Supporting information

Solid-phase synthesis of branched oligonucleotides containing the dCyd341 interstrand crosslink DNA lesion.

Marisa L. Taverna Porro, Christine Saint-Pierre, Didier Gasparutto and Jean-Luc Ravanat*

Univ. Grenoble Alpes, CEA, CNRS, IRIG-SyMMES UMR 5819, F-38054 Grenoble Cedex 9.

* To whom correspondence should be addressed. Tel +33 4 38 78 47 97; Email: jravanat@cea.fr

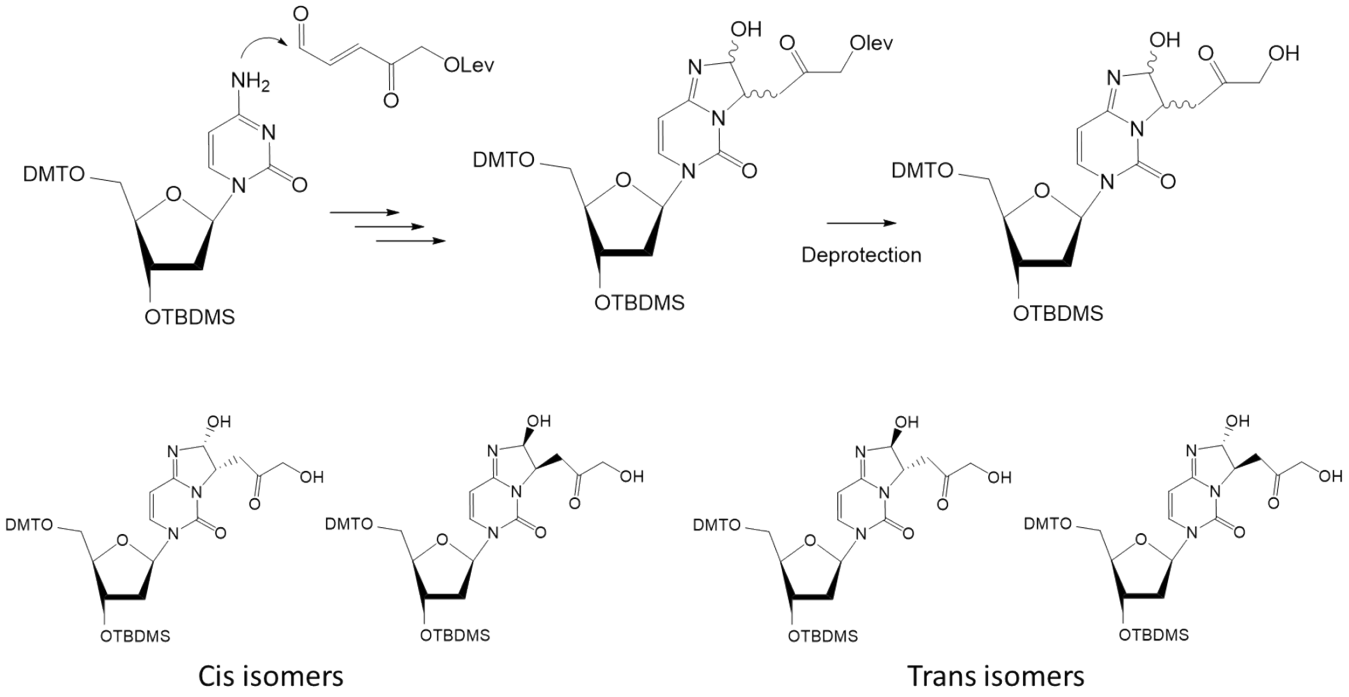


Figure S1. Chemical structures of all possible isomers of **dCyd341**. The formation of the oxadiazabicyclo(3.3.0)octaimine adducts (dCyd341) is initiated by the reaction of the exocyclic amino group of the cytosine base with the reactive aldehyde (upper panel). dCyd341 exists as four potential isomers, two *cis* and two *trans* (lower panel).

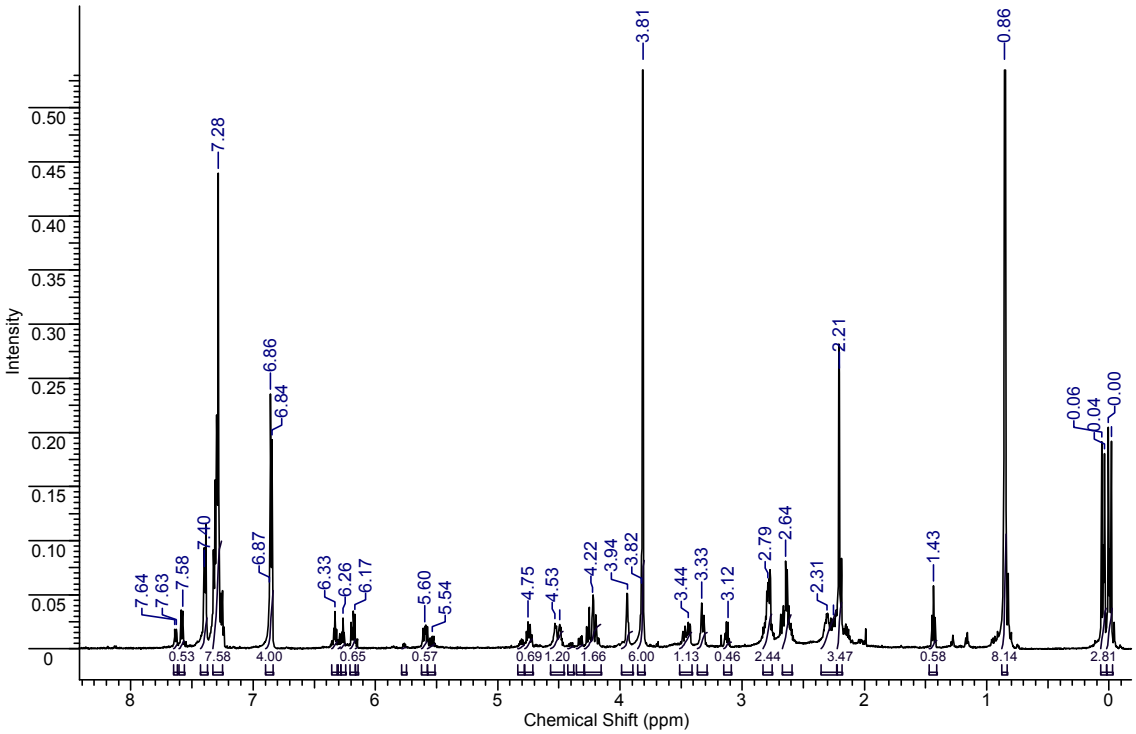


Figure S2. 600 MHz ^{1}H NMR spectrum of compound **6.**

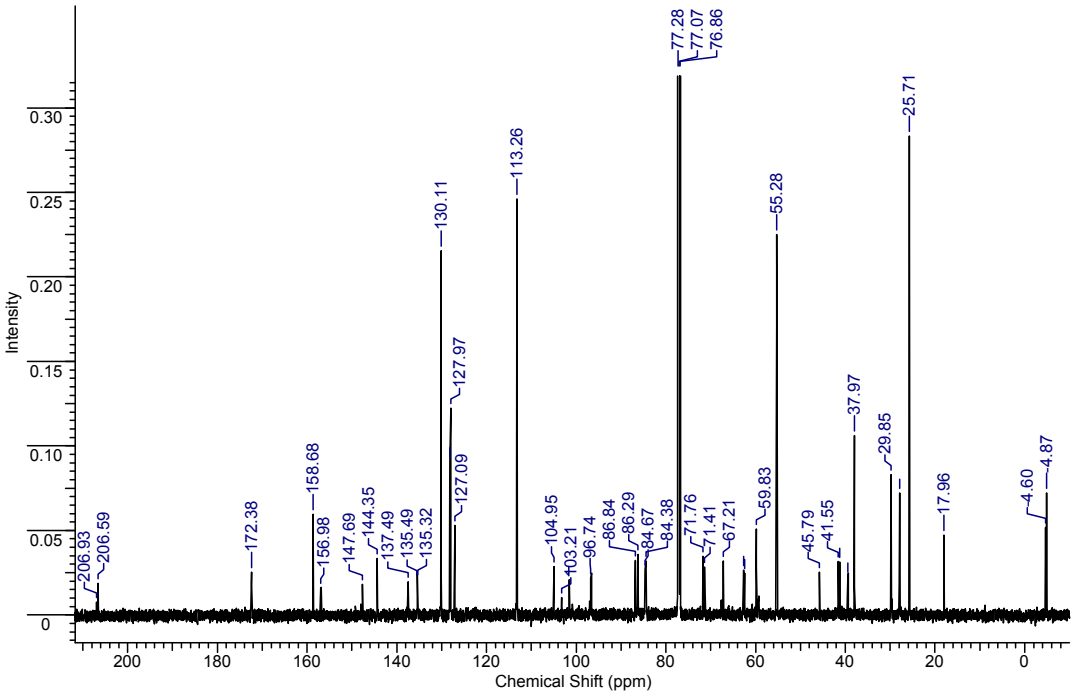


Figure S3. 150 MHz ^{13}C NMR spectrum of compound **6.**

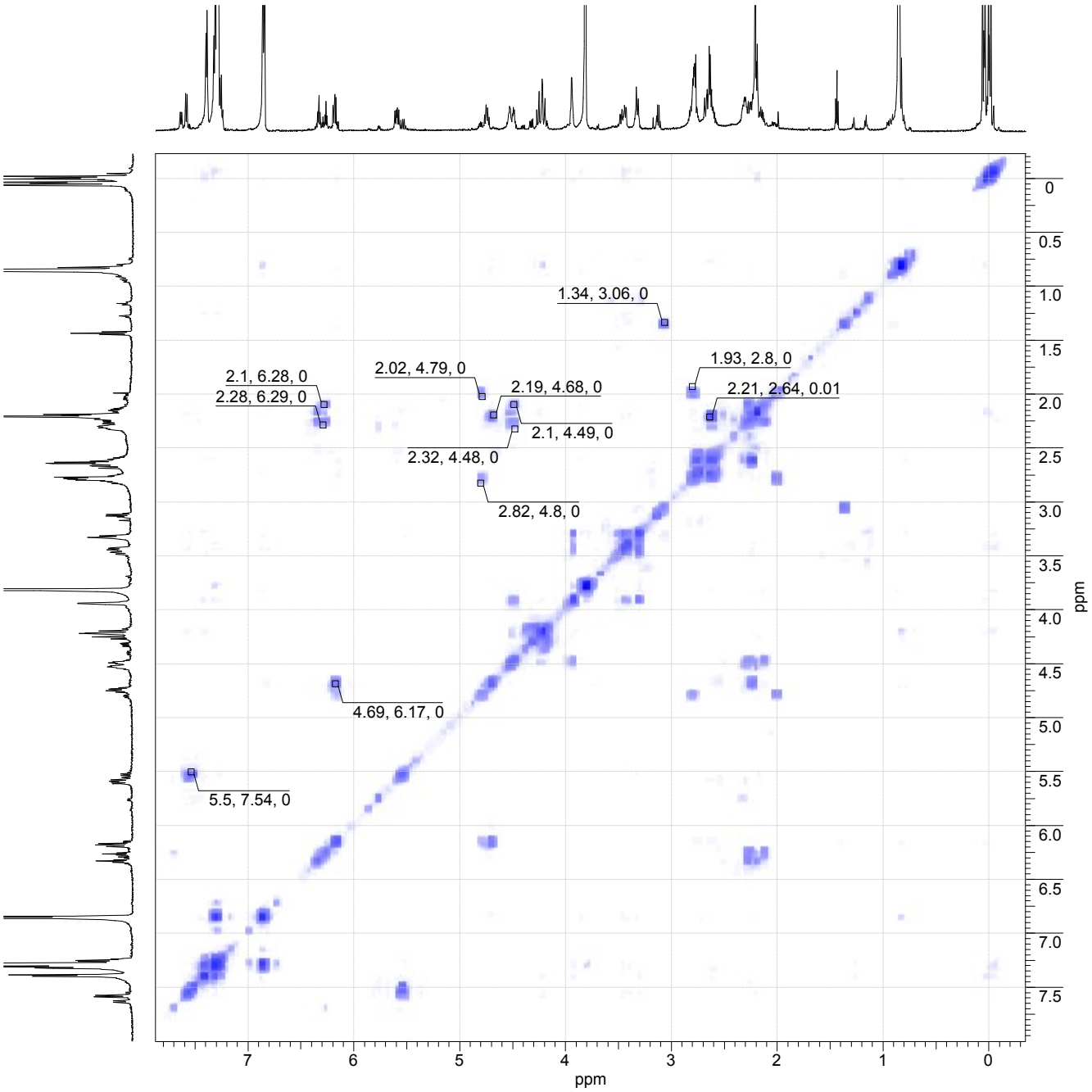


Figure S4. ^{1}H-^{1}H COSY NMR correlation spectroscopy 2D NMR spectrum of compound **6**.

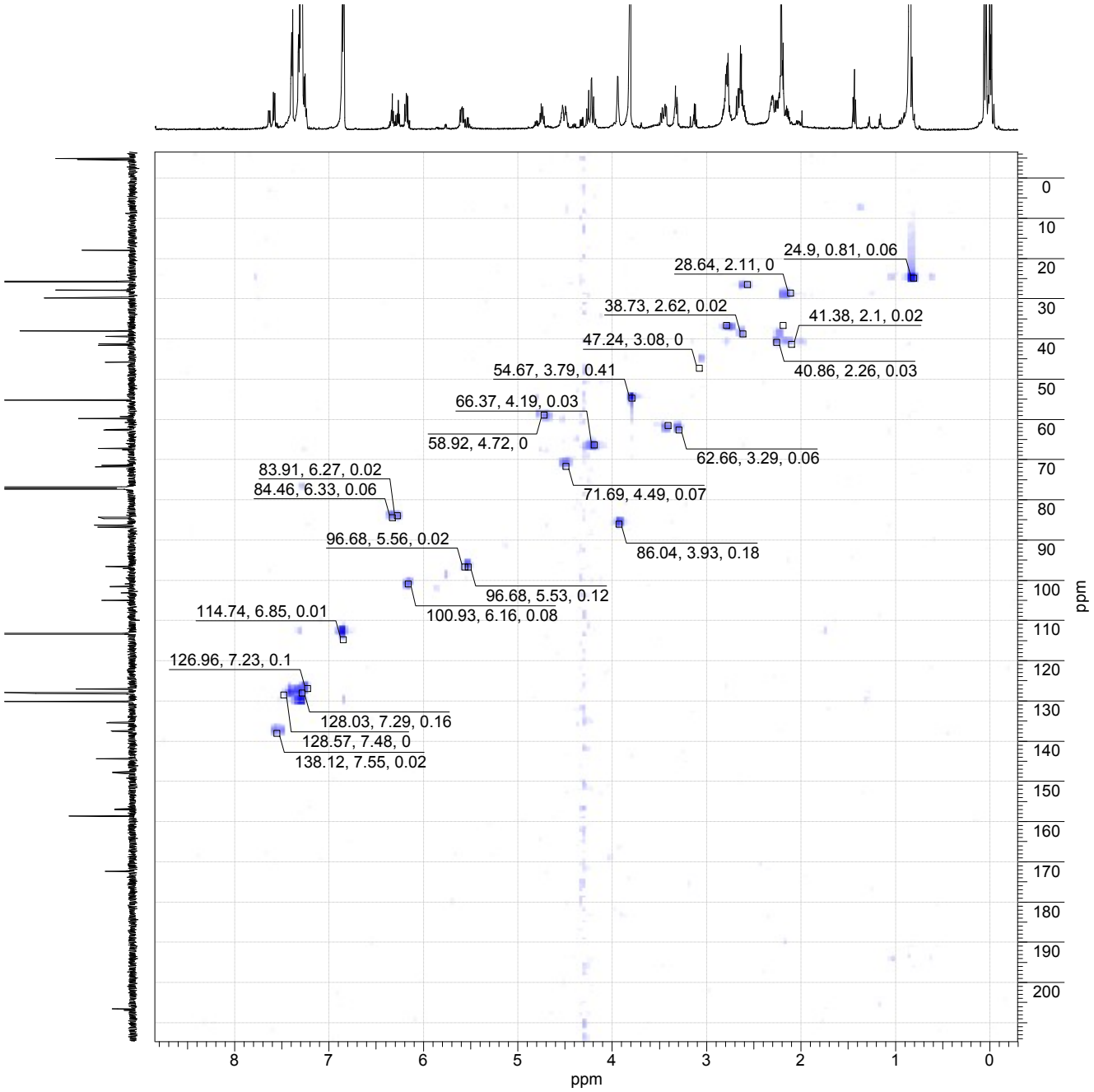


Figure S5. ^{1}H-^{13}C HSQC NMR correlation spectroscopy 2D NMR spectrum of compound **6**.

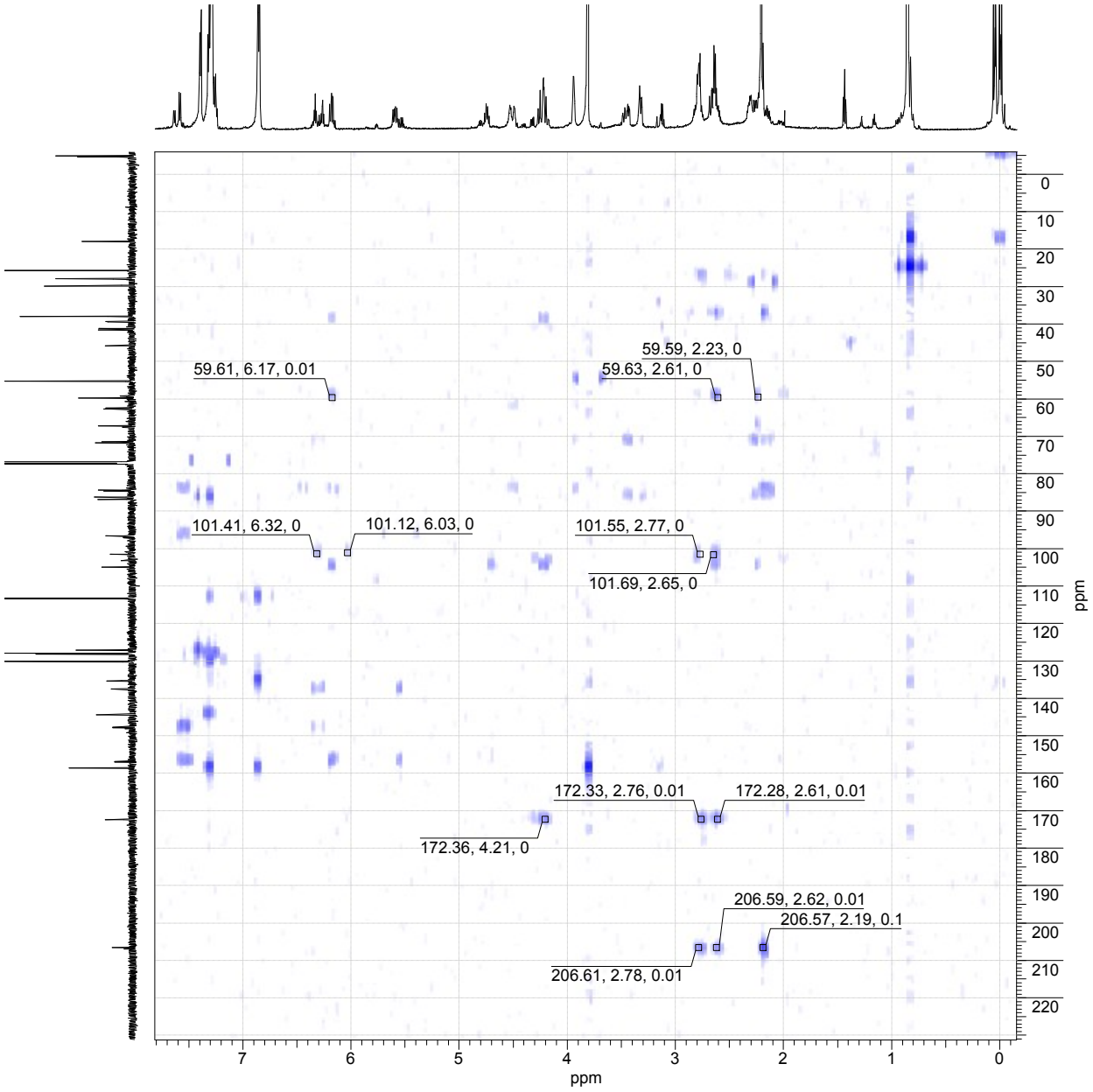
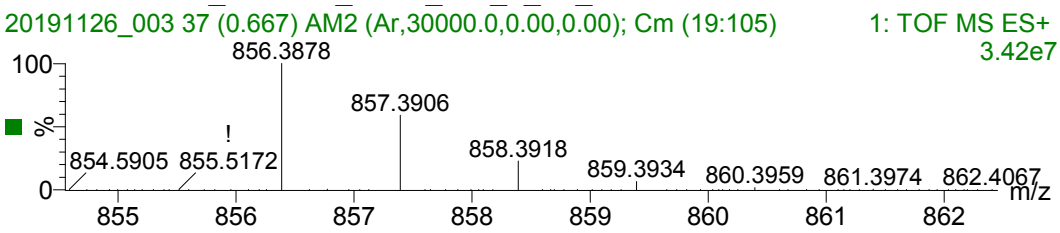


Figure S6. ^{1}H-^{13}C HMBC NMR correlation spectroscopy 2D NMR spectrum of compound **6**.



Compound #	Proposed molecular formula	Theoretical $[M+H]^+$	Experimental $[M+H]^+$	Δ (ppm)
#6	$C_{46}H_{57}N_3O_{11}Si$	856,3840	856,3878	4

Figure S7. High Resolution Mass spectrum (HRMS) of compound **6.**

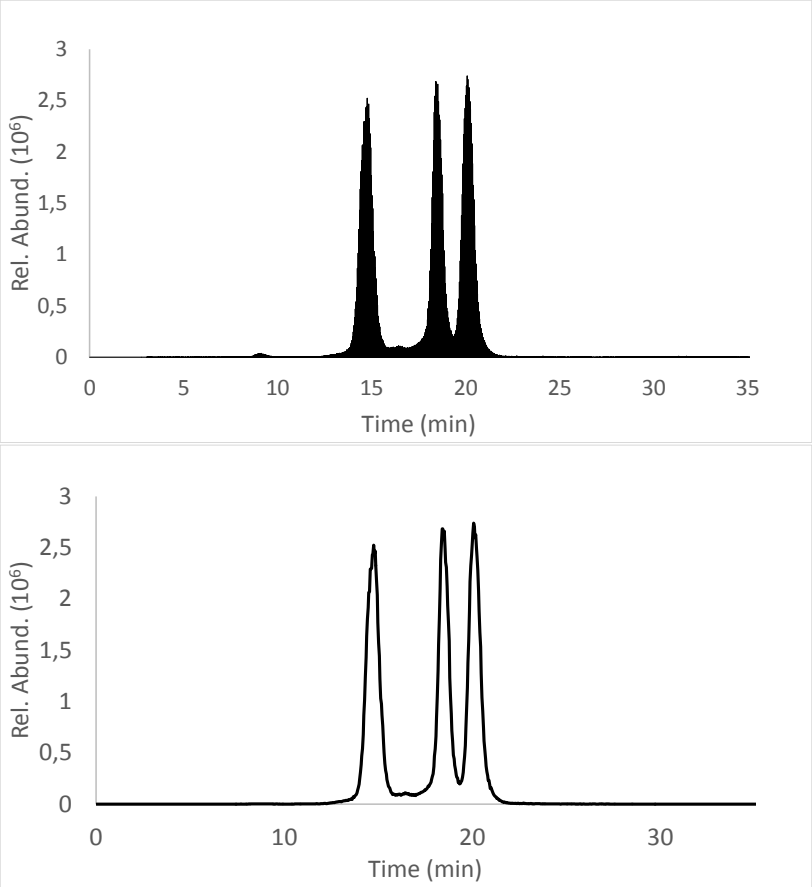


Figure S8. HPLC/MS-MS analysis of alkaline treatment of **6**. The upper part shows the TIC (total ion chromatogram) while the lower part shows the extracted ion chromatogram of m/z = 342. The three characteristic peaks of the four isomers of dCyd341 are observed (the two fastest isomers co-eluted together in a single peak).

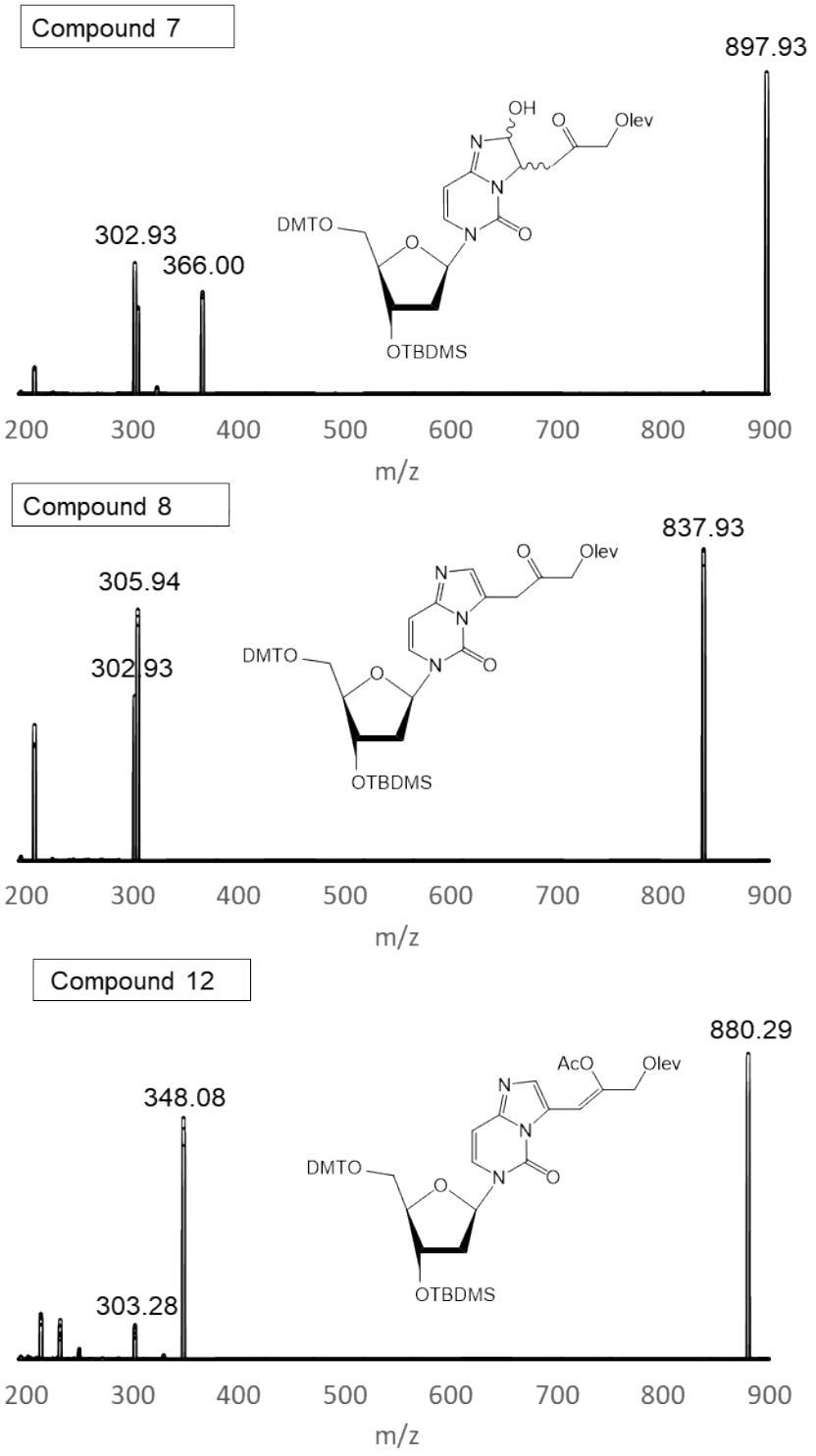
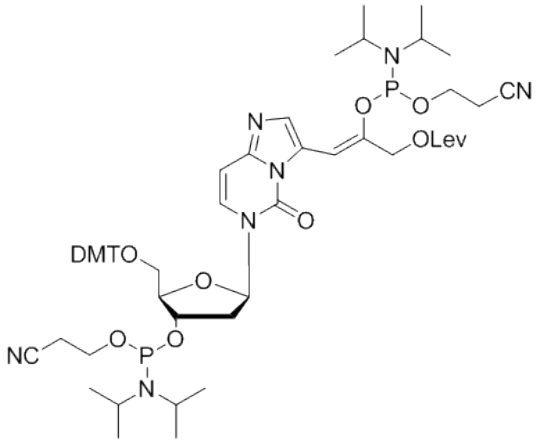


Figure S9. MS/MS spectra of compounds **7**, **8** and **12** obtained by collision-induced fragmentation of their parent ions m/z = 898, 838, 880 for **7**, **8** and **12**, respectively.



13

Figure S10. Chemical structure of compound **13.**

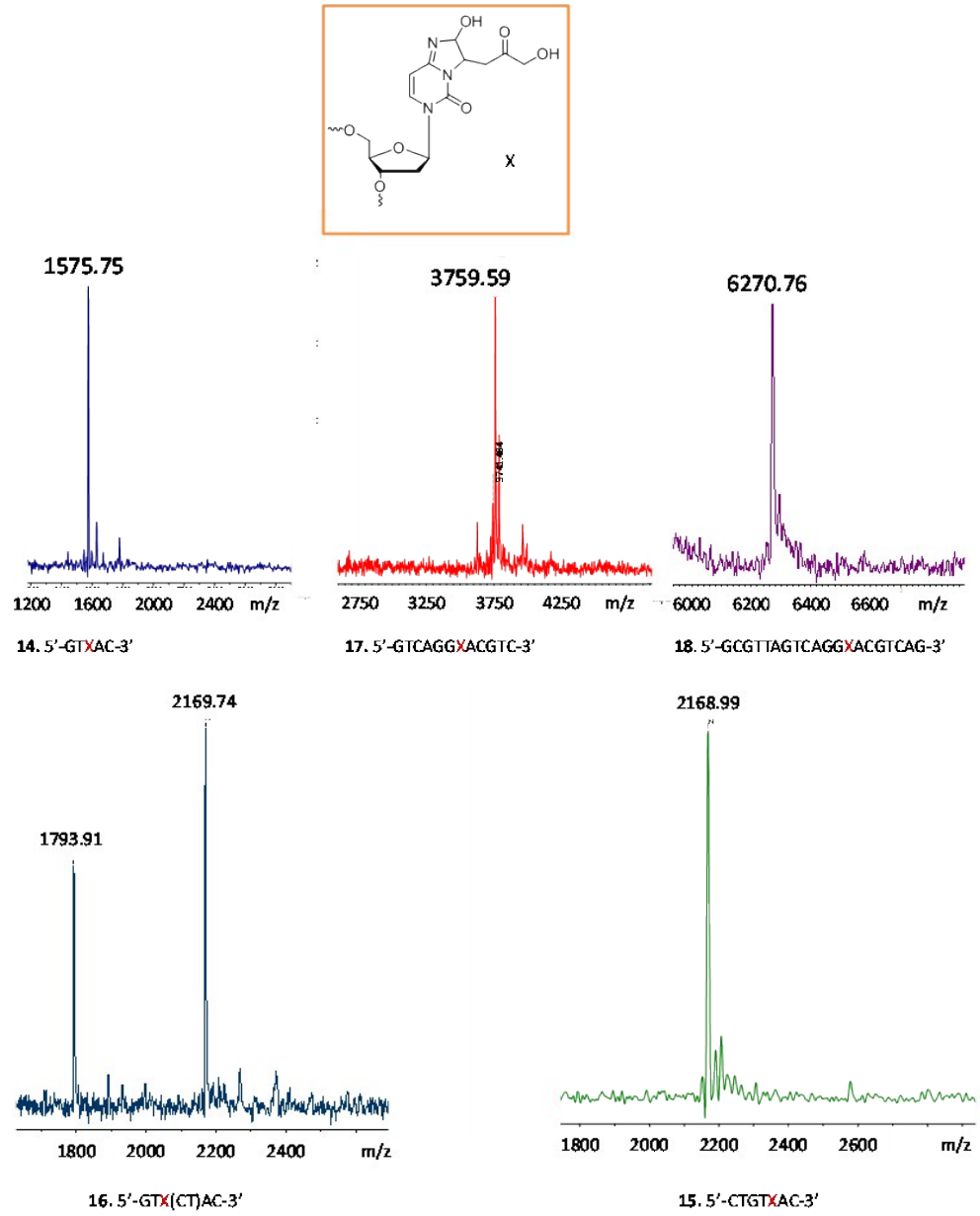


Figure S11. MALDI-TOF mass spectra of linear (**14**, **15**, **17** and **18**) and branched (**16**) oligonucleotides. Additional peak at m/z = 1794 observed for the analysis of **16** was not attributed.

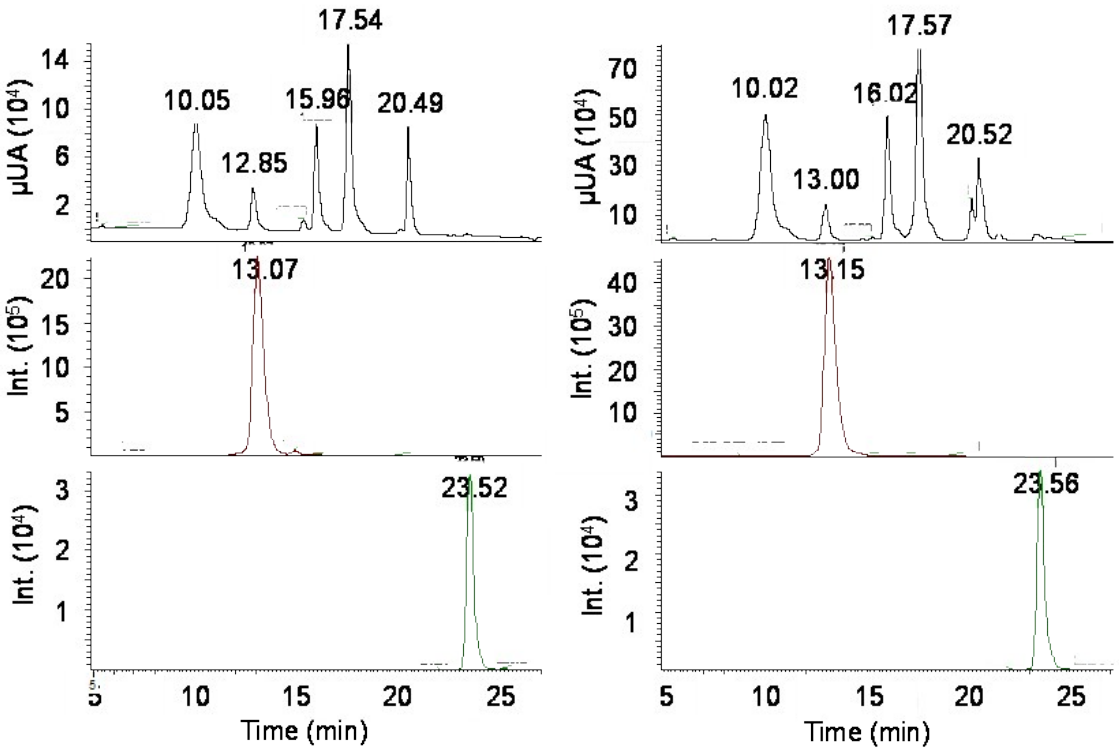
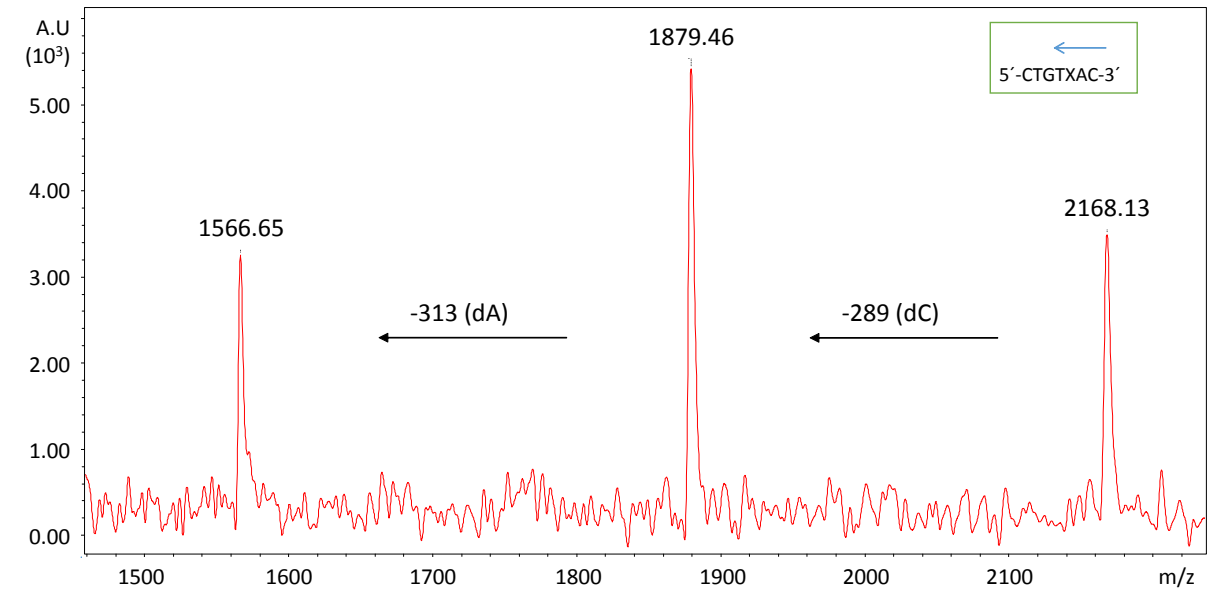


Figure S12. HPLC/MS-MS analysis of the complete enzymatic digestion of oligonucleotides **15** (linear, left panel) and **16** (branched, right panel).The upper chromatogram show the UV detection of all 4 nucleosides plus the modified nucleoside dCyd341 (RT 13 min). The middle show the ion extracted chromatogram of m/z 342 (dCyd341) while the lower show the extraction ion chromatogram of m/z 324 (dehydration product of dCyd341). The only detected modified nucleoside in the UV chromatogram is the dCyd341 indicating that formation of the dehydration product is negligible.

A



B.

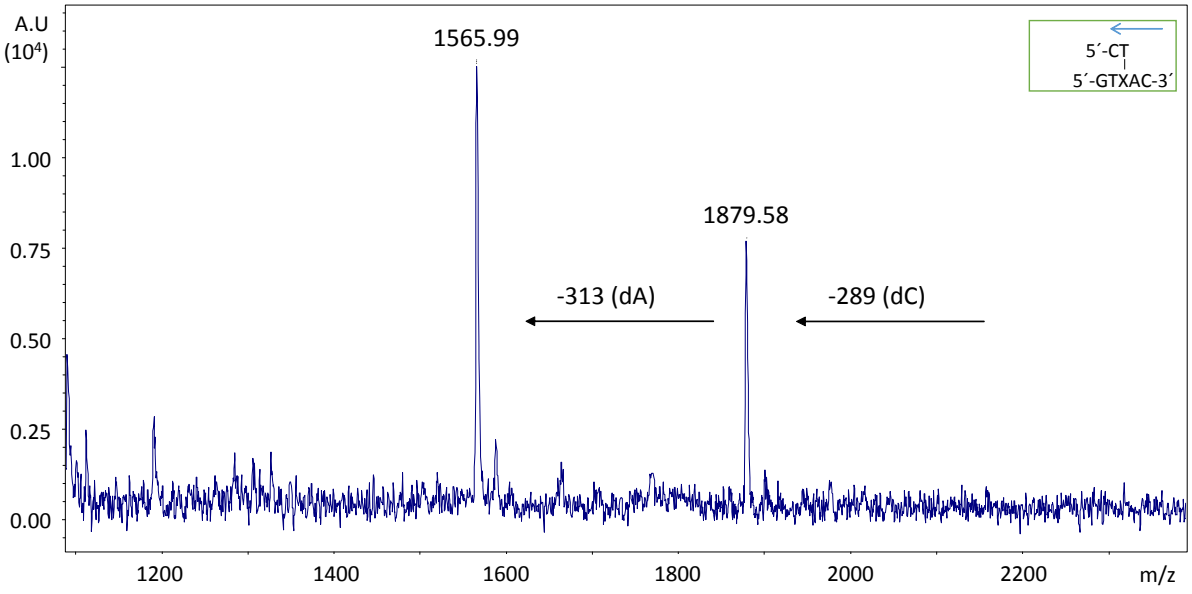
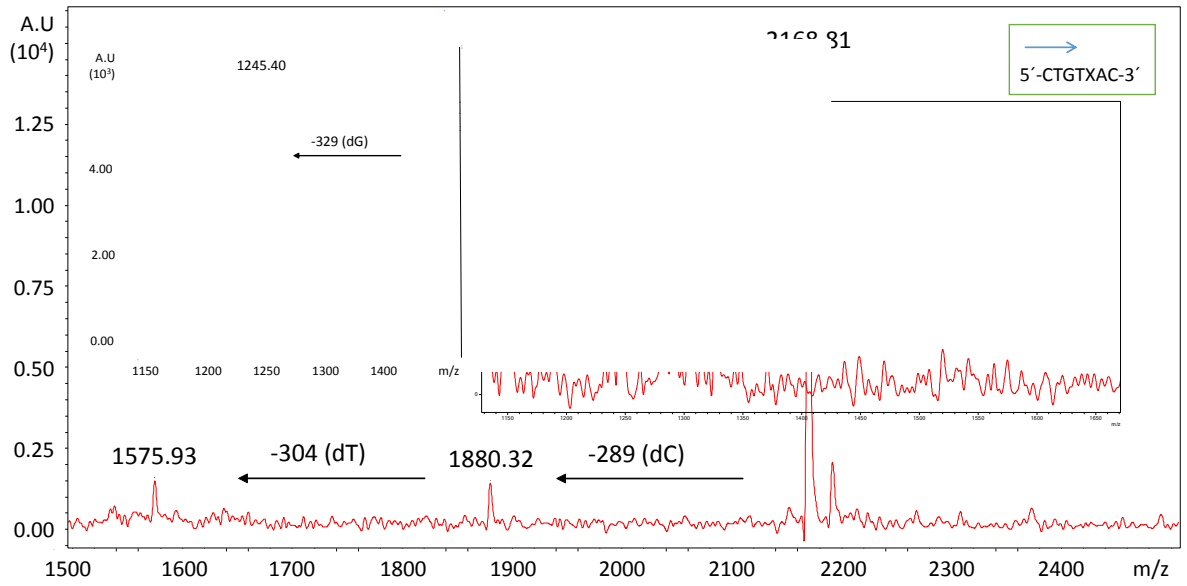


Figure S13. MALDI-TOF/MS analysis of 3´-5´exonuclease digestion of: A. Non-branched oligonucleotide **15** ; B. Branched oligonucleotide **16**. Spectrum A was obtained after 15 min digestion, and spectrum B was obtained after 5 min digestion.

A.



B.

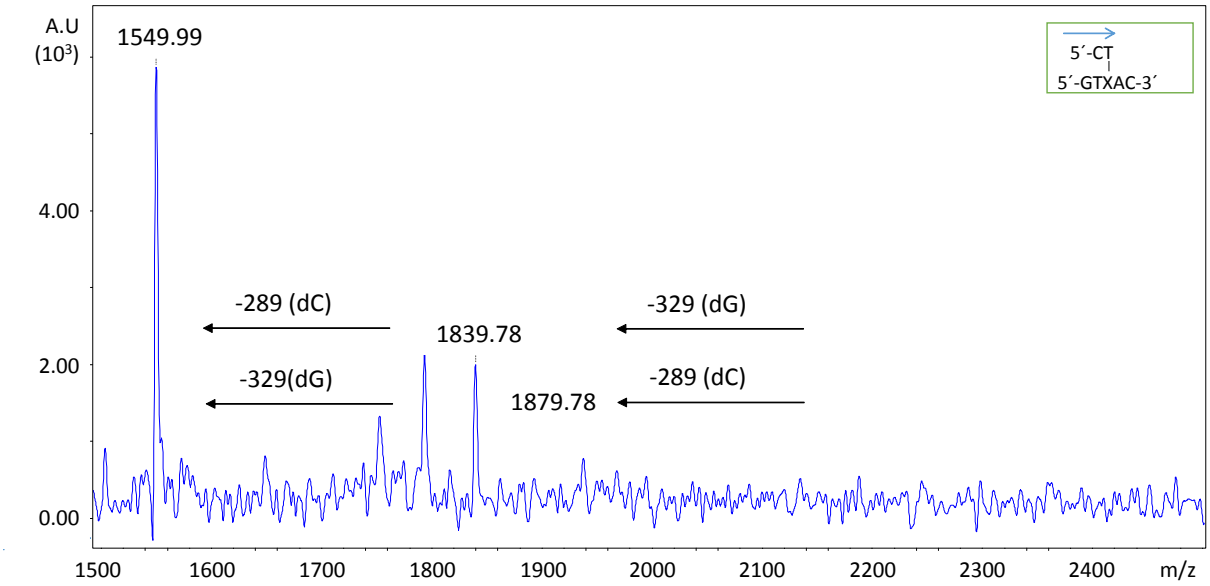


Figure S14. MALDI-TOF/MS analysis of 5´-3´exonuclease digestion of: A. Non-branched oligonucleotide **15** ; B. Branched oligonucleotide **16**. Spectrum A was obtained after 15 min digestion, and spectrum B was obtained after 5 min digestion.

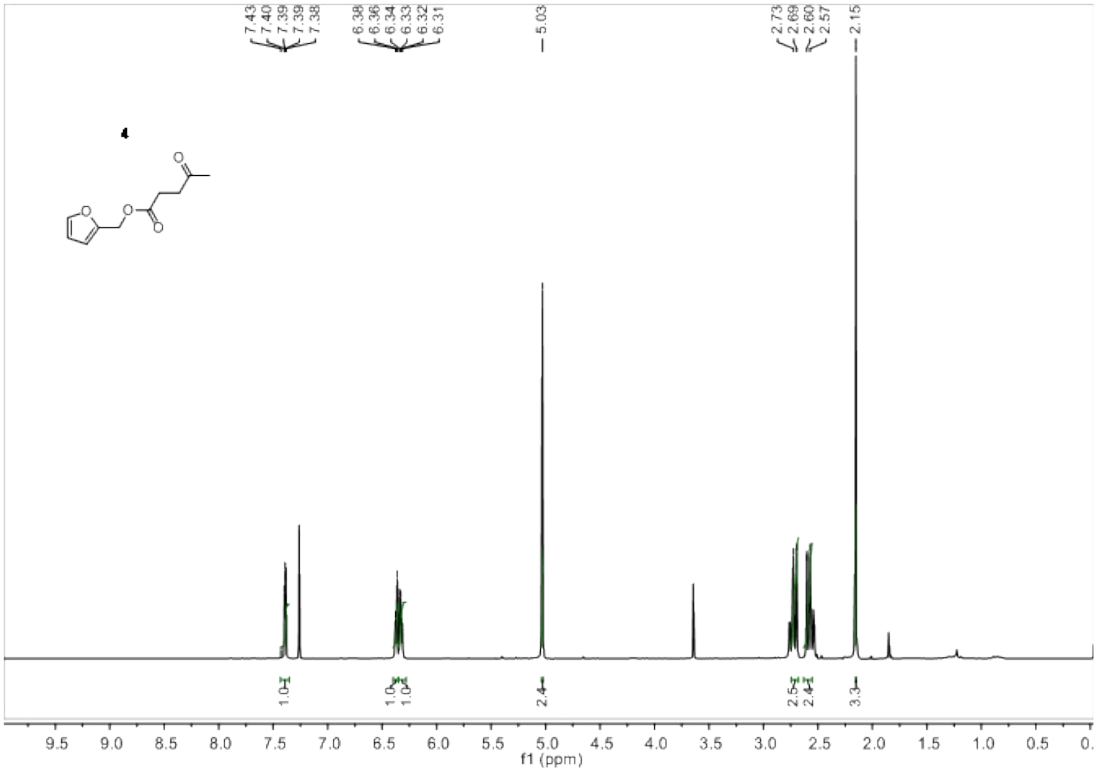


Figure S15.200 MHz ^{1}H NMR spectrum of compound **4.**

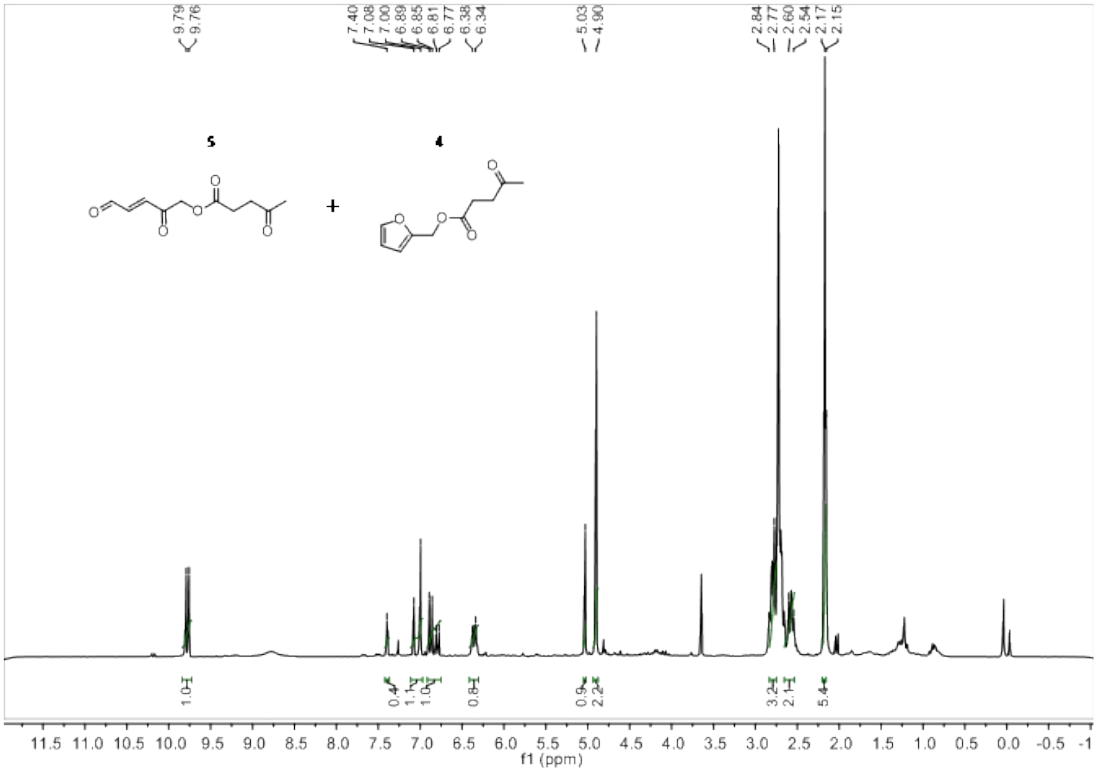


Figure S16. 200 MHz ^{1}H NMR spectrum of compound **4/5.**

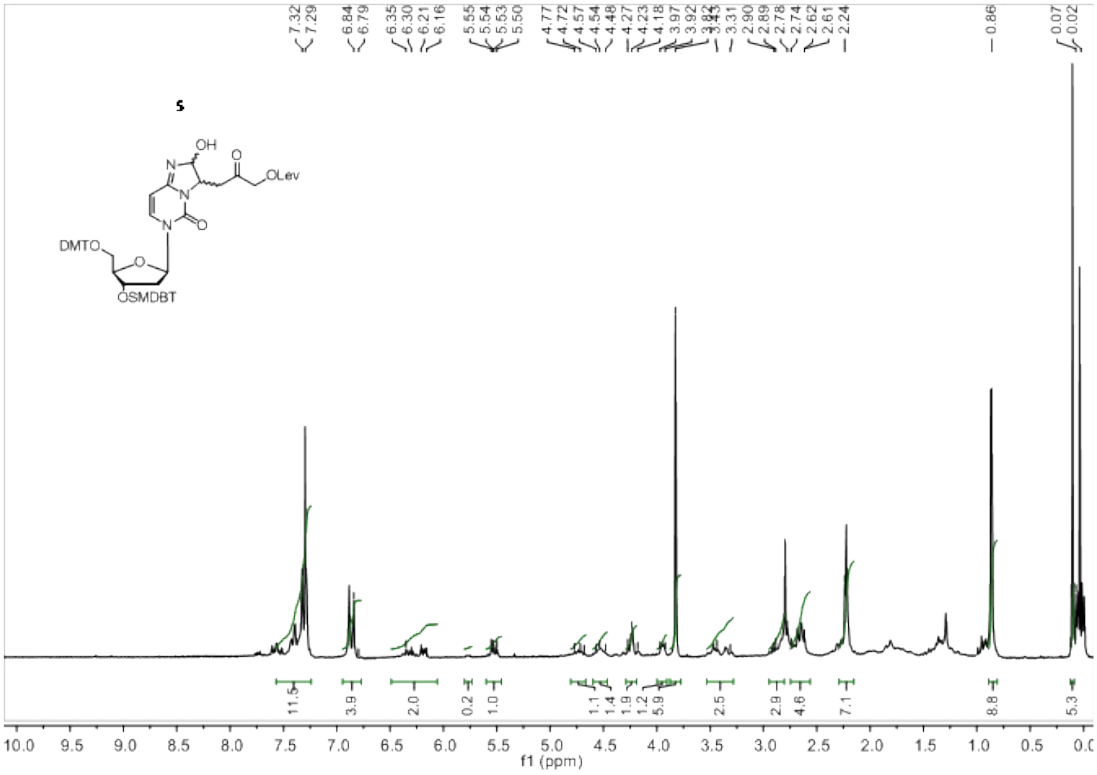


Figure S17.[1]H NMR spectrum of compound **6.**

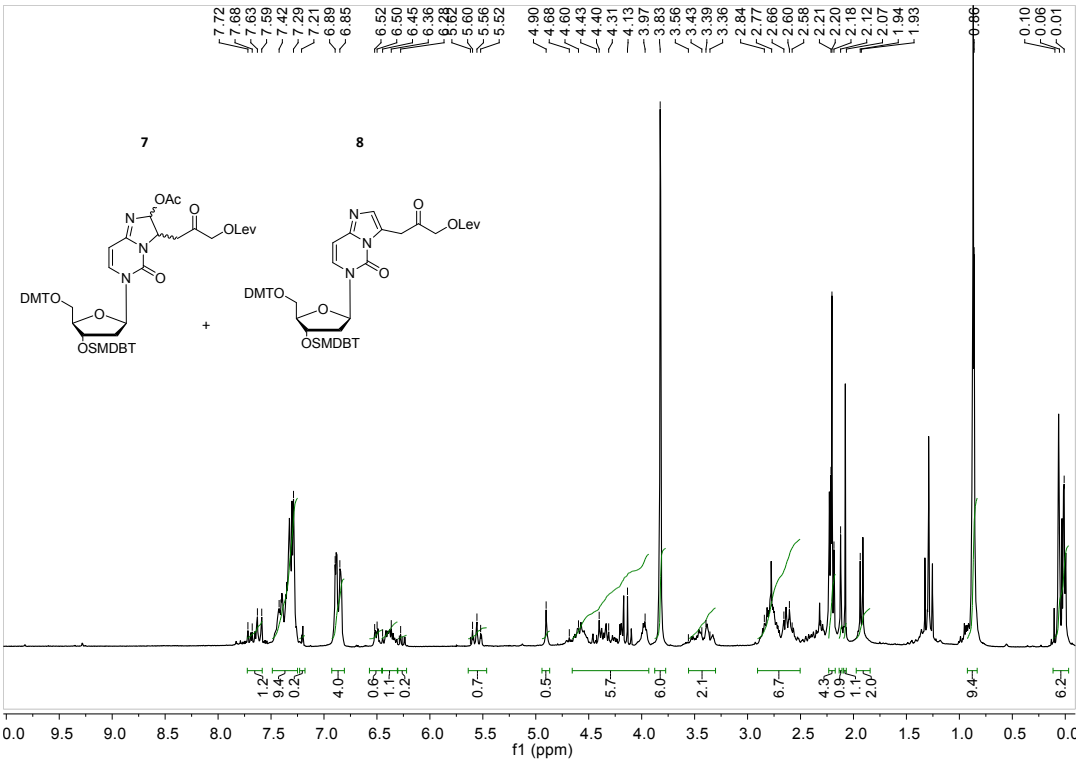


Figure S18. 200 MHz ^{1}H NMR spectrum of compound **7/8.**

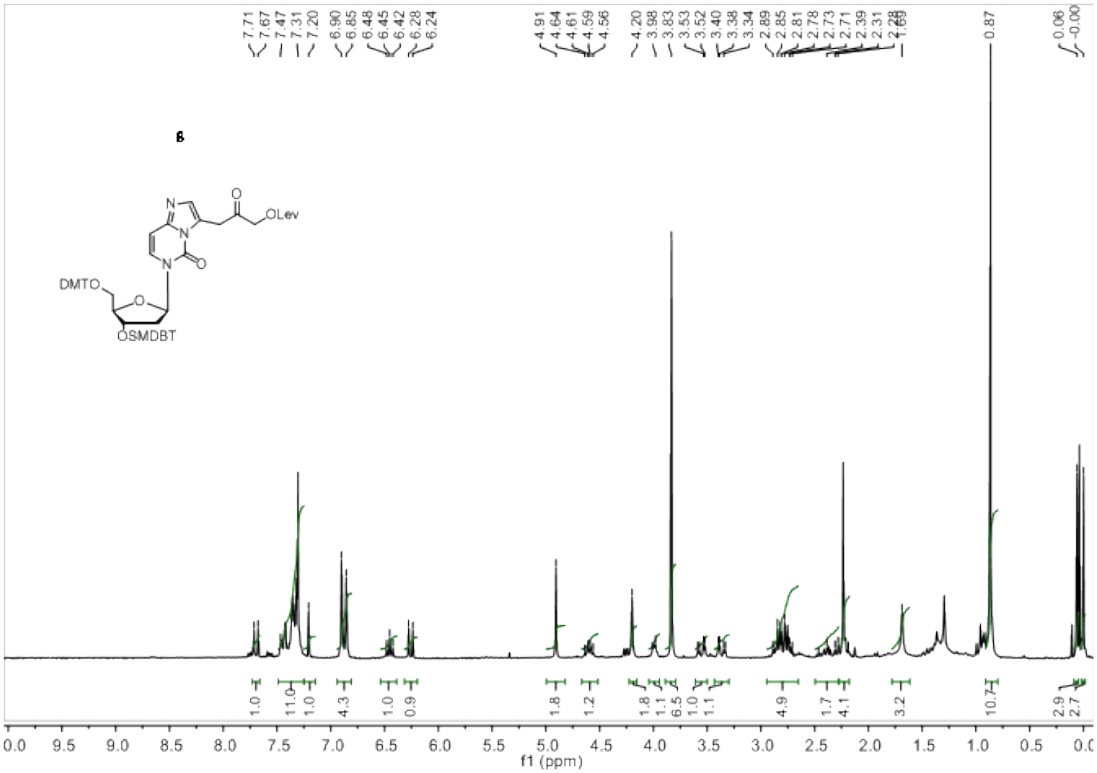


Figure S19. 200 MHz ^{1}H NMR spectrum of compound **8.**

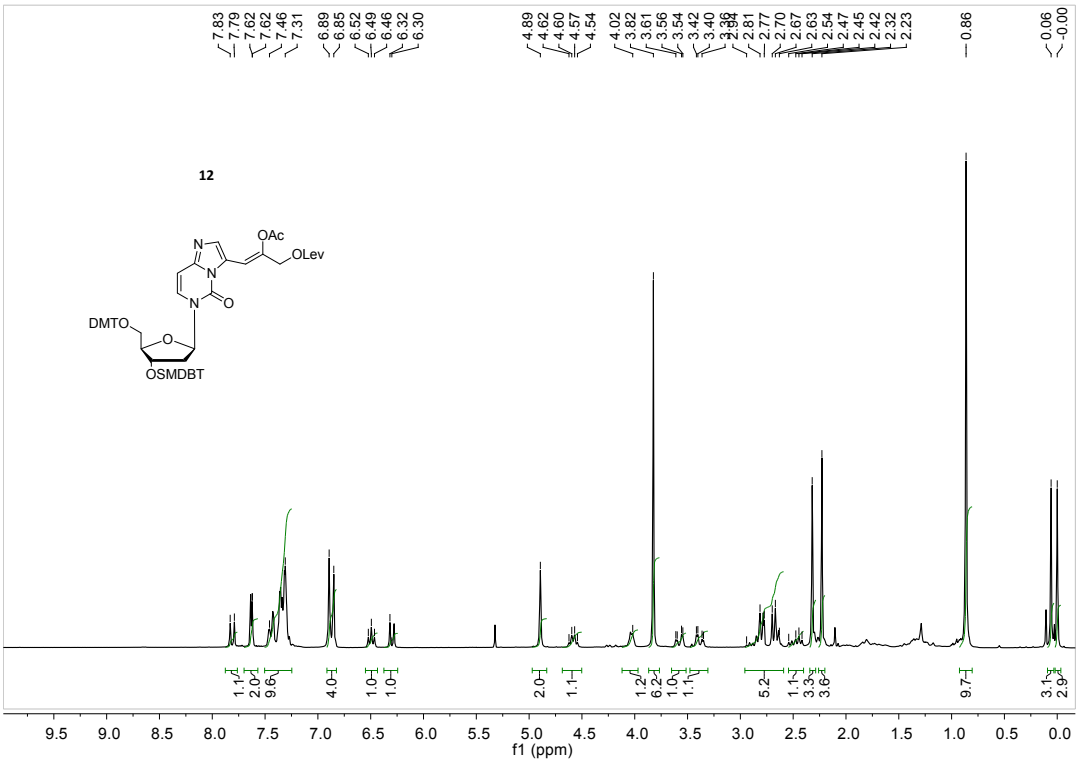


Figure S20. 200 MHz ^{1}H NMR spectrum of compound **12.**

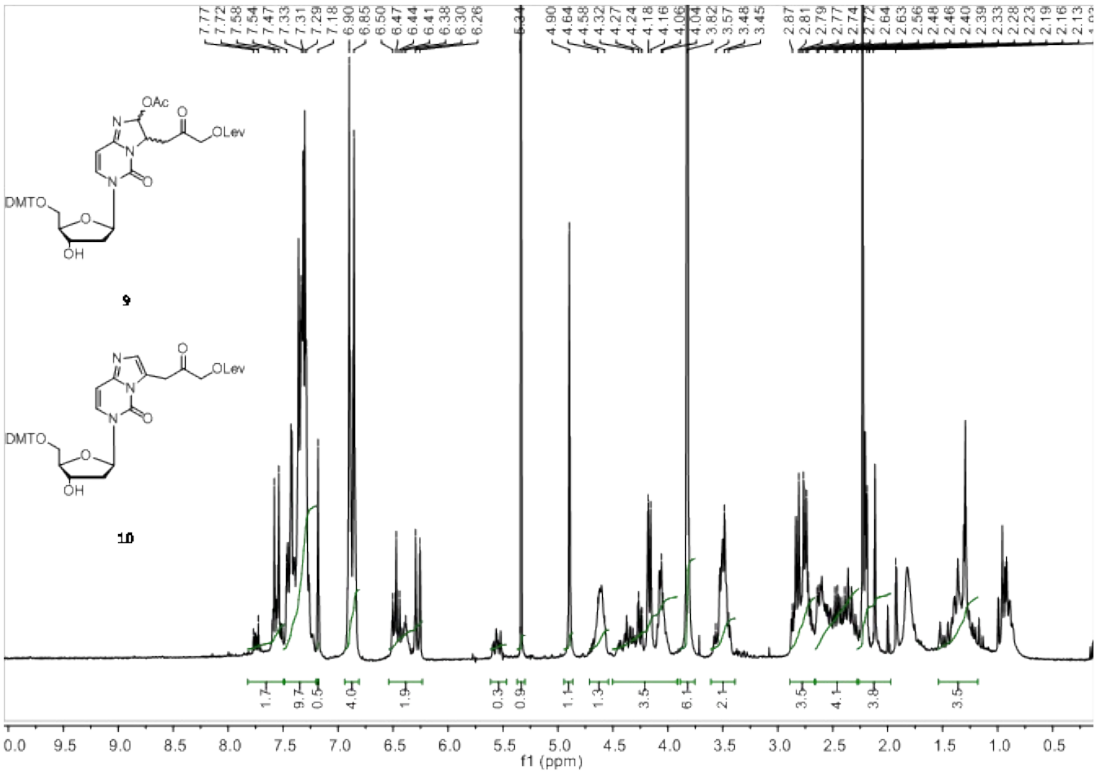


Figure S21. 200 MHz ^{1}H NMR spectrum of compound **9/10.**

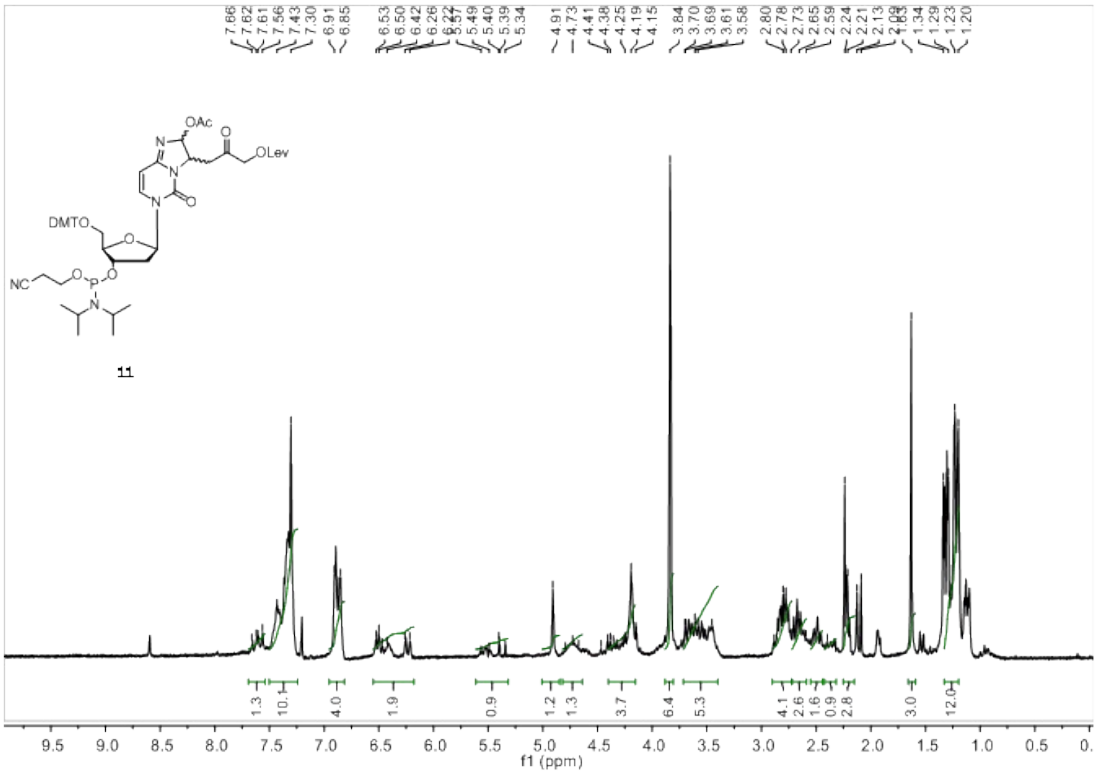


Figure S22. 200 MHz ^{1}H NMR spectrum of compound **11.**

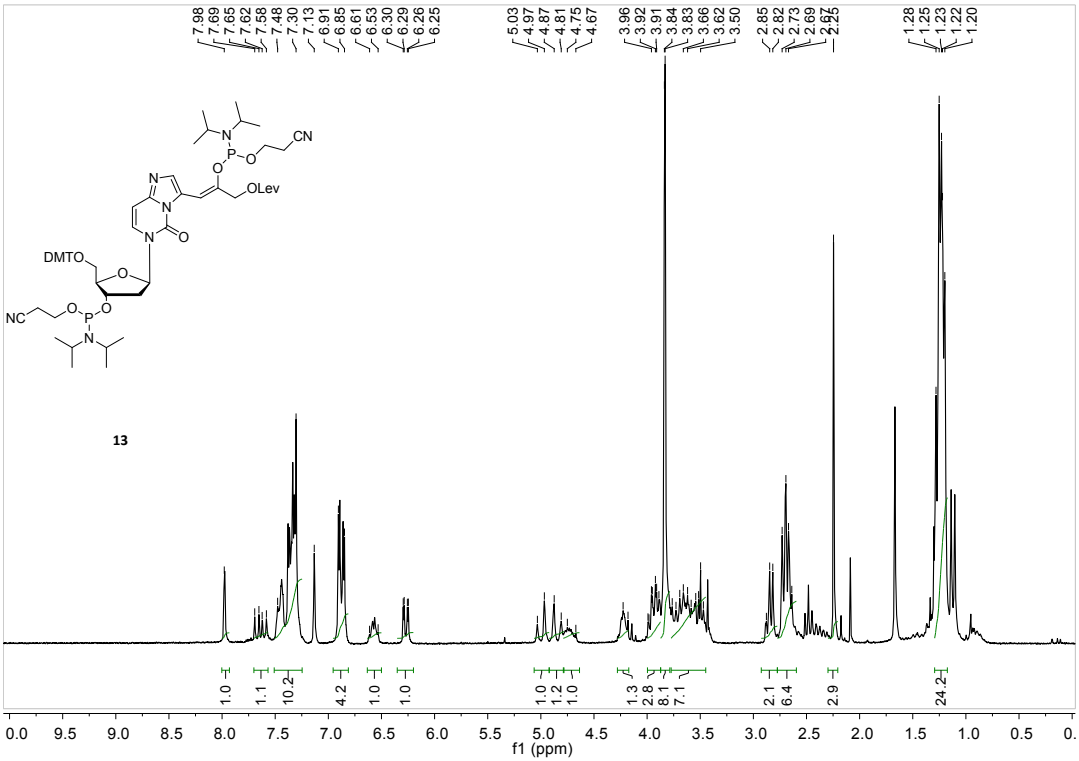


Figure S23. 200 MHz ^{1}H NMR spectrum of compound **13.**