(13.8)
3′ to 3′+1[d]	35.7(5.1)	36.7(8.3)	34.3(5.5)	35.4(5.3)	34.3(5.5)	33.1(6.7)

[a] The helical twist was evaluated over all MD production replicas. Base pair vectors were defined using the C1′ atoms of each base in the pair. The helical twist was calculated as the cross angle between the base pair vectors using `vectormath`. [b] 5′-1 to 5′ is defined as the G5–C6 containing base pair step. [c] 5′ to 3′ is defined as the base pairs flanking G7 or the probe. [d] 3′ to 3′+1 is defined as the C8–C9 containing base pair step.

Table S10. Average per-residue RMSF (standard deviation in parentheses) for Narl12 DS calculated over all MD production simulation replicas.

Residue	Canonical	THF	THF-Ind6HI	THF-An6HI	C-Ind6HI	C-An6HI
C1	18.7(0.5)	18.5(0.2)	18.6(0.6)	19.2(0.3)	19.5(0.1)	19.3(0.4)
T2	16.2(0.5)	15.9(0.2)	16.0(0.5)	16.5(0.2)	16.8(0.2)	16.6(0.4)
C3	13.4(0.3)	13.0(0.1)	13.3(0.4)	13.6(0.1)	13.7(0.2)	13.5(0.4)
G4	10.4(0.0)	10.2(0.1)	10.6(0.4)	10.6(0.2)	10.5(0.2)	10.3(0.5)
G5	8.4(0.1)	8.3(0.2)	8.8(0.2)	8.6(0.2)	8.1(0.2)	8.1(0.5)
C6	7.6(0.1)	7.6(0.4)	7.9(0.1)	7.8(0.2)	6.8(0.3)	7.1(0.4)
G7[a]	7.1(0.1)	6.5(0.5)	6.1(0.1)	6.5(0.1)	6.0(0.2)	6.5(0.1)
C8	8.7(0.1)	8.9(0.0)	8.1(0.0)	8.4(0.0)	8.8(0.1)	8.8(0.1)
C9	10.6(0.0)	10.8(0.1)	10.3(0.3)	10.4(0.1)	10.9(0.2)	10.5(0.2)
A10	12.8(0.1)	12.6(0.0)	12.8(0.5)	12.9(0.2)	13.0(0.4)	12.6(0.7)
T11	15.6(0.1)	15.1(0.1)	15.6(0.6)	15.7(0.4)	15.7(0.4)	15.4(0.9)
C12	18.4(0.4)	18.0(0.1)	18.6(0.8)	18.8(0.6)	18.6(0.4)	18.8(0.9)
G13	18.5(0.4)	18.6(0.1)	18.0(0.6)	18.6(0.4)	18.7(0.1)	18.7(0.4)
A14	16.3(0.3)	16.3(0.1)	15.8(0.4)	16.3(0.5)	16.2(0.2)	16.5(0.3)
T15	13.4(0.2)	13.3(0.2)	13.3(0.4)	13.5(0.5)	13.2(0.2)	13.8(0.4)
G16	10.2(0.1)	9.9(0.2)	10.7(0.3)	10.6(0.4)	10.1(0.1)	10.6(0.5)
G17	7.6(0.1)	7.0(0.2)	8.4(0.2)	8.2(0.2)	7.6(0.3)	7.9(0.5)
C18[b]	6.3(0.1)	8.6(0.3)	11.2(0.1)	10.4(0.3)	9.6(0.7)	9.9(0.4)
G19	6.0(0.1)	6.3(0.2)	6.1(0.1)	6.0(0.0)	6.8(0.3)	6.5(0.1)
C20	8.2(0.1)	8.5(0.1)	8.2(0.2)	8.1(0.2)	8.6(0.1)	8.4(0.2)
C21	10.7(0.0)	10.8(0.1)	10.8(0.3)	10.7(0.4)	10.8(0.2)	10.7(0.3)
G22	13.3(0.1)	13.4(0.1)	13.4(0.5)	13.4(0.4)	13.3(0.3)	13.2(0.4)
A23	16.4(0.1)	16.2(0.1)	16.4(0.7)	16.5(0.5)	16.2(0.4)	16.1(0.7)
G24	19.1(0.2)	18.7(0.1)	18.9(0.8)	19.1(0.4)	19(0.4)	18.9(1.1)

[a] G7 is mutated to either Ind6HI or An6HI as indicated. [b] C18 is mutated to THF in systems indicated.

References:

[1] X. Lv, J. Liu, Y. Liu, Y. Zhao, Y.-Q. Sun, P. Wang, W. Guo, *Chemical Communications.* **2011**, *47*, 12843.

[2] H. Watanabe, M. Ono, H. Saji, *Chemical Communications.* **2015**, *51*, 17124–17127.

[3] O. V. Vinogradova, E. A. Filatova, N. V. Vistorobskii, A. F. Pozharskii, I. V. Borovlev, Z. A. Starikova, *Russian Journal of Organic Chemistry.* **2006**, *42*, 338–348.

[4] J. H. Brannon, D. Magde, *Journal of Physical Chemistry.* **1978**, *82*, 705–709.

[5] T. Karstens, K. Kobs, *Journal of Physical Chemistry.* **1980**, *84*, 1871–1872.

[6] P. M. Punt, G. H. Clever, *Chemical Science.* **2019**, *10*, 2513–2518.

[7] A. Alenaizan, J. L. Barnett, N. V. Hud, C. D. Sherrill, A. S. Petrov, *Nucleic Acids Research.* **2021**, *49*, 79-89.

[8] R. D. a. T. A. K. a. J. M. Millam, **2019**.

[9] W. L. DeLano, *CCP4 Newsletter on protein crystallography* **2002**, *40*, 82-92.

[10] R. Galindo-Murillo, J. C. Robertson, M. Zgarbová, J. Šponer, M. Otyepka, *Journal of chemical theory and computation.* **2016**, *12*, 4114-4127.

[11] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, W. M. Skiff, *Journal of the American Chemical Society.* **1992**, *114*, 10024-10035.

[12] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, Wallingford, CT, **2016**.

[13] E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, *Nucleic Acids Research.* **2011**, *39*, W511-W517.

[14] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *Journal of Computational Chemistry.* **2004**, *25*, 1157-1174.

[15] D. Case, I. Ben-Shalom, S. Brozell, D. Cerutti, T. Cheatham III, *University of California, San Francisco*.

[16] T. Peng, T.-M. Chang, X. Sun, A. V. Nguyen, L. X. Dang, *Journal of Molecular Liquids.* **2012**, *173*, 47-54.

[17] J. D. Schmit, N. L. Kariyawasam, V. Needham, P. E. Smith, *Journal of Chemical Theory and Computation.* **2018**, *14*, 1823-1827.

[18] V. Kräutler, W. F. van Gunsteren, P. H. Hünenberger, *Journal of Computational Chemistry.* **2001**, *22*, 501-508.

[19] R. L. Davidchack, R. Handel, M. Tretyakov, *The Journal of Chemical Physics.* **2009**, *130*, 234101.

[20] T. J. Matray, E. T. Kool, *Journal of the American Chemical Society.* **1998**, *120*, 6191-6192.