Supporting Information

for "Facile Photocyclization Chemistry of 5-Phenylthio-2'-deoxyuridine in Duplex DNA" by

Yu Zeng, Huachuan Cao and Yinsheng Wang* (*Org. Lett.*, Vol. 8, 2006, ol060659v)

Table of Contents

Experimental Procedures

Materials

All chemicals were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise specified. Calf intestinal alkaline phosphatase was from New England Biolabs (Beverly, MA). Snake venom phosphodiesterase, calf spleen phosphodiesterase and nuclease P1 were obtained from US Biological (Swampscott, MA). The reagents used for solid-phase DNA synthesis were purchased from Glen Research Inc. (Sterling, VA). Dinucleoside monophosphates and ODNs containing a 5-phenylthio-2'-deoxyuridine were synthesized on a Beckman Oligo 1000S DNA synthesizer (Fullerton, CA). The nucleobase deprotection was carried out in 29% ammonia at room temperature for 48 hrs.

Synthesis and Characterization of Compounds (Scheme 1)

Preparation of 2 (Scheme 1). To a 50-mL flask were added freshly cut sodium (345 mg, 15 mmol) and thiophenol (11 mL). The solution was stirred at room temperature under Ar atmosphere until all the Na was dissolved. A solution of DMSO (7.33 mL) containing 5-bromo-2'-deoxyuridine (921 mg, 3 mmol) was slowly added to the stirred solution over 15 min. The solution was stirred for another 15 min, and the flask was immersed in an oil bath (65°C) for 20 h during which the solution was continuously stirred. The solution was then cooled to room temperature, and white precipitate was formed. The precipitate was dried under reduced pressure, and the resulting product was dissolved in a 20-mL solution containing CH_3OH and CH_2Cl_2 (1:1, v/v) and taken up with silica gel (10 g). The slurry was then dried under reduced pressure, and the resulting silica gel was loaded onto the top of a silica gel column and the desired compound was eluted from the column with a step gradient of 0-15% CH_3OH in CH_2Cl_2. The appropriate fractions were pooled, and the solvent was evaporated to give white powder

(479 mg, 47.5%). Compound **2**: ESI-MS: *m/z* 336.9 $[M + H]^+$, 358.9 $[M + Na]^+$. ^1H NMR (DMSO-d_6): δ 11.65 (s, 1H, NH), 8.43 (s, 1H, H-6), 7.21-7.32 (m, 5H, H-PhS), 6.17 (t, 1H, H-1'), 5.26 (s, 1H, 3'-OH), 5.08 (s, 1H, 5'-OH), 4.23 (m, 1H, H-3'), 3.81 (m, 1H, H-4'), 3.58 (m, 2H, H-5' and H-5''), 2.10 (m, 2H, H-2' and H-2'').

Preparation of 3 (Scheme 1). Compound **2** (672 mg, 2 mmol) was evaporated twice with dry pyridine (5 mL) and redissolved in a fresh portion of dry pyridine (13 mL), to which solution were quickly added 4,4'-dimethoxytrityl chloride (DMTrCl) (814 mg, 2.4 mmol) and 4-(dimethylamino)pyridine (DMAP) (13 mg, 0.11 mmol). To the resulting mixture was then slowly added TEA (0.46 mL, 3.30 mmol), and the solution was stirred at room temperature for 3 h. After the reaction was shown to be complete by TLC (95/5, $CHCl_3/CH_3OH$, v/v), the solution was cooled to 5°C and methanol (2 mL) was added. The solvent was then removed *in vacuo* and the residue was loaded onto a silica gel column and eluted with 0-5% methanol in $CHCl_3$/TEA (99/1, v/v). Appropriate fractions were pooled and solvent was evaporated to leave compound **3** as white foam (726 mg, 56.8%). Compound **3**: ESI-MS: *m/z* 661.1 $[M + Na]^+$. ^1H NMR (DMSO-d_6): δ 11.66 (br), 8.17 (s, 1H, H-6), 7.05-7.20 (m, 13H, aromatic H on DMT), 6.80 (m, 4H, H-PhS), 6.19 (t, 1H, H-1'), 5.27 (s, 1H, 3'-OH), 4.32 (m, 1H, H-3'), 3.91 (m, 1H, H-4'), 3.69 (s, 6H, H-OCH_3), 3.07-3.20 (m, 2H, H-5' and H-5"), 2.28 (m, 2H, H-2' and H-2").

Preparation of 4 (Scheme 1). Compound **3** (300 mg, 0.47 mmol) was added to a 50-ml flask and evaporated twice with dry CH_2Cl_2 (5 mL). The resulting residue was again dissolved in dry CH_2Cl_2 (3.5 mL) and the solution was kept under an argon atmosphere. To the flask was then added dry diisopropylethylamine (DIEA) (0.21 mL, 1.17 mmol), followed by dropwise addition of 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (0.16 mL, 0.71 mmol). After stirring at room temperature for 15 min, another portion of DIEA (0.21 mL) was added and the solution

was stirred for another 15 min. The reaction mixture was cooled in an ice bath, to which solution was then added CH_3OH (1 mL). The resulting solution was quickly extracted with EtOAc (20 mL) and the EtOAc layer was washed twice with saturated $NaHCO_3$ (10 mL) and once with saturated NaCl (10 mL). The organic layer was then dried with anhydrous Na_2SO_4 and evaporated to dryness to give **4** in white foam. Compound **4**: ESI-MS: *m/z* 838.8 $[M+H]^+$.

UV Irradiation

An aqueous solution of dinucleoside monophosphate with an OD_{260} of 0.4 was dispersed in a quartz tube and irradiated at room temperature with 254-nm UV light from a TLC lamp (UVGL-58, UVP Inc., Upland, CA) for 30 min. The solution was exposed to air during irradiation, and, after irradiation, the mixture was dried by using a Savant Speed-vac (ThermoSavant, Holbrook, New York). The dried residue was redissolved in water for HPLC analysis. For the time-course studies, 60 nmol of d(^{PhS}UG) was dissolved in 6 mL of water and dispersed in quartz tube with a diameter of 1 cm, and the solution was irradiated with the same TLC lamp for 30 sec, 2 min, 5 min, 10 min, or 25 min. The light intensity at the position of the quartz tube was measured to be 0.72 mW/cm^2.

For the irradiation of duplex ODNs, d(ATGGCA^{PhS}UGCT) (50 nmol) was annealed with its complementary strand in a buffer containing 50 mM NaCl and 50 mM phosphate (pH 6.8). The irradiation was then carried out under the same conditions as described above.

Enzymatic Digestion

A 50-nmol photoirradiation mixture was dried by Speed-vac. Five units of nuclease P1, 0.005 unit of calf spleen phosphodiesterase, 10 μL of 300 mM sodium acetate (pH 5.0), and water were added to make the final volume of the solution 250 μL. The digestion was carried out at 37°C for 6 h. The resulting solution was dried, and the dried residue was redissolved in a

250-μL solution containing 50 mM Tris-HCl (pH 8.6), 200 units of calf intestinal phosphatase, and 0.05 unit of snake venom phosphodiesterase. The digestion was conducted at 37 °C for 4 h, and the digestion mixture was extracted with an equal volume of chloroform to remove the enzymes. The aqueous layer was dried and redissolved in water for LC-MS/MS analysis.

HPLC

The HPLC separation was performed on a Surveyor system (ThermoFinnigan, San Jose, CA) with a photodiode array (PDA) detector, which was set at 260 nm for monitoring the effluents. A 4.6 × 250 mm Apollo reverse-phase C18 column (5 μm in particle size and 300 Å in pore size, Alltech Associates Inc., Deerfield, IL) was used, and the flow rate was 0.8 mL/min.

A gradient of acetonitrile in 10 mM ammonium formate (pH 6.3), where acetonitrile increased from 0 to 3% in 5 min and then from 3 to 9% in 35 min, was employed for the separation of the irradiation mixtures of d(^{PhS}UG). Different buffer solutions were used for the purification of the synthetic mixture of ODNs and the UV irradiation mixture of ODNs. In this respect, the composition of buffer A was 50 mM TEAA (pH 6.5), and buffer B contained 50 mM TEAA and acetonitrile (70/30, v/v). The gradient program for the mobile phase was: 0 min, 0% B; 5 min, 20% B; 45 min, 50% B; 50 min, 100% B.

For NMR analysis, the products were further desalted by using the Apollo column. After the sample was loaded, the column was washed with water for 20 min, and the analyte was eluted from the column with an equal-volume solvent mixture of methanol and water.

NMR and Mass Spectrometry

All NMR spectra were acquired on a Varian Unity Inova 500 MHz instrument (Palo Alto, CA). The residual proton signal of the solvent serves as internal reference.

Electrospray ionization-mass spectrometry (ESI-MS) and tandem MS (MS/MS) experiments were carried out on an LCQ Deca XP ion-trap mass spectrometer (ThermoFinnigan, San Jose, CA). An equal-volume solvent mixture of acetonitrile and water was used as the carrier and electrospray solvent, and a 2-μL aliquot of 1 μM sample solution was injected in each run. The spray voltages were 4.5 and 3.0 kV for experiments in the positive- and negative-ion modes, respectively. The mass width for precursor ion selection in MS/MS and multi-stage MS modes was 3 *m/z* units. Each spectrum was obtained by averaging approximately 50 scans, and the time for each scan was 0.1 s.

LC-MS/MS

A Zorbax SB-C18 column (0.5 × 150 mm, Agilent Technologies, Palo Alto, CA) with a particle size of 5 μm was used for the separation of the above enzymatic digestion mixture. An Agilent 1100 capillary HPLC pump (Agilent Technologies) was employed for LC-MS/MS experiments. A 100-min gradient of 0-35% acetonitrile in 20 mM ammonium formate was employed and the flow rate was 6.0 μL/min. The effluent was coupled directly to the LCQ Deca XP ion-trap mass spectrometer.

*Estimation of Quantum Yield for the Formation of d(^{PhS}U*G) [calculations were based on results from 10-min irradiation mixture of d(^{PhS}UG), see Figure S9e for the HPLC trace]:*

Energy of light delivered to the sample solution: 0.72 mW/cm^2*(7.0 cm^2)*(10 min*60 sec/min) = 3.0 *10^3 mJ = 3.0 J

The extinction coefficient of $d^{PhS}U$ was determined to be 8600 L/mol/cm at 260 nm and 9400 L/mol/cm at 254 nm.

$A = \varepsilon bc = (9400\ L/mol/cm)*(0.79\ cm)*[60*10^{-9}\ mol/(6*10^{-3}\ L)] = 0.074$

Percent transmission (%T) = $1/10^{A} = 84\%$

Energy absorbed by the $d^{PhS}U$ component = 3.0 J *(1-%T) = 3.0 J*(1-0.84) = 0.48 J

Number of photons absorbed by the $d^{PhS}U$ component (n) = $0.48/(6.63*10^{-34}*3.0*10^{8}/254*10^{-9})$ = $6.1*10^{17}$ photons

Number of molecules of $d(^{PhS}U*G)$ formed: $20*10^{-9}$ mol*$6.02*10^{23}$ molecule/mol = $1.2*10^{16}$ molecules

Estimated Quantum Yield = $1.2*10^{16}/6.1*10^{17} = 0.020$

Table S1. Results for exact mass measurement for the $[M-H]^-$ ions of d(^{PhS}UG) and d(^{PhS}U*G) as well as the $[M-4H]^{4-}$ ions of d(ATGGCA^{PhS}UGCT) and d(ATGGCA^{PhS}U*GCT).

Ion identity	Calculated mass (*m/z*)	Measured mass (*m/z*)	Deviation (ppm)
d(^{PhS}UG)	664.1227	664.1230	0.5
$[M-H]^-$		664.1232	0.8
		664.1230	0.5
d(^{PhS}U*G)	662.1070	662.1057	-2.0
$[M-H]^-$		662.1056	-2.1
		662.1053	-2.6
d(ATGGCA^{PhS}UGCT)	782.8761	782.8792	4.0
$[M-4H]^{4-}$		782.8797	4.6
		782.8799	4.9
d(ATGGCA^{PhS}U*GCT)	782.3722	782.3772	6.4
$[M-4H]^{4-}$		782.3776	6.9
		782.3756	4.3

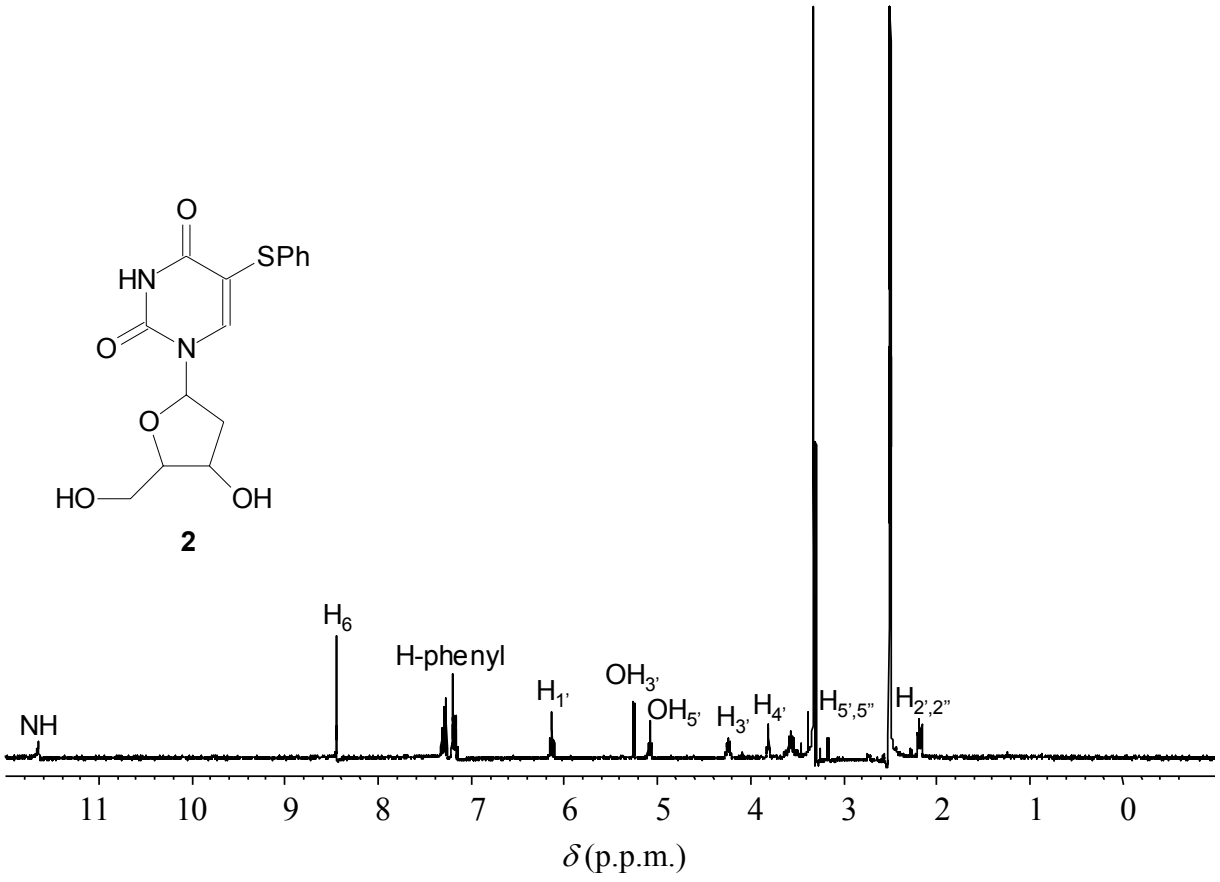

Figure S1. ^{1}H-NMR spectrum of 5-phenylthio-2'-deoxyurdidine (DMSO-d_6, 500-MHz, 25°C).

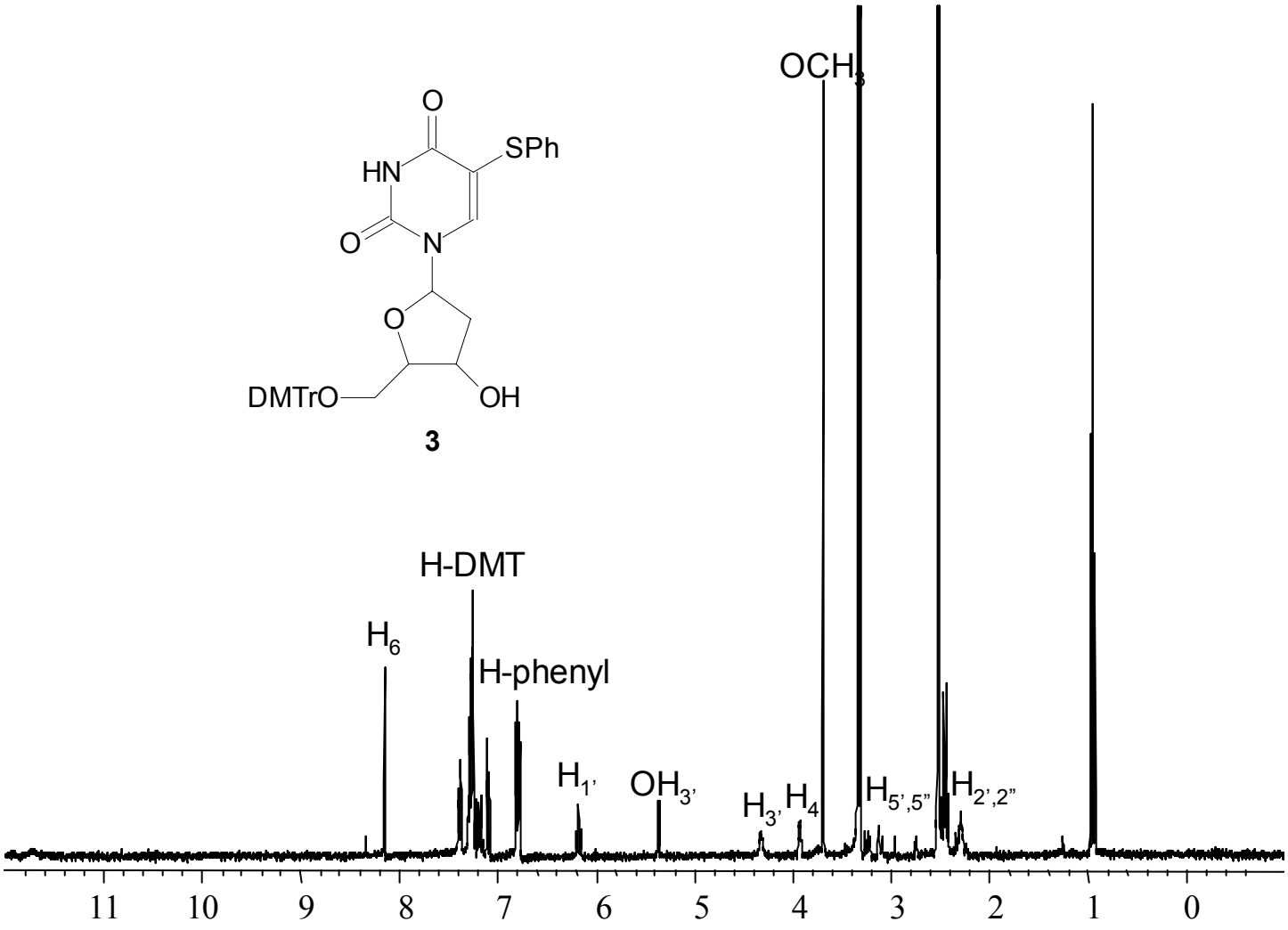

Figure S2. ^{1}H-NMR spectrum of compound **3** (DMSO-d_6, 500-MHz, 25°C).

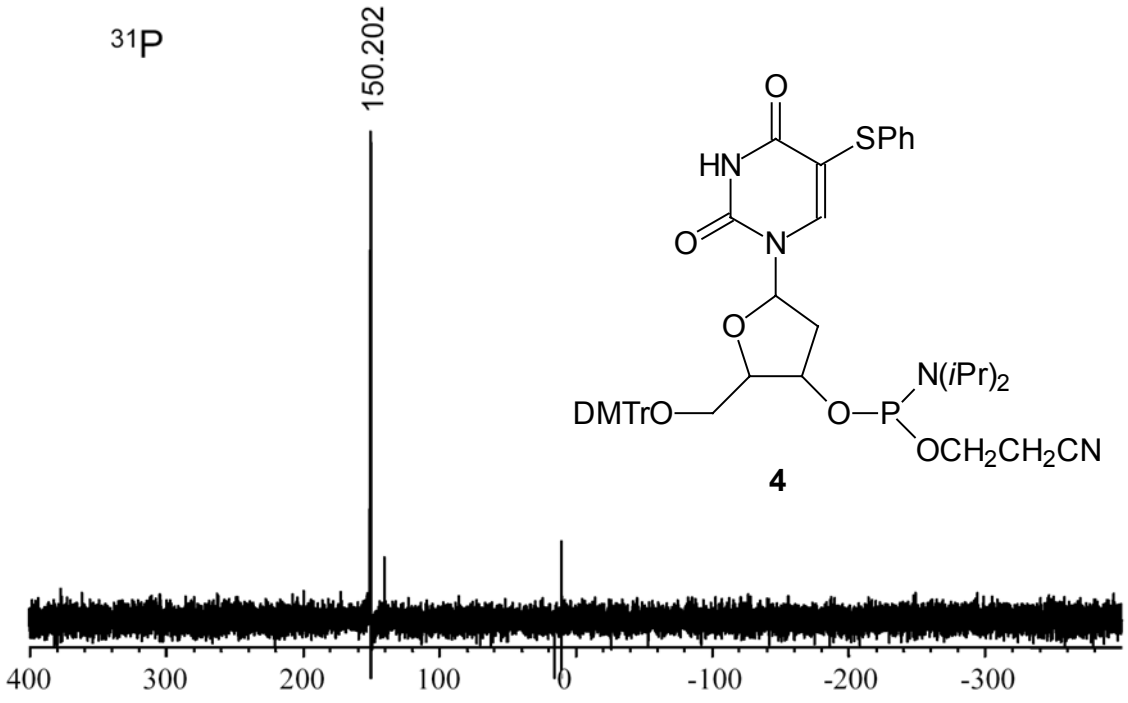

Figure S3. ^{31}P-NMR spectrum of compound **4** (DMSO-d_6, 500-MHz, 25°C).

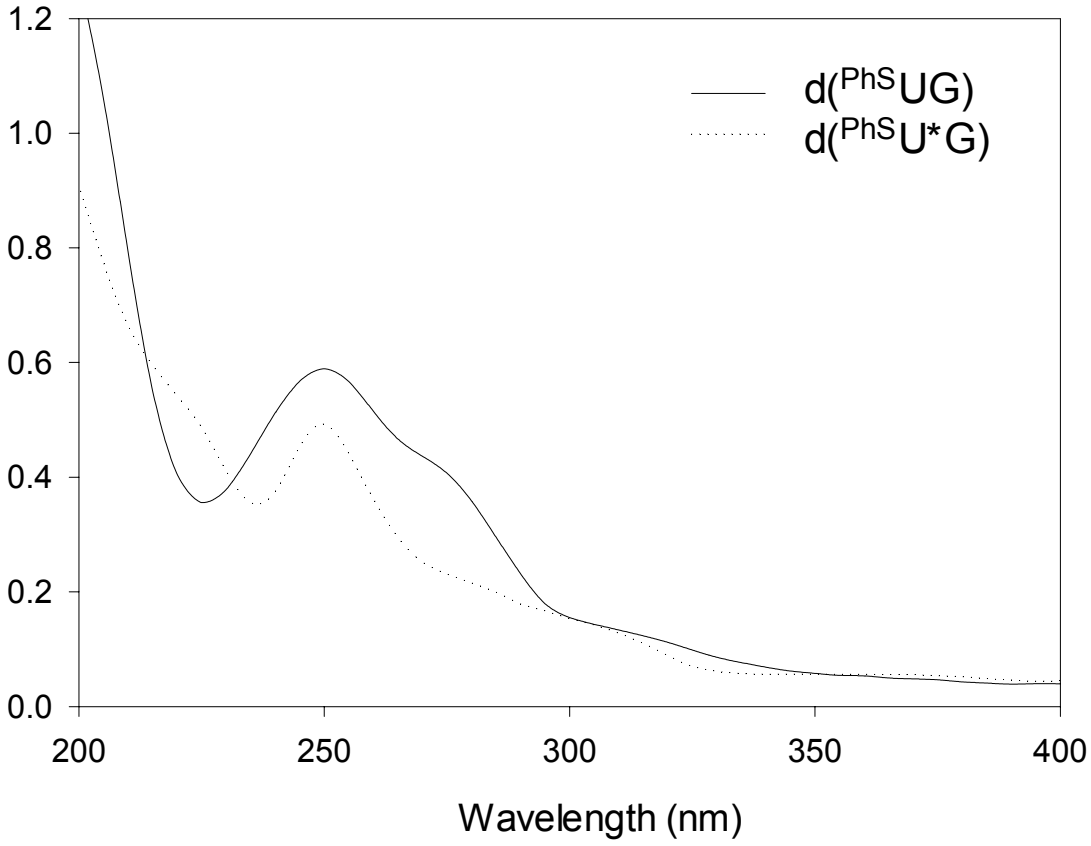

Figure S4. UV absorption spectra of d(^{PhS}UG) and d(^{PhS}U*G).

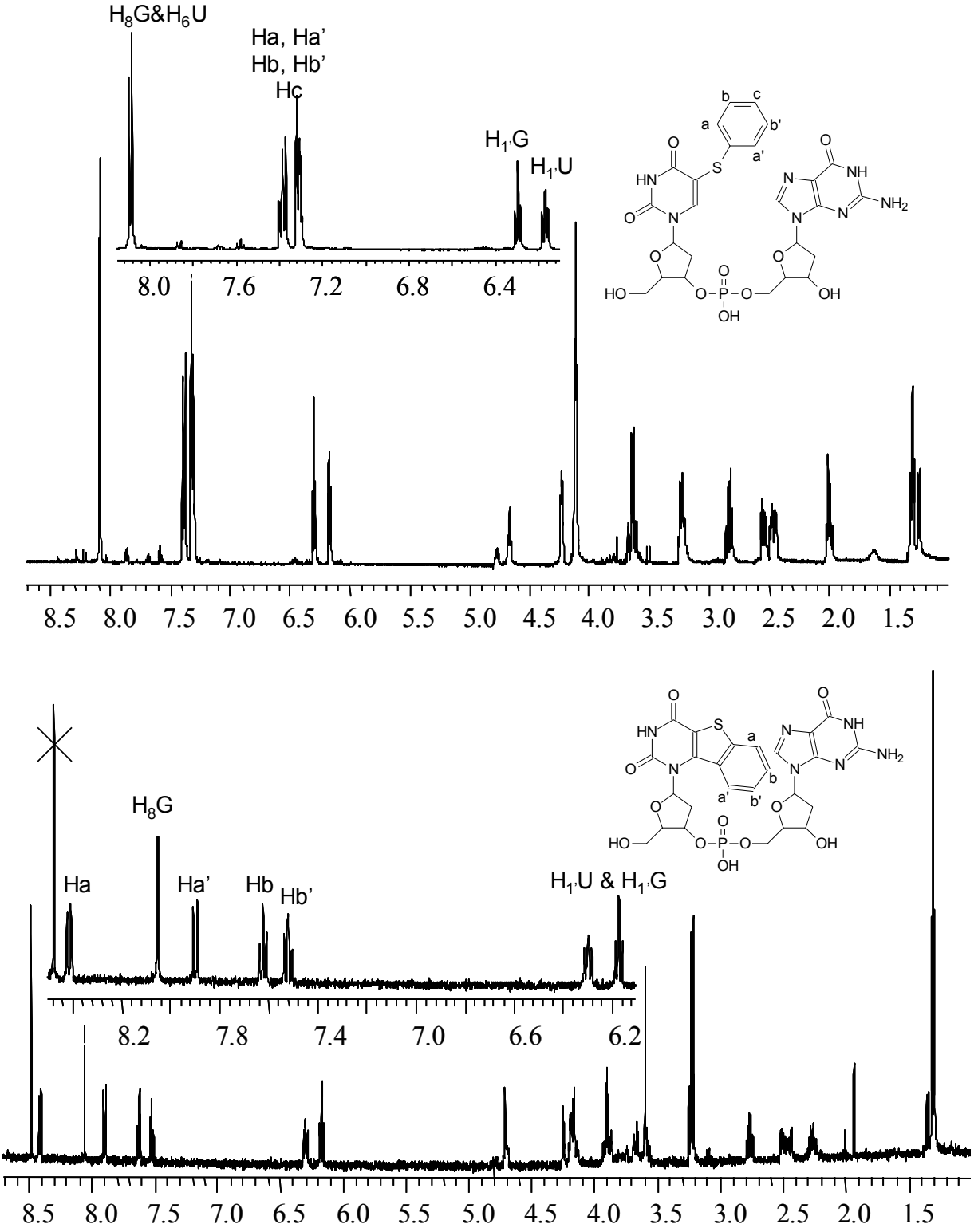

Figure S5. ^{1}H-NMR spectra of d(^{PhS}UG) (top) and d(^{PhS}U*G) (bottom, D_2O, 500 MHz, 25 °C).

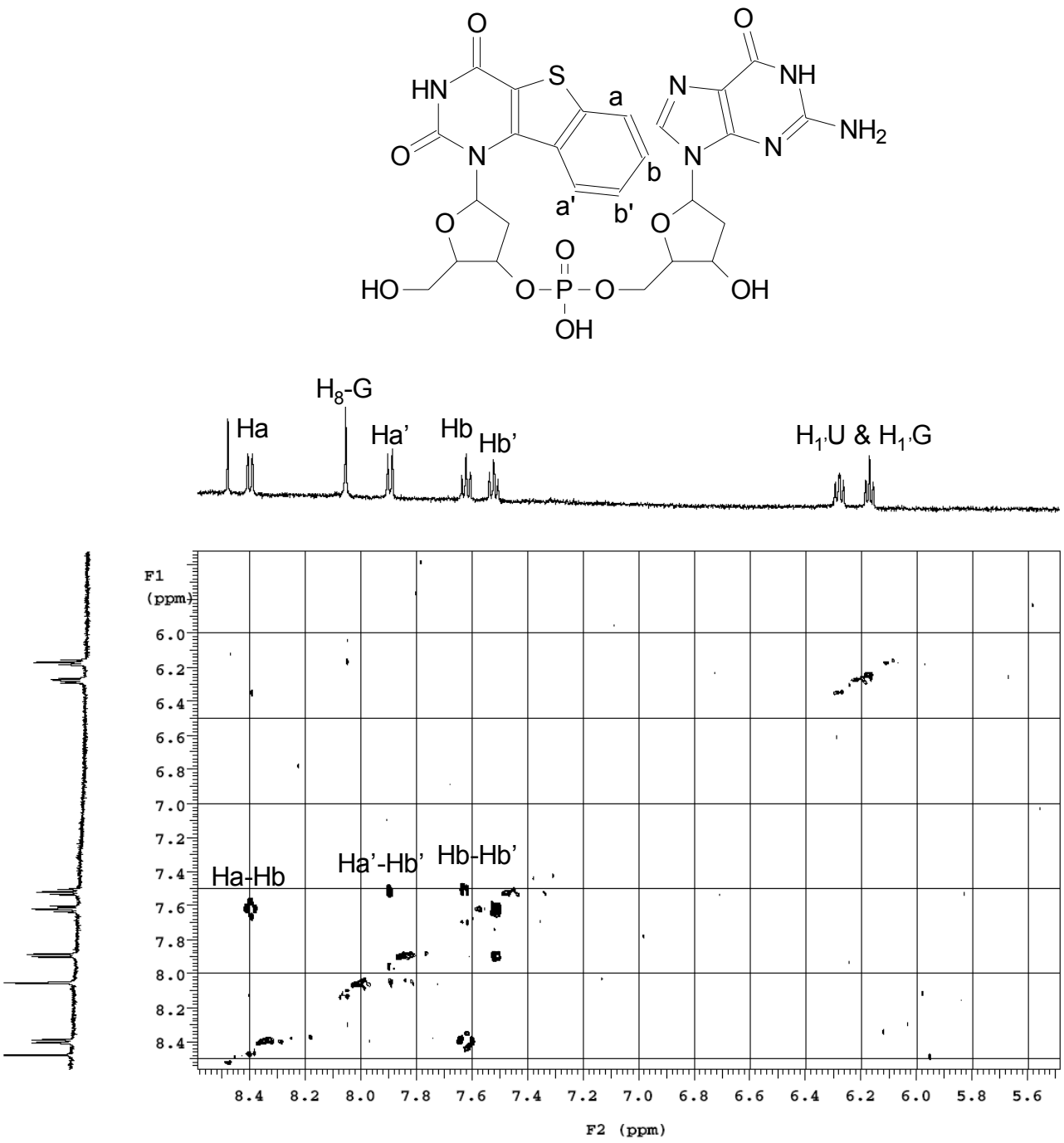

Figure S6. A portion of 2-D NOESY spectrum (500 MHz, 25°C, D_2O) of d(^{PhS}U*G) showing correlations among the aromatic protons of the phenyl moiety.

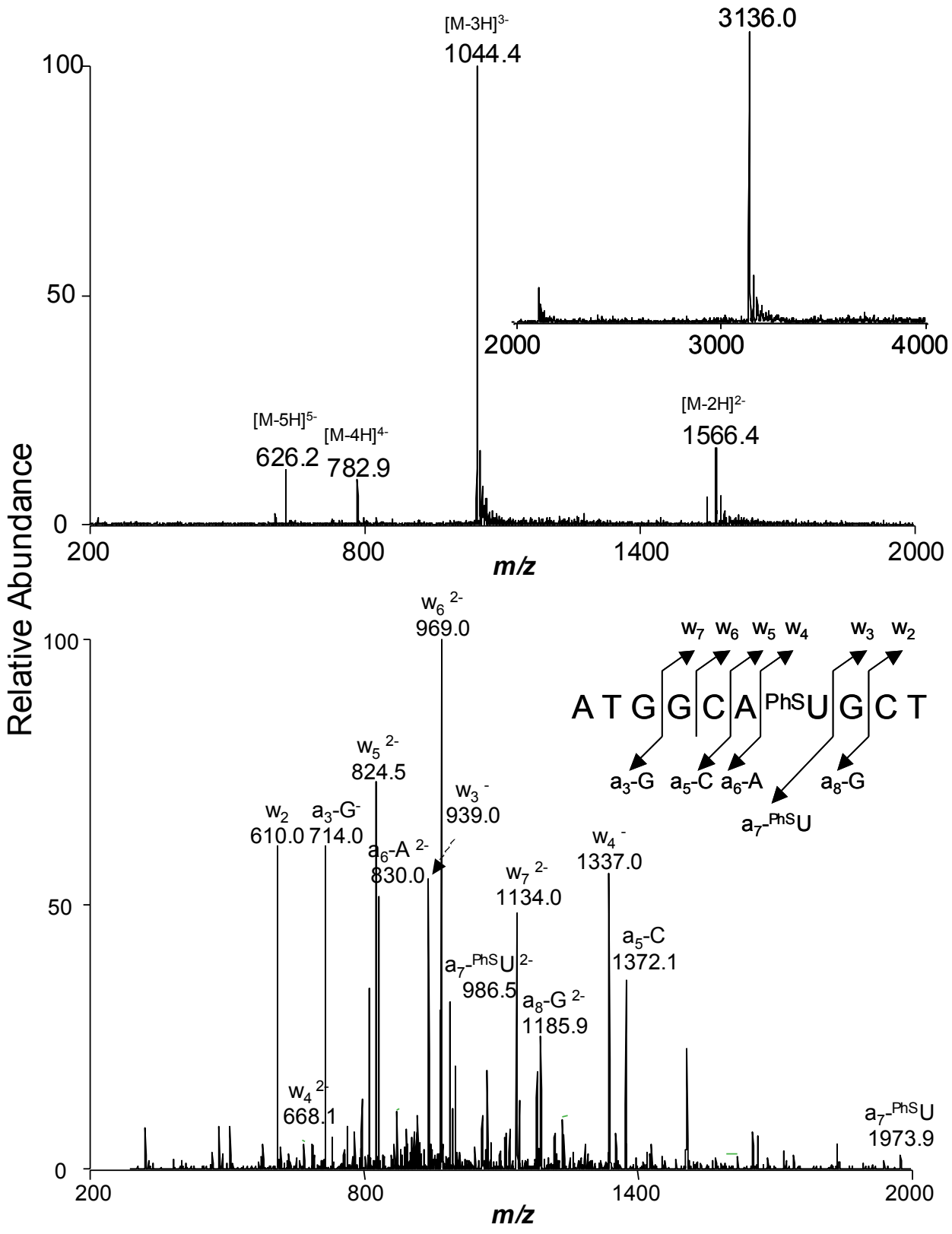

Figure S7. Negative-ion ESI-MS (top) and the product-ion spectrum of the $[M - 3H]^{3-}$ ion (*m/z* 1044.4) of d(ATGGCA^{PhS}UGCT) (bottom). Shown in the inset of the top panel is the deconvoluted spectrum for the ODN.

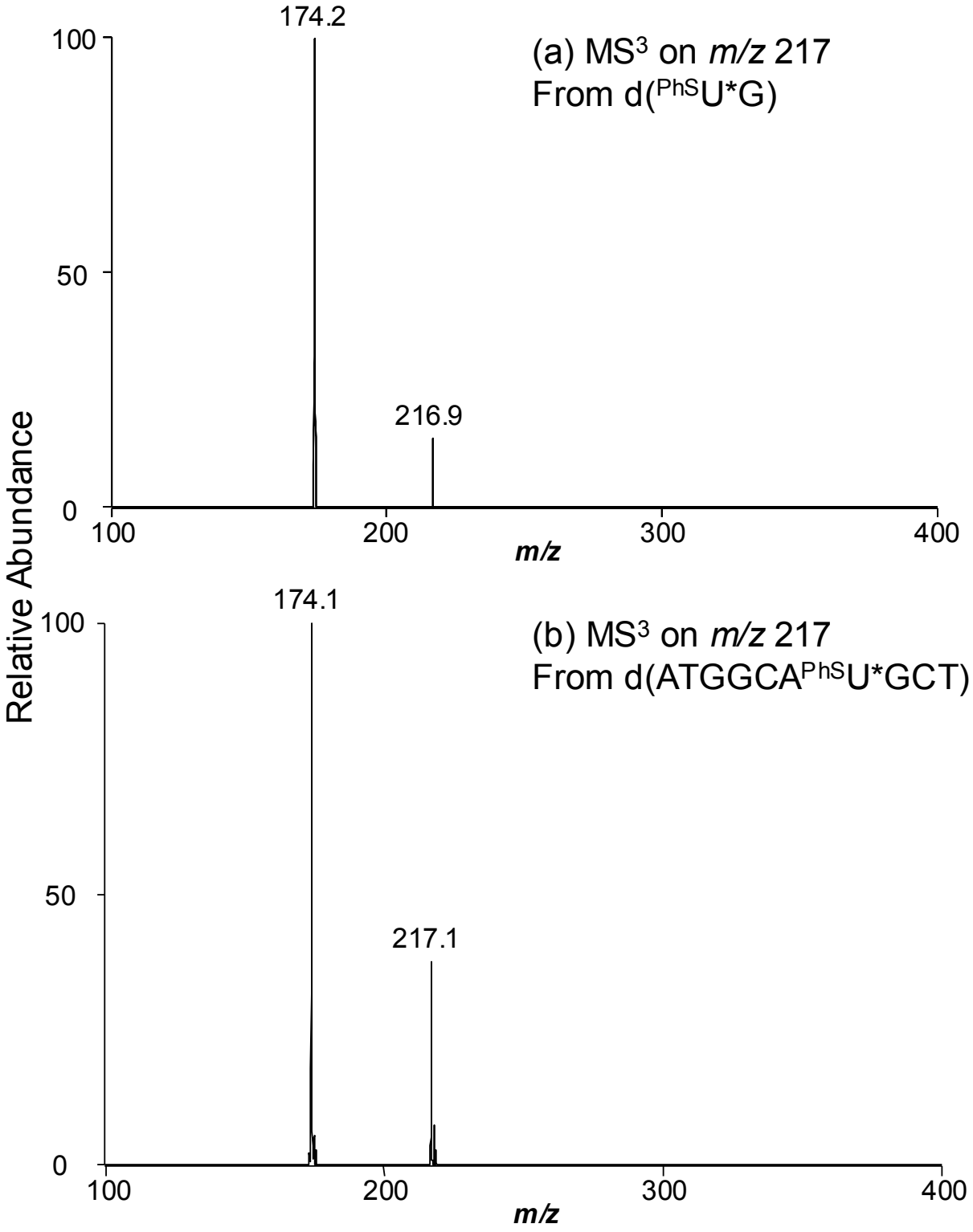

Figure S8. Product-ion spectra of the ion of *m/z* 217 (MS^3) observed in the product-ion spectra of the $[M - H]^-$ ion of d(^{PhS}U*G) (a, See Figure 1b for MS/MS) and the $[M - 5H]^{5-}$ ion of d(ATGGCA^{PhS}U*GCT) (b, see Figure 3b for MS/MS).

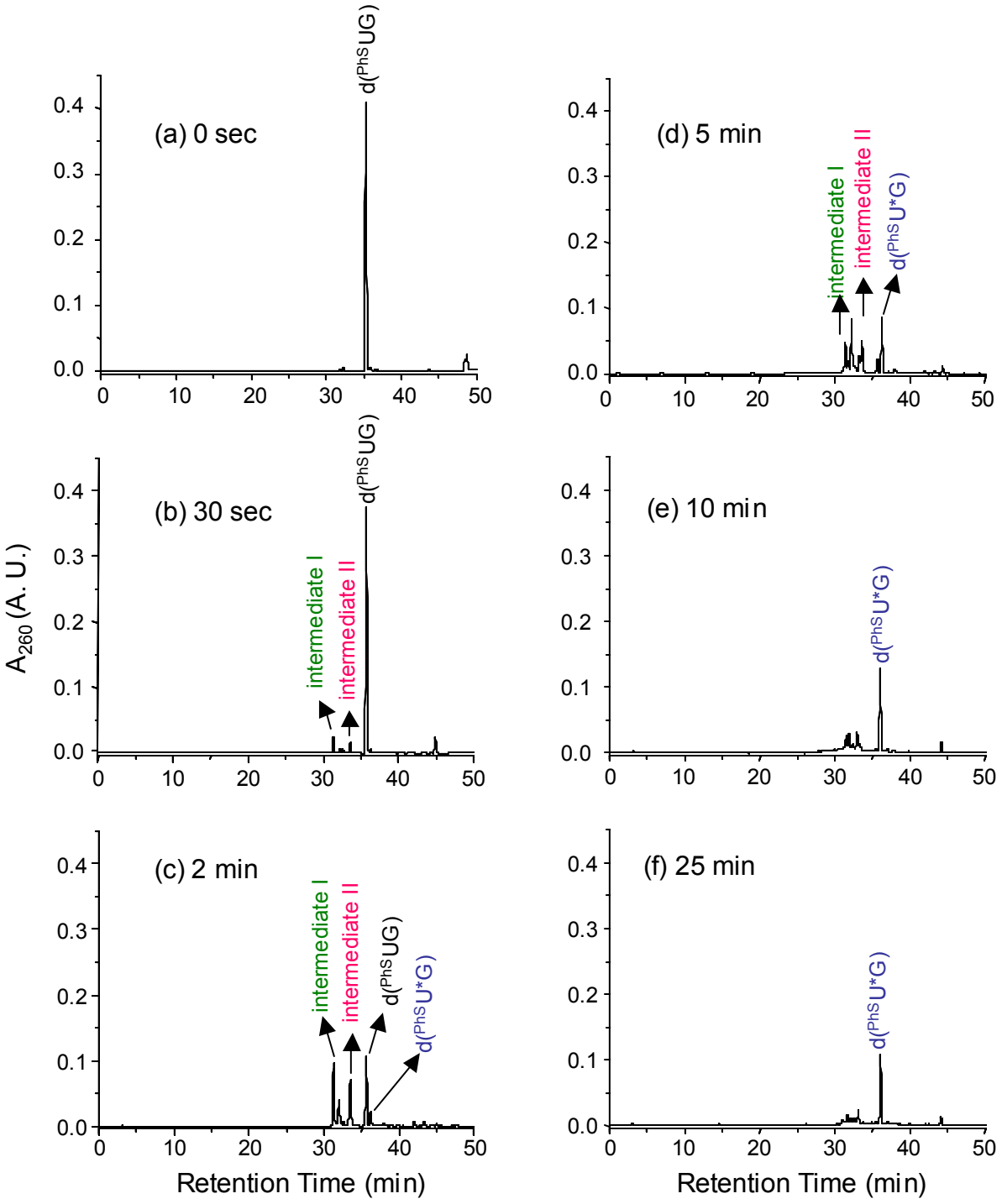

Figure S9. HPLC traces for the separations of the d(^{PhS}UG) after being irradiated with 254-nm for 0 sec (a); 30 sec (b); 2 min (c); 5 min (d); 10 min (e); and 25 min (f). The same amount of irradiation mixture was injected for each analysis, and the HPLC traces were plotted in the same scale.

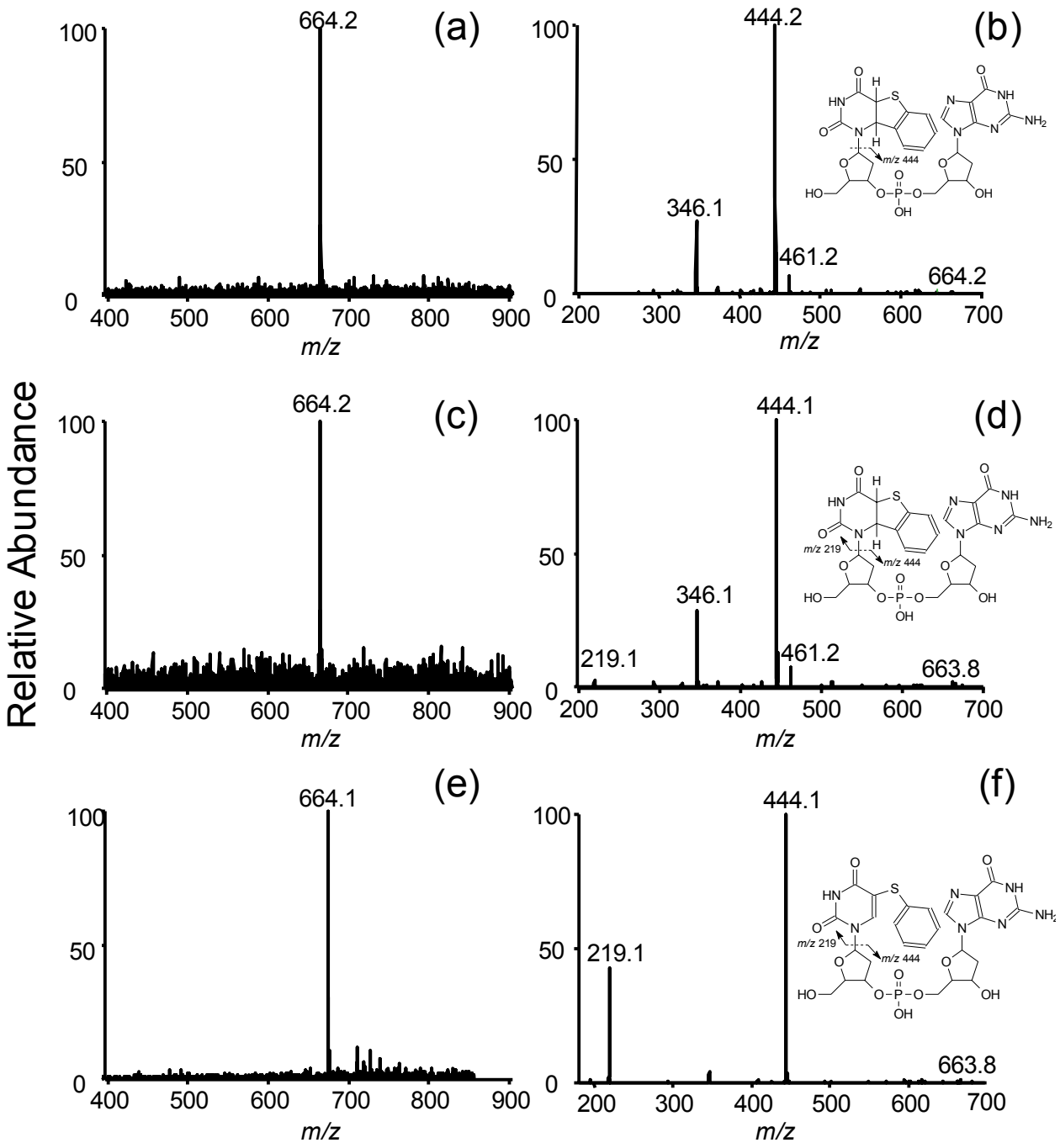

Figure S10. Negative-ion ESI-MS (left) and MS/MS (right) of intermediate I (a&b, see Figure S9), intermediate II (c&d, see Figure S9), and the starting d(^{PhS}UG) (e&f). The markedly decreased formation of the ion of *m/z* 219.1 for intermediates I and II is in line with the loss of aromaticity of the uracil component.

Scheme S1