Structure Determination of an Inter-strand-type *cis-anti* Cyclobutane Thymine Dimer Produced in High Yield by UVB Light in an Oligodeoxynucleotide at Acidic pH

Dian G. T. Su, Jeffrey L.-F. Kao, Michael L. Gross and John-Stephen A. Taylor

Department of Chemistry, Washington University, St. Louis, MO 63130

Contents:

Experimental

Photoreversion of minor I and minor II. The UVB irradiation mixture of d(GTATCATGAGGTGC) produced at pH 3.6 was irradiated with 254 nm light for 30 min. HPLC assay showed that both minor I and minor II disappeared and were presumably converted to the parent ODN (Figure S4), suggesting that the two minor photoproducts are cyclobutane dimers.

NP1-Coupled MALDI-TOF MS Assay of minor I and minor II. Nuclease P1 (NP1) digestion was carried out on the minor I and minor II photoproducts followed by MALDI-TOF-MS with default calibration. For the minor I photoproduct a molecular ion $[M+H]^+$ was detected at *m/z* 1247.4 corresponding to pd(T[A])=pd(T[C]) and for the minor II photoproduct a molecular ion was detected at *m/z* 1287.4 corresponding to pd(T[A])=pd(T[G]). Further MALDI-TOF-MS/MS experiments confirmed the assignment of the minor I digestion product by the presence of $[Cyt+H]^+$ at *m/z* 112.1 and $[Ade+H]^+$ at *m/z* 136.1 (Figure S5A). Likewise, assignment of the minor II digestion product was confirmed by the presence of $[Ade+H]^+$ at *m/z* 136.1 and $[Gua+H]^+$ at *m/z* 152.1 (Figure S5B). The results suggest that the thymine dimer of minor I is formed between T2 and T4 and that of minor II is between T2 and T7 or T2 and T12.

DNA Ladder Sequencing of photoproduct I and minor photoproduct II with BIMP. Bovine intestinal mucosa phosphodiesterase (BIMP) was used for 3'→5' sequencing to locate the 3'-thymine of the photodimer in both photoproduct I and minor II. In a typical sequencing experiment, the control or photodamaged ODN (150 pmol) in 7 μL Milli-Q water was mixed with 1 μL BIMP in 10 mM Tris-HCl buffer with 50% Glycerol (pH 7.5) and digested at 37 ºC. Aliquots were removed, spotted and finally analyzed by MALDI-TOF-MS as described for SVP sequencing. As a result, photoproduct minor II was seen having exactly the same digestion ladder pattern (Figure S6) as photoproduct I (Figure S2A) and could be explained in the same way (Figure S2B). After sequential removal of C14 to G8, exonuclease degradation was inhibited at T7. A mass loss of 366.1 u was observed immediately after loss of G8, which corresponds to addition of water by hydrolysis of the T4C5 phosphodiester bond by its minor endonuclease activity and then loss of T4 by its exonuclease activity along with 5'-HPO_3 of C5 by the 5'-phosphatase activity that must contaminate the commercial preparation of BIMP. The enzyme then excised A3 by its 3'-exonuclease activity. Finally the degradation terminated at *m/z* 1127.2 after a loss of and dC5 by its endonuclease activity, and 5'-HPO_3 of A6, indicating that the dimer of minor II is formed between T2 and T7.

Cyclobutane dimer stereochemistry of minor I and minor II. Hydrolysis of the minor I and minor II photoproducts with hydrogen fluoride in pyridine released products that had the same retention time as the authentic *cis-syn* thymine dimer and as the authentic *trans-anti* thymine dimer, respectively (Figure S7). The molecular ions $[M+H]^+$ of the hydrolysates of both the minor I and minor II

photoproducts were detected at *m/z* 253.1 by ESI-MS. ESI-MS/MS of the hydrolysate of the minor I photoproduct gave the same base peak of *m/z* 210.0 as with that of authentic Thy[*c*,*s*]Thy (Figure S8A). ESI-MS/MS of the hydrolysate of the minor II photoproduct, however, gave a base peak of *m/z* 208.0 and another intense peak at *m/z* 127.1 that was also observed for authentic Thy[*t*,*a*]Thy (Figure S8B).

Assignment of proton chemical shifts of the parent ODN and photoproduct I by TOCSY and NOESY spectra. Cytosines were identified first by strong H5-H6 cross peaks (F, G, H, I of Figure S10 and Figure S12A). The unique AH8 (i) to CH1' (i-1) NOE cross peaks of A6-C5 (E and J, Figure S10A and Figure S12A) were then used to start the sequential analysis. The base and sugar moieties are connected through intra-residue H6/H8 (i)-H1'/H2'2" (i) cross-relaxation peaks while inter-residue H6/H8 (i)-H1'/H2'H2" (i)-H6/H8 (i+1) pathways yield sequential assignment along the strand (Figure S10 and Figure S12). Pyrimidine 5-CH_3's of the parent ODN were assigned by the intra-residue H8-5CH_3 cross peaks and were also confirmed by the unique A3-T4 and A6-T7 inter-residue AH8 (i)-T5CH_3 (i+1) NOE (Figure S10B). The assignment of H3' and H4' sugar protons were made from the combined H3' (i)-p-H4' (i+1) backbone ^{1}H-^{31}P correlation (Figure S11A and Figure S13A) with the sugar H1' to H4' spin propagation in TOCSY (Figure S11B and Figure S13B). For instance, by tracing TOCSY T7-H3' cross peak into the ^{1}H-^{31}P correlated spectrum ($7_{3'}$ of Figure S11) can lead to the observation of the G8-H4' ($8_{4'}$ of Figure S11). The resonance of G8-H3' ($8_{3'}$ of Figure S11B) can then be assigned by further tracing the G8-H4' back into the TOCSY spectrum. Comparing Figure S11A and Figure S13A revealed that unusually downfield shifted ^{31}P resonances at -2.7 and -2.15 ppm were observed in G1-p-T2 and A6-p-T7, respectively. Upfield shifted of T2, T7-H3' (at 4.56 and 4.62 ppm) and downfield shifted of G8-H1' (at 5.64 ppm) were also observed in Figure S13B suggesting that the formation of T2=T7 dimer is indeed the cause of the chemical shift changes in the neighboring residues.

Table S1. Proton and ^{31}P chemical shifts of the parent ODN and photoproduct I in D_2O at 25 ^{o}C.

		G1	T2	A3	T4	C5	A6	T7	G8	A9	G10	G11	T12	G13	C14
H6/H8	parent	7.85	7.37	8.34	7.36	7.53	8.34	7.22	7.78	8.08	7.75	7.86	7.29	7.94	7.76
	cis-anti	7.87	3.59	8.18	7.69	7.52	8.16	3.62	7.84	8.12	7.74	7.86	7.31	7.94	7.75
H5/CH_3	parent		1.67		1.66	5.91		1.64					1.63		5.87
	cis-anti		1.29		1.85	5.95		1.38					1.65		5.87
H1’	parent	6.08	6.02	6.28	6.04	5.97	6.27	5.86	5.45	5.95	5.64	5.94	5.96	6.07	6.21
	cis-anti	6.22	5.94	6.33	6.34	6	6.22	6.01	5.64	5.97	5.67	5.97	6.04	6.11	6.24
H2’2’	parent	2.63	1.95	2.77	2.11	1.79	2.74	1.73	2.4	2.52	2.42	2.6	1.91	2.61	2.25
			2.27		2.34	2.26		2.11				2.66	2.29	2.72	2.39
	cis-anti	2.66	1.91	2.79	2.31	1.45	2.65	1.81	2.45	2.55	2.41	2.58	1.93	2.63	2.25
		2.81	1.96		2.51	2.13	2.79	1.9				2.67	2.29	2.73	2.37
H3’	parent	4.85	4.81	4.97	4.8	4.75	4.98	4.72	4.82	4.92	4.89	4.94	4.79	4.95	4.55
	cis-anti	4.96	4.56	5.04	4.87	4.68	5.03	4.62	4.87	4.93	4.89	4.93	4.77	4.93	4.54
H4’	parent	4.23	4.18	4.4	4.23	4.15	4.39	4.1	4.17	4.17	4.27	4.34	4.16	4.33	4.13
	cis-anti	4.27	4.12	4.44	4.39	4.15	4.36	4.13	4.23	4.32	4.27	4.32	4.15	4.32	4.12
^{31}P		δ_{1-2}	δ_{2-3}	δ_{3-4}	δ_{4-5}	δ_{5-6}	δ_{6-7}	δ_{7-8}	δ_{8-9}	δ_{9-10}	δ_{10-11}	δ_{11-12}	δ_{12-13}	δ_{13-14}	
	parent	-3.08	-3.06	-3.24	-2.99	-3.13	-3.24	-3.31	-3.11	-3.22	-3.05	-3.16	-3.18	-2.82	
	cis-anti	-2.7	-2.83	-3.04	-2.78	-3.04	-2.14	-3.21	-3.22	-3.23	-3.03	-3.14	-3.18	-2.81	

Figure S1. Exonuclease-coupled MS ladder sequencing of the parent ODN. A) BSP digestion (5'→3') coupled MALDI sequencing at room temperature for 0.5, 1.5, and 5.5 min (top to bottom). B) SVP digestion (3'→5') coupled MALDI sequencing at 37° for 1, 2.5, and 5.5 min (top to bottom). An asterisk refers to a doubly charged ion.

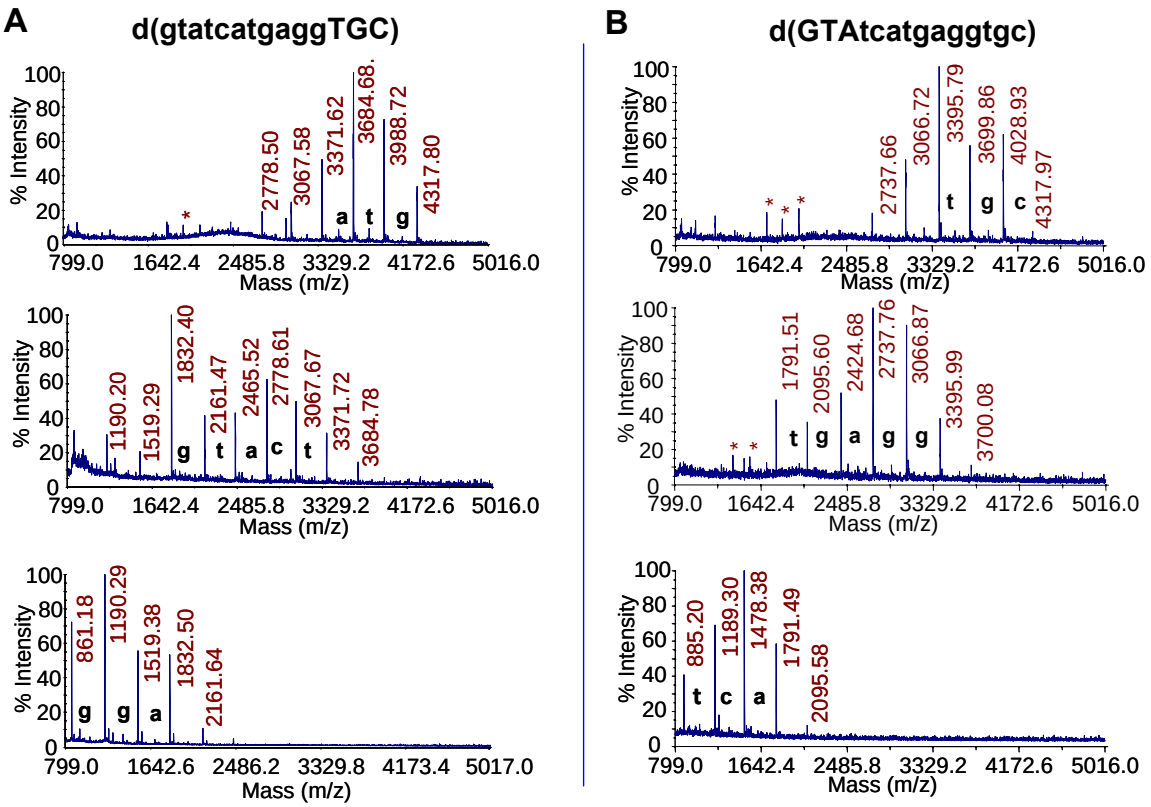

Figure S2. Bovine intestinal mucosa phosphodiesterase (BIMP) sequencing of photoproduct I. A) BIMP digestion (3'→5') coupled MALDI sequencing of photoproduct I at 37° for 0.5 min, 7 min, and 110 min (top to bottom). The mass loss of 289.05, 304.05, 313.06, and 329.05 u corresponding to the excision of pdC, pT, pdA, and pdG, respectively (+H_2O-pdN). Nucleotides that were enzymatically removed are represented by small-case letters. Right: proposed enzymatic cleavage pathway by BIMP for photoproduct I.

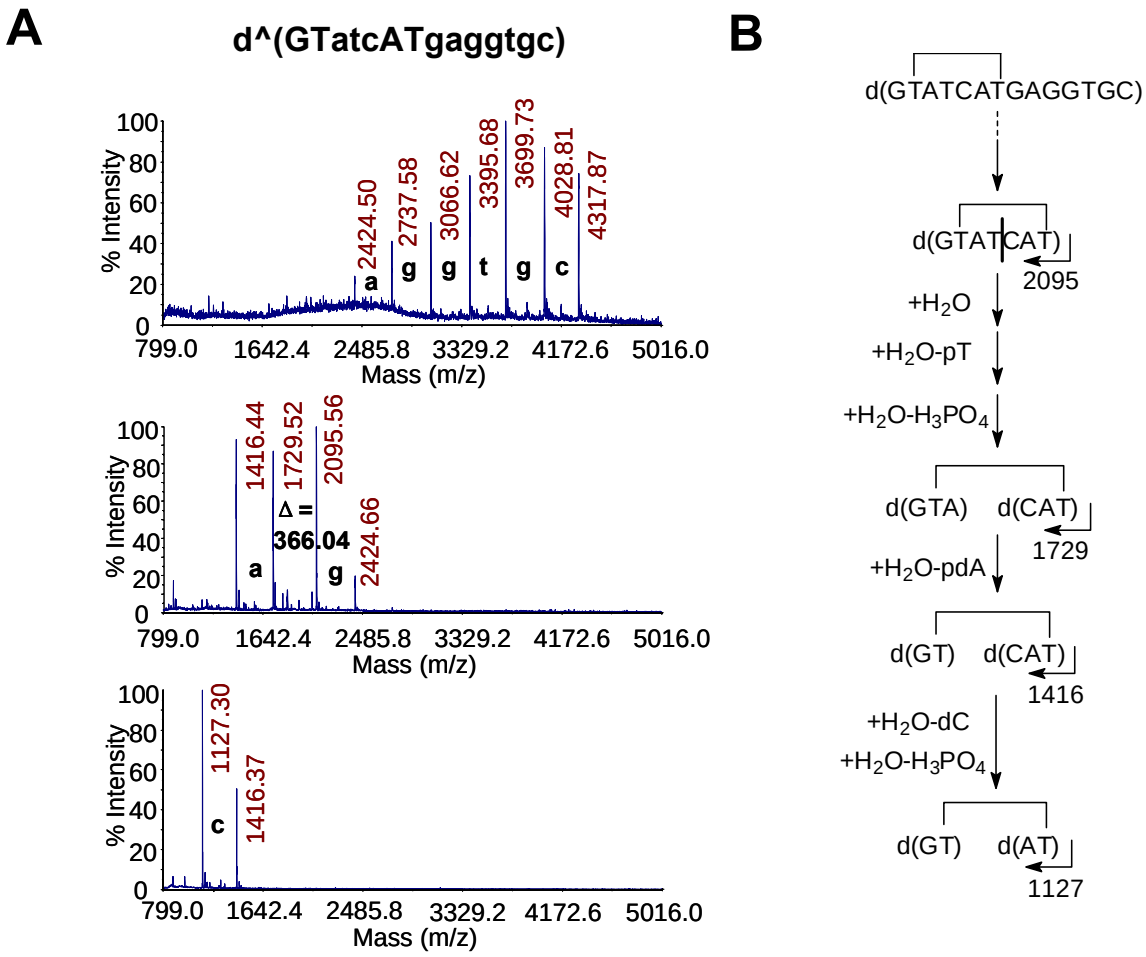

Figure S3. ESI-MS/MS characterization of the cyclobutane thymine dimers. A) Thy[*c*,*a*]Thy dimer from photoproduct I and B) Thy[*c*,*s*]Thy dimer from d(GTAT[*c*,*s*]TATGAGGTGC).

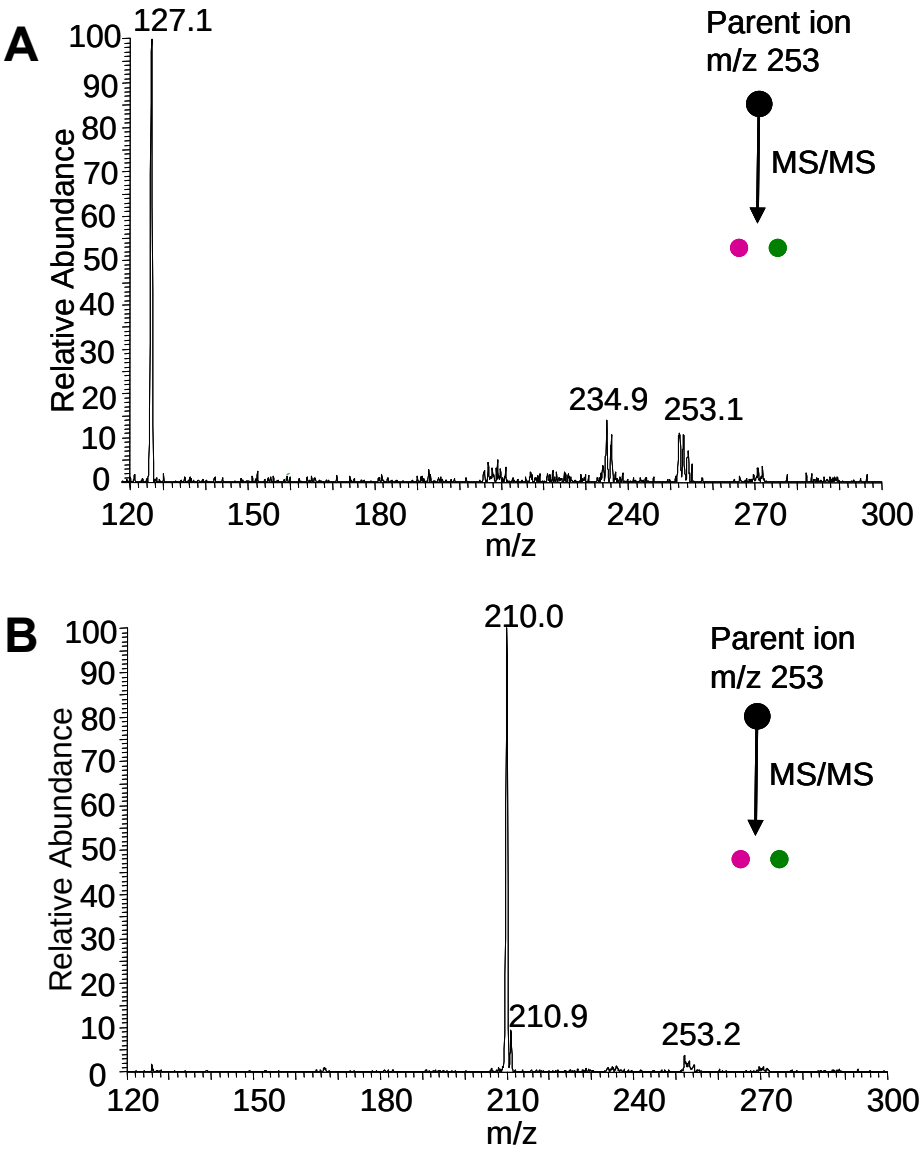

Figure S4. Photoreversal of UVB irradiation products to the parent ODN. Reverse phase HPLC analysis (method B) of the photoproducts of d(GTATCATGAGGTGC) produced at pH 3.6 to the parent ODN with 254 nm light: A) UVB irradiation mixture of d(GTATCATGAGGTGC) produced at pH 3.6, and B) UVB irradiation mixture of d(GTATCATGAGGTGC) produced at pH 3.6 after 30 min irradiation with 254 nm light.

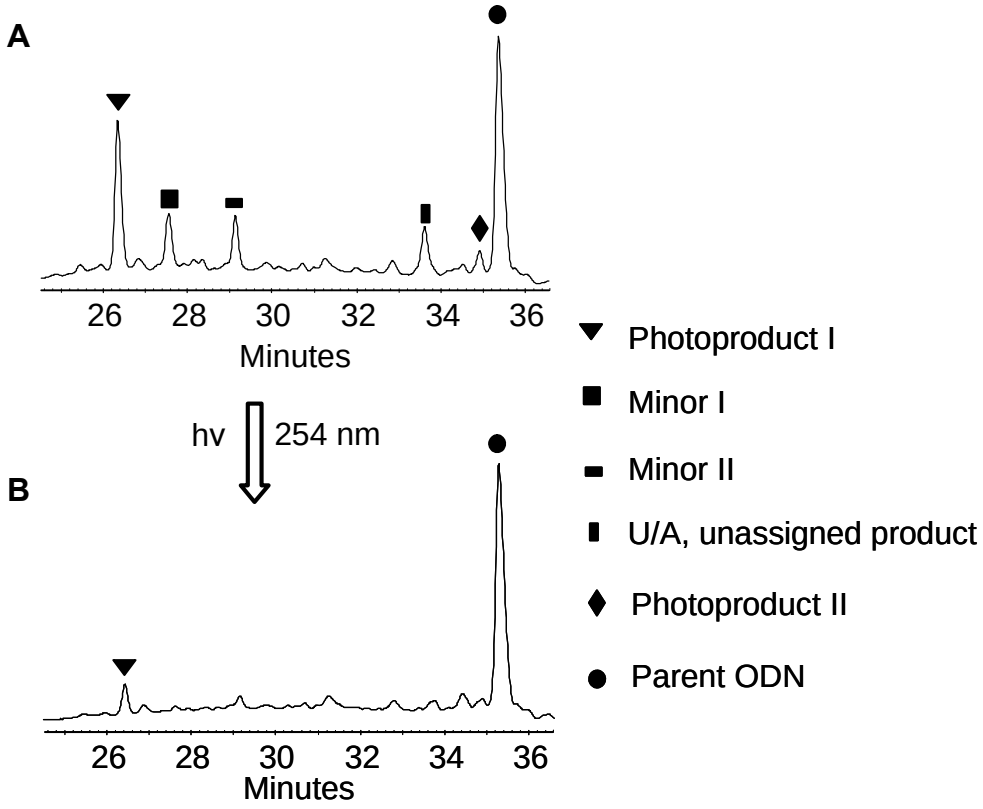

Figure S5. Mass spectrometry analysis of NP1 digestion product of minor I and minor II. MALDI-TOF-MS/MS spectra of: A) the parent ion $[M+H]^+$ (*m/z* 1247.4) of HPLC purified ultimate NP1 digestion product of minor I, *c,s*-pd(T[A])=pd(T[C]), showing loss of $[Ade+H]^+$ at 136 and $[Cyt+H]^+$ at 112, and B) the parent ion $[M+H]^+$ (*m/z* 1287.4) of HPLC purified ultimate NP1 digestion product of minor II, *t,a*-pd(T[A])=pd(T[G]) showing the same fragments as for photoproduct I (Figure 5).

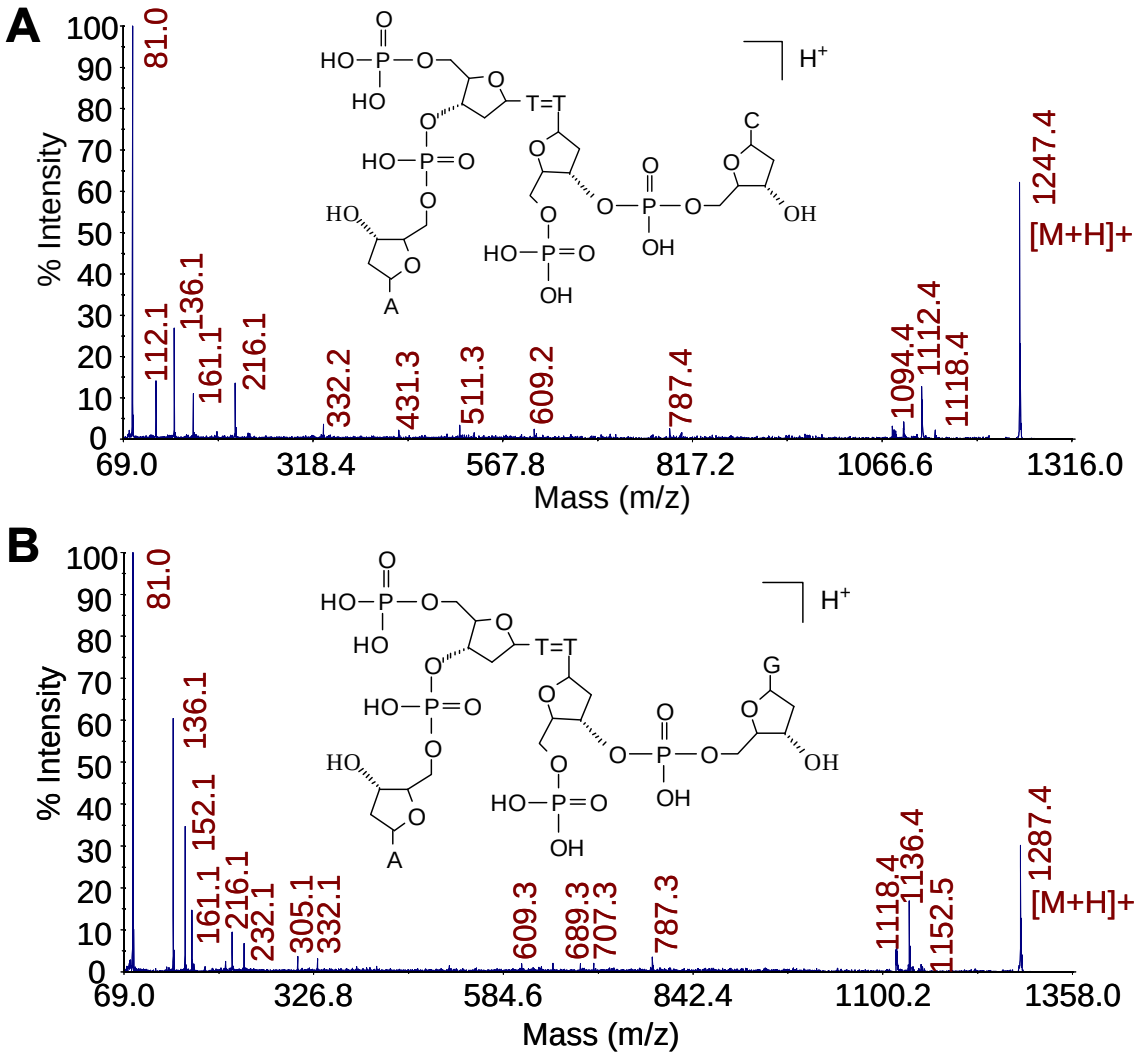

Figure S6. BIMP digestion (3'→5') coupled MALDI sequencing of minor photoproduct II. Minor photoproduct II was treated with BIMP at 37° for A) 1 min, B) 7 min, and C) 110 min. The mass loss of 289.05, 304.05, 313.06, and 329.05 u corresponds to the excision of pdC, pT, pdA, and pdG, respectively (+H_2O-pdN). Nucleotides that were enzymatically removed are represented by small-case letters.

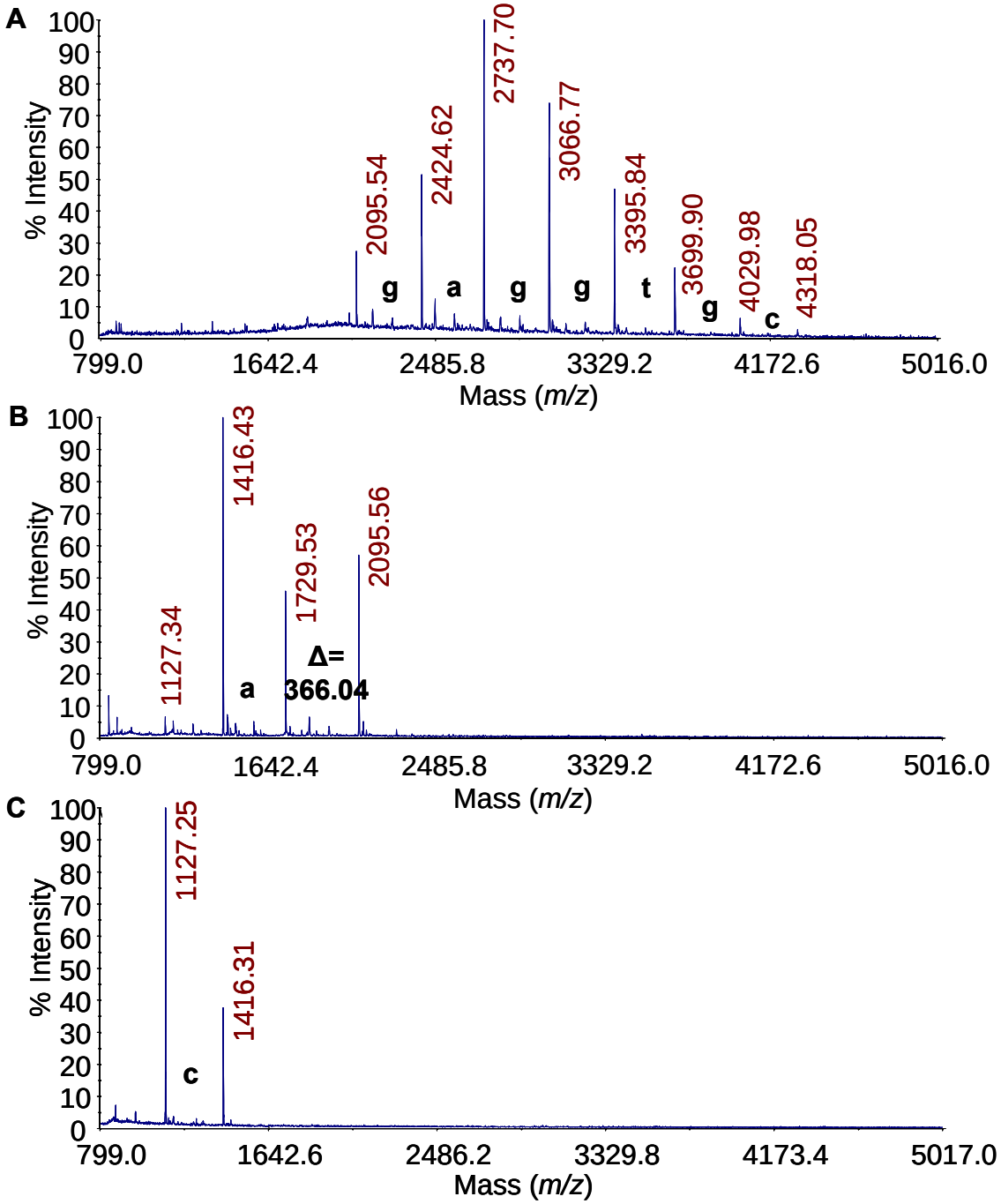

Figure S7. Thymine dimer stereochemistry of minor photoproducts I and II by chemical correlation. Comparison of cyclobutane dimers of thymine with the acid hydrolysis products of minor I and minor II photoproducts by reverse phase HPLC (method D). HPLC traces of: A) the hydrolysates of minor I, B) the product mixture from the irradiation of thymine, C) the hydrolysates of minor II.

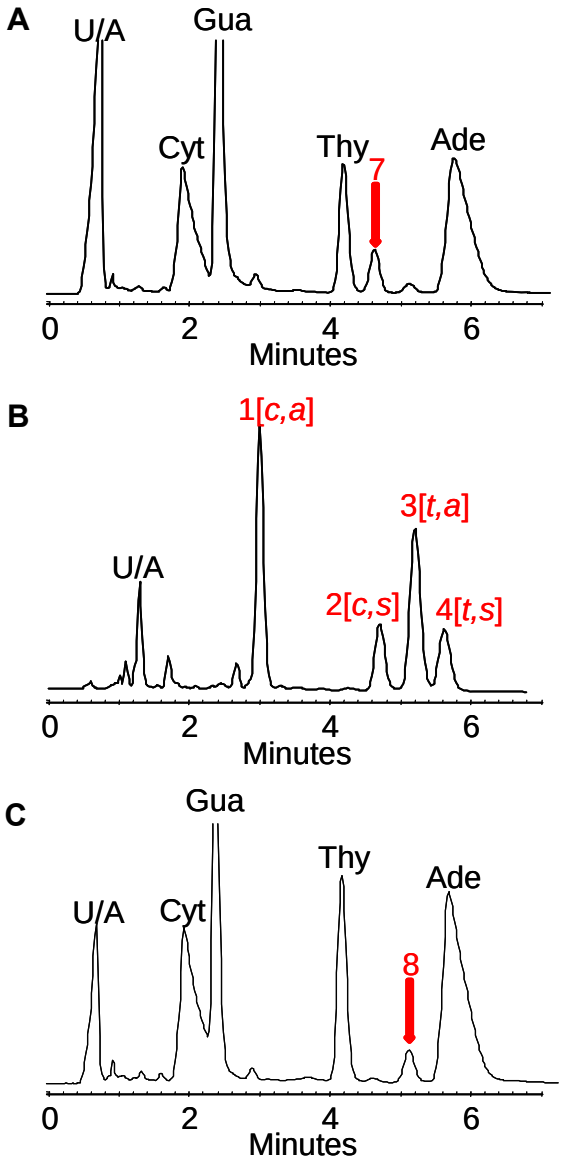

Figure S8. ESI-MS/MS characterization of the cyclobutane thymine dimers. A) Thy[*c*,*s*]Thy dimer from minor I. B) Thy[*t*,*a*]Thy dimer from minor II.

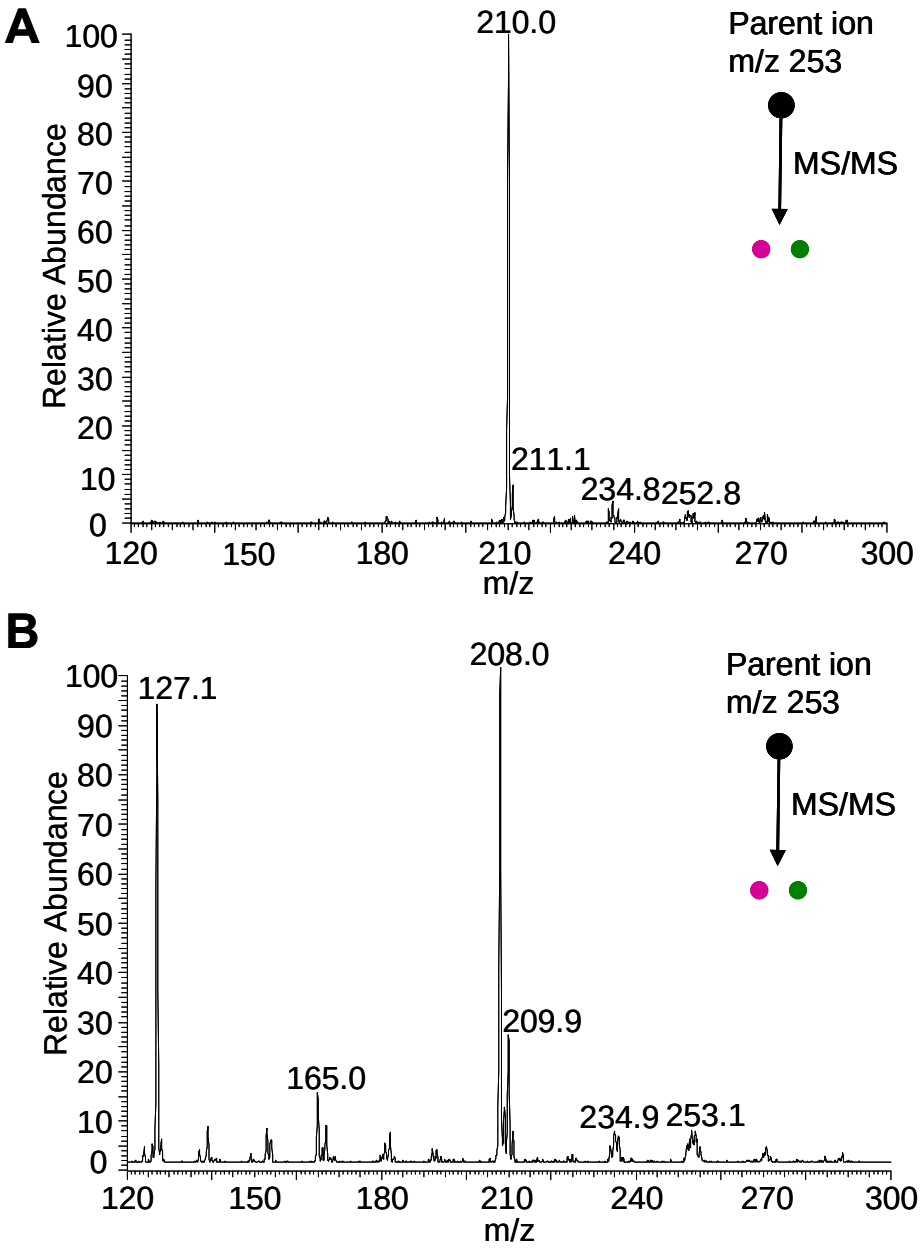

Figure S9. Formation and identification of the T2[*c*,*a*]T7 product in the mut6 sequence. Reverse phase HPLC analysis with method B of the UVB irradiation mixture of A) d(GTATCGTGAGGTGC) at pH 4.8, and B) MALDI-TOF-MS/MS assay of HPLC purified ultimate NP1 digestion product of T2[*c*,*a*]T7 non-adjacent photoproduct of d(GTATCGTGAGGTGC).

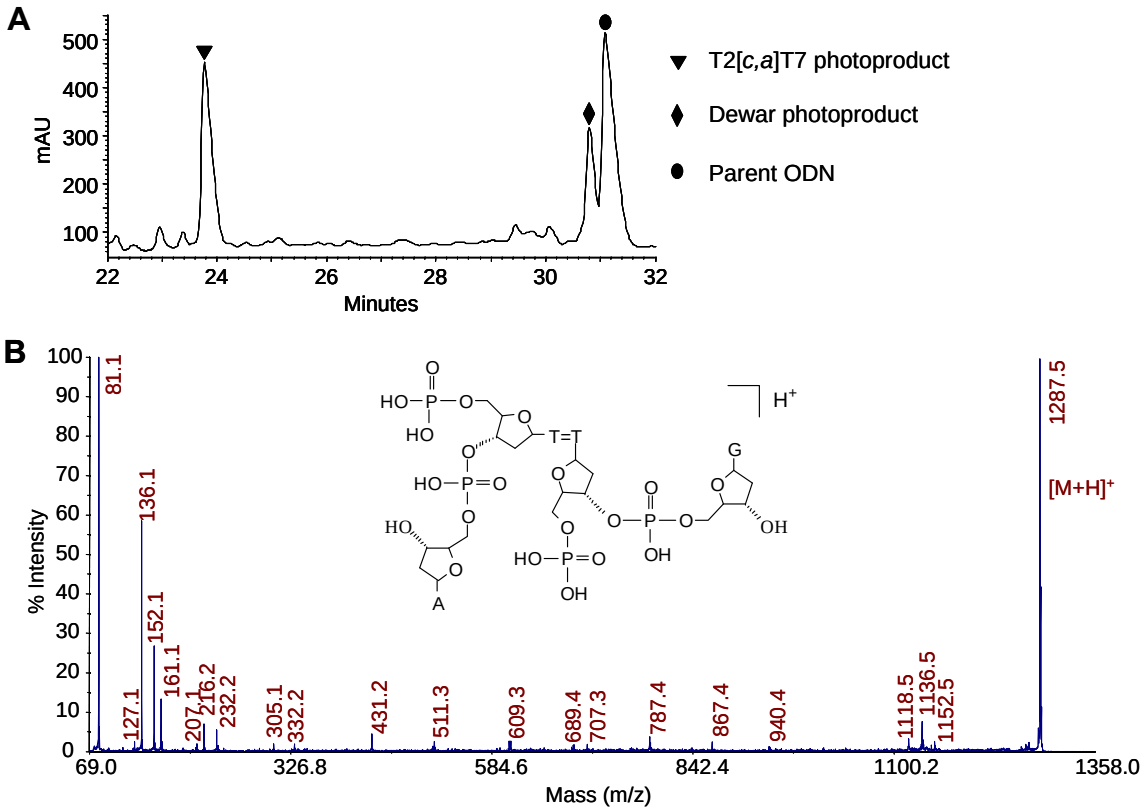

Figure S10. NOESY subspectra of the parent ODN. Expansion from the aromatic H6/H8 region to (A) H1'/H5 and (B) H2'2"/TCH_3 region of 600 MHz NOESY spectrum of the parent ODN 14-mer in "100%" D_2O at 25 °C. Sequence of A3 through A9 is connected by line with number that represents (A) H6/8 (i)-H1'(i), and (B) H6/8 (i)-H2'2"/TCH_3 (i) crosspeaks.

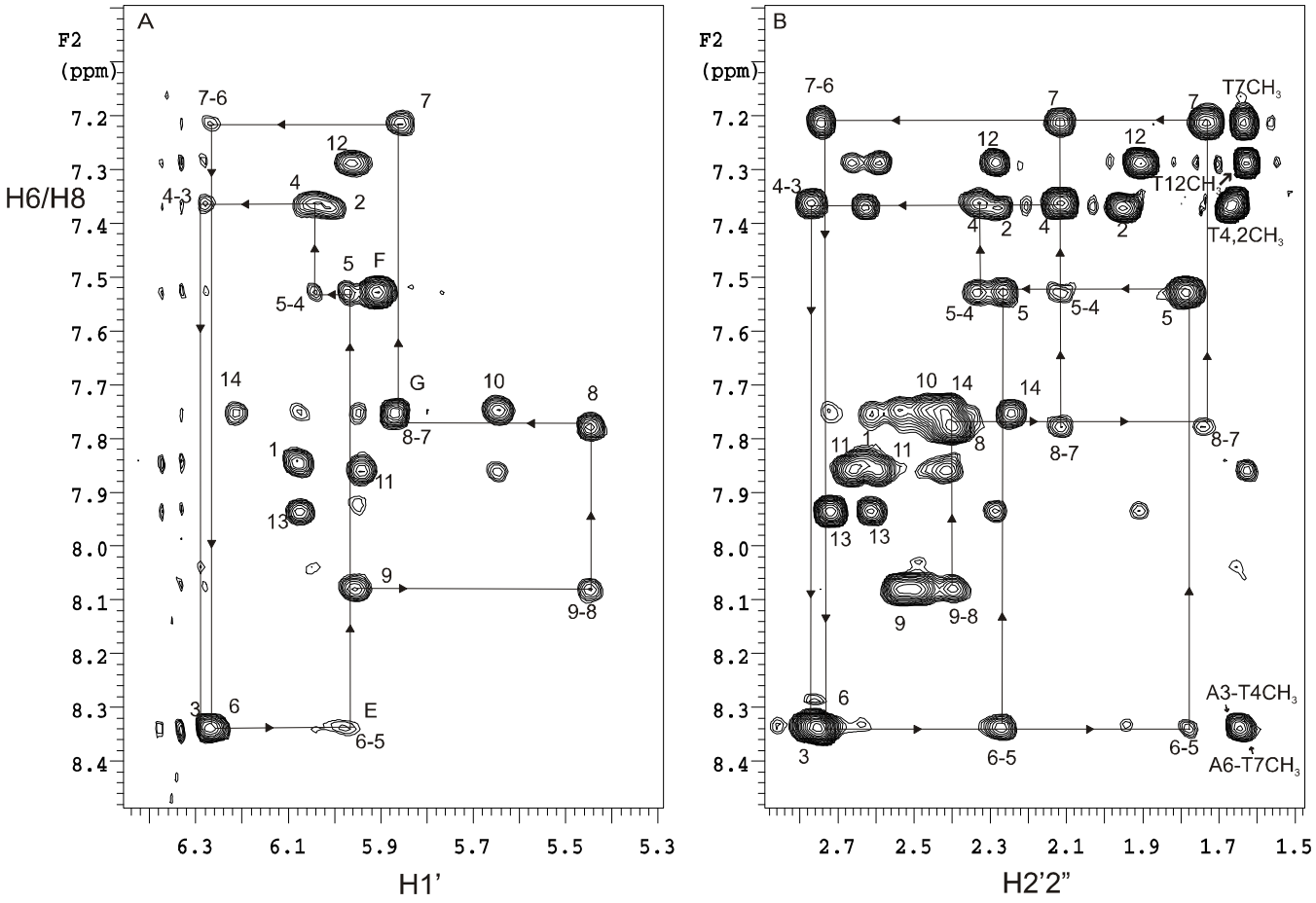

Figure S11. ^{1}H-^{31}P HSQC and TOCSY subspectra of the parent ODN. A) ^{1}H-^{31}P HSQC spectrum of the parent ODN 14-mer in "100%" D_2O. Horizontal line connections represent H3' (i)-p-H4' (i+1) correlation. B) Expanded H3'/H4' region of TOCSY spectrum. Horizontal line connections represent H3'-H4' spin propagation.

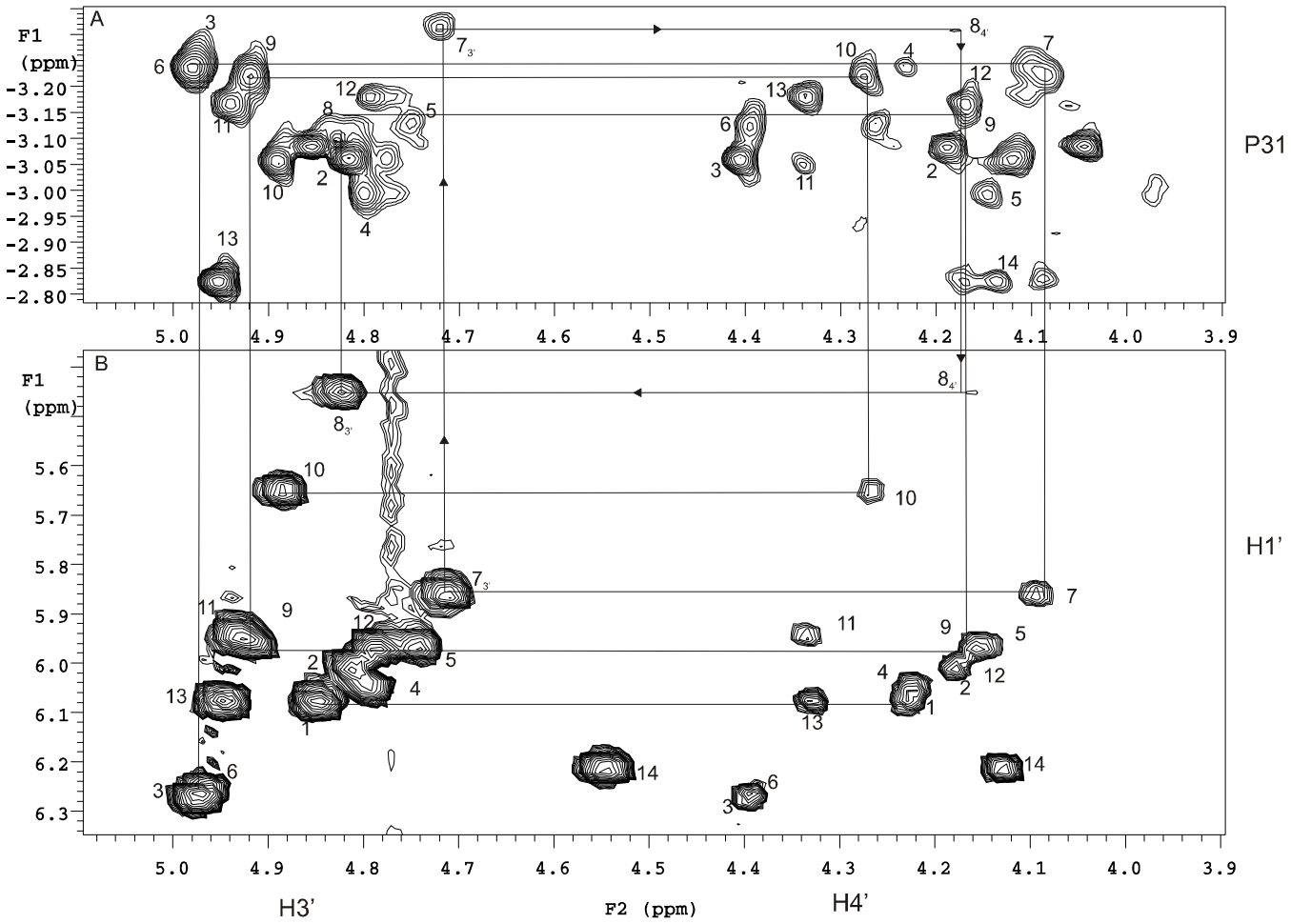

Figure S12. NOESY subspectra for photoproduct I. Expansion from the aromatic H6/H8 region to H1'/H5 and H2'2"/T-5CH_3 region of 600 MHz NOESY spectrum of photoproduct I 14-mer at 25 °C. Sequence is connected by line with number represents. A) H6/H 8 (i)-H1' (i), and B) H6/H8 (i)-H2'2"/T-5CH_3 (i) cross peaks.

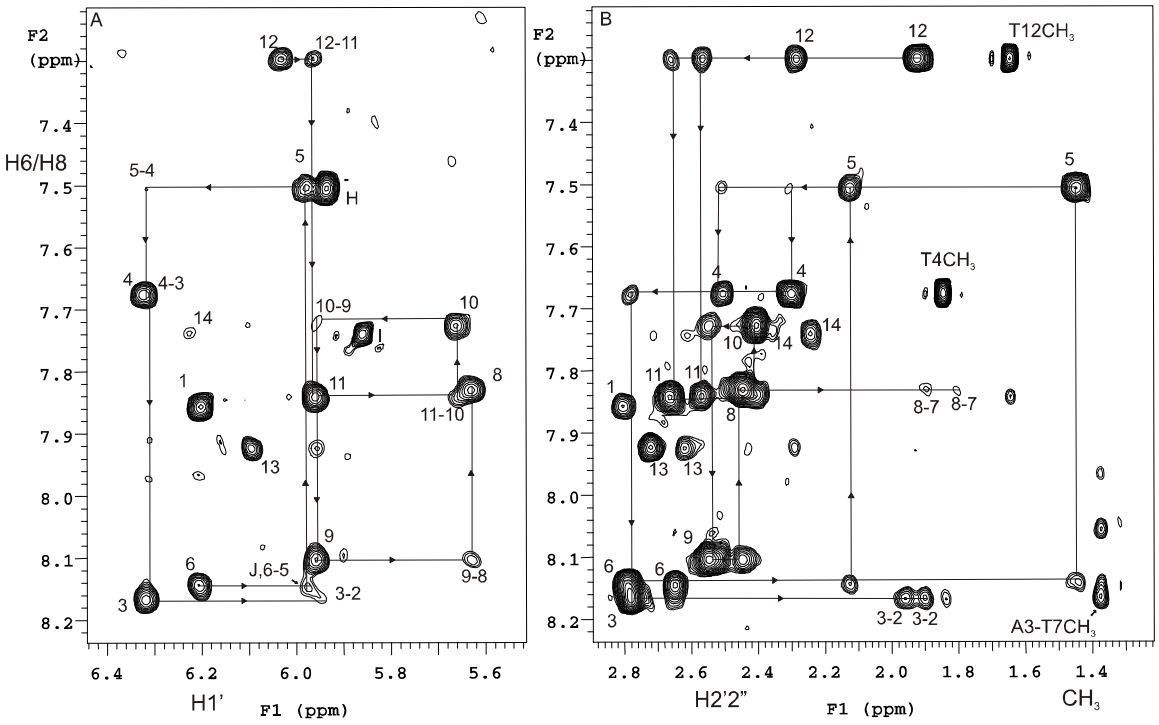

Figure S13. ^{1}H-^{31}P HSQC and TOCSY subspectra of photoproduct I. A) ^{1}H-^{31}P HSQC spectrum of photoproduct I in "100%" D_2O. Horizontal line connections represent H3'(i)-p-H4'(i+1) correlation, B) Expanded H3'/H4' region of TOCSY spectrum. Horizontal line connections represent H3'-H4' spin propagation.