Triplet-Sensitized Cyclobutane Pyrimidine Damage and Crosslinks in DNA: Filling the Triplet Energy Gap Between Xanthone and Thioxanthone

Sebastian Häcker,[a] Julian Amir Moghtader,[b] Christoph Kerzig,[b]

and Hans-Achim Wagenknecht*[a]

[a]Institute of Organic Chemistry, Karlsruhe Institute of Technology (KIT), Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany.

[b]Department of Chemistry, Johannes Gutenberg University Mainz, Duesbergweg 10-14, 55128 Mainz.

*Corresponding Author E-Mail: Wagenknecht@kit.edu

Supporting Information

Table of Contents

1. General information

Materials

All chemicals were purchased from Abovchem, Aldrich, Alfa Aesar,ABCR, BLDpharm, Fluka, Sigma-Aldrich, TCI. Unmodified and ATTO550-modified oligonucleotides were obtained from Metabion. Flash chromatography was performed using silica gel 60 (43 – 60 µm) from Aldrich. Doubly distilled water (ddH_2O) was collected from a Merck Milli-Q Direct 8 system.

Argon (5.0, purity: 99,999 %) was used for air- and moisture-sensitive reactions as well as for removing dissolved oxygen prior to all irradiation experiments.

NMR spectroscopy

^{1}H NMR (400 MHz), ^{13}C NMR (101 MHz), and ^{31}P NMR (162 MHz) spectra were measured on a Bruker Avance 400. The chemical shifts in the ^{1}H and ^{13}C NMR spectra are reported in parts per million (ppm) relative to tetramethylsilane as an internal standard. The chemical shifts in the ^{31}P NMR spectra are also reported in ppm relative to the deuterated solvent. The coupling constant (*J*) is given in Hertz (Hz), and the multiplicity of signals are reported as follows: s (singlet), d (doublet), t (triplet), m (multiplet), dt (doublet of triplets), td (triplet of doublets), dd (doublet of doublets), tt (triplet of triplets) and ddd (doublet of doublet of doublets).

Mass Spectrometry

Mass spectrometry was performed on a Finnigan MAT 95 with FAB and EI as ionization methods. ESI mass spectrometry was performed on a Thermo Fisher Scientific Q Exactive (Orbitrap). Oligonucleotides were identified by MALDI mass spectrometry using an AXIMA Confidence spectrometer from Shimadzu. The matrix contained 3-hydroxypicolinic acid (in MeCN/ddH_2O 1:1) and diammonium hydrogen citrate (0.44 M in ddH_2O) in a 9:1 ratio.

HPLC

A Thermo Scientific Dionex Ultimate 3000 HPLC system, equipped with an autosampler, pump module, column oven, multi-diode array detector, fluorescence detector and fraction collector, was used for the semi-preparative purification of oligonucleotides as well as for their analytical characterization. A Supercosil LC 318 column (25 cm x 10 mm, 5 µm) from VDS Optilab was used for the semi-preparative purification. A VDSpher OptiBio Pur 300 S18-SE from VDS Optilab was used for the analytical determination of the oligonucleotides.

Optical Spectroscopy

Absorption spectra and melting temperatures of DNA double strands (2.5 µM DNA, 250 mM NaCl, 10 mM Na-P_i buffer, 10-90 °C, 0.5 °C/min, step width 0.5 °C) were recorded with a Cary 3500 Multicell UV/Vis spectrometer.

2. DNA synthesis and purification

Reagents and CPG (1 µmol) were purchased from ABI and GlenResearch. Oligonucleotide synthesis was carried out under an argon atmosphere using an H-6 DNA/RNA synthesizer from K&A Laborgeräte. The coupling protocol was modified for the incorporation of thioxanthone- and naphthalene-phosphoramidites (Table S2). The modified oligonucleotides were synthesized as trityl-on oligonucleotides and pre-purified using Glen-Pak DNA Purification Cartidges before purification by semi-preparative HPLC with the following conditions: mobile phase A = NH_4OAc buffer (50 mM, pH 7), mobile phase B = acetonitrile, flow rate 2.5 mL/min, UV/Vis detection at 260, 280 and 310/385 nm. After purification, the oligonucleotides were lyophilized, and their concentrations were determined by absorbance at 260 nm using a Nanodrop ND-1000 spectrophotometer. Analytical characterization of purified oligonucleotides was performed via RP-HPLC under the following conditions: mobile phase A = NH_4OAc buffer (50 mM, pH 7), mobile phase B = acetonitrile (gradient: 1-45% B), flow rate = 1.0 mL/min, UV/Vis detection at 260, 280 and 310/385 nm.

Table S1: Coupling protocol with the modified phosphoramidites **14** and **18**.

	Time [0.1 s]	**Source**	**Mixed**	**Destin**	***S.Col.Ptr.***	**Lag time [s]**	***Branch***
				BRANCH			
1	4	TET		COL	ON		
2	6	AMD	TET	COL			
3					ON		
4						99	
5						99	
6						99	
7						99	
8						99	
9						99	
10						99	
11						99	
12	6	AMD	TET	COL	ON		
13					ON		
14						99	
15						99	
16						99	
17						99	
18						99	
19						99	
20						99	
21						99	
22	4	TET		COL	ON		
23					ON		
24	20	ACN		M_W			
25	20	GAS		M_W			

3. Irradiation Experiments and Analysis

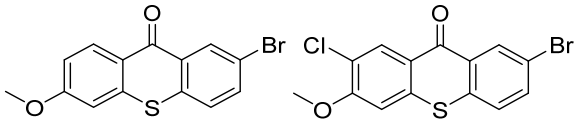
Irradiation experiments were performed with two Nichia UVA-LEDs (NCSU033B) at 365 and 385 nm under the following conditions: 2.50 µM DNA, 10 mM Na-P_i buffer, 250 mM NaCl in ddH_2O (1 ml) at 10 °C after degassing under an argon atmosphere. HPLC analysis was performed at Thermo Scientific Dionex UltiMate on a RP VDSpher OptiBio PUR 300 C18-SE 250 ×4.6 mm column using the following conditions (Table S2). Solution A= NH4OAc buffer (50 mM; pH = 6.5) and B = acetonitrile.

Table S2: HPLC method for the analysis of DNA damage. A = NH_4OAc buffer (50 mM; pH = 6.5) and B = acetonitrile.

DNA	Gradient [% B in A]	Run time [min]	Temperature [°C]	λ [nm]
DNA1-0-1, DNA2-0-1	1-45	75	40	260, 290, 554, 365/385
DNA2-2-3	1-42	75	40	260, 290, 554, 365/385

4. Synthesis

Synthesis of **5** and **6**



5 **6**

In a three-necked flask equipped with a reflux condenser and thermometer, 1.00 g of **3** (5.43 mmol, 1.00 eq.) and 12.6 µl of DMF (11.9 mg, 0.163 mmol, 0.0300 eq.) were dissolved in 10.0 ml of bromobenzene and heated to 113 °C. After 30 min, 1.97 ml of thionyl chloride (3.23 g, 27.2 mmol, 5.00 eq.) was added. The solution was then cooled to 5 °C, and 0.482 ml of sulfuryl chloride (0.806 g, 5.97 mmol, 1.10 eq.) was added over 30 min. After gas evolution ceased, the solution was heated to 60 °C until gas evolution ceased again. After cooling to 20 °C, 1.59 g of aluminum chloride (11.9 mmol, 2.20 eq.) was slowly added. The reaction mixture was stirred for 30 min, then heated to 60 °C. After cooling to 20 °C, the precipitate was filtered off and washed with hexane. The filter cake was added to 30 ml of 5 M NaOH and heated to 80 °C for 1 hour. After cooling to 20 °C, the precipitate was filtered, washed with water, and dried in vacuo. The crude product was recrystallized from ethyl acetate, yielding 1.03 g of compound 6 (2.89 mmol, 53%) as a yellowish solid as the main product, and 0.227 g of compound 5 (0.707 mmol, 13%) as a colorless solid as a side product.

Analysis of **5**:

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.72 (d, *J* = 2.3 Hz, 1H, C*H*$_{Ar}$), 8.54 (d, *J* = 9.0 Hz, 1H, C*H*$_{Ar}$), 7.68 (dd, *J* = 8.5, 2.3 Hz, 1H, C*H*$_{Ar}$), 7.42 (d, *J* = 8.5 Hz, 1H, C*H*$_{Ar}$), 7.06 – 7.03 (m, 1H, C*H*$_{Ar}$), 6.96 (d, *J* = 2.5 Hz, 1H, C*H*$_{Ar}$), 3.93 (s, 3H, C*H*$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ 162.83, 134.90, 132.41, 132.18, 127.35, 120.24, 115.44, 108.22, 55.78.

HR-MS (ESI, $[M+H]^+$) m/z: found: 320.9578; calc.: 320.9579.

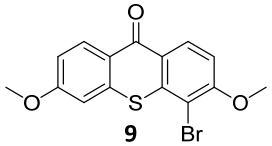
Analysis of **6**:

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.73 (d, *J* = 2.3 Hz, 1H, C*H*$_{Ar}$), 8.61 (s, 1H, C*H*$_{Ar}$), 7.72 (dd, *J* = 8.5, 2.3 Hz, 1H, C*H*$_{Ar}$), 7.45 (d, *J* = 8.6 Hz, 1H, C*H*$_{Ar}$), 6.98 (s, 1H, C*H*$_{Ar}$), 4.03 (s, 3H, C*H*$_3$).

^{13}C NMR (126 MHz, $CDCl_3$) δ 177.19, 158.44, 137.59, 135.32, 132.65, 131.74, 130.33, 127.57, 123.26, 123.21, 120.72, 107.34, 56.85.

HR-MS (ASAP, $[M+H]^+$) m/z: found: 354.9166; calc.: 354.9190.

Synthesis of **9**



9

0.300 g of thioxanthone **8** (1.10 mmol, 1.00 eq.) was stirred in 12.0 ml of acetic acid for 1 hour. Then, 0.400 g of benzyltrimethylammonium tribromide (1.10 mmol, 1.00 eq.) and 0.200 g of zinc chloride (1.10 mmol, 1.00 eq.) were added, and the mixture was stirred for 5 hours. The suspension was then heated to 80 °C for 1 hour. After cooling to room temperature, the precipitate was filtered off, washed with 12.0 ml of methanol, and dried. The product was recrystallized from ethyl acetate, yielding 0.200 g of **9** (0.569 mmol, 43%) as a yellowish solid.

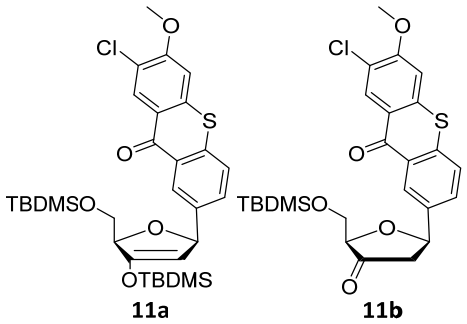
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.64 (d, *J* = 9.0 Hz, 1H, CH_{Ar}), 8.50 (d, *J* = 9.7 Hz, 1H, CH_{Ar}), 7.09 (d, *J* = 9.0 Hz, 1H, CH_{Ar}), 7.05 – 7.02 (m, 2H, CH_{Ar}), 4.04 (s, 3H, CH_3), 3.93 (s, 3H, CH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 178.77, 162.78, 158.91, 140.09, 139.73, 131.92, 131.31, 124.88, 122.03, 115.54, 110.03, 108.60, 107.73, 56.83, 55.89.

HR-MS (ESI, $[M+H]^+$) m/z: found: 350,9682; calc.: 350,9685.

The ^{1}H NMR spectrum contains 20% of the by-product 2-bromo-3,6-dimethoxythioxanthone, which could not be separated.

Synthesis of **11**



Under argon atmosphere 0.546 g **6** (1.54 mmol, 1.00 eq.), 1.06 g **10** (3.07 mmol, 2.00 eq.), and 0.459 ml of triethylamine (0.335 g, 3.31 mmol, 2.15 eq.) were dissolved in 10 ml of DMF. The reaction mixture was then degassed three times using the freeze-pump-thaw method. Subsequently, 141 mg of Pd_2dba_3 (0.154 mmol, 0.100 eq.) and 218 mg of Q-Phos (0.307 mmol, 0.200 eq.) were added, and the reaction was stirred at 75 °C for 72 hours. The solvent was removed under reduced pressure, and the crude product was purified by column chromatography (SiO_2, hexane/EtOAc, 10:1). This yielded 400 mg of **11a** (0.646 mmol, 42%) and 67.0 mg of **11b** (0.133 mmol, 9%) as colorless solids, with a total yield of 51%.

Analytics of 11a

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.56 (s, 1H, CH_{Ar}), 8.43 (d, J = 1.9 Hz, 1H, CH_{Ar}), 8.02 (dd, J = 8.3, 2.0 Hz, 1H, CH_{Ar}), 7.47 (d, J = 8.3 Hz, 1H, CH_{Ar}), 6.90 (s, 1H, CH_{Ar}), 5.83 (dd, J = 4.1, 1.5 Hz, 1H, C*H*), 4.83 – 4.81 (m, 1H, C*H*), 4.64 – 4.59 (m, 1H, C*H*), 3.97 (s, 3H, CH_3), 3.93 (dd, J = 11.4, 2.3 Hz, 1H, CH_2), 3.81 (dd, J = 11.3, 3.8 Hz, 1H, CH_2), 0.95 (s, 9H, CH_3), 0.89 (s, 9H, CH_3), 0.23 (s, 3H, CH_3), 0.21 (s, 3H, CH_3), 0.06 (s, 3H, CH_3), 0.05 (s, 3H, CH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 178.15, 157.95, 151.45, 142.30, 137.90, 135.94, 132.06, 131.46, 128.28, 127.66, 126.16, 123.50, 122.64, 107.22, 102.02, 84.42, 84.23, 64.20, 56.69, 27.01, 26.15, 25.70, 18.70, 18.19, -4.76, -4.87, -5.18.

HR-MS (ESI, $[M+H]^+$) m/z: found: 619.2133; calc.: 619.2131.

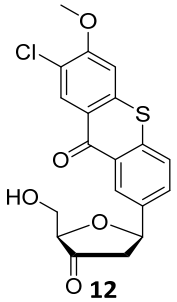
Analytics of 11b

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.61 (s, 1H, CH_{Ar}), 8.58 – 8.54 (m, 1H, CH_{Ar}), 7.70 (dd, J = 8.4, 2.0 Hz, 1H, CH_{Ar}), 7.60 (d, J = 8.4 Hz, 1H, CH_{Ar}), 6.98 (d, J = 1.5 Hz, 1H, CH_{Ar}), 5.75 (dd, J = 8.6, 7.0 Hz, 1H, C*H*), 4.26 – 4.22 (m, 1H, C*H*), 4.07 (dd, J = 11.0, 2.2 Hz, 1H, CH_2), 4.02 (s, 3H, CH_3), 4.01 – 3.95 (m, 1H, CH_2), 2.93 (dd, J = 18.1, 7.0 Hz, 1H, CH_2), 2.55 (dd, J = 18.1, 8.7, 1H, CH_2), 0.92 (s, 9H, CH_3), 0.10 (s, 3H, CH_3), 0.08 (s, 3H, CH_3).

^{13}C NMR (101 MHz, $CDCl_3$) δ 214.57, 158.24, 140.57, 137.85, 136.31, 131.65, 129.81, 128.82, 128.38, 128.33, 127.22, 126.63, 107.33, 81.89, 78.63, 64.98, 56.81, 45.24, 25.99, -5.32, -5.53.

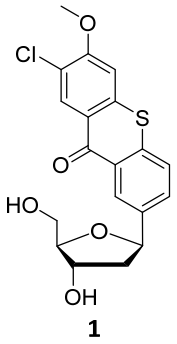
HR-MS (ESI, $[M+H]^+$) m/z: found: 505.1273; calc.: 505.1266.

Synthesis of **12**



Under argon atmosphere, 311 mg of **11** (0.502 mmol, 1.00 eq.) was dissolved in 10 ml of THF and cooled to 0 °C. Subsequently, 0.652 ml of $Et_3N \cdot 3HF$ (0.645 g, 4.00 mmol, 8.00 eq.) was added, and the mixture was stirred at 0 °C for 15 minutes. The reaction was then allowed to proceed overnight at room temperature. The solvent was removed under reduced pressure, and the crude product was used without further purification.

Synthesis of **1**



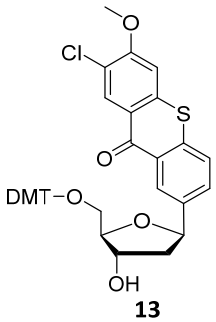
Under an argon atmosphere, 100 mg of **12** (0.256 mmol, 1.00 eq.) was dissolved in 8.00 ml AcOH/MeCN (1:3) and cooled to 0 °C. Subsequently, 434 mg of sodium triacetoxyborohydride (2.05 mmol, 8.00 eq.) was added, and the reaction was stirred at room temperature for 1 hour. The reaction was then quenched by adding 8 ml EtOH/H_2O (1:1). The solvent was removed under reduced pressure, and the crude product was purified by column chromatography (SiO_2, gradient: DCM → DCM/MeOH 20:1). Yielding 50.0 mg of **1** (0.146 mmol, 39%) as a colorless solid over two steps.

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.69 (s, 1H, CH_{Ar}), 7.63 (s, 1H, CH_{Ar}), 7.00 – 6.96 (m, 1H, CH_{Ar}), 6.84 (d, J = 8.3 Hz, 1H, CH_{Ar}), 6.49 (s, 1H, CH_{Ar}), 4.43 – 4.38 (m, 1H, C*H*), 3.65 – 3.57 (m, 1H, C*H*), 3.28 – 3.23 (m, 1H, C*H*), 3.21 (s, 3H, CH_3), 2.99 – 2.94 (m, 1H, CH_2), 2.93 – 2.89 (m, 1H, CH_2), 1.98 – 1.90 (m, 1H, CH_2), 1.17 – 1.13 (m, 1H, CH_2).

^{13}C NMR (101 MHz, CD_3OD) δ 150.22, 134.44, 130.48, 127.73, 122.23, 122.20, 120.00, 117.98, 117.90, 114.63, 114.28, 99.56, 78.33, 70.78, 53.85, 47.99, 35.16.

HR-MS (ESI, $[M+H]^+$) m/z: found: 393.0557; calc.: 393.0558.

Synthesis of **13**



13

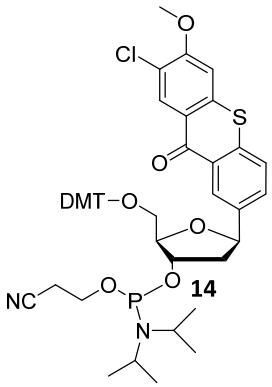
Under argon atmosphere, 100 mg of compound **1** (0.255 mmol, 1.00 eq.) were dissolved in 10 ml of pyridine. Subsequently, 104 mg of 4,4'-Dimethoxy-triphenylchlormethane (0.306 mmol, 1.20 eq.) were slowly added and the mixture was stirred overnight at room temperature. The solvent was removed under reduced pressure, and the crude product was purified by column chromatography (SiO_2, hexane/EE 3:1, 0.1% Et_3N). 58.5 mg of compound **13** (0.0841 mmol, 33%) were obtained as a colorless solid.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.62 (s, 1H, CH_{Ar}), 8.54 (d, J = 2.1 Hz, 1H, CH_{Ar}), 7.74 (dd, J = 8.4, 2.1 Hz, 1H, CH_{Ar}), 7.57 (d, J = 8.3 Hz, 1H, CH_{Ar}), 7.32 – 7.27 (m, 5H, CH_{Ar}), 7.19 – 7.15 (m, 4H, CH_{Ar}), 6.98 (s, 1H, CH_{Ar}), 6.85 – 6.80 (m, 4H, CH_{Ar}), 5.25 (t, J = 7.5 Hz, 1H, CH), 4.53 (td, J = 6.8, 5.3 Hz, 1H, CH), 4.19 – 4.14 (m, 1H, CH), 4.03 (s, 3H, CH_3), 3.89 – 3.82 (m, 1H, CH_2), 3.80 (s, 6H, CH_3), 3.79 – 3.74 (m, 1H, CH_2), 2.77 (dt, J = 13.3, 6.8 Hz, 1H, CH_2), 2.14 – 2.06 (m, 1H, CH_2).

^{13}C NMR (101 MHz, $CDCl_3$) δ 158.79, 139.60, 131.65, 129.86, 129.28, 128.85, 128.01, 127.91, 127.23, 126.94, 126.47, 113.32, 107.33, 85.85, 79.25, 73.36, 62.88, 56.81, 55.41, 43.81.

HR-MS (ESI, $[M+H]^+$) m/z: found: 695.1848; calc.: 695.1865.

Synthesis of **14**

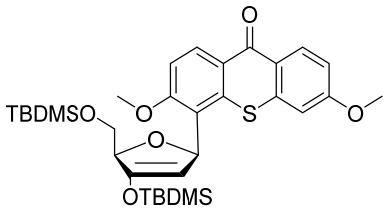


Under argon atmosphere, 83.0 mg of compound **13** (0.119 mmol, 1.00 eq.) were dissolved in 10.0 ml DCM. Subsequently, 71.4 µL of diisopropylethylamine (54.3 mg, 0.417 mmol, 3.50 eq.) were added, and the solution was stirred for 10 minutes at room temperature. Next, 39.9 µL of 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite (42.4 mg, 0.179 mmol, 1.50 eq.) were added, and the reaction mixture was stirred for 3 hours at room temperature. The reaction mixture was purified by column chromatography (SiO_2, DCM/acetone 50:1, 0.1% Et_3N). 50.0 mg of compound **14** (0.0558 mmol, 47%) were obtained as a colorless solid.

^{31}P NMR (162 MHz, $CDCl_3$) δ 148.26, 147.91.

HR-MS (ESI, $[M+H]^+$) m/z: found: 895.2934; calc.: 895.2943.

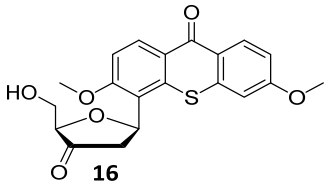
Synthesis of **15**



15

Under argon atmosphere, 3.10 g of **9** (4.507 mmol, 1.00 eq.), 3.10 g of **10** (9.01 mmol, 2.00 eq.), and 1.35 ml of triethylamine (1.00 g, 9.69 mmol, 2.15 eq.) were dissolved in 10 ml of DMF, and the reaction mixture was degassed three times using the freeze-pump-thaw method. Subsequently, 0.400 g of Pd_2dba_3 (0.450 mmol, 0.100 equiv.) and 0.600 g of Q-Phos (0.901 mmol, 0.200 eq.) were added, and the reaction was stirred at 70 °C for 72 hours. The solvent was removed under reduced pressure, and the crude product was purified by column chromatography (SiO_2, hexane/EtOAc, 25:1 → 10:1). Due to residual impurities, the product was used without additional purification.

Synthesis of **16**



16

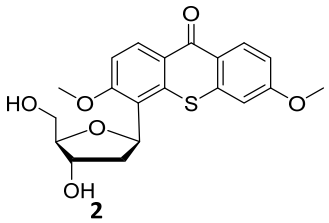
Under argon atmosphere, 1.01 g of **15** (1.64 mmol, 1.00 eq.) was dissolved in 15 ml of THF. The reaction mixture was cooled to 0 °C, and 1.2 ml of $Et_3N{\cdot}HF$ (1.19 g, 7.36 mmol, 4.50 eq.) was added dropwise. The reaction was stirred overnight at room temperature. Subsequently, the solvent was removed under reduced pressure, and the crude product was purified by column chromatography (SiO_2, gradient: DCM → DCM/MeOH 15:1). Over two steps, 644 mg of **16** (1.67 mmol, 37%) was obtained as a colorless solid.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.70 (d, *J* = 9.1 Hz, 1H, CH_{Ar}), 8.49 (d, *J* = 9.0 Hz, 1H, CH_{Ar}), 7.14 (d, *J* = 9.1 Hz, 1H, CH_{Ar}), 7.03 (dd, *J* = 9.0, 2.4 Hz, 1H, CH_{Ar}), 6.94 (d, *J* = 2.6 Hz, 1H, CH_{Ar}), 6.04 (dd, *J* = 10.2, 7.1 Hz, 1H, C*H*), 4.12 (t, *J* = 3.4 Hz, 1H, C*H*), 4.04 – 4.00 (m, 2H, CH_2), 3.98 (s, 3H, CH_3), 3.92 (s, 3H, CH_3), 3.20 (dd, *J* = 18.4, 10.2 Hz, 1H, CH_2), 2.79 (dd, *J* = 18.4, 7.0 Hz, 1H, CH_2).

^{13}C NMR (101 MHz, $CDCl_3$) δ 213.46, 178.81, 162.80, 160.82, 138.54, 133.21, 131.79, 124.20, 122.26, 120.35, 115.49, 110.47, 108.46, 106.32, 82.11, 72.45, 61.66, 56.20, 55.91, 40.46.

HR-MS (ESI, $[M+H]^+$) m/z: found: 387.0902; calc.: 387,0897.

Synthesis of **2**



2

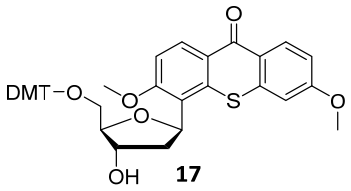
Under argon atmosphere, 0.397 g of **16** (1.03 mmol, 1.00 eq.) was dissolved in 24 ml of AcOH/MeCN (1:3) and cooled to 0 °C. Subsequently, 1.74 g of sodium triacetoxyborohydride (8.21 mmol, 8.00 eq.) was added, and the reaction was stirred at room temperature for 1 hour. The reaction was then quenched by adding 24 mL of EtOH/H_2O (1:1). The solvent was removed under reduced pressure, and the crude product was purified by column chromatography (SiO_2, gradient: DCM → DCM/MeOH 10:1). 333 mg of **2** (0.857 mmol, 83%) was obtained as a colorless solid.

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.47 (d, J = 9.0 Hz, 1H, CH_{Ar}), 8.30 (d, J = 9.0 Hz, 1H, CH_{Ar}), 7.32 (d, J = 9.1 Hz, 1H, CH_{Ar}), 7.28 (d, J = 2.5 Hz, 1H, CH_{Ar}), 7.11 (dd, J = 9.0, 2.5 Hz, 1H, CH_{Ar}), 5.76 – 5.71 (m, 1H, CH), 5.18 (s, 1H, OH), 4.76 (s, 1H, OH), 4.34 – 4.23 (m, 1H, CH), 3.93 (s, 3H, CH_3), 3.92 (s, 3H, CH_3), 3.80 – 3.75 (m, 1H, CH), 3.75 – 3.61 (m, 2H, CH_2), 2.45 (dd, J = 6.4, 4.0 Hz, 1H, CH_2), 1.95 – 1.89 (m, 1H, CH_2).

^{13}C NMR (101 MHz, DMSO) δ 177.67, 162.30, 160.27, 139.57, 136.82, 131.15, 130.60, 122.78, 122.11, 121.25, 115.59, 110.86, 108.54, 88.37, 73.16, 72.44, 62.15, 56.17, 56.00, 38.16.

HR-MS (ESI, $[M+H]^+$) m/z: found: 389.1046; calc.: 389.1053.

Synthesis of **17**



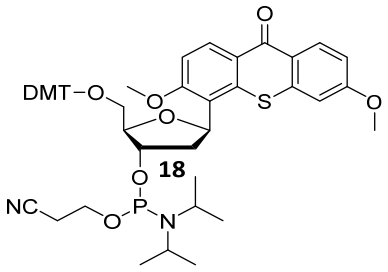
Under argon atmosphere 333 mg of **2** (0.857 mmol, 1.00 eq.) was dissolved in 15 ml of pyridine. Subsequently, 378 mg of 4,4'-dimethoxytriphenylchloromethane (1.11 mmol, 1.30 eq.) was added slowly, and the reaction was stirred overnight at room temperature. The solvent was removed under reduced pressure, and the crude product was purified by column chromatography (SiO_2, gradient: DCM → DCM/acetone 20:1, 0.1% Et_3N). 418 mg of **17** (0.605 mmol, 71%) was obtained as a colorless solid.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.55 (d, *J* = 9.0 Hz, 1H, CH_{Ar}), 8.34 (d, *J* = 9.0 Hz, 1H, CH_{Ar}), 7.52 – 7.45 (m, 2H, CH_{Ar}), 7.39 – 7.32 (m, 4H, CH_{Ar}), 7.22 – 7.13 (m, 2H, CH_{Ar}), 7.14 – 7.04 (m, 1H, CH_{Ar}), 6.94 (d, *J* = 9.1 Hz, 1H, CH_{Ar}), 6.82 (dd, *J* = 9.0, 2.5 Hz, 1H, CH_{Ar}), 6.76 – 6.68 (m, 4H, CH_{Ar}), 6.45 (d, *J* = 2.5 Hz, 1H, CH_{Ar}), 5.84 (dd, *J* = 10.2, 6.4 Hz, 1H, C*H*), 4.60 – 4.49 (m, 1H, C*H*), 3.95 (q, *J* = 4.7 Hz, 1H, C*H*), 3.73 (s, 3H, CH_3), 3.64 (s, 3H, CH_3), 3.64 (s, 3H, CH_3), 3.48 – 3.39 (m, 2H, CH_2), 3.33 (s, 3H, CH_3), 2.58 (ddd, *J* = 13.2, 10.3, 7.5 Hz, 1H, CH_2), 2.02 (ddd, *J* = 13.2, 6.5, 2.4 Hz, 1H, CH_2).

^{13}C NMR (101 MHz, $CDCl_3$) δ 179.20, 162.31, 162.25, 160.21, 158.73, 158.64, 158.62, 145.17, 139.60, 136.42, 136.15, 132.06, 131.24, 130.32, 130.23, 129.26, 128.46, 128.37, 128.05, 127.96, 127.89, 127.19, 126.96, 123.90, 122.17, 121.71, 115.37, 113.31, 113.27, 113.14, 109.77, 108.06, 86.58, 86.43, 74.38, 74.18, 63.77, 55.93, 55.37, 55.32, 39.03.

HR-MS (ESI, $[M+H]^+$) m/z: found: 691.2342; calc.: 691.2360.

Synthesis of **18**



Under argon atmosphere 333 mg of **17** (0.605 mmol, 1.00 eq.) was dissolved in 20.0 ml of DCM in a preheated round-bottom flask. Subsequently, 361 µL of diisopropylethylamine (274 mg, 2.12 mmol, 3.50 eq.) was added, and the solution was stirred at room temperature for 10 minutes. Then, 203 µL of 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite (215 mg, 0.908 mmol, 1.50 eq.) was added, and the reaction mixture was stirred for 3 hours at room temperature. The crude product was purified by column chromatography (SiO_2, DCM/acetone 50:1, 0.1% Et_3N). 379 mg of **18** (0.425 mmol, 70%) was obtained as a colorless solid.

^{31}P NMR (162 MHz, $CDCl_3$) δ 149.03, 147.97.

HR-MS (ESI, $[M+H]^+$) m/z: found: 891.3451; calc.: 891.3439.

5. Images of ^{1}H, ^{13}C, ^{31}P NMR and MS spectra

Compound **5**

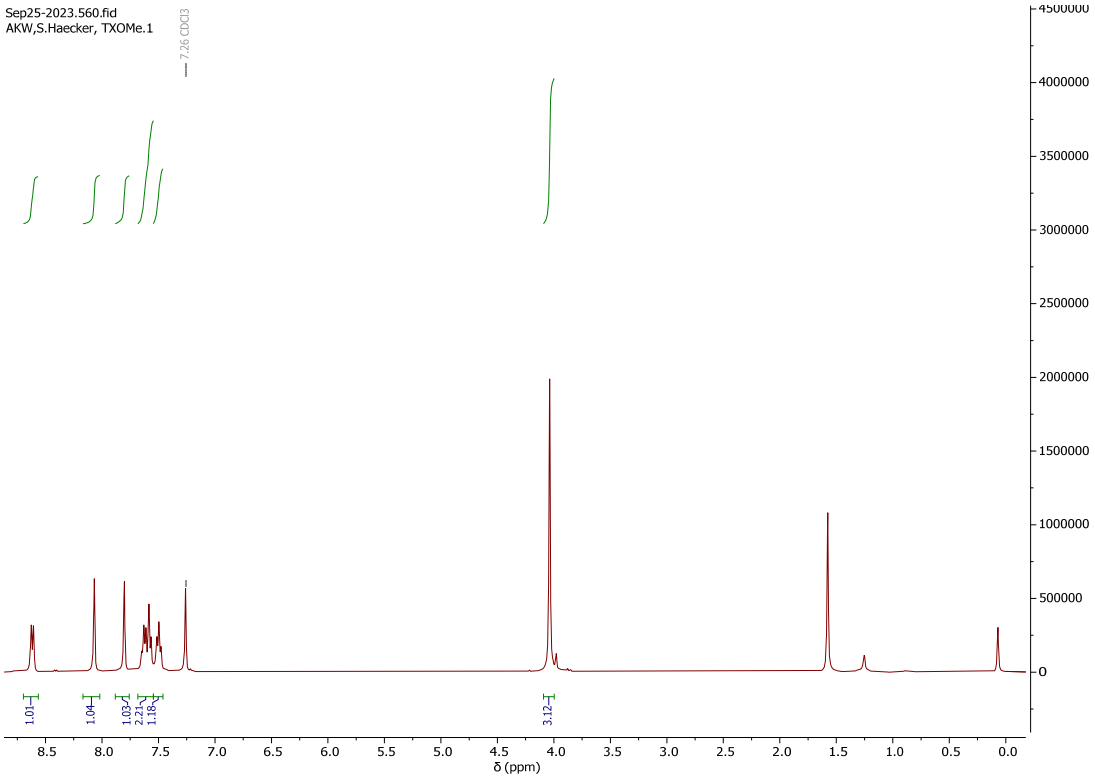


Figure S1: ^{1}H-NMR (400 MHz, $CDCl_3$) of **5**.

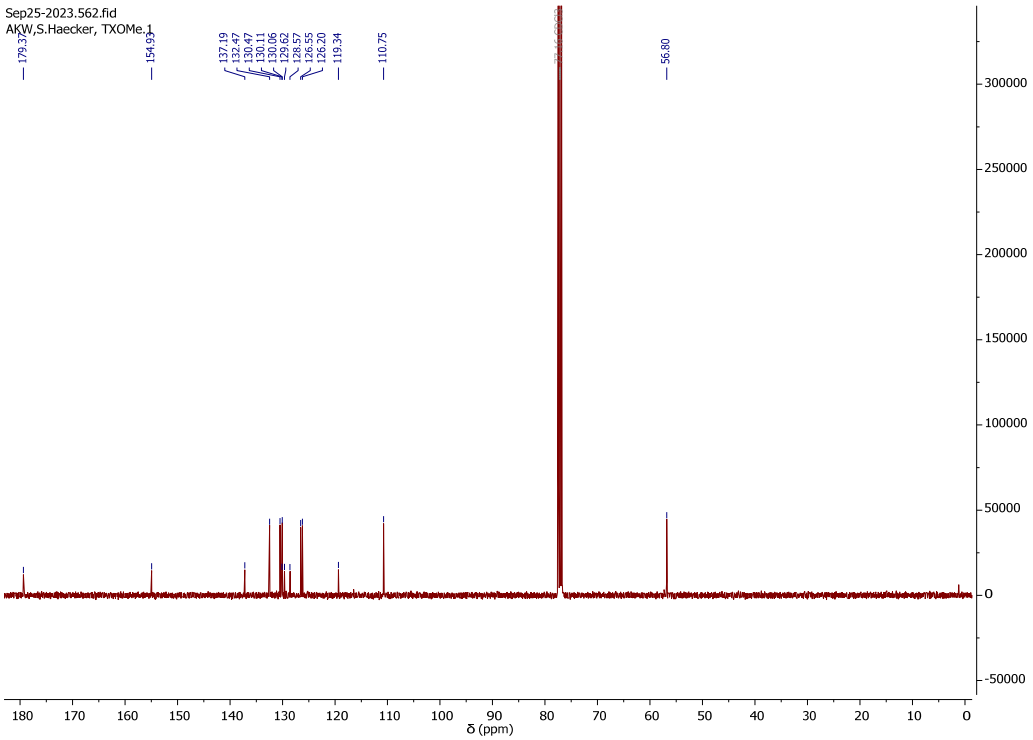


Figure S2: ^{13}C-NMR (101 MHz, $CDCl_3$) of **5**.

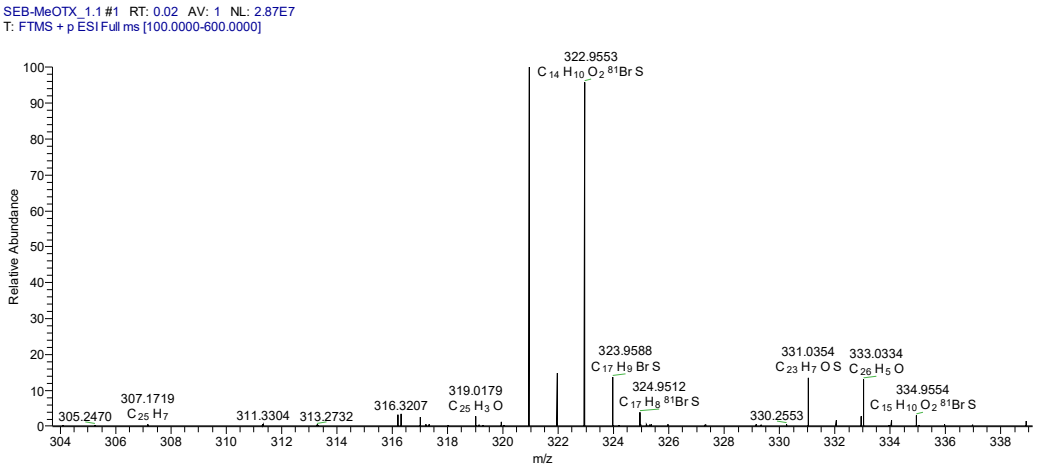


Figure S3: HR-ESI-MS of **5**.

Compound **6**

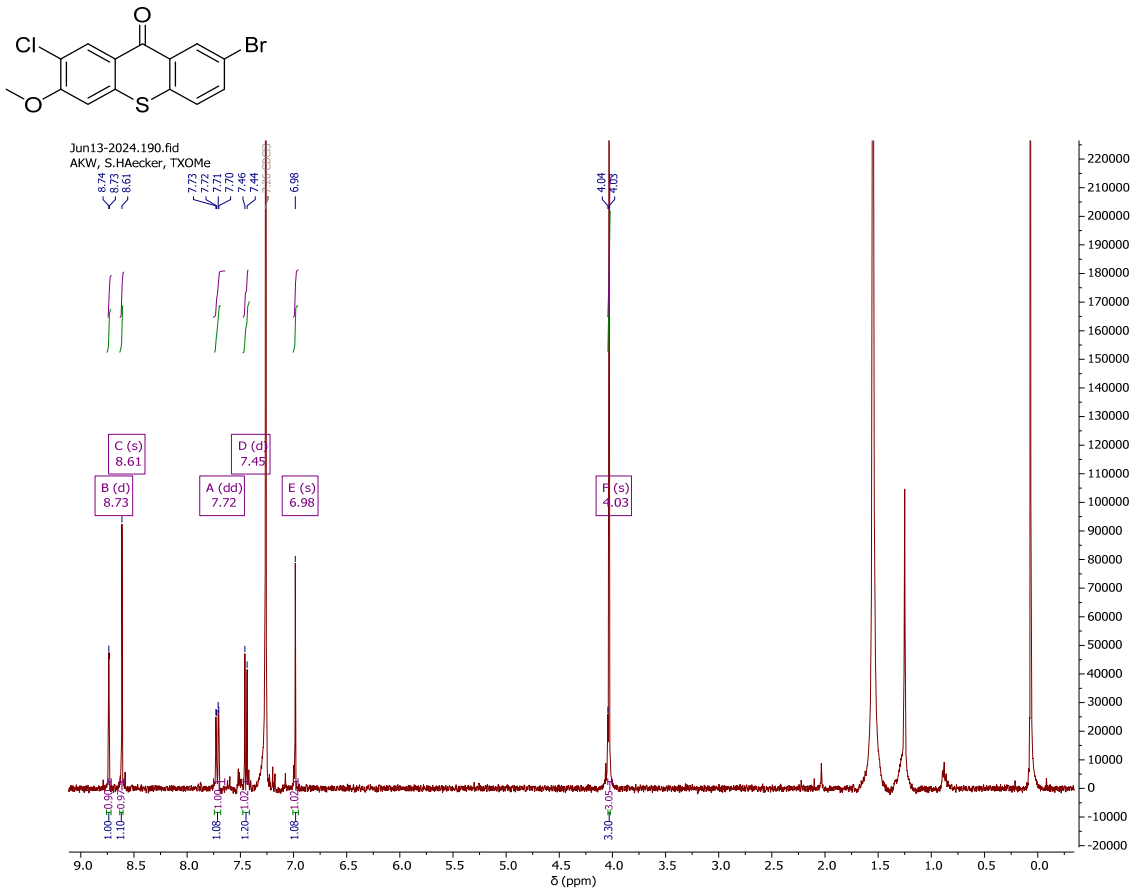


Figure S4: ^{1}H-NMR (400 MHz, $CDCl_3$) of **6.**

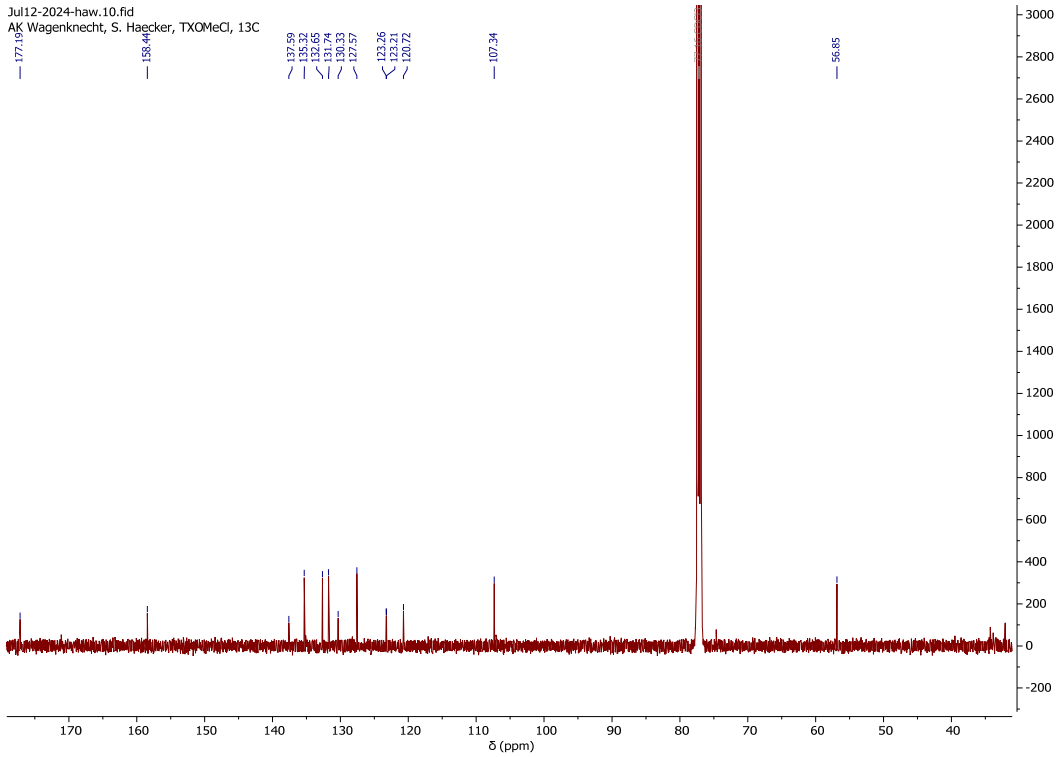


Figure S5: ^{13}C-NMR (101 MHz, $CDCl_3$) of **6.**

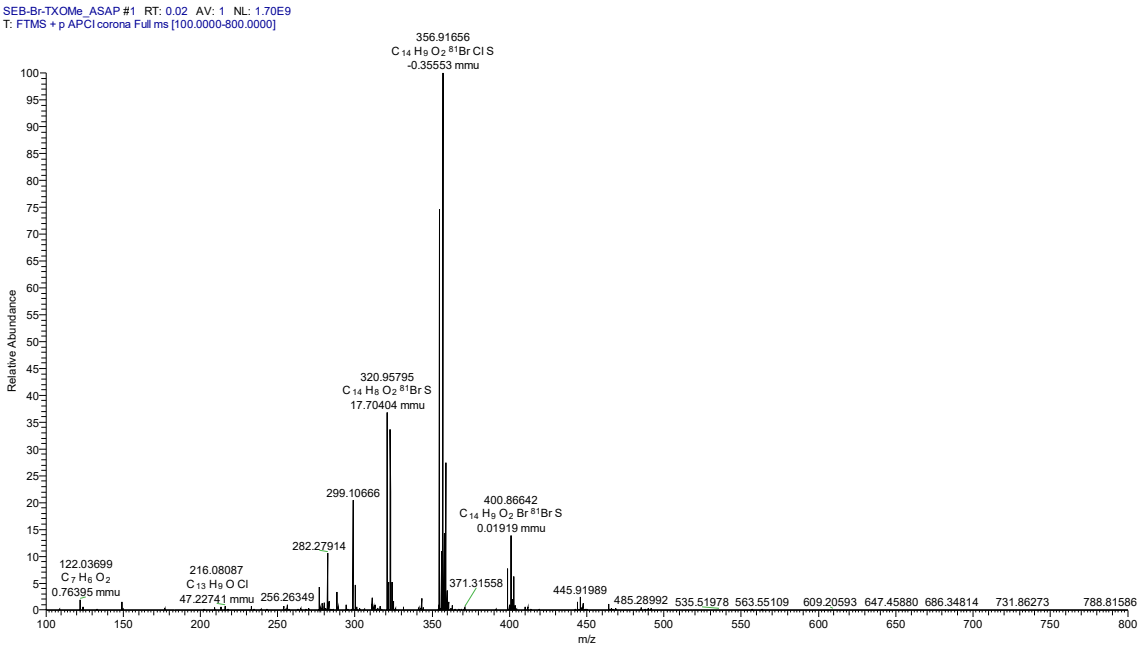
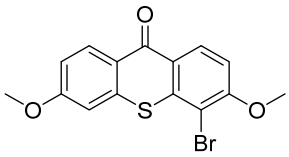


Figure S6: HR ASAP analysis of **6**.

Compound 9



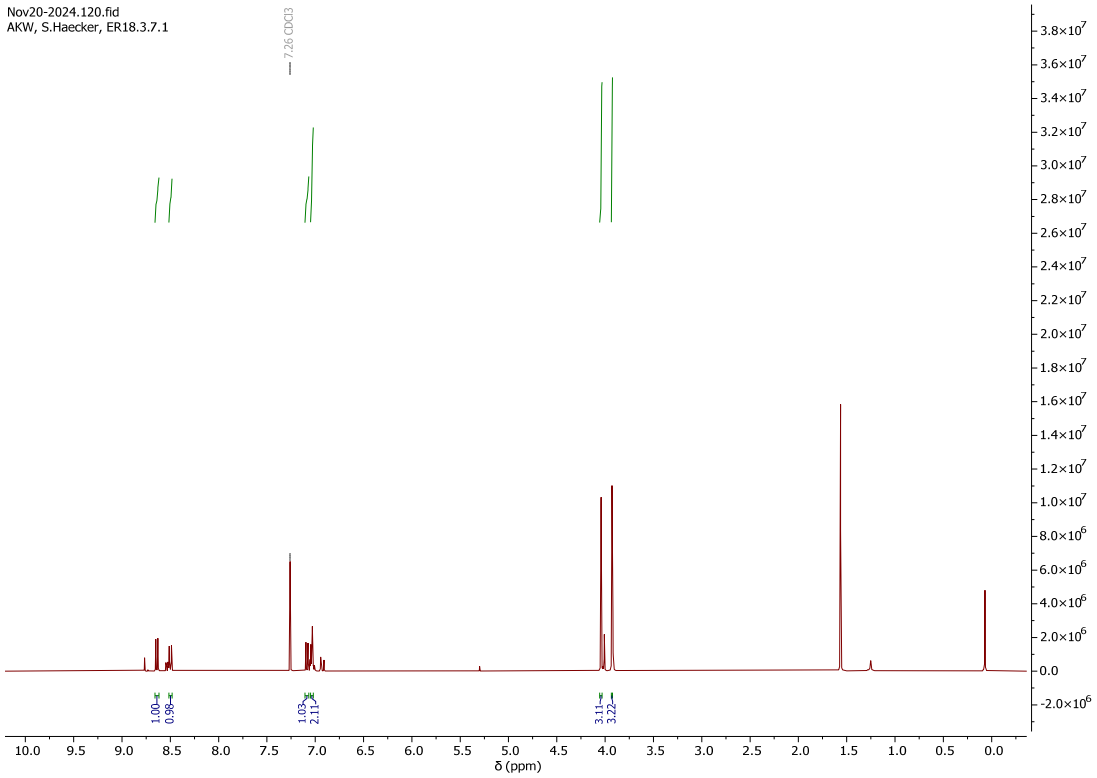


Figure S7: ^{1}H-NMR (400 MHz, $CDCl_3$) of **9**.

The ^{1}H-NMR spectrum contains 20% of the byproduct 2-Brom-3,6-Dimethoxythioxanthon.

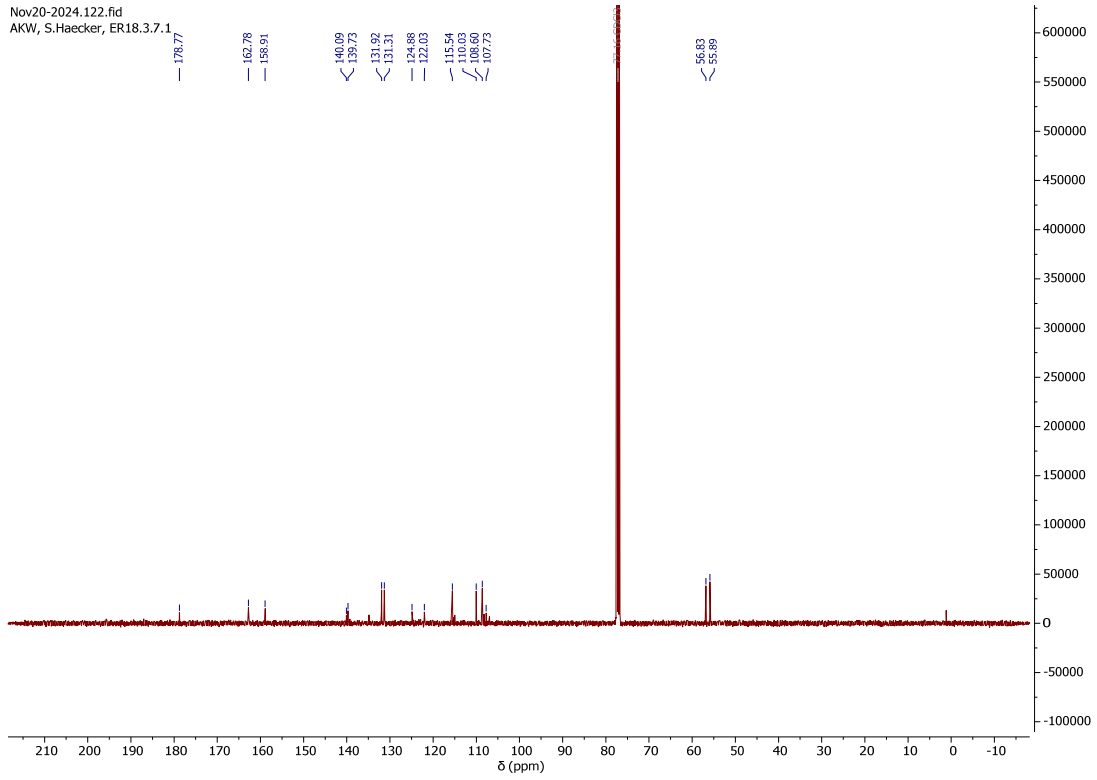


Figure S8: ^{13}C-NMR (101 MHz, $CDCl_3$) of **9**.

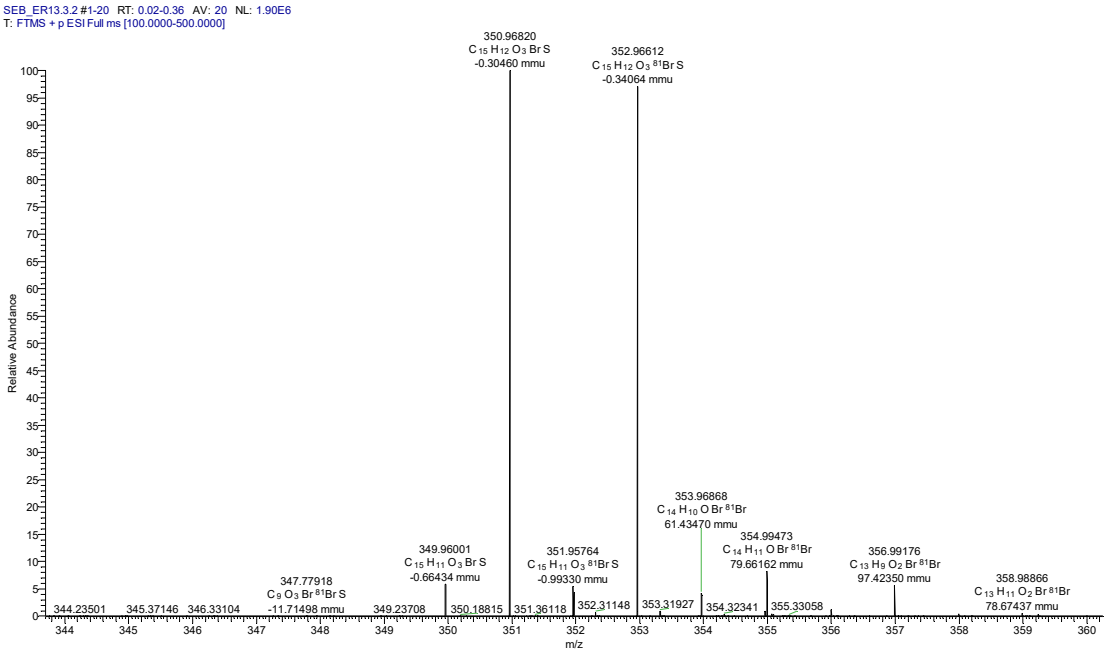


Figure S9: HR ESI-MS analysis of **9**.

Compound 11a

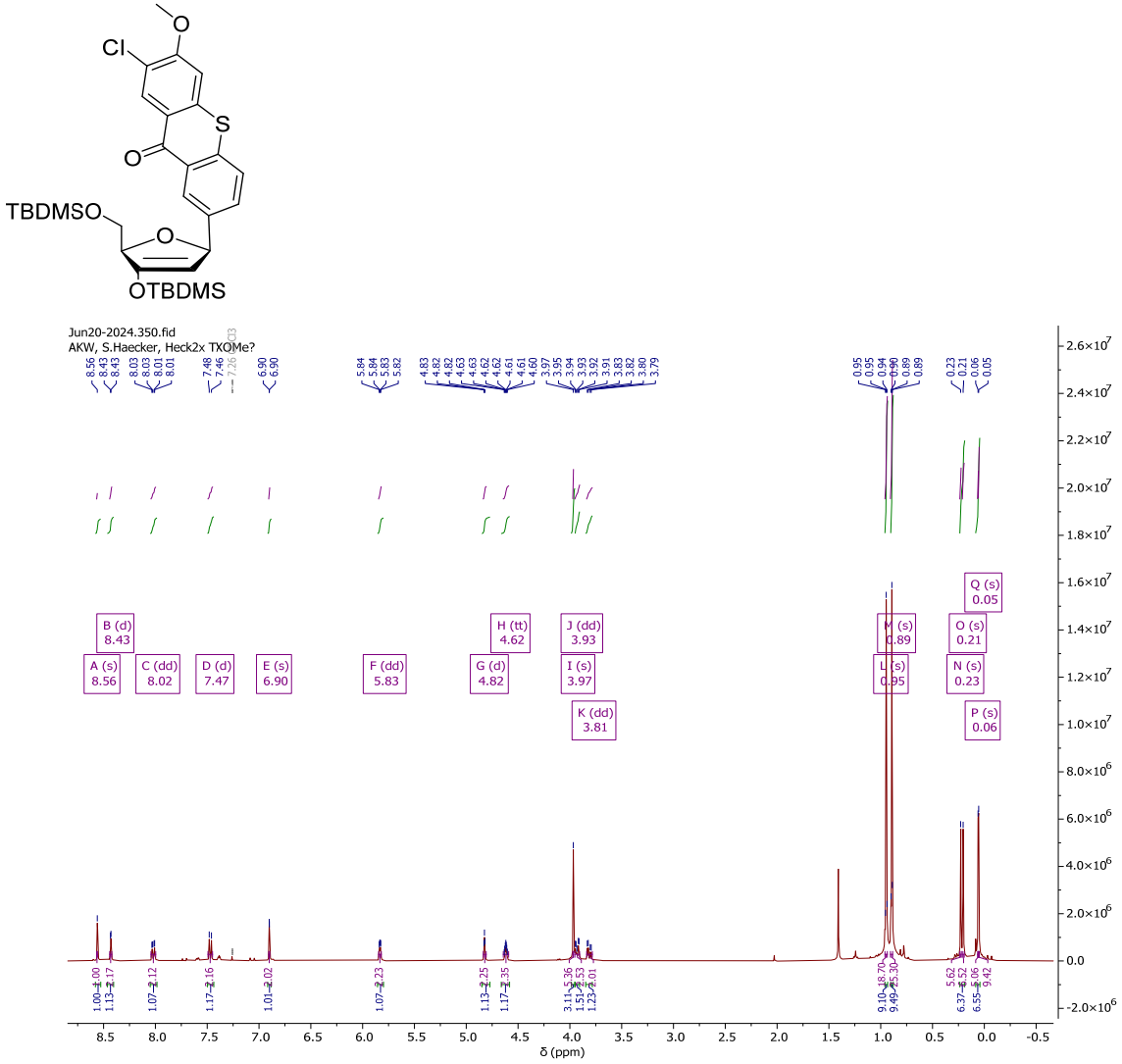


Figure S10: ^{1}H-NMR (400 MHz, $CDCl_3$) of **11a**.

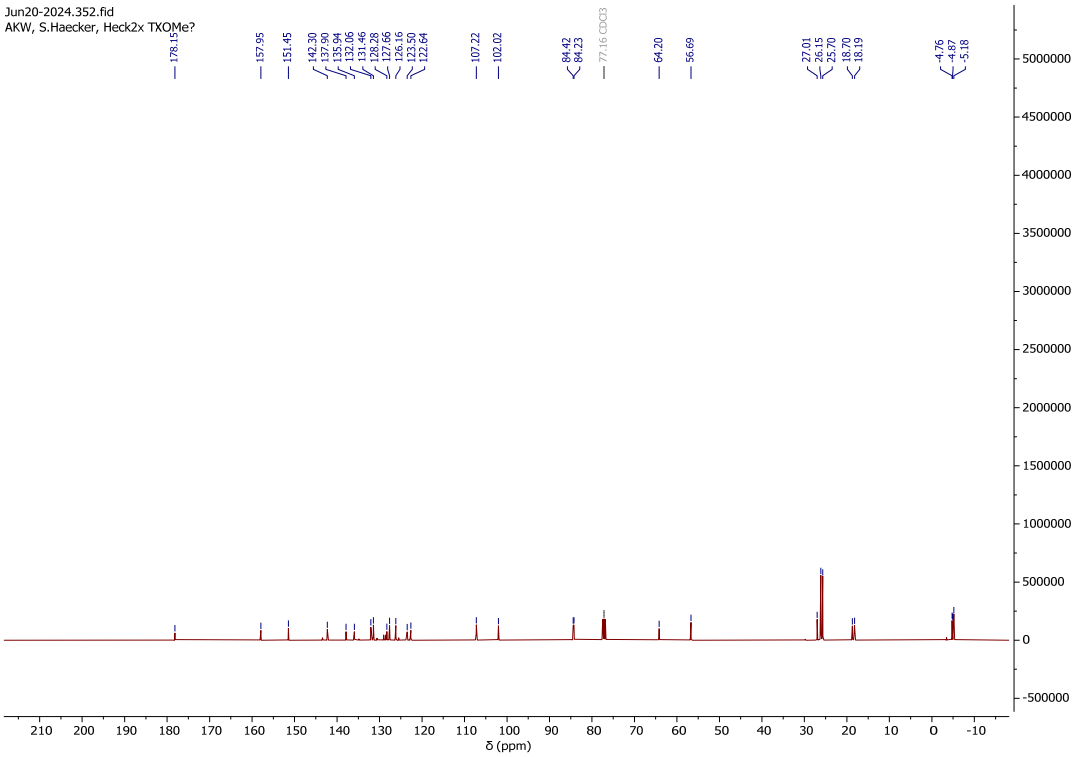


Figure S11: ^{13}C-NMR (101 MHz, $CDCl_3$) of **11a**.

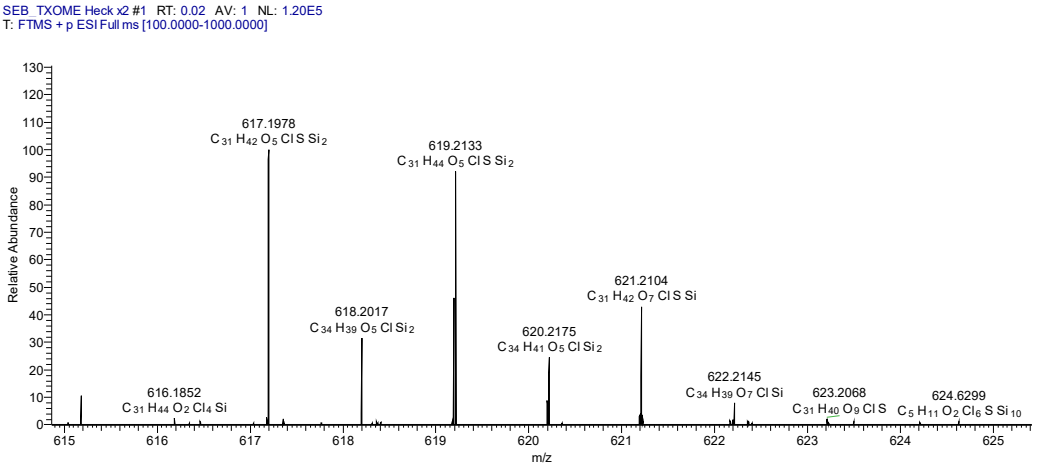


Figure S12: HR ESI-MS analysis of **11a**.

Compound **11b**

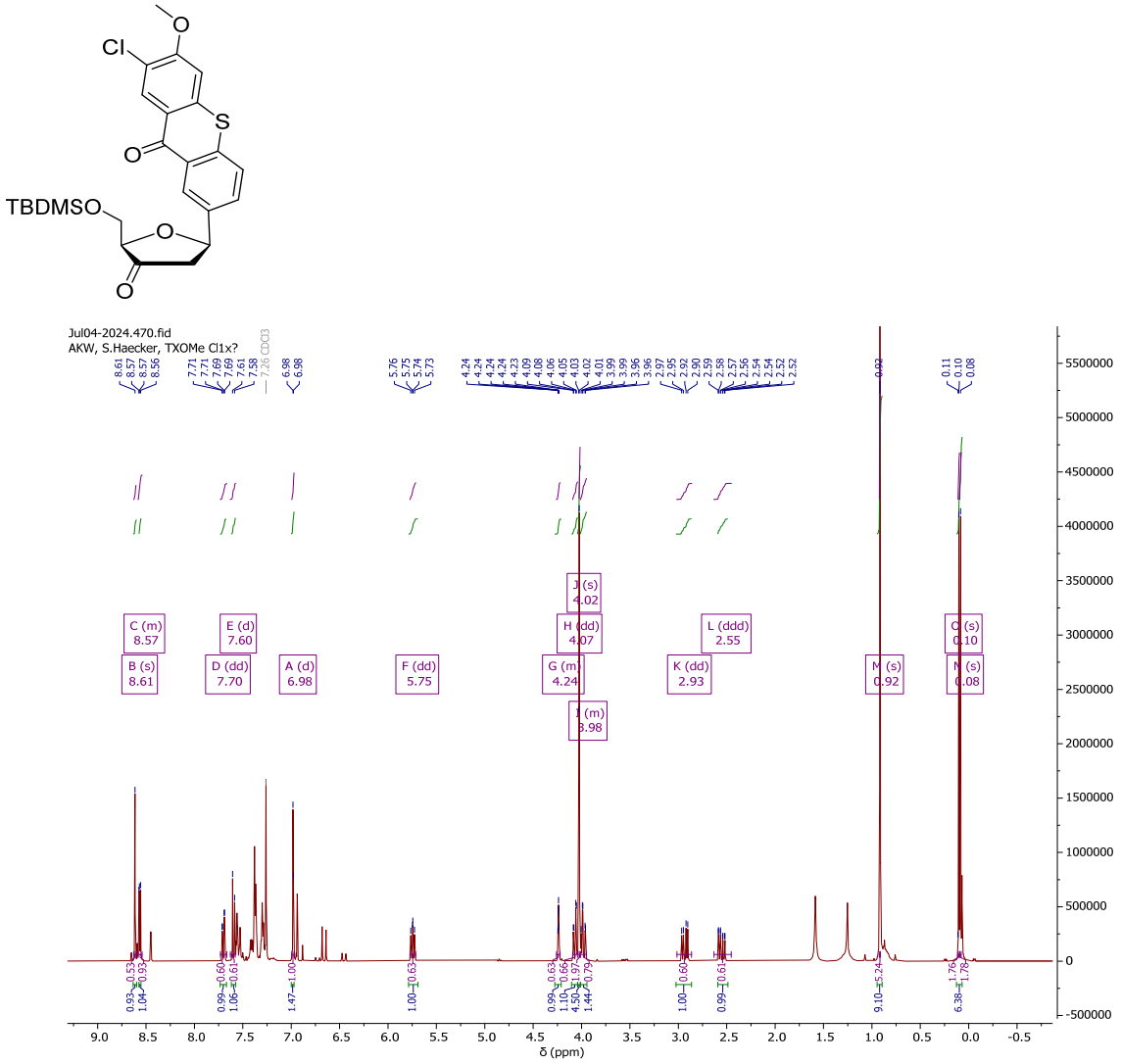


Figure S13: ^{1}H-NMR (400 MHz, $CDCl_3$) of **11b**.

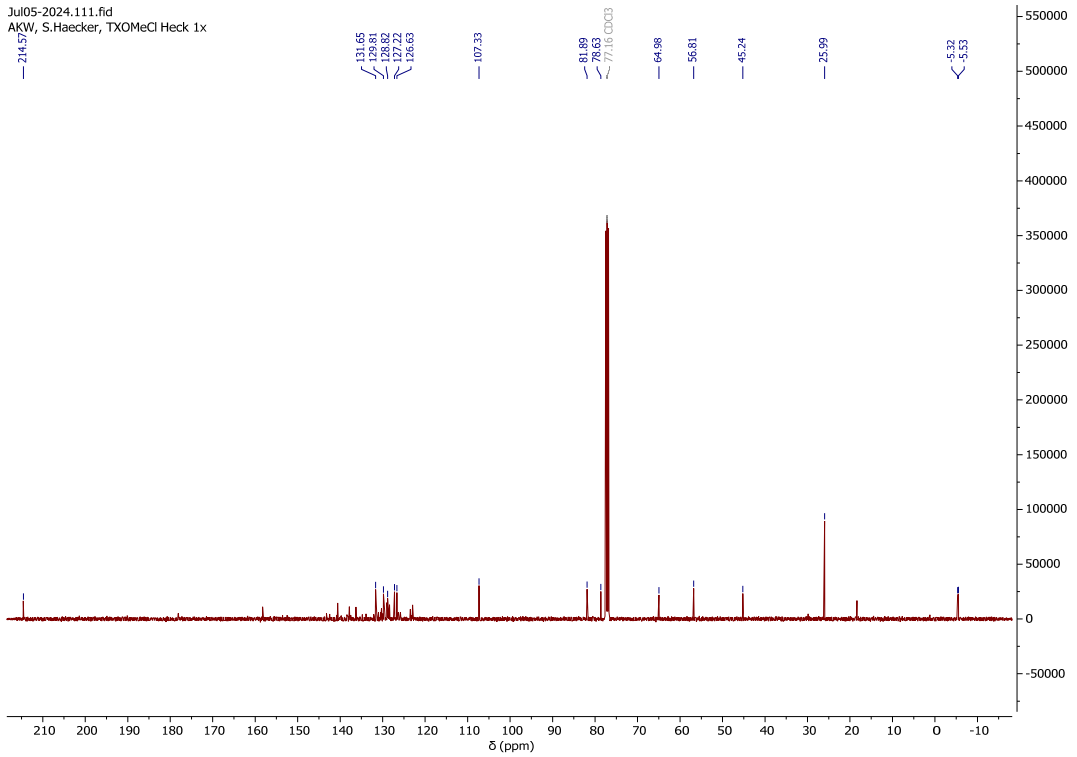


Figure S14: ^{13}C-NMR (101 MHz, $CDCl_3$) of **11b**.

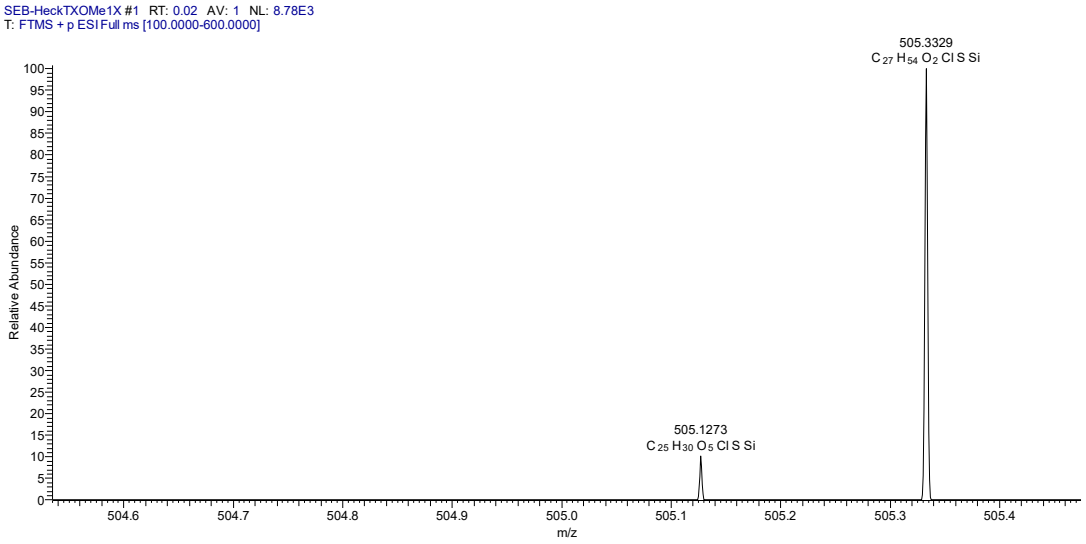
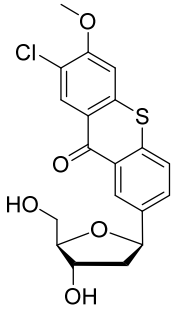


Figure S15: HR ESI-MS of **11b**.

Compound **1**



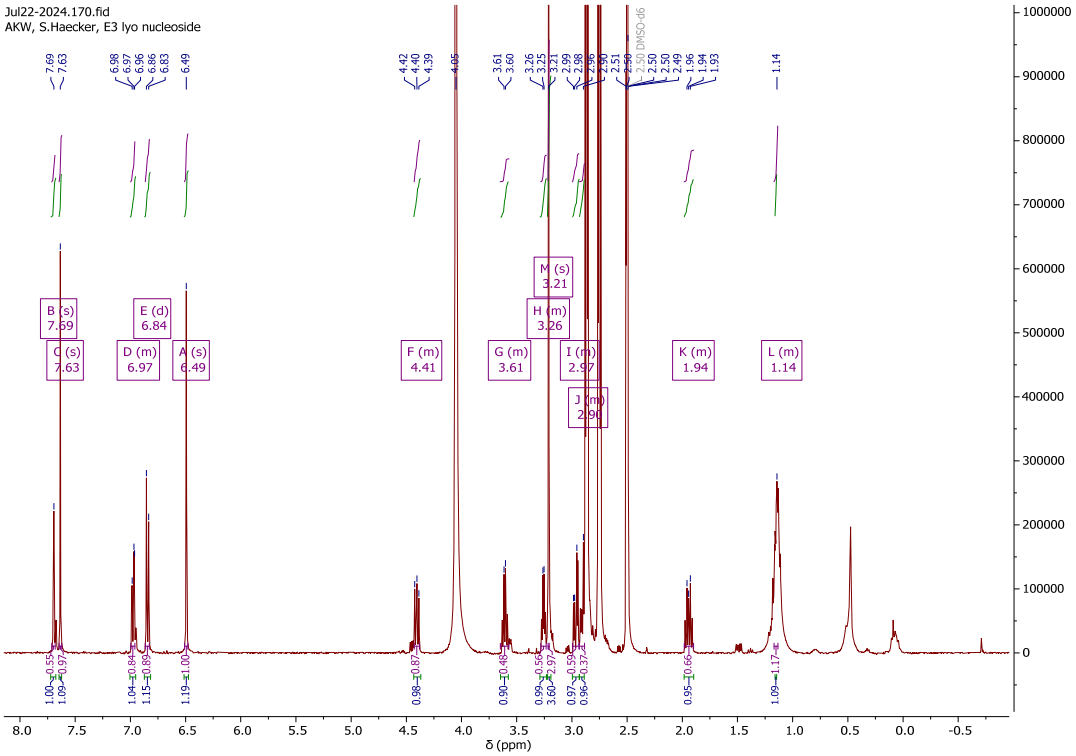


Figure S16: ^{1}H-NMR (400 MHz, DMSO-d6) of **1**.

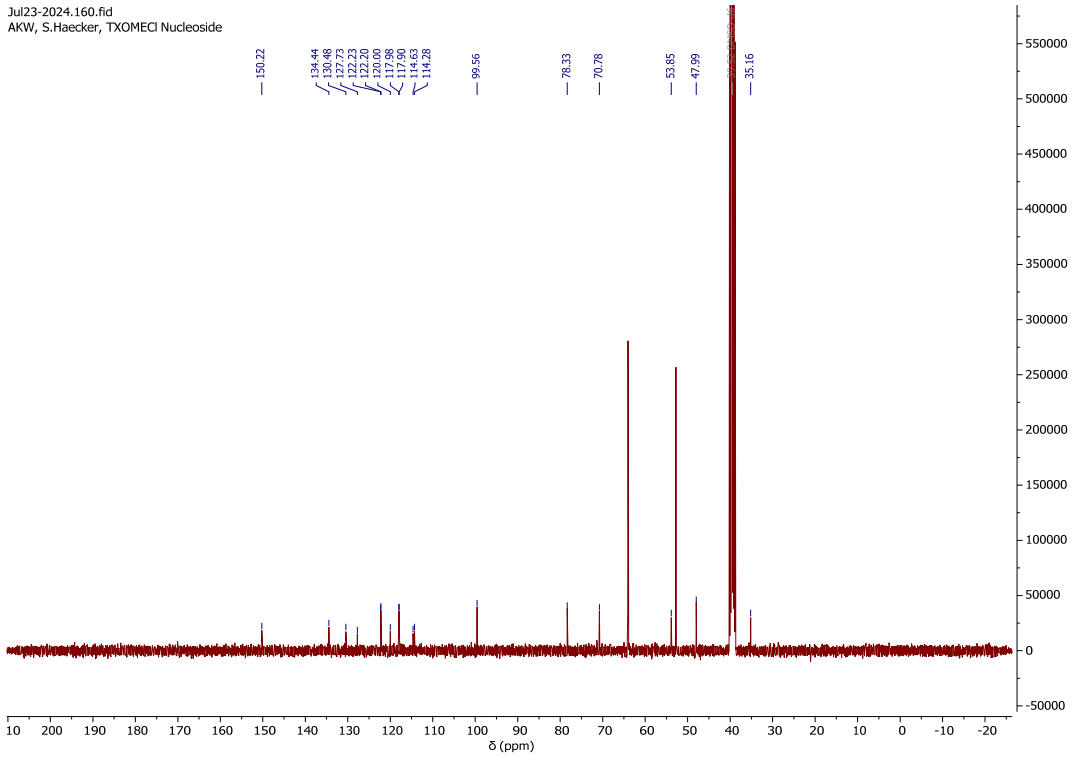


Figure S17: ^{13}C-NMR (101 MHz, DMSO-d6) of **1**.

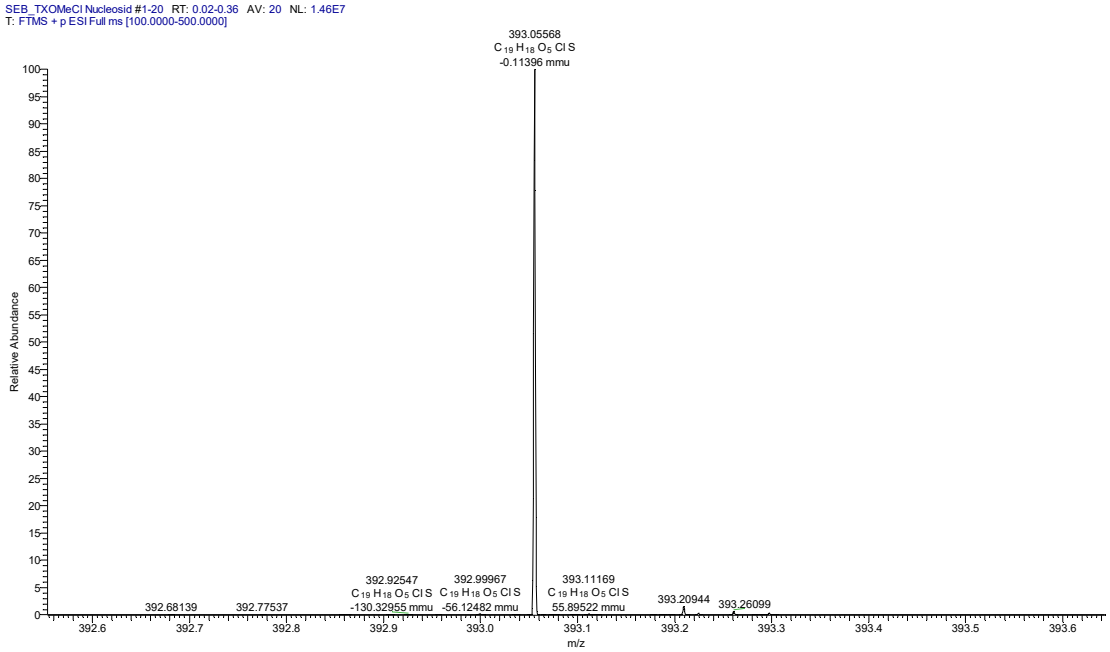


Figure S18: HR ESI-MS analysis of **1**.

Compound **13**

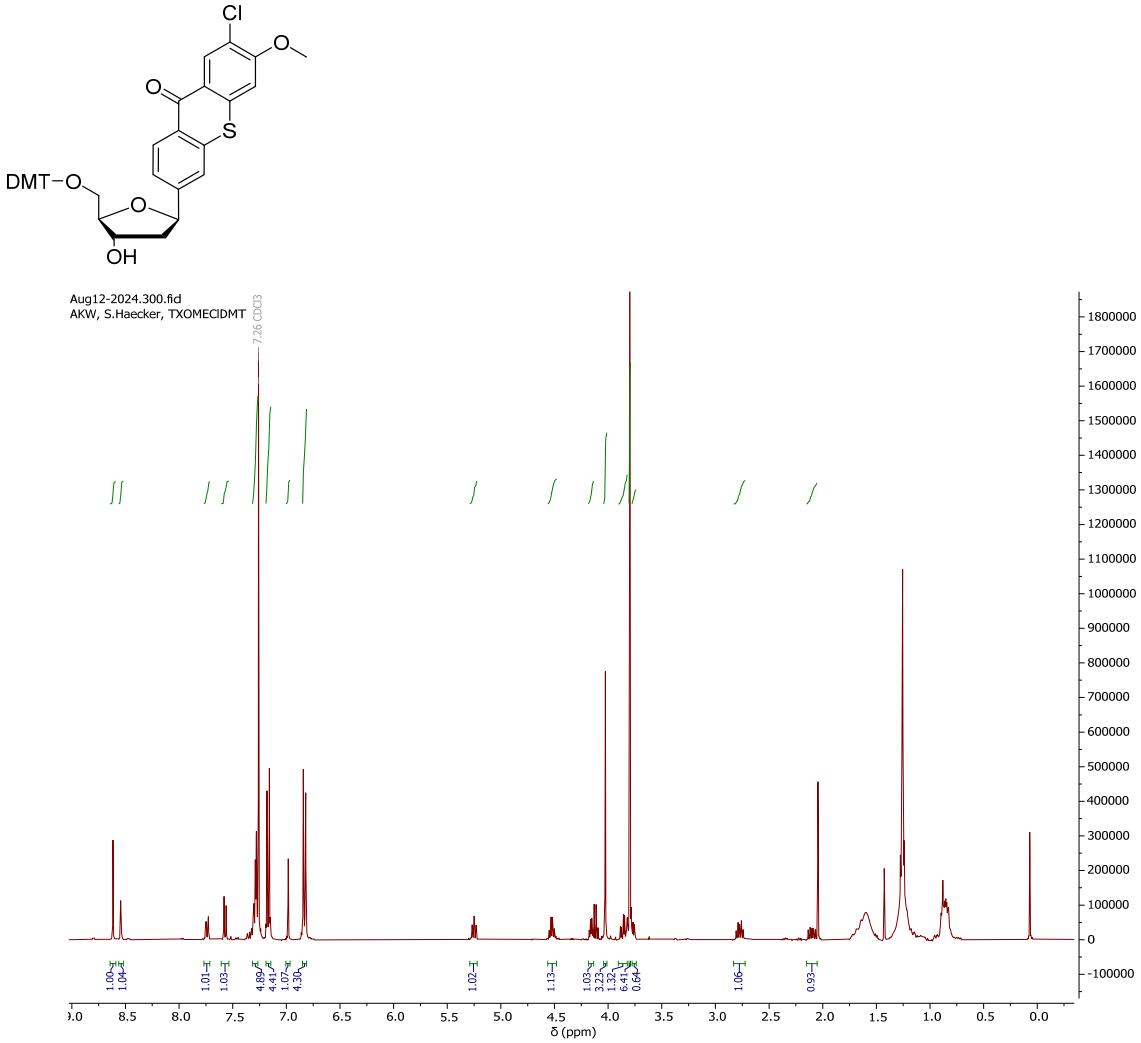


Figure S19: ^{1}H-NMR (400 MHz, $CDCl_3$) of **13**.

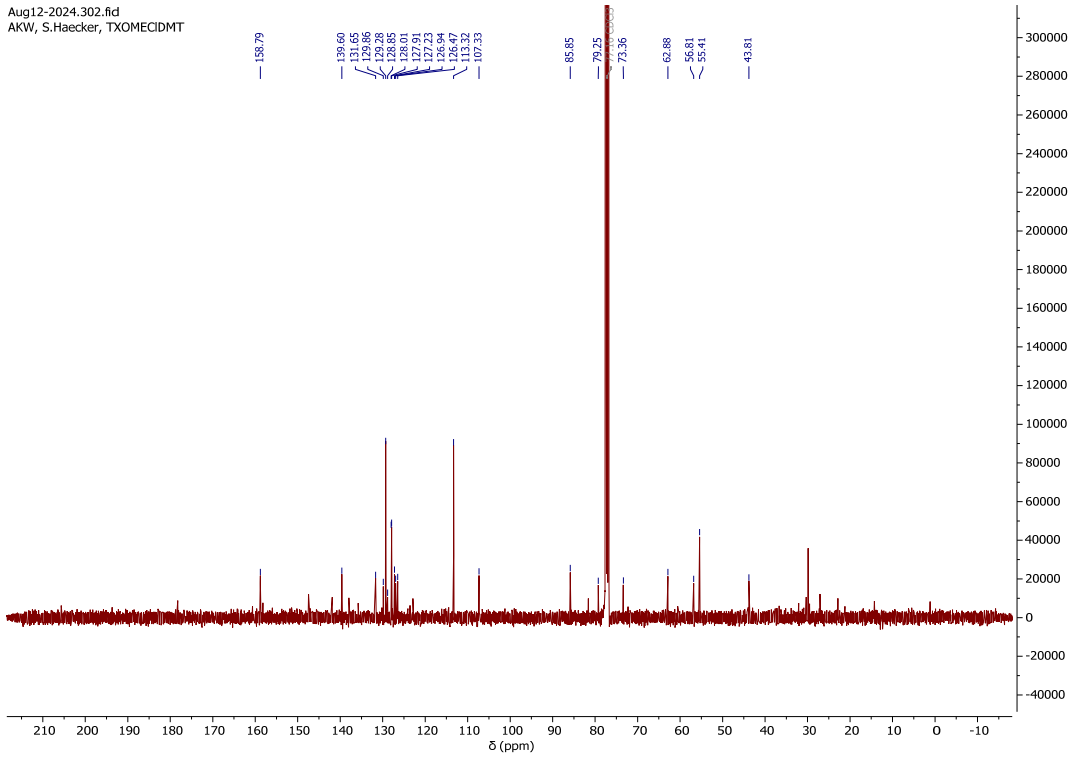


Figure S20: ^{13}C-NMR (101 MHz, CDCl3) of **13**.

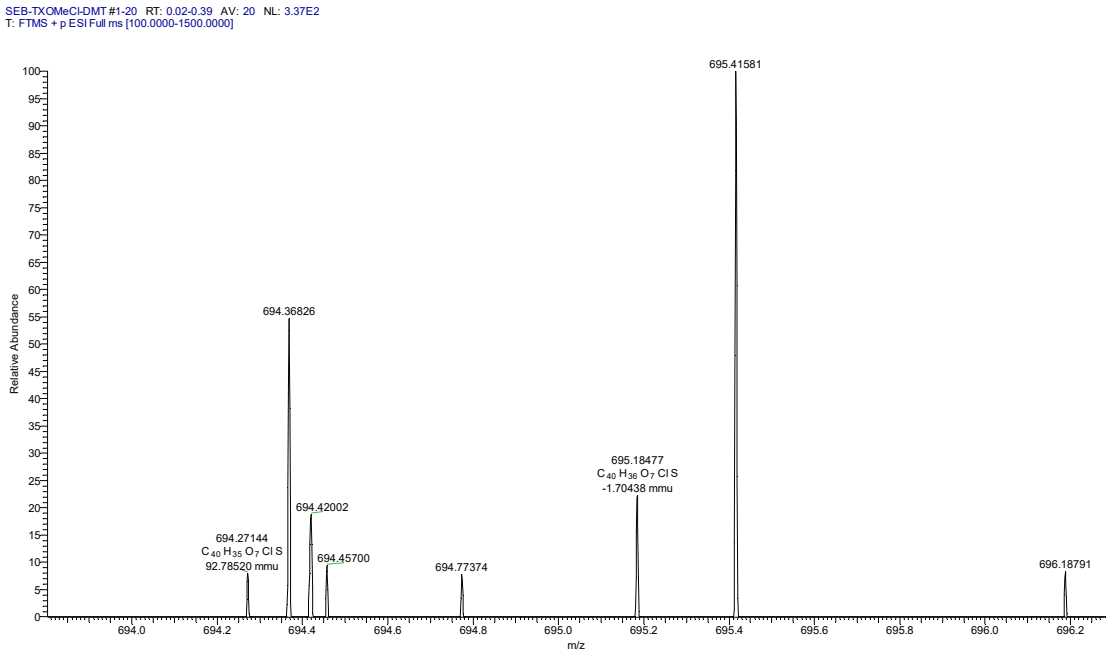


Figure S21: HR ESI-MS analysis of **13**.

Compound **14**

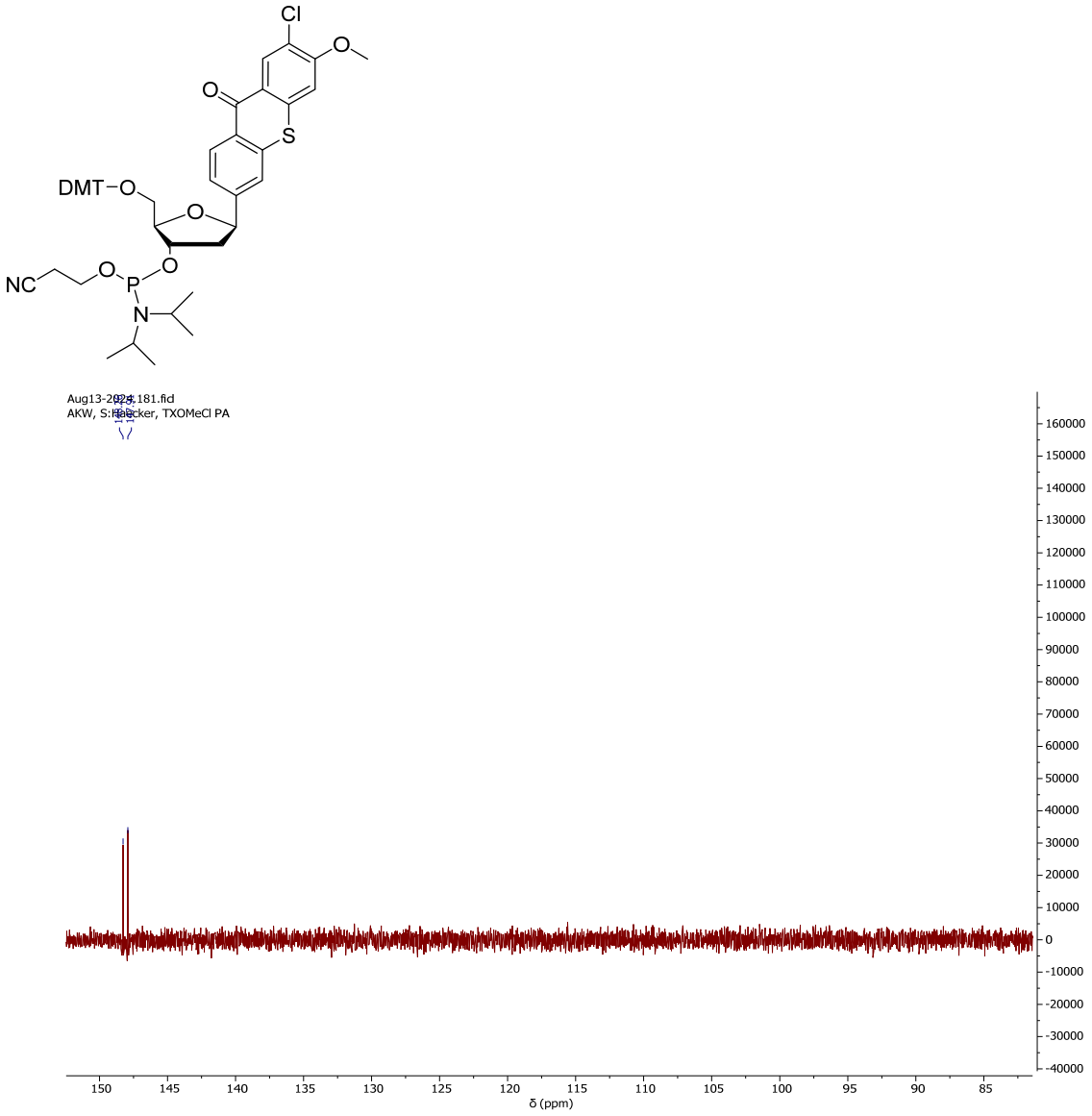


Figure S22: ^{31}P-NMR (162 MHz, $CDCl_3$) of **14**.

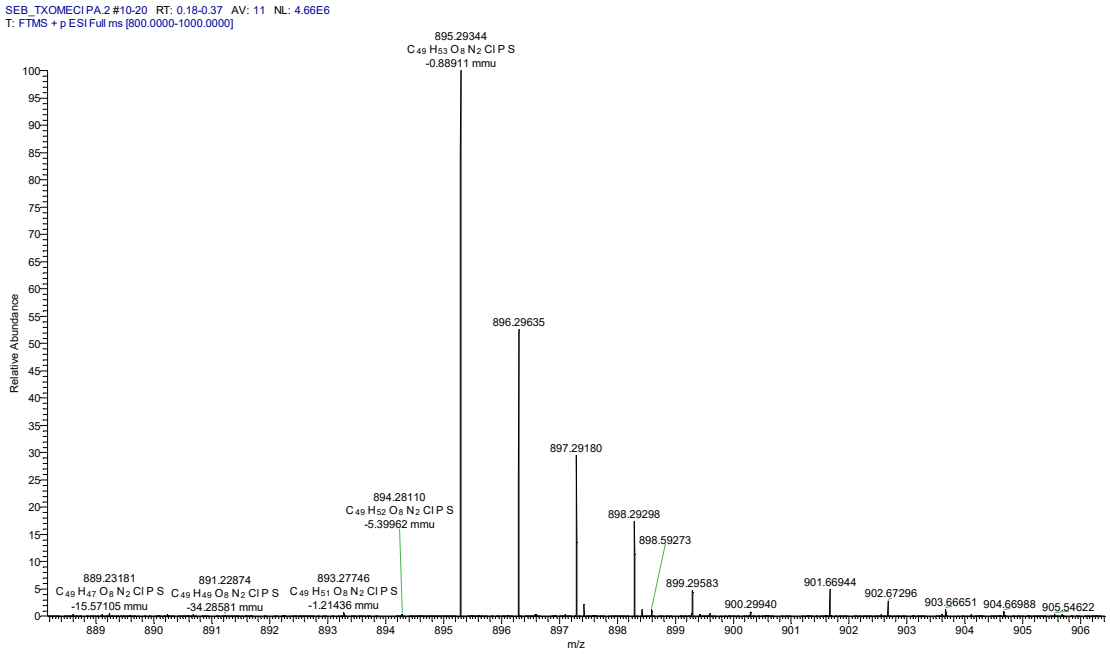
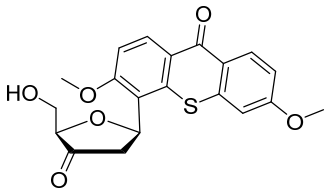


Figure S23: HR ESI-MS analysis of **14**.

Compound 16



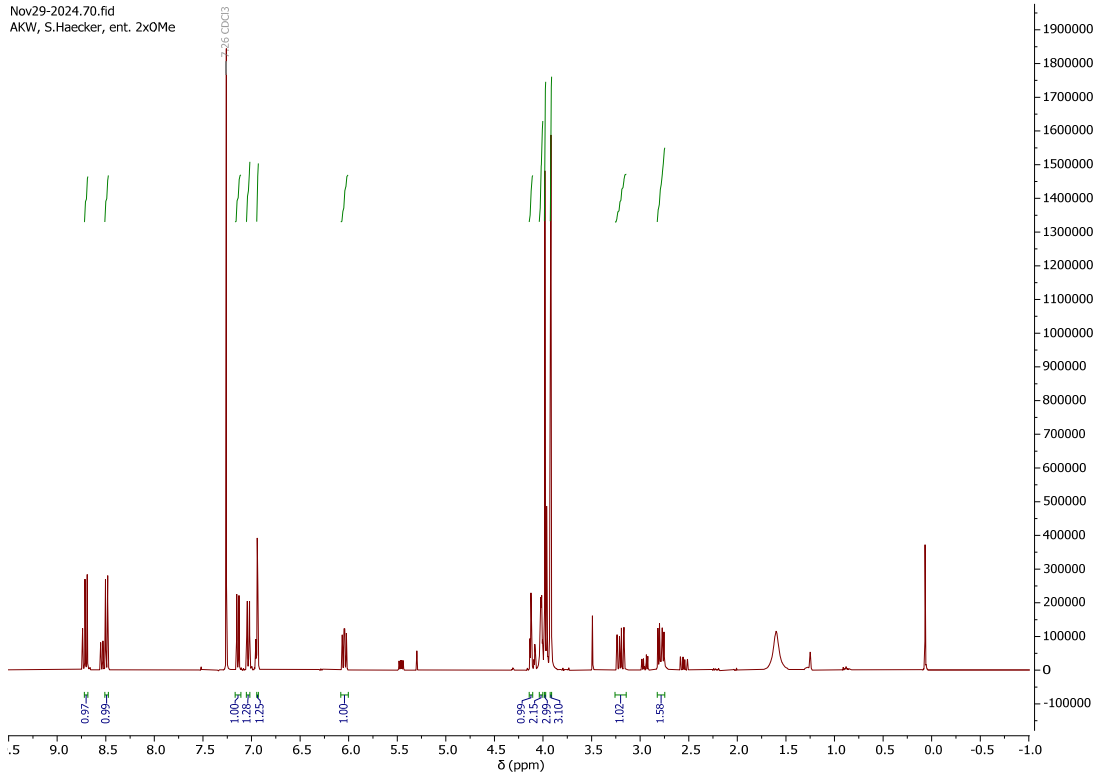


Figure S24: ^{1}H-NMR (400 MHz, $CDCl_3$) of **16**.

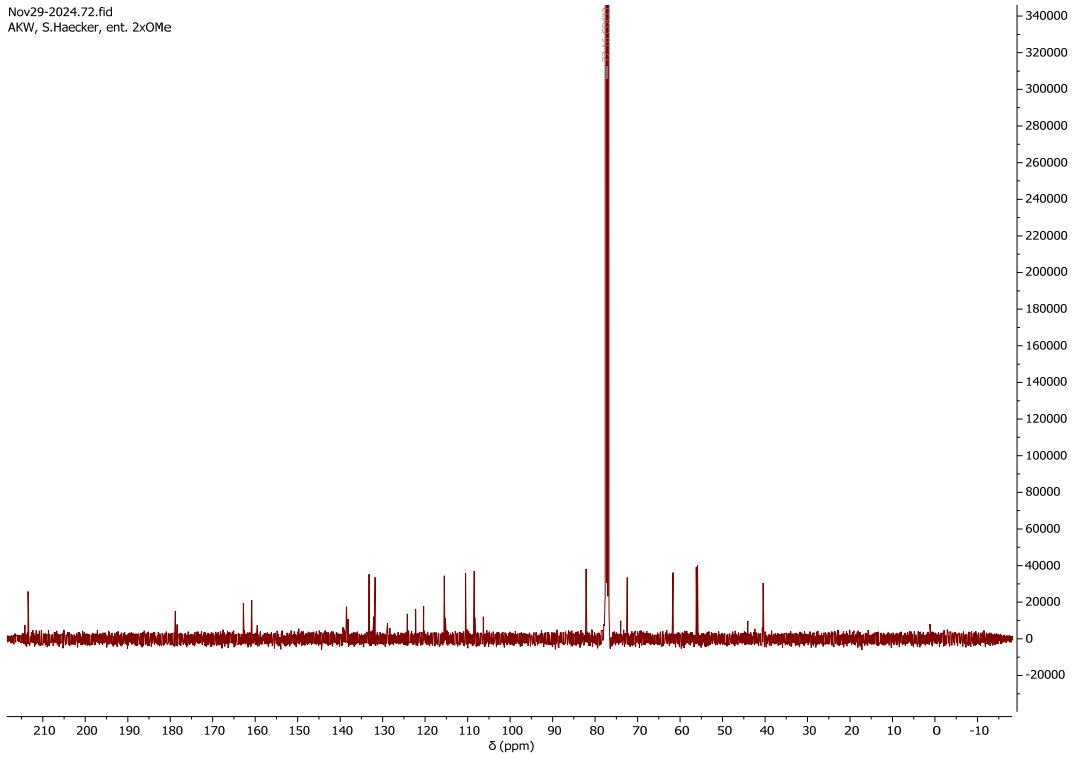


Figure S25: ^{13}C-NMR (101 MHz, $CDCl_3$) of **16**.

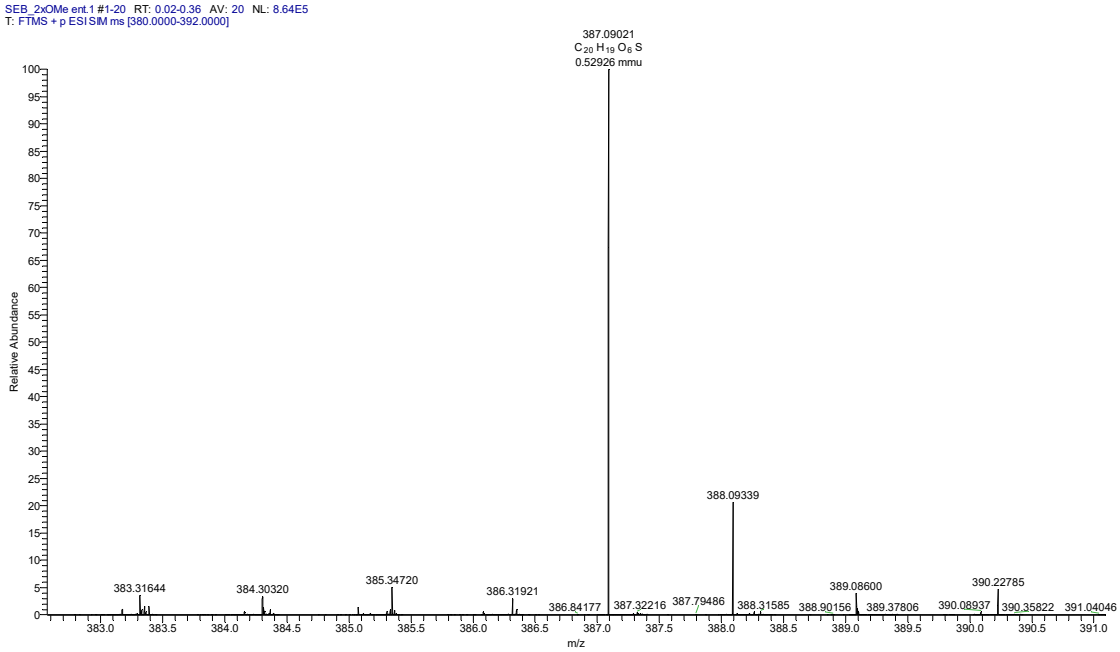
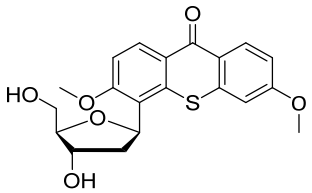


Figure S26: HR ESI-MS analysis of **16**.

Compound 2



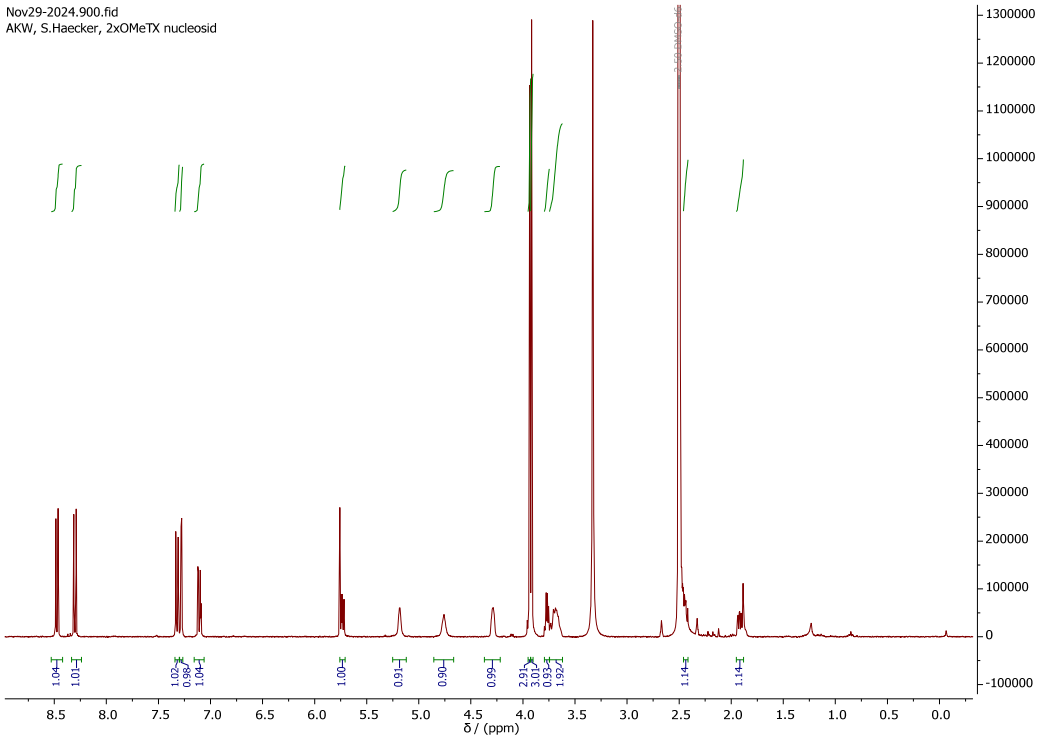


Figure S27: ^{1}H-NMR (400 MHz, DMSO-d6) of **2**.

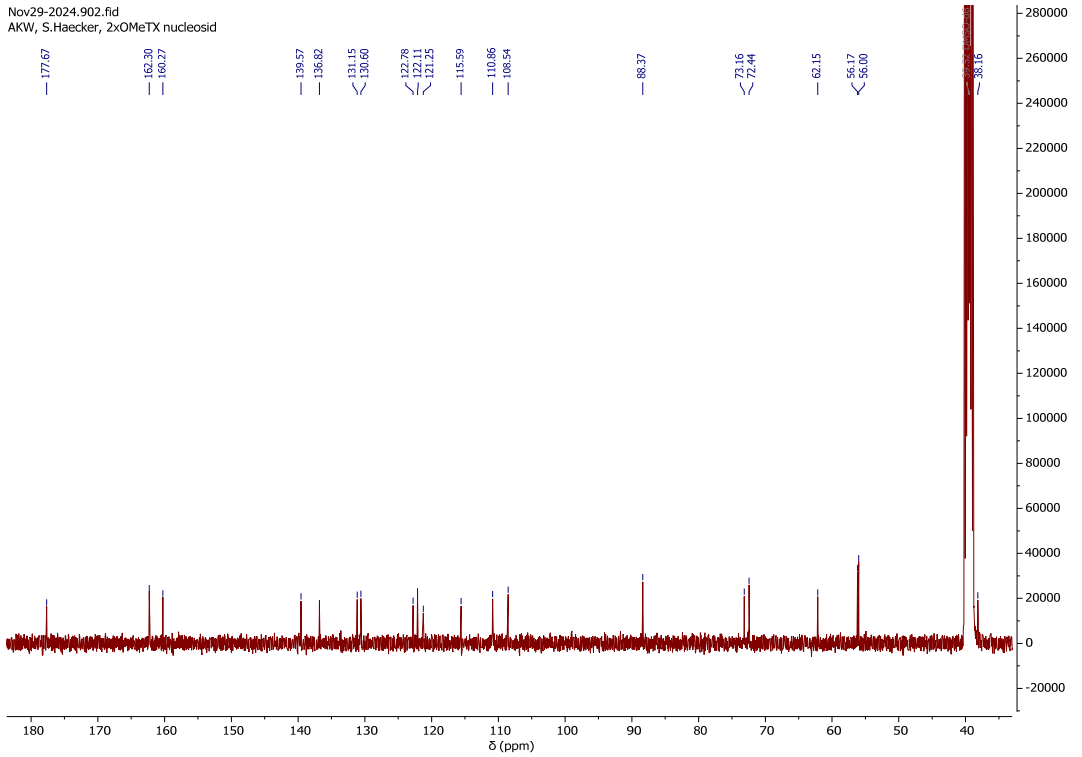


Figure S28: ^{13}C-NMR (101 MHz, in DMSO-d6) of **2**.

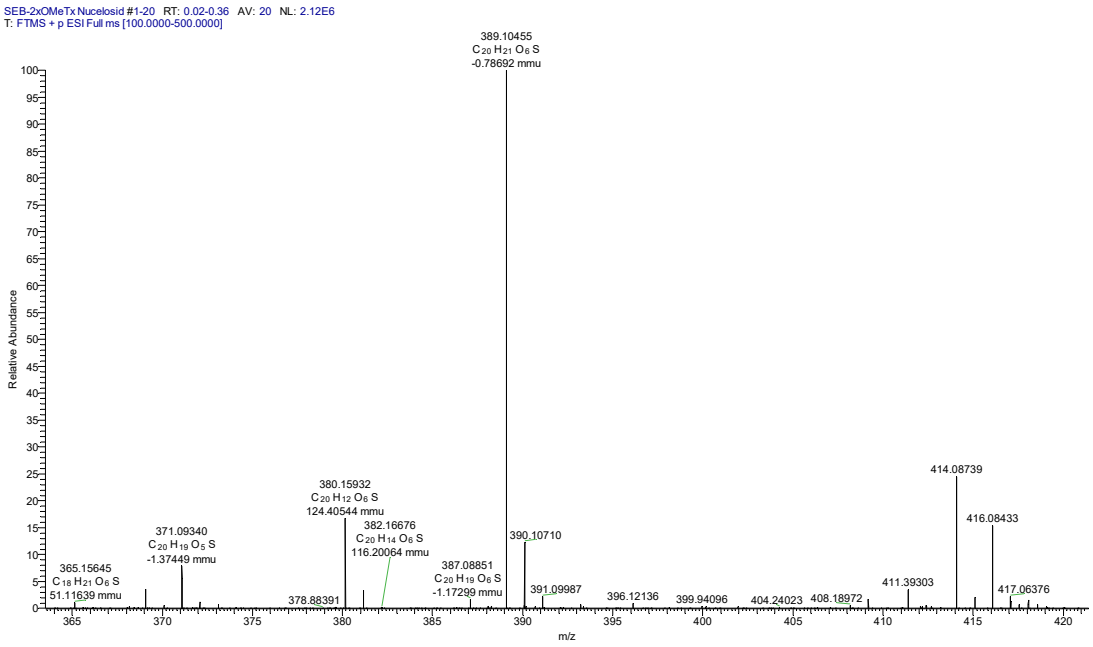


Figure S29: HR ESI-MS analysis of **2**.

Compound **17**

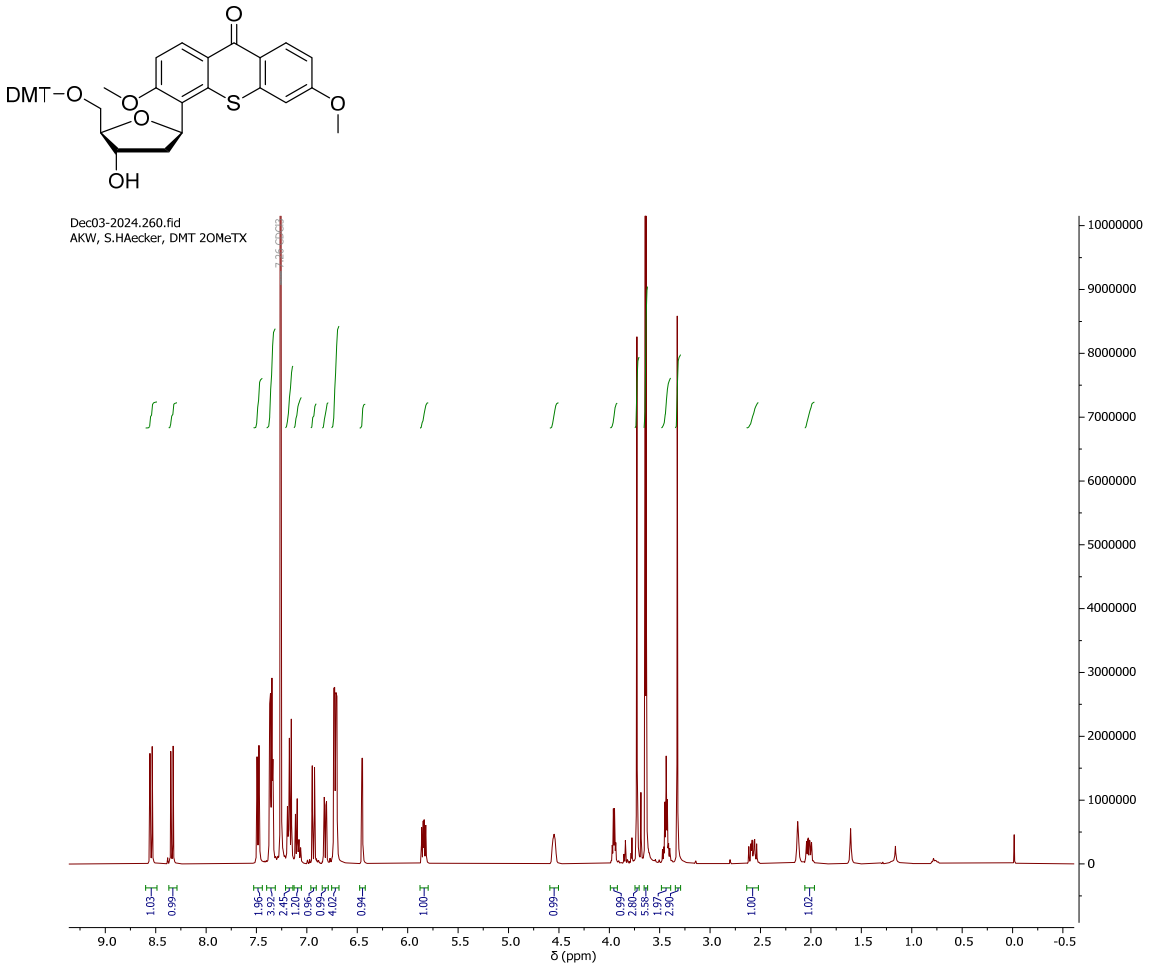


Figure S30: ^{1}H-NMR (400 MHz, $CDCl_3$) of **17**.

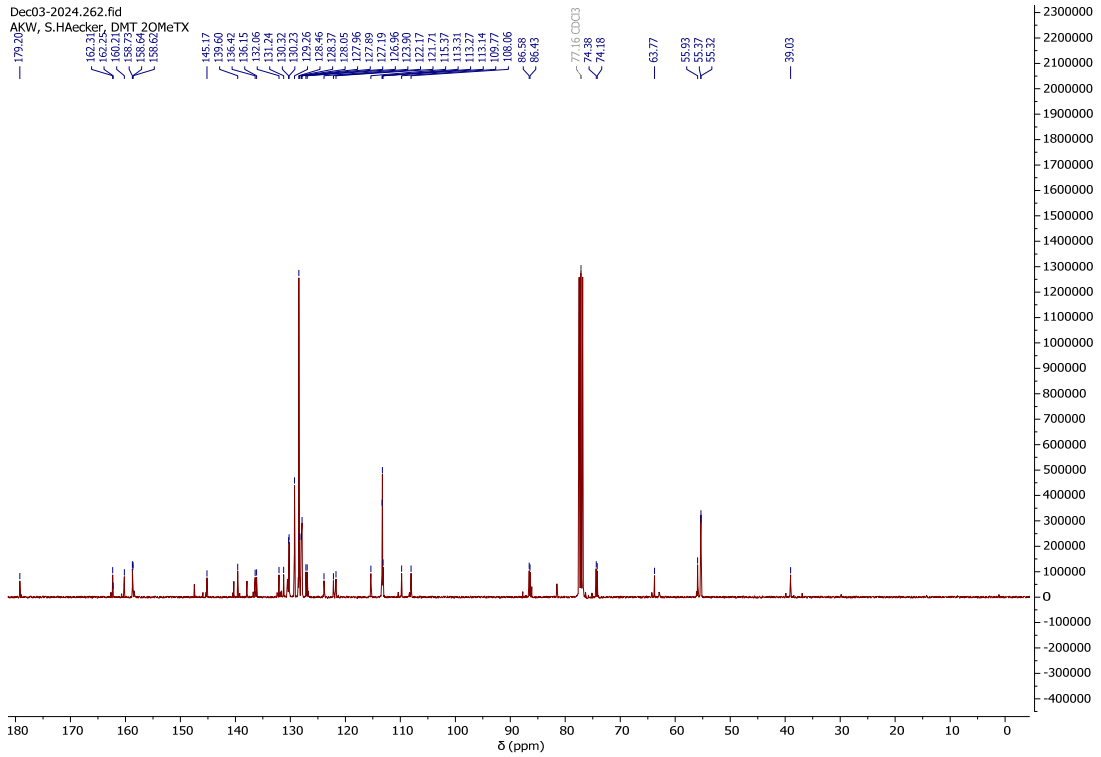


Figure S31: ^{13}C-NMR (101 MHz, $CDCl_3$) of **17**.

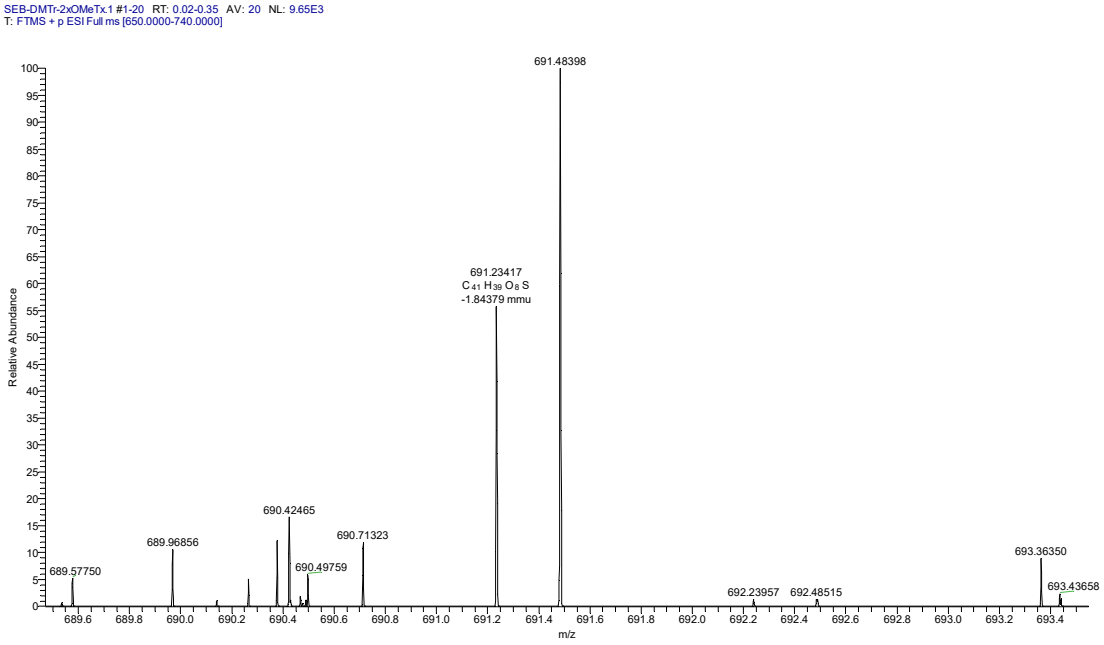
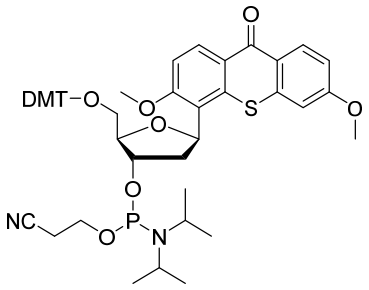


Figure S32: HR ESI-MS analysis of **17**.

Compound 18



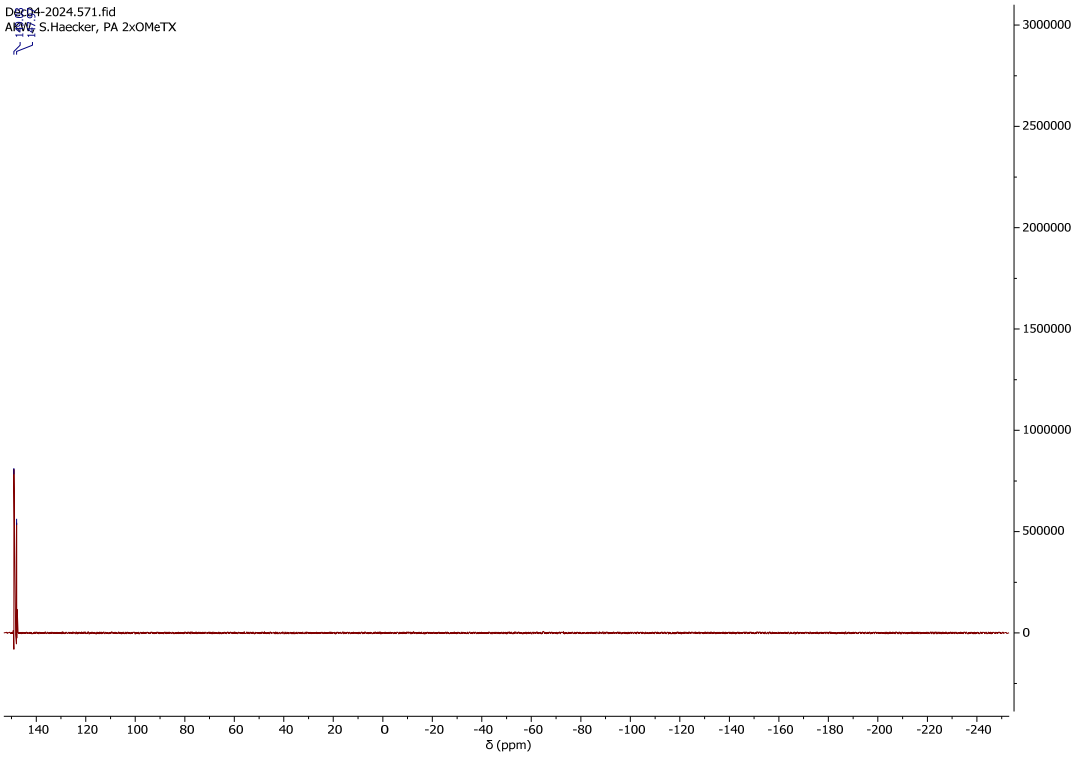


Figure S33: ^{31}P-NMR (162 MHz, $CDCl_3$) of **18**.

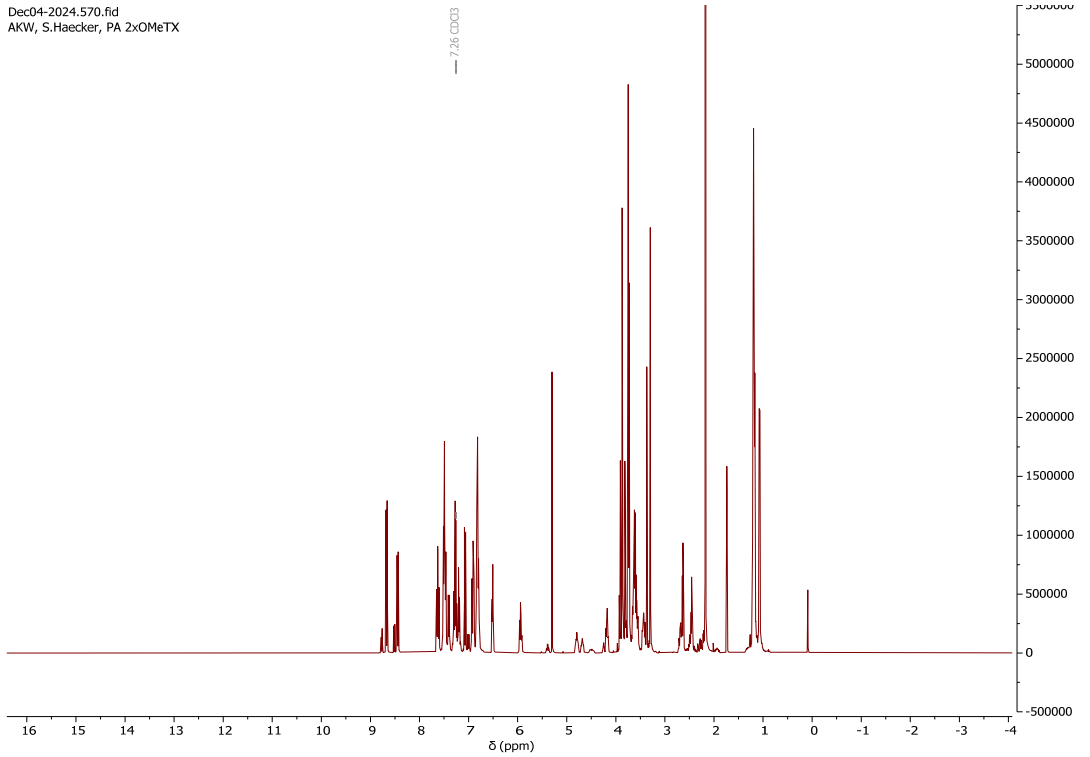


Figure S34: ^{1}H-NMR (400 MHz, $CDCl_3$) of **18**.

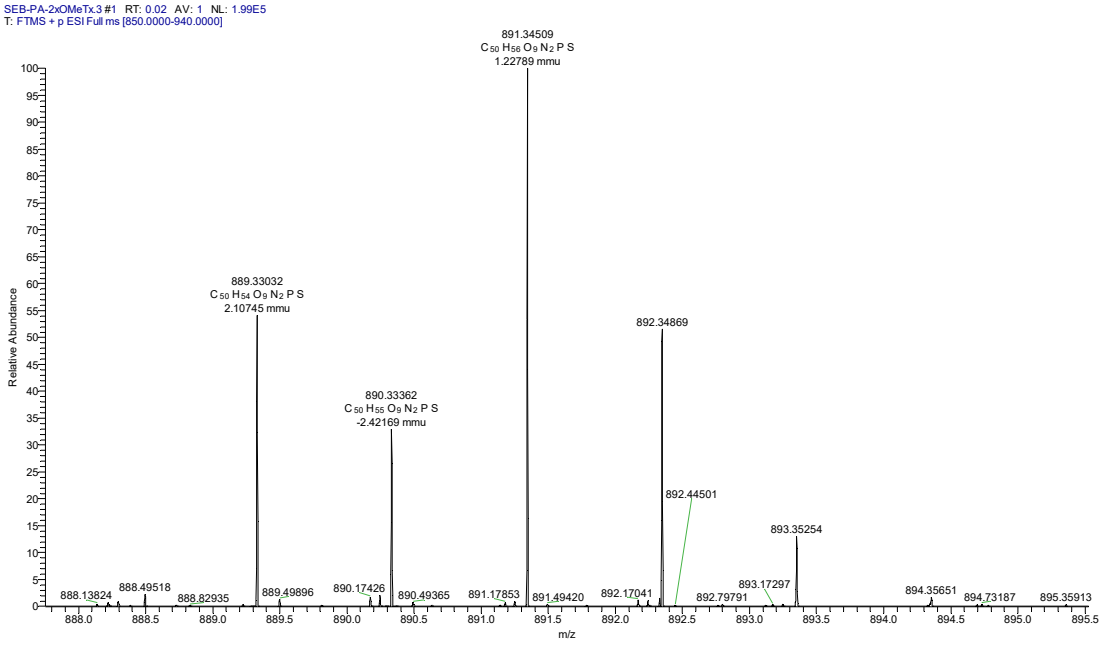


Figure S35: HR ESI-MS analysis of **18**.

6. DNA analyses and melting temperatures

The extinction coefficients ε_{260} of the **X**-modified oligonucleotides were determined using the following equation in $M^{-1} \cdot cm^{-1}$:

$$\varepsilon_{260} = \varepsilon_{260}X + (nA^* \varepsilon_A(15400) + nT^* \varepsilon_T(8800) + nG^* \varepsilon_G(11700) + nC^* \varepsilon_C(7300))^* 0.9.$$

Table S3: Extinction coefficients at 260 nm for the nucleotides and the synthesized C-nucleosides.

Nukleotid bzw. Chromophor Nukleosid	**ε_{260} [$L \cdot mol^{-1}cm^{-1}$]**
A	15400
T	8800
G	11700
C	7300
1	13100
2	35080

Table S4: Extinction coefficients and MALDI masses of synthesized X-modified single-stranded DNA (DNAX-n). *m/z: m/3 determined by LC-ESI-MS.

DNAX-n	ε_{260} [$M^{-1} \cdot cm^{-1}$]	Mass calculated [Da]	Mass found [Da]
DNA1-0	164750	4748	4751
DNA1-1	186530	5365	5365
DNA2-0	186730	4744	4755
DNA2-1	208510	5361	5367
DNA2-2	230290	5978	5979
DNA2-3	252070	6596	6619
DNA2-2-G-C	223000	5954	1987*

Table S5: DNA Melting Temperatures of Hybridized DNA Architectures: DNAX-n.

DNAX-n	Melting temperature [°C]	
	T_{m1}	**T_{m2}**
DNA1-0	15	47
DNA1-1	11	46
DNA1-0	11	42
DNA2-1	11	44
DNA2-2	12	47
DNA2-3	12	50
DNA2-2-G-C	13	50

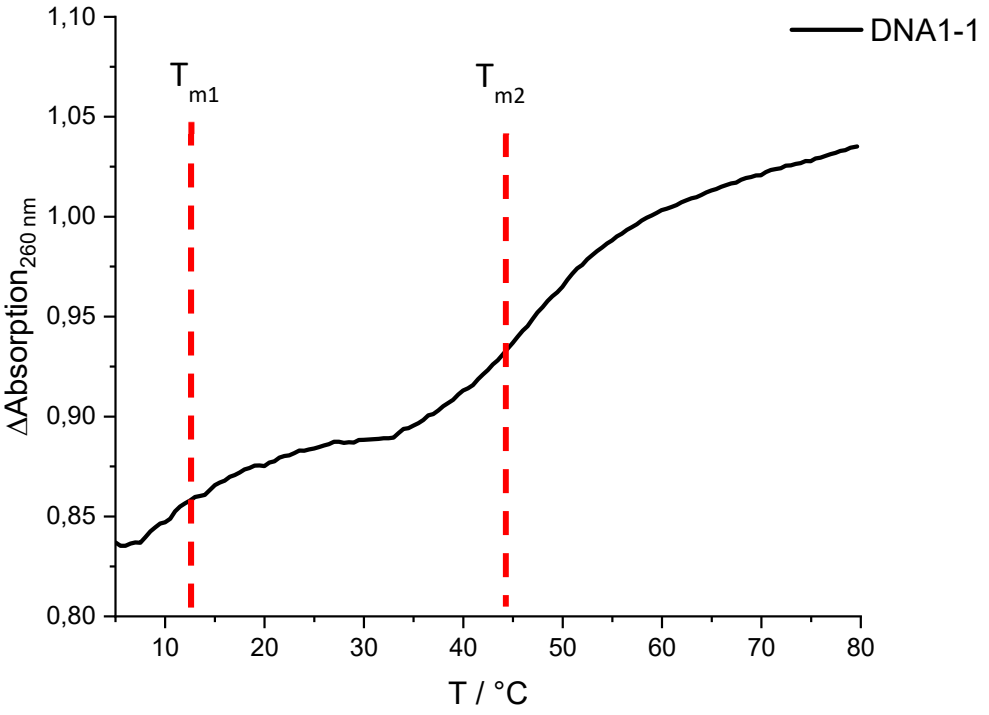


Figure S36: Exemplary melting curve of **DNA1-1** (2.5 µM DNA, 250 mM NaCl, 10 mM Na-Pi buffer at 10 °C).

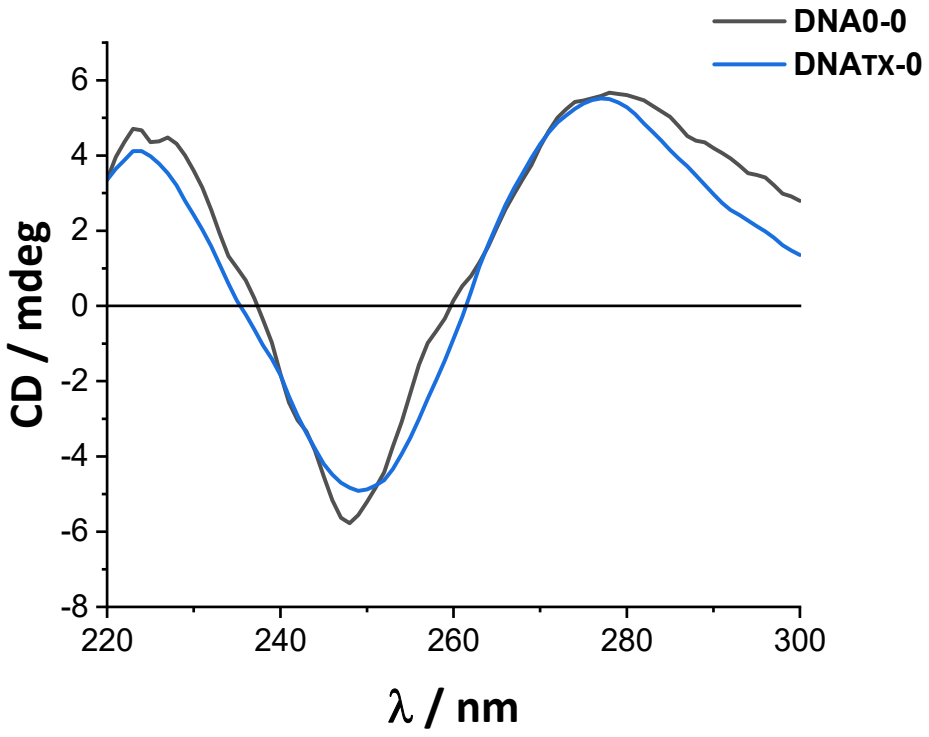


Figure S37: CD spectrum of **DNA0-0**, and **DNATX-0** (2.5 µM DNA, 250 mM NaCl, 10 mM Na-Pi buffer at 10 °C).

7. Images of MS and HPLC analyses of the DNA

DNA1-0

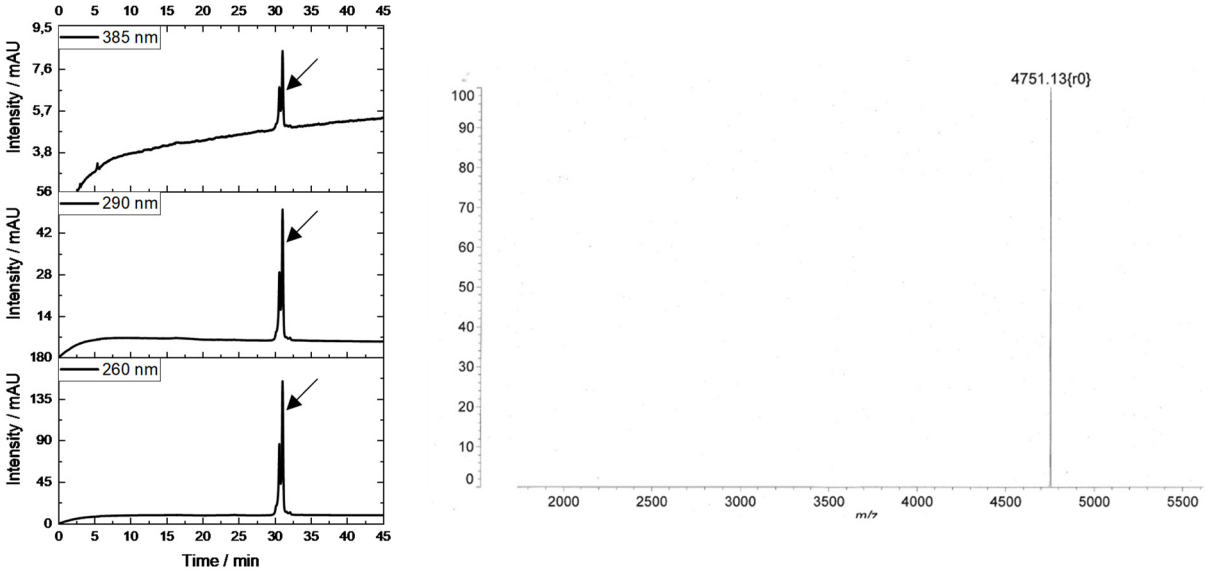


Figure S38: Left: Analytical HPLC chromatogram of DNA1-0, detected at λ = 385 nm, 290 nm, and 260 nm. Right: MALDI-TOF-MS spectrum of DNA1-0.

DNA1-1

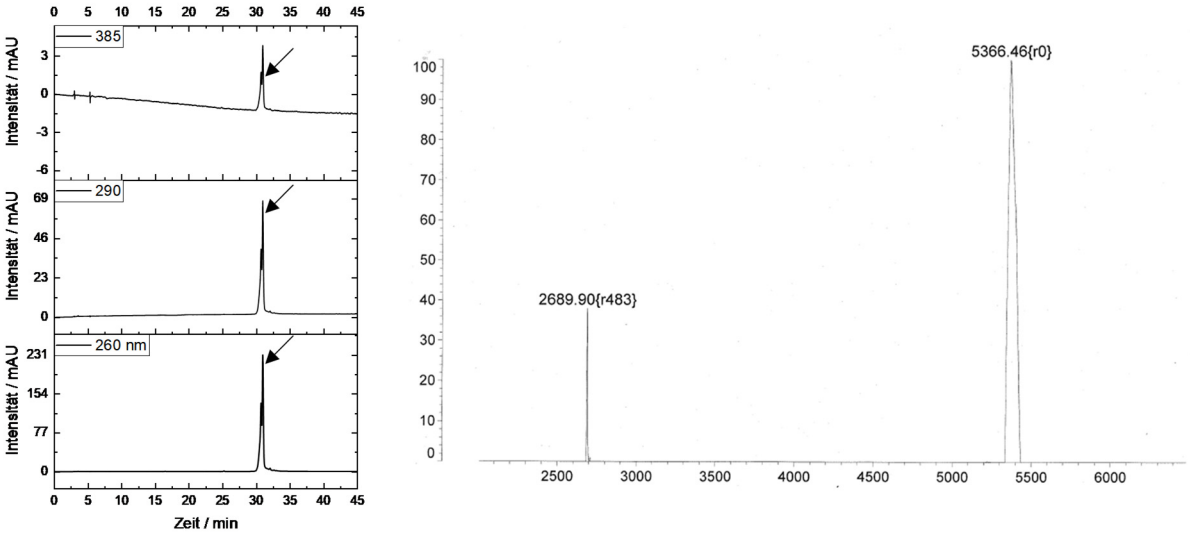


Figure S39: Left: Analytical HPLC chromatogram of DNA1-1, detected at λ = 385 nm, 290 nm, and 260 nm. Right: MALDI-TOF-MS spectrum of DNA1-1.

DNA2-0

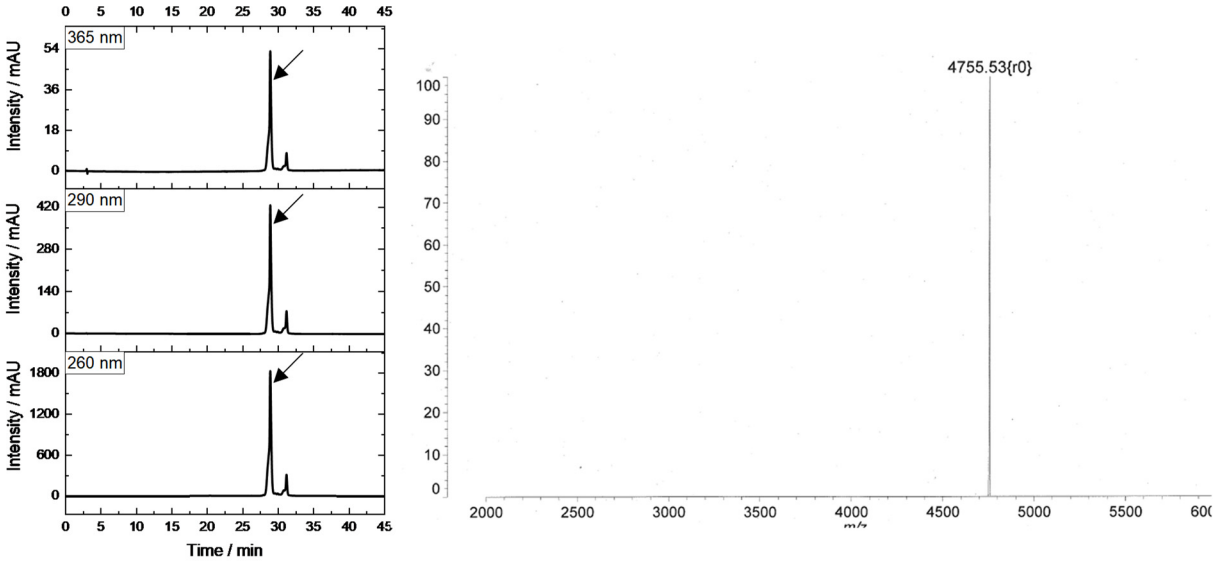


Figure S40: Left: Analytical HPLC chromatogram of DNA2-0, detected at λ = 365 nm, 290 nm, and 260 nm. Right: MALDI-TOF-MS spectrum of DNA2-0.

DNA2-1

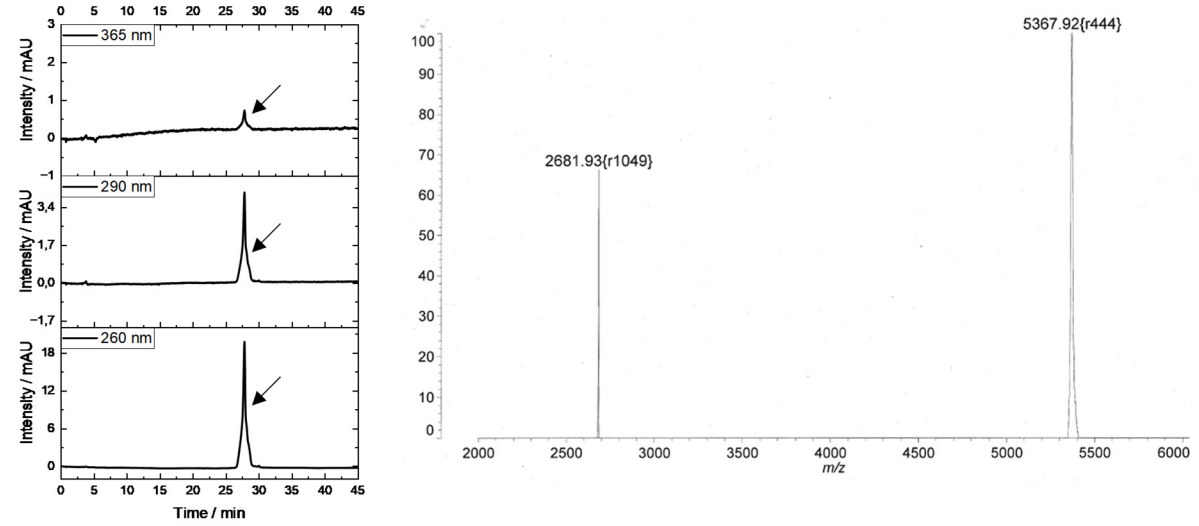


Figure S41: Left: Analytical HPLC chromatogram of DNA2-1, detected at λ = 365 nm, 290 nm, and 260 nm. Right: MALDI-TOF-MS spectrum of DNA2-1.

DNA2-2

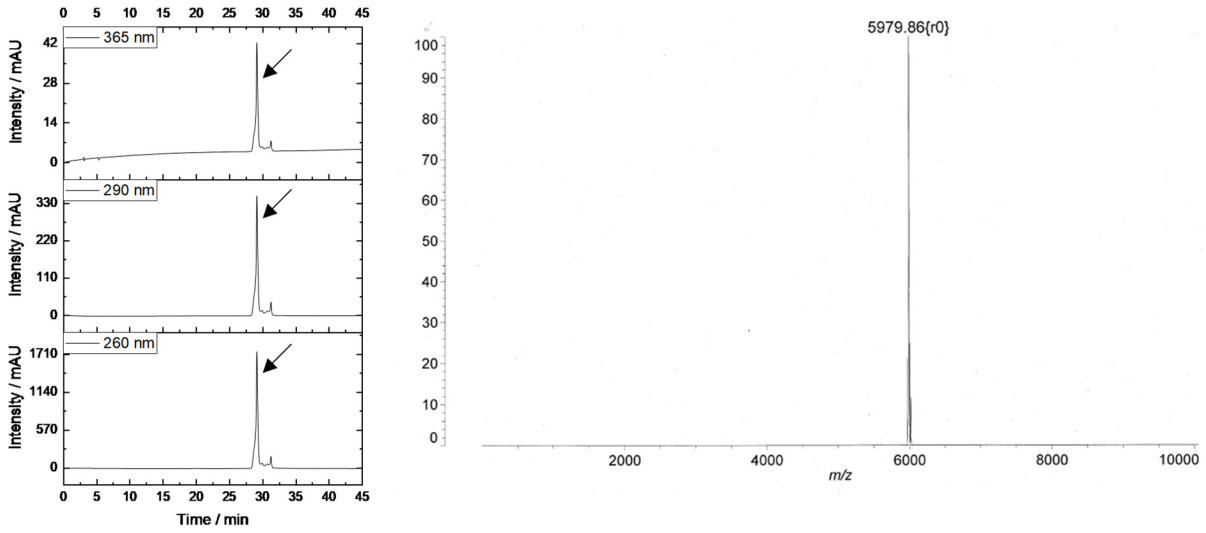


Figure S42: Left: Analytical HPLC chromatogram of DNA2-2, detected at λ = 365 nm, 290 nm, and 260 nm. Right: MALDI-TOF-MS spectrum of DNA2-2.

DNA2-3

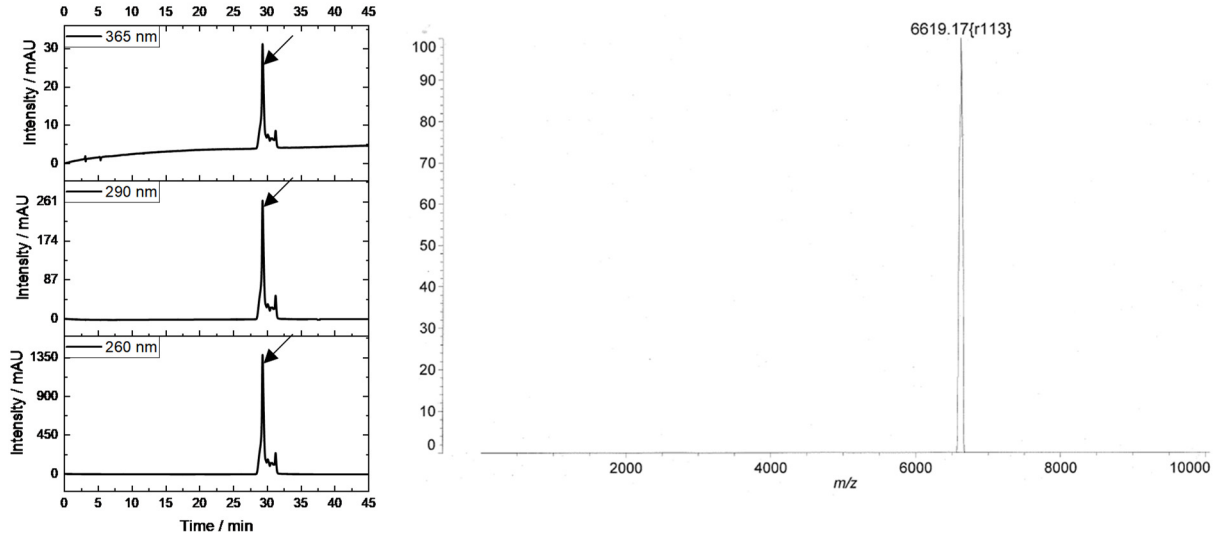


Figure S43: Left: Analytical HPLC chromatogram of DNA2-3, detected at λ = 365 nm, 290 nm, and 260 nm. Right: MALDI-TOF-MS spectrum of DNA2-3.

DNA2-2-G-C

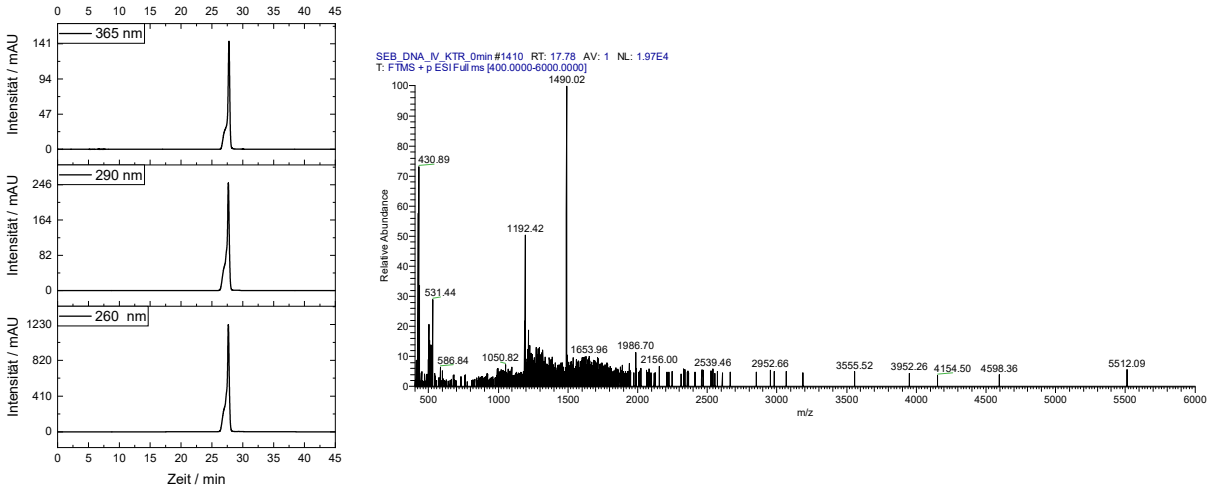


Figure S44: Left: Analytical HPLC chromatogram of DNA2-2-G-C, detected at λ = 365 nm, 290 nm, and 260 nm. Right: Determined Mass of the purified Oligonucleotide via ESI-MS [ESI, M^+] m/z: m/3: found 1987; calculated: 1986; m/4: found: 1490; calculated: 1490.

HPLC Analysis of the irradiation experiments

As a negative control, the counterbase thymine was used instead of the photosensitizer and the DNA architecture DNA0-0 was exposed to a 385 nm and 365 nm LED. No DNA damage was observed.

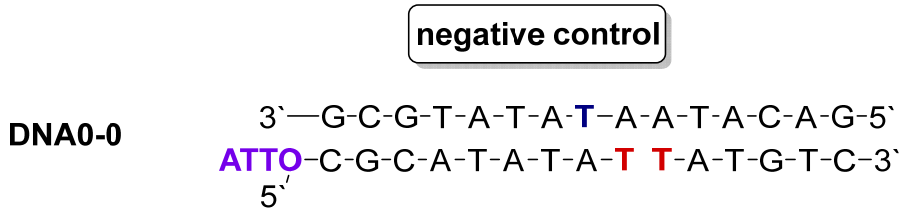


Figure S45: Sequence of the negative control which was exposed with 365/385 nm LED.

Irradiation of negative control DNA0-0

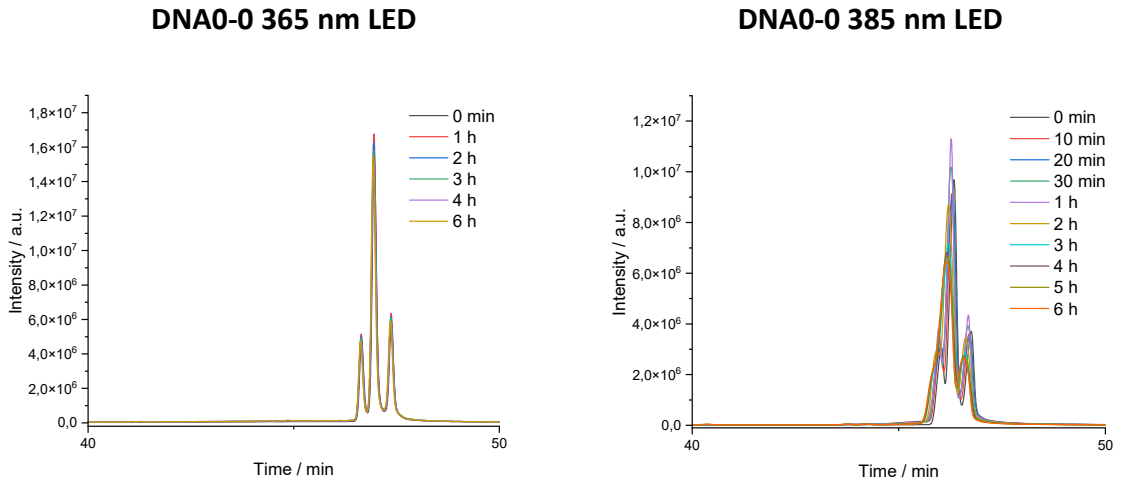


Figure S46: Left: HPLC emission spectrum of the DNA architecture DNA0-0 after different exposure times with a 365 nm LED. Right: HPLC emission spectrum of the DNA architecture DNA0-0 after different exposure times with a 385 nm LED. The signal loss during irradiation by the 385 nm LED can be attributed to the photochemical bleaching of the fluorescent ATTO marker which depends on the excitation wavelength.

The CPD and ICL yields were determined as the ratio of the integrals in the HPLC emission spectrum:

CPD (%) = [Integral (blue) / Integral (blue + red + green)] * 100.
ICL (%) = [Integral (green) / Integral (green + blue + red)] * 100.

Emission HPLC analysis of Irradiation experiment DNA1-0:

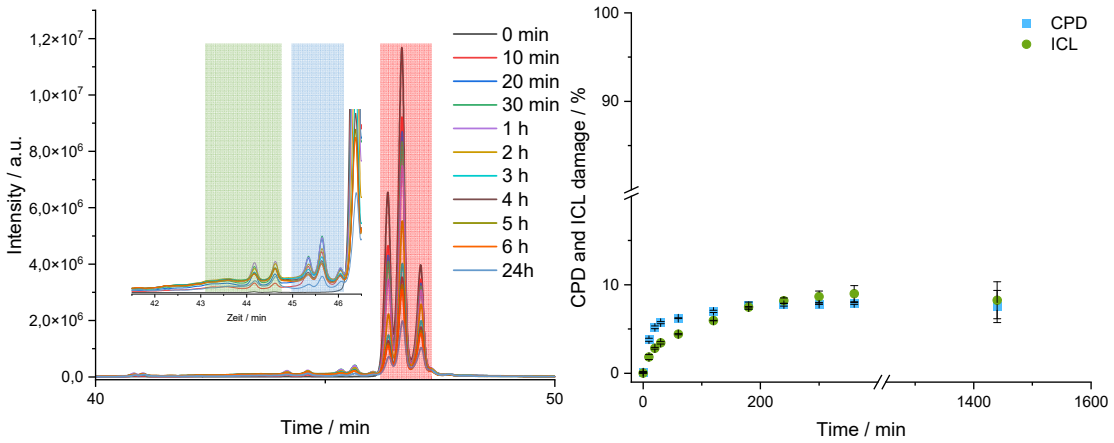


Figure S47: Left: HPLC emission spectrum of the DNA architecture DNA1-0 after different exposure times. Right: CPD (blue) and ICL (green) yields in percent over time.

Emission HPLC analysis of Irradiation experiment DNA1-1:

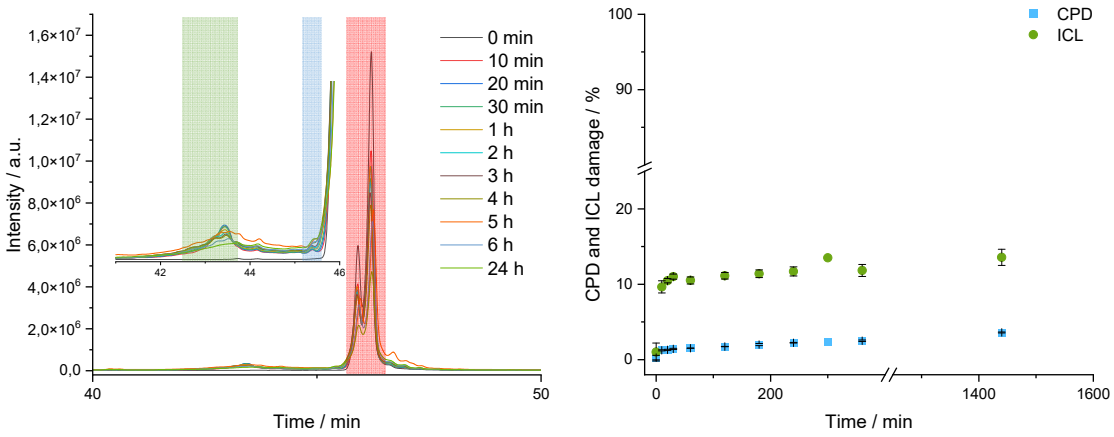


Figure S48: Left: HPLC emission spectrum of the DNA architecture DNA1-1 after different exposure times. Right: CPD (blue) and ICL (green) yields in percent over time.

Emission HPLC analysis of Irradiation experiment DNA2-0:

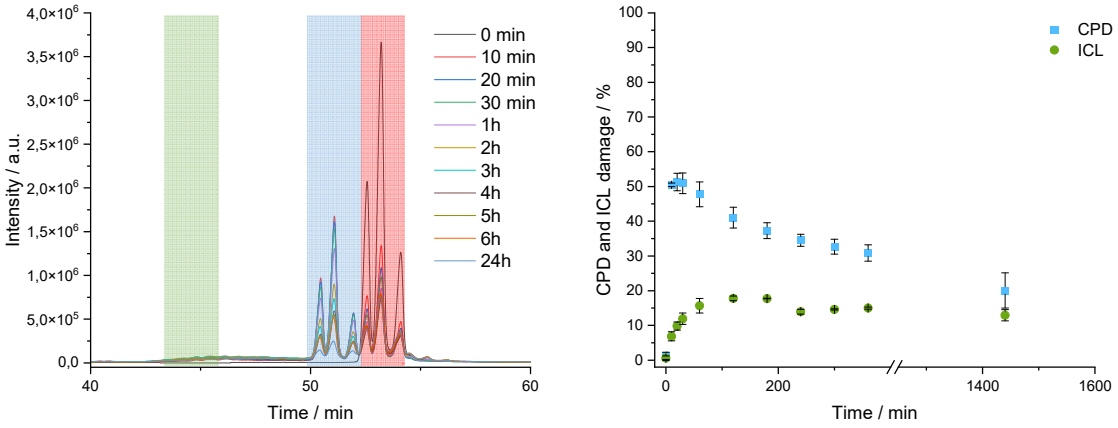


Figure S49: Left: HPLC emission spectrum of the DNA architecture DNA2-0 after different exposure times. Right: CPD (blue) and ICL (green) yields in percent over time.

Emission HPLC analysis of Irradiation experiment DNA2-1:

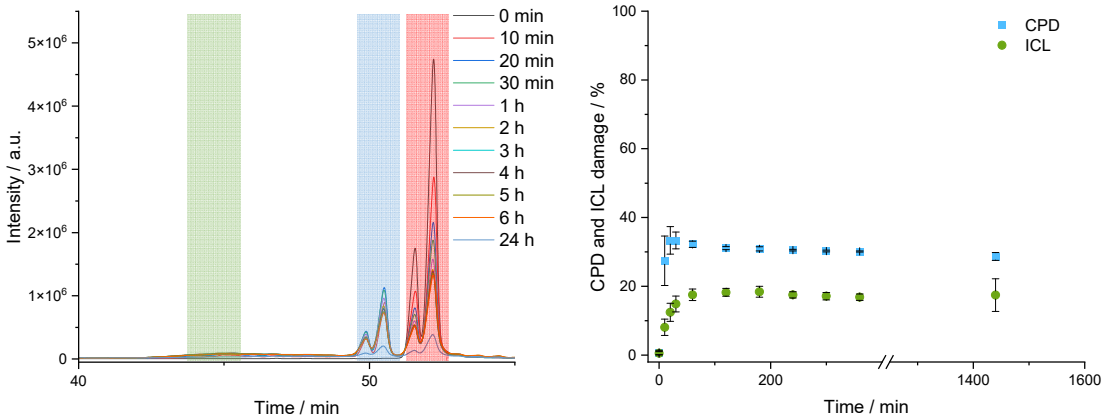


Figure S50: Left: HPLC emission spectrum of the DNA architecture DNA2-1 after different exposure times. Right: CPD (blue) and ICL (green) yields in percent over time.

Emission HPLC analysis of Irradiation experiment DNA2-2:

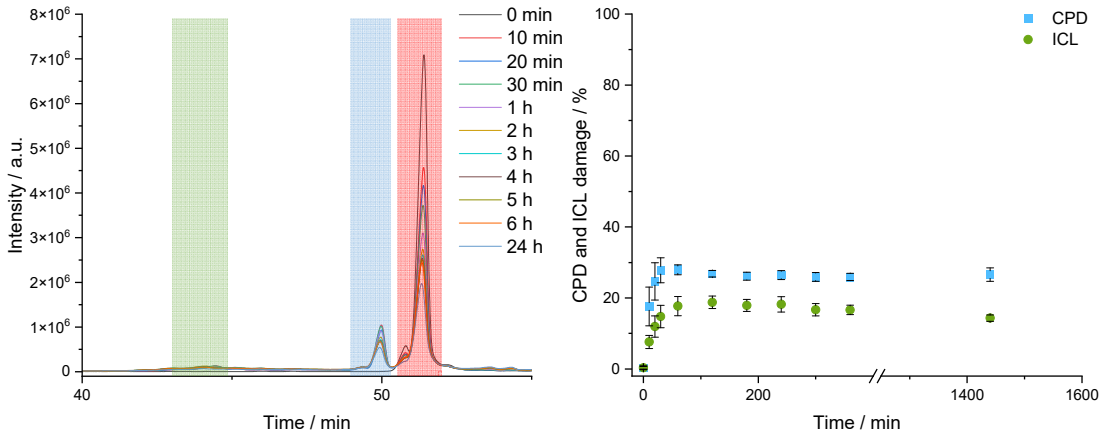


Figure S51: Left: HPLC emission spectrum of the DNA architecture DNA2-2 after different exposure times. Right: CPD (blue) and ICL (green) yields in percent over time.

Emission HPLC analysis of Irradiation experiment DNA2-3:

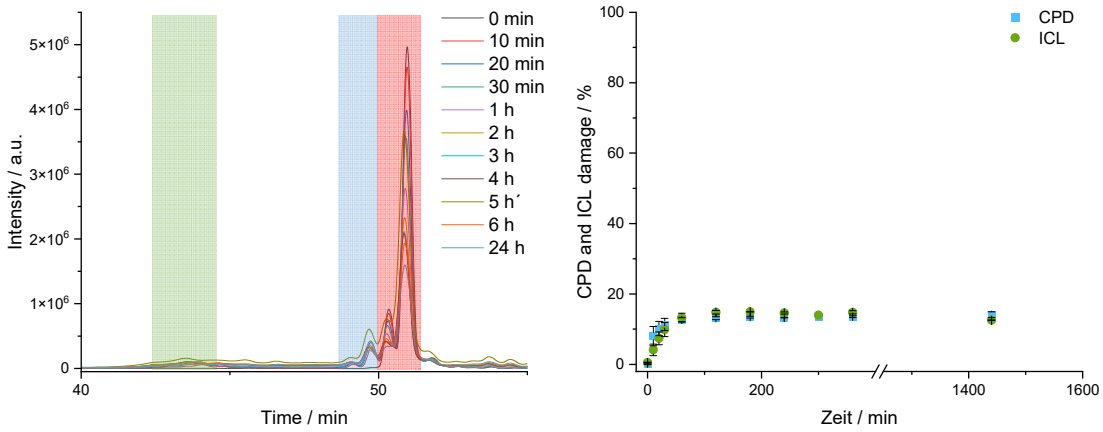


Figure S52: Left: HPLC emission spectrum of the DNA architecture DNA2-3 after different exposure times. Right: CPD (blue) and ICL (green) yields in percent over time.

Emission HPLC analysis of Irradiation experiment DNA2-2-G-C:

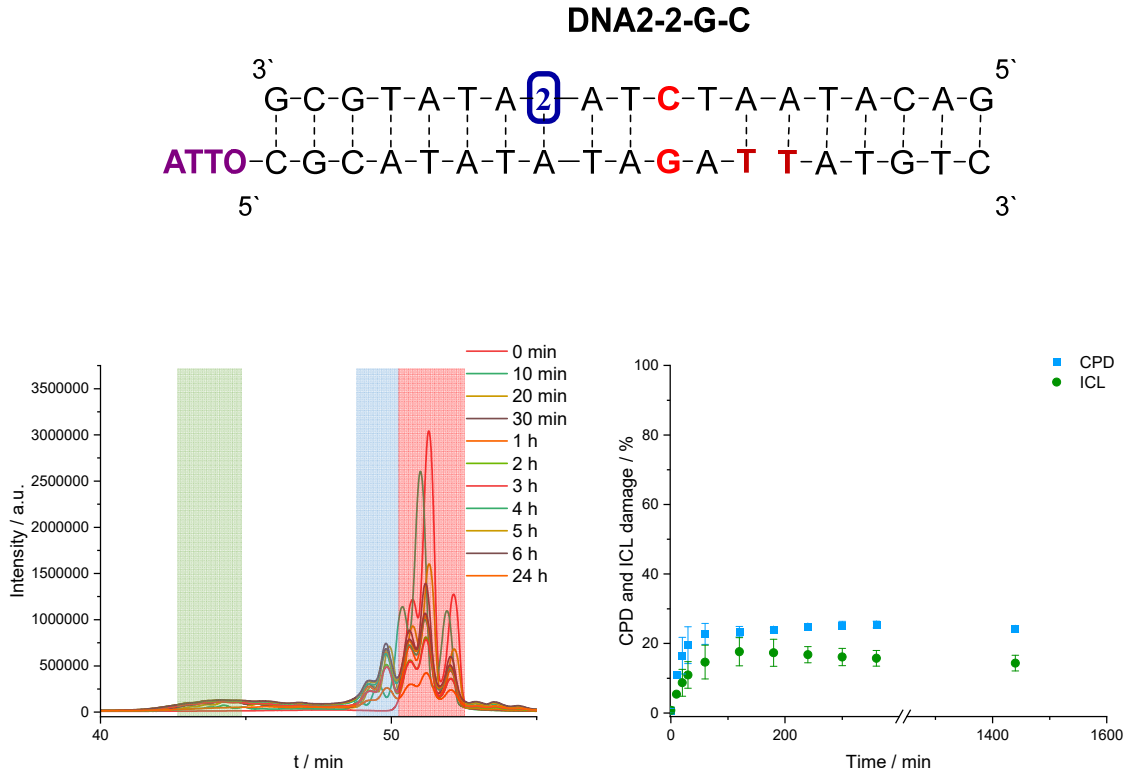


Figure S53: Left: HPLC emission spectrum of the DNA architecture DNA2-2-G-C after different exposure times. Right: CPD (blue) and ICL (green) yields in percent over time.

Images of LC-MS

DNA1-0 after 0 min Irradiation:

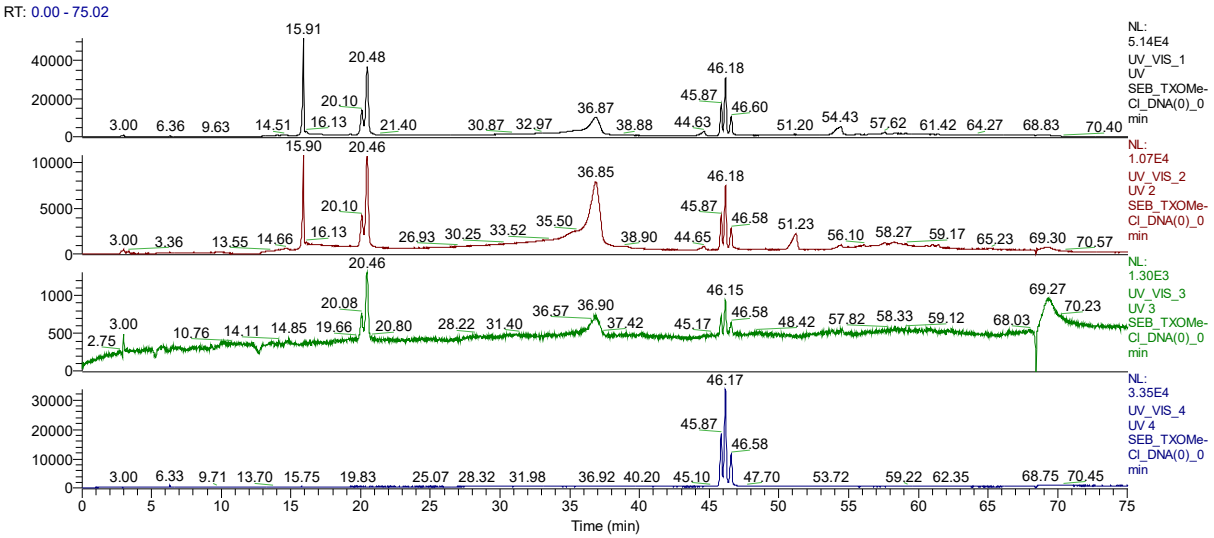


Figure S54: LC-MS chromatogram overview of the different wavelengths (260 nm, 290 nm, 365 nm, 554 nm) after 0 min of irradiation.

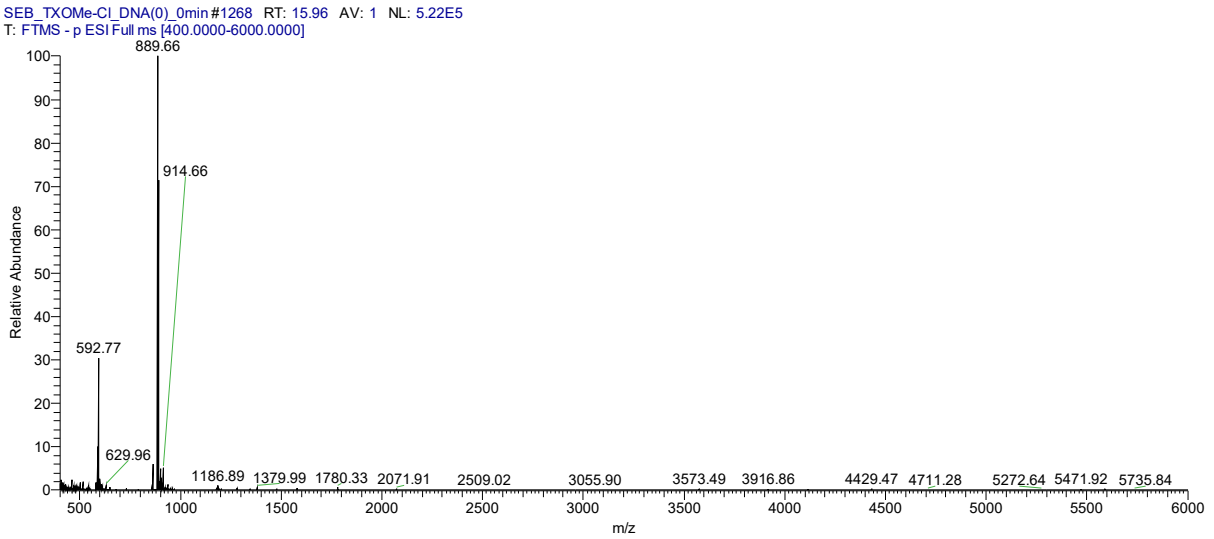


Figure S55: LC-MS analysis of the complementary strand. RT = 15.96 min: [ESI, M]+ m/z: found: 1780; calculated: 1781; m/2: found: 890; calculated: 891.

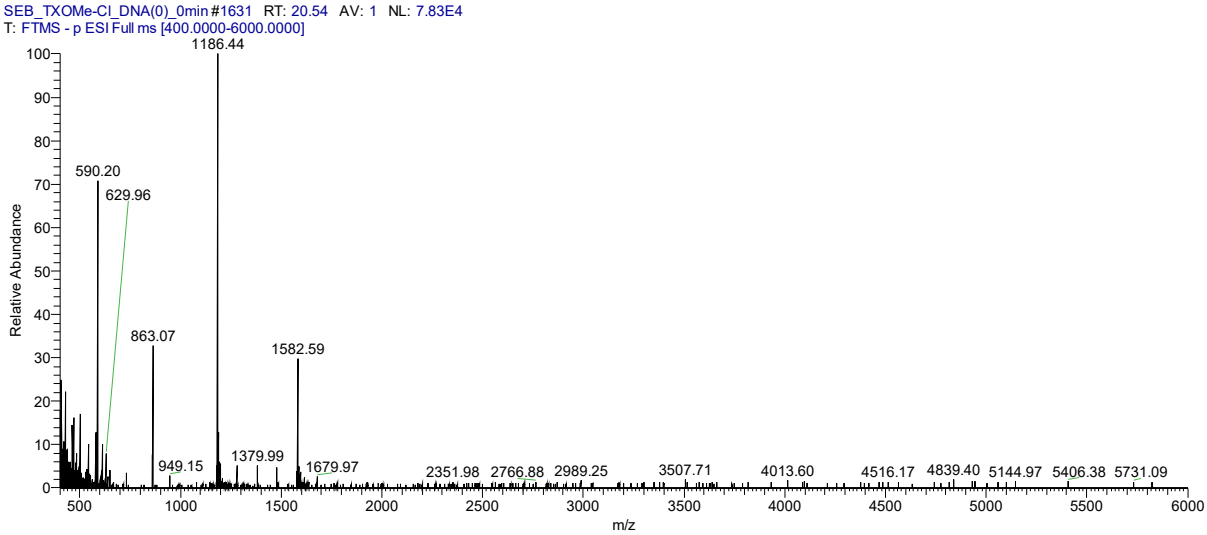


Figure S56: LC-MS analysis of the template strand. RT = 20.54 min: [ESI, M]+ m/z: m/3: found: 1583; calculated: 1583; m/4: found: 1186; calculated: 1187.

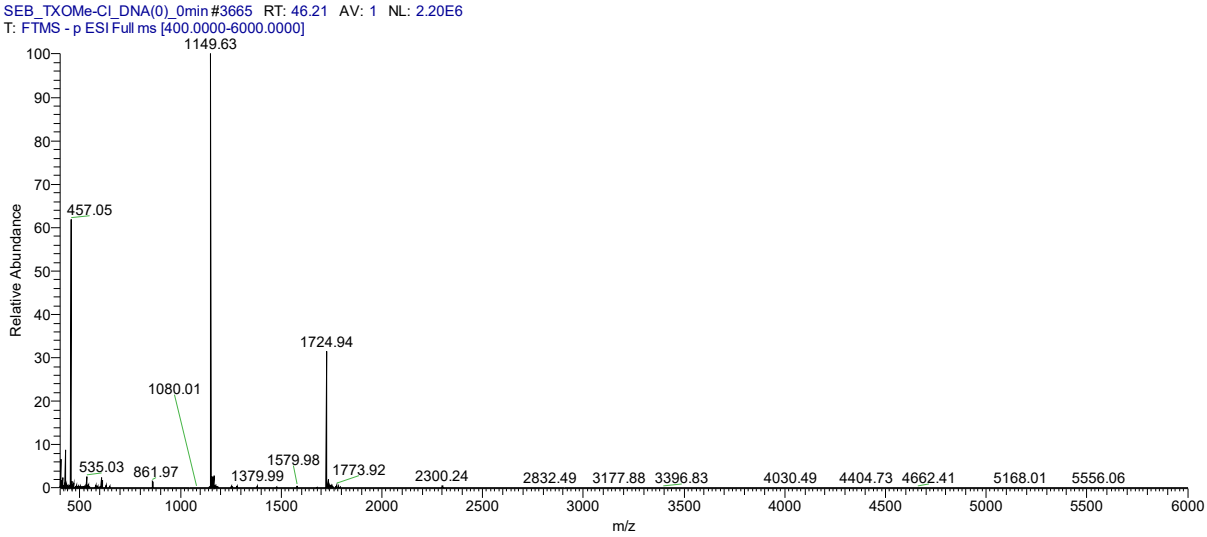


Figure S57: LC-MS analysis of the attostrand (AS). RT = 46.21 min: [ESI, M]+ m/z: m/2: found: 1725; calculated: 1726; m/3: found: 1150; calculated: 1151; m/4: found: 862; calculated: 863.

DNA1-0 after 6 h Irradiation

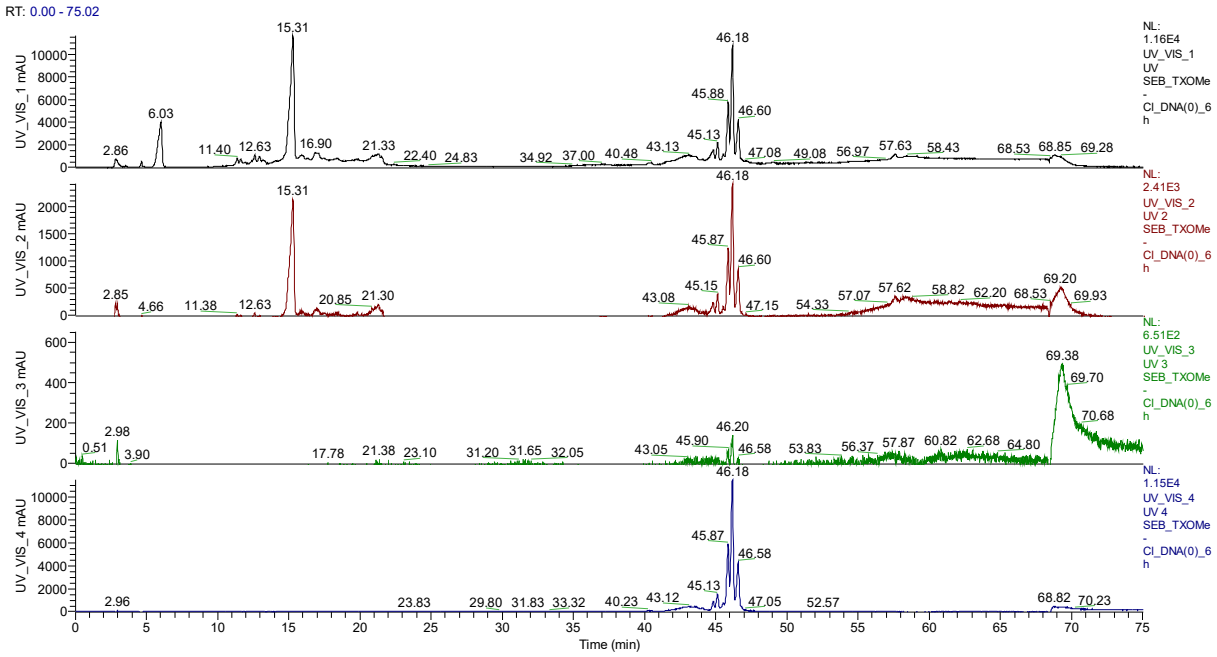


Abbildung S1: LC-MS chromatogram overview of the different wavelengths (260 nm, 290 nm, 365 nm, 554 nm) after 6 h of irradiation.

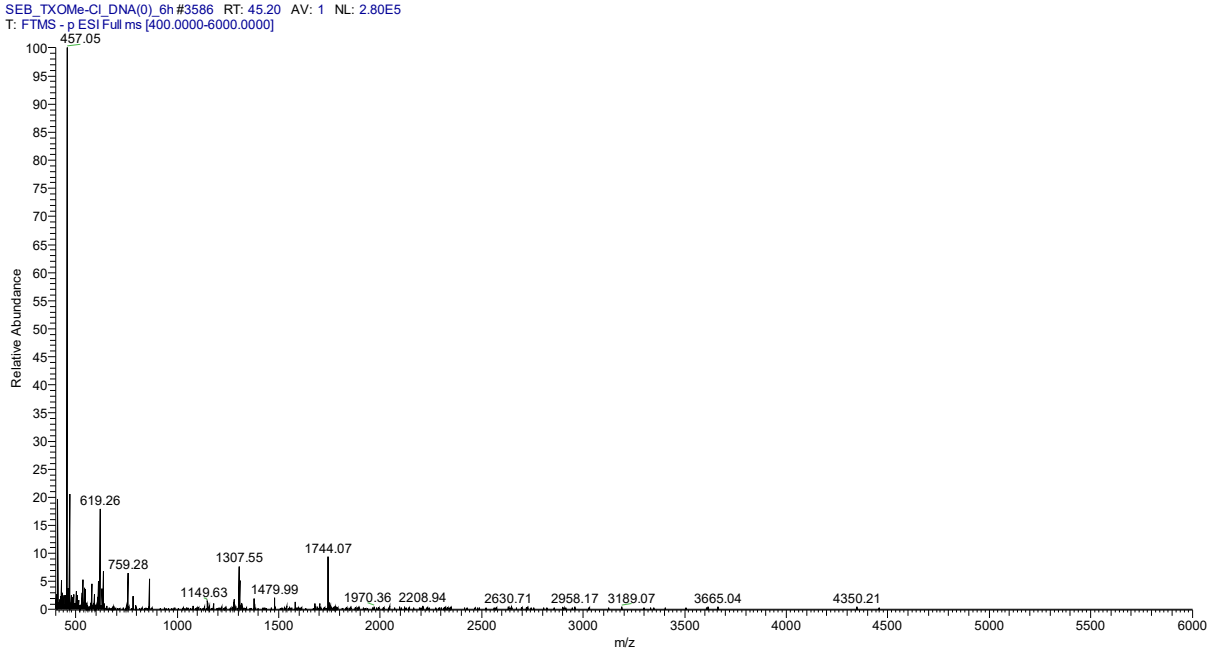


Figure S58: LC-MS analysis of the CPD damage. RT = 45.20 min: [ESI, M]+ m/z: m/3: found: 1744; calculated: 1744; m/4: found: 1308; calculated: 1308.

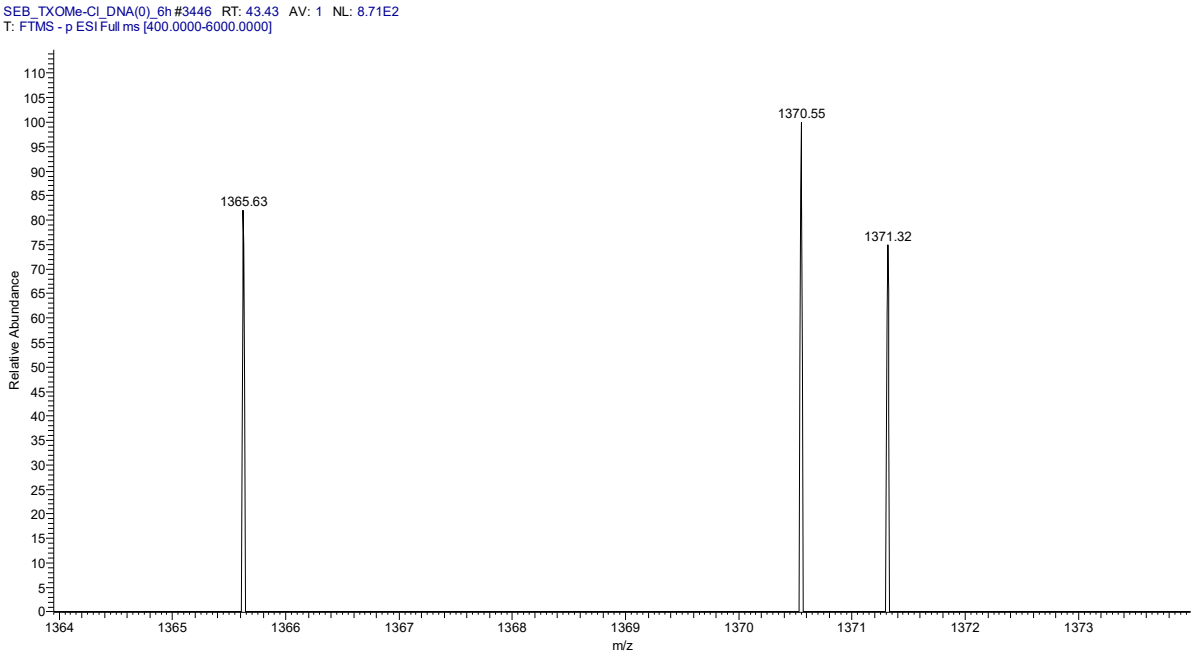


Figure S59: LC-MS analysis of the ICL damage. RT = 43.43 min: [ESI, M]+ m/z: m/6: found: 1366; calculated: 1367.

DNA2-2 after 0 min Irradiation

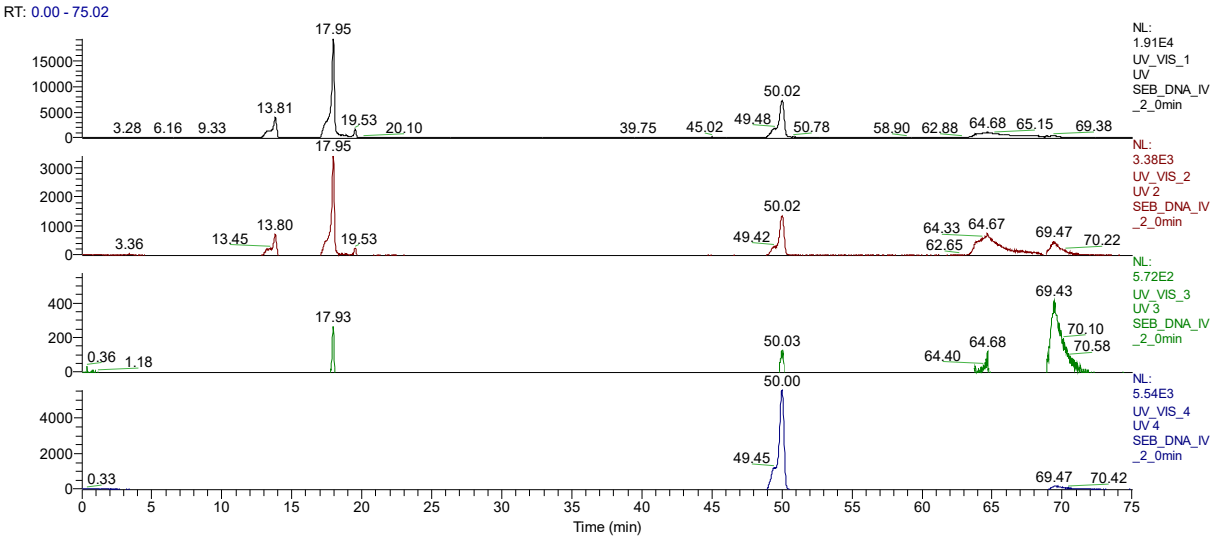


Figure S60: LC-MS chromatogram overview of the different wavelengths (260 nm, 290 nm, 365 nm, 554 nm) after 0 min of irradiation.

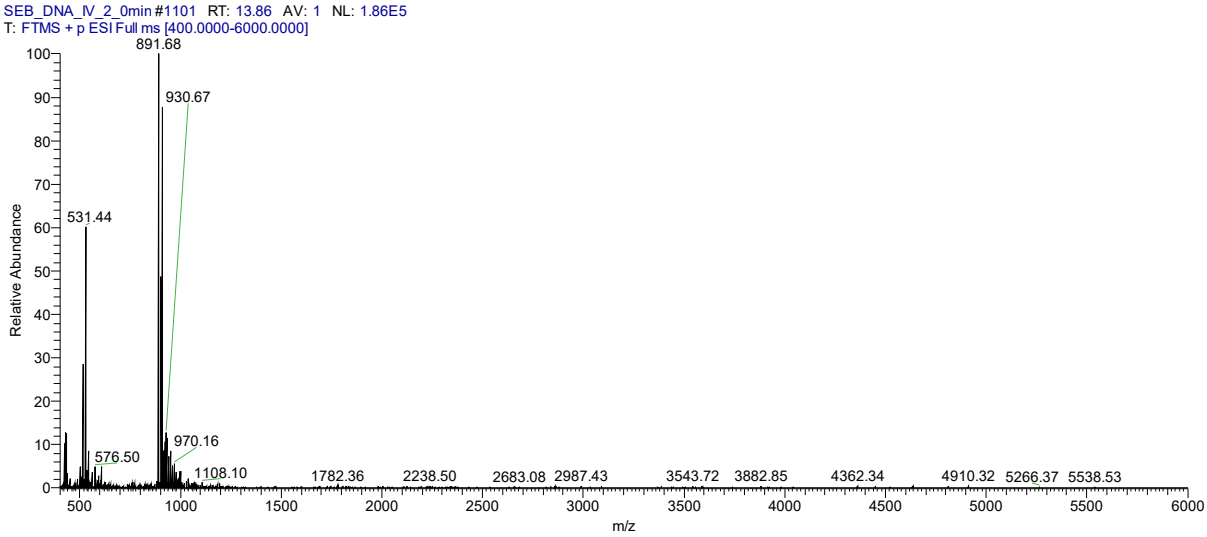


Figure S61: LC-MS analysis of the complementary strand. RT = 13.86 min: [ESI, M]+ m/z: found: 1782; calculated: 1781; m/2: found: 892; calculated: 891.

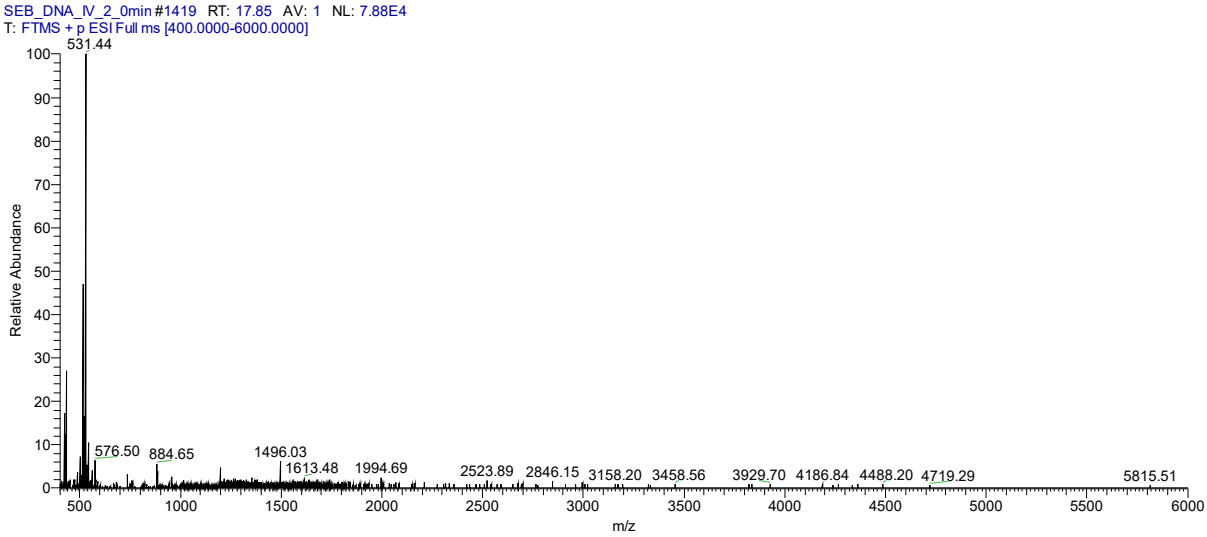


Figure S62: LC-MS analysis of the template strand. RT = 17.85 min: [ESI, M]+ m/z: m/3: found: 1995; calculated: 1993; m/4: found: 1496; calculated: 1495.

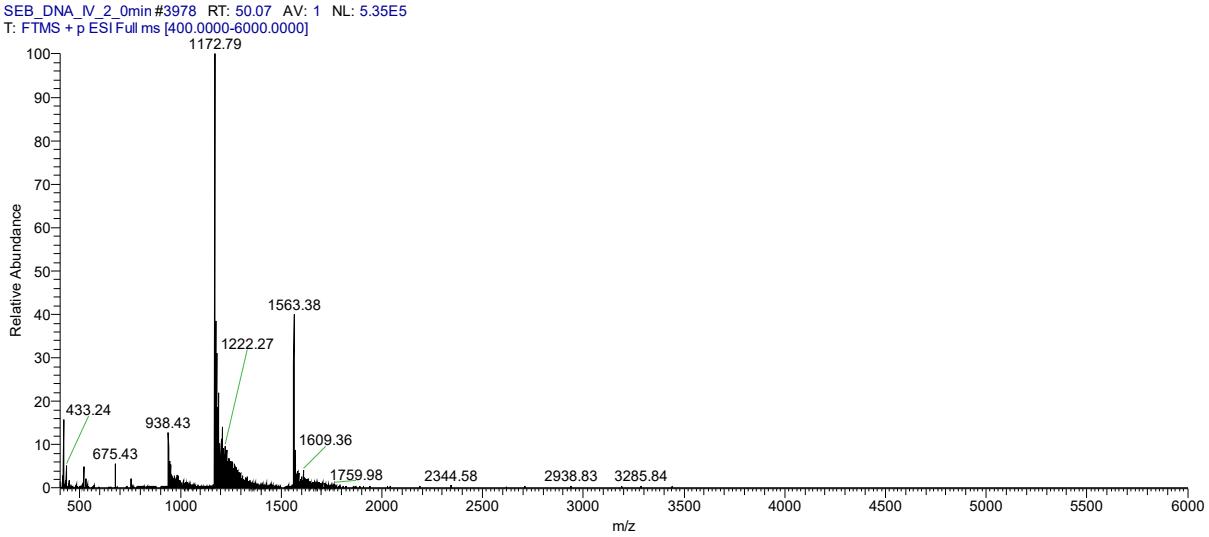


Figure S63: LC-MS analysis of the attostrand (AS). RT = 50.07 min: [ESI, M]+ m/z: m/2: found: 2345; calculated: 2344; m/3: found: 1563; calculated: 1562; m/4: found: 1173; calculated: 1172.

DNA2-2 after 30 min Irradiation

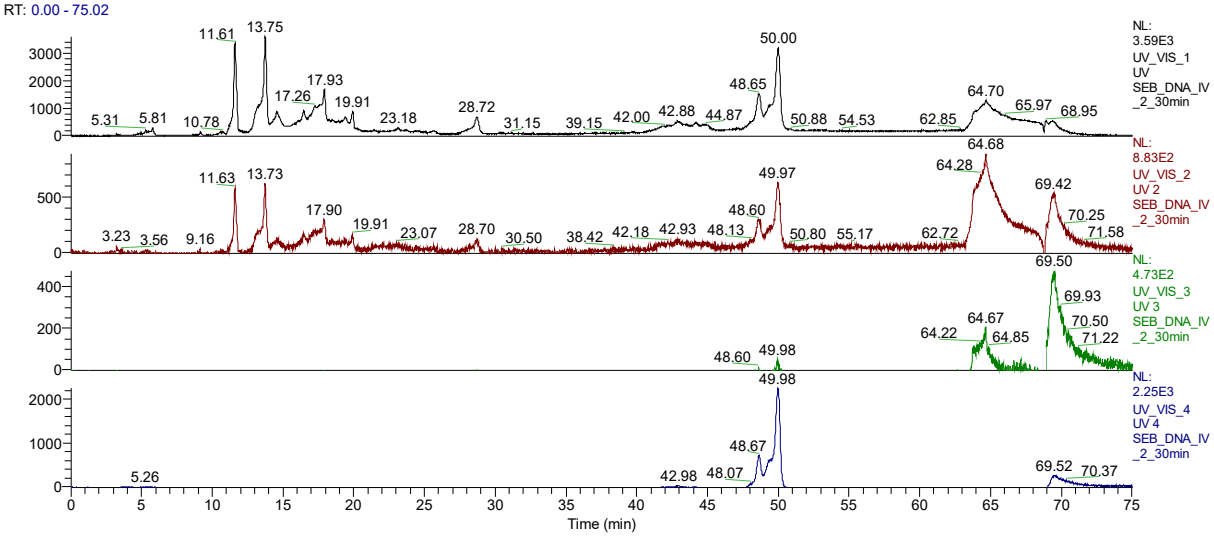


Figure S64: LC-MS chromatogram overview of the different wavelengths (260 nm, 290 nm, 365 nm, 554 nm) after 30 min of irradiation.

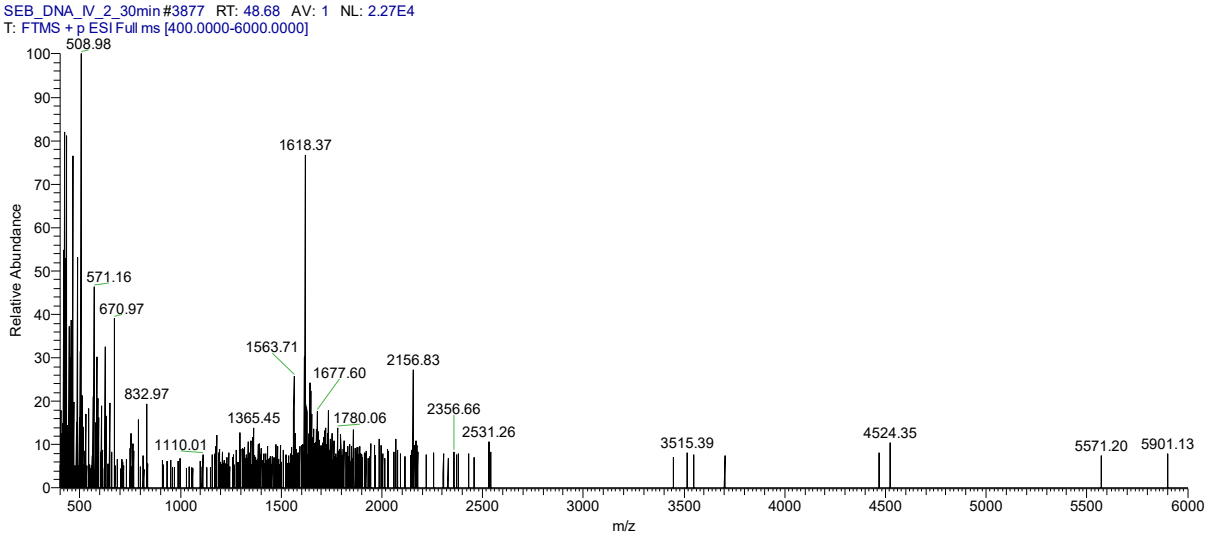


Figure S65: LC-MS analysis of the CPD damage. RT = 48.68 min: [ESI, M]+ m/z: m/3: found: 2157; calculated: 2156; m/4: found: 1618; calculated: 1617.

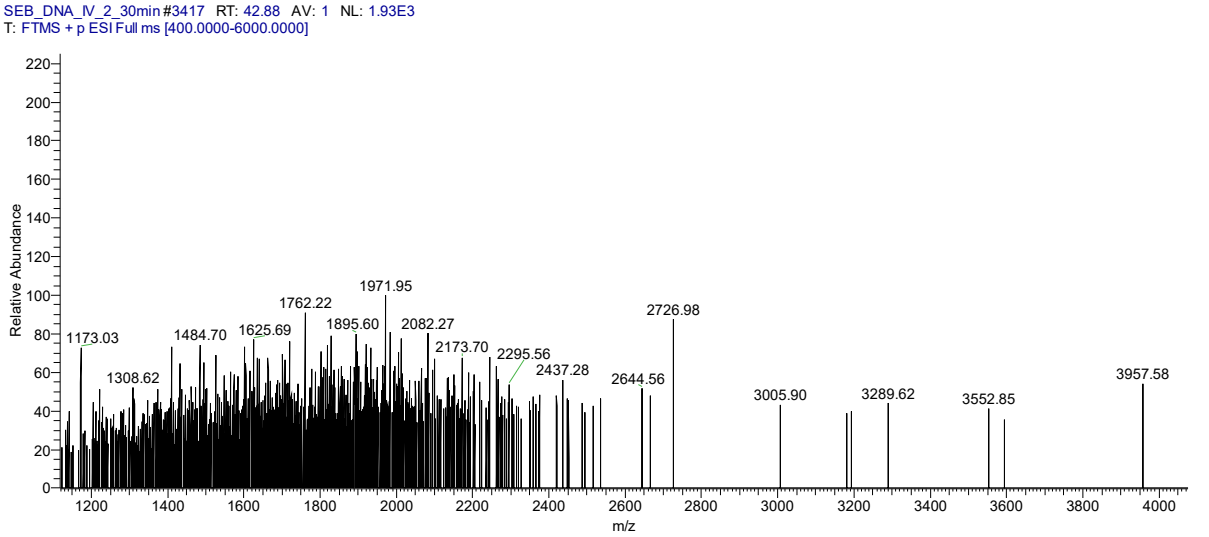


Figure S66: LC-MS analysis of the ICL damage. RT = 42.88 min: [ESI, M]+ m/z: m/3: found: 3553; calculated: 3555; m/4: found: 2665; calculated: 2666; m/5: found: 2134; calculated: 2133; m/6: found: 1779; calculated: 1778.

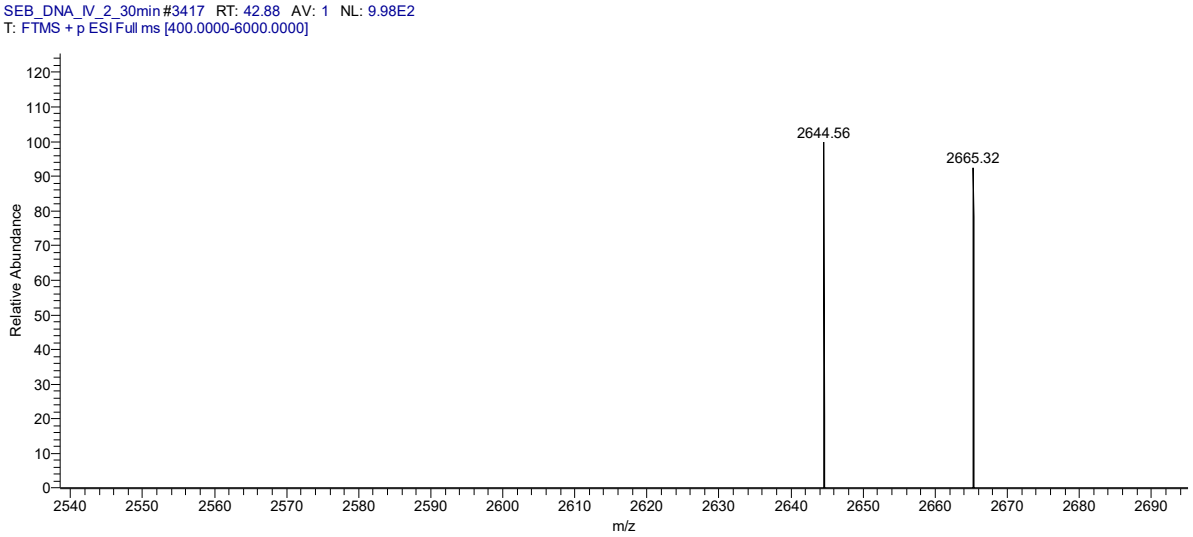


Figure S67: Zoom: LC-MS analysis of the ICL damage. RT = 42.88 min: [ESI, M]+ m/z: m/4: found: 2665; calculated: 2666.

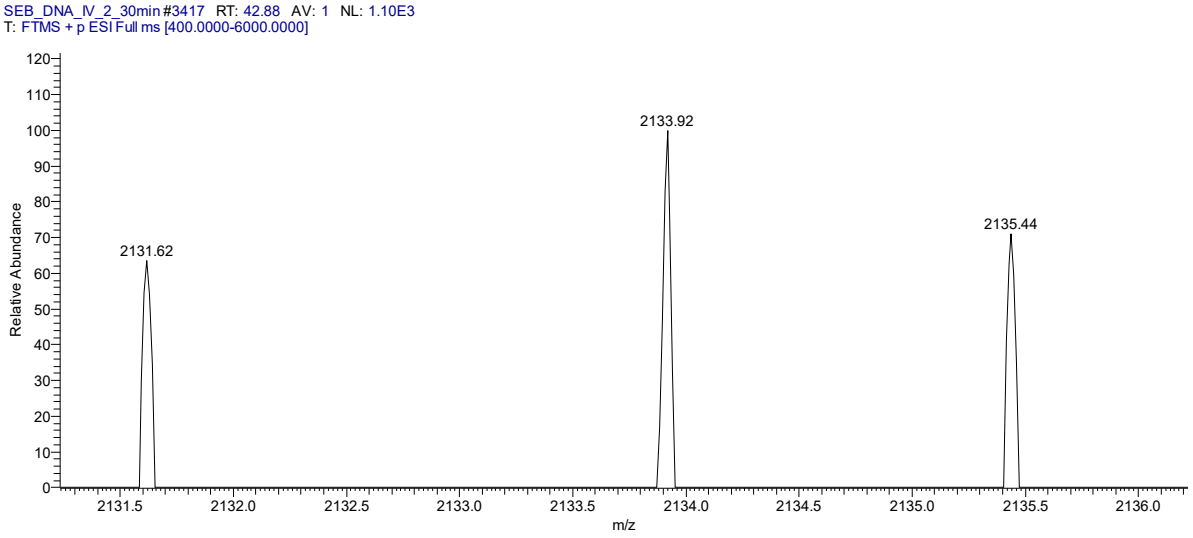


Figure S68: Zoom: LC-MS analysis of the ICL damage. RT = 42.88 min: [ESI, M]+ m/z: m/5: found: 2134; calculated: 2133.

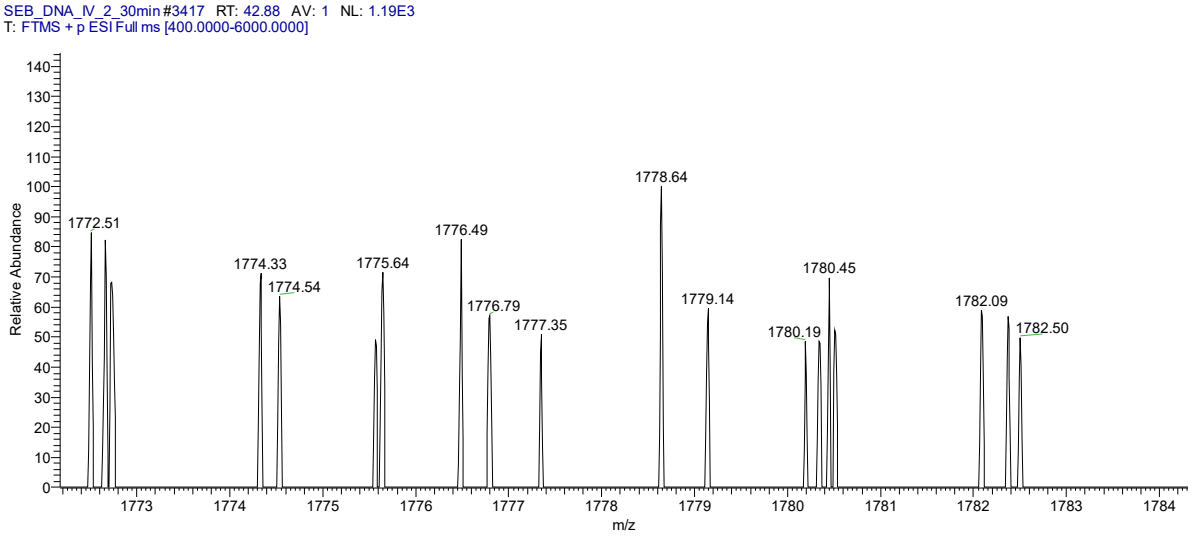


Figure S69: Zoom: LC-MS analysis of the ICL damage. RT = 42.88 min: [ESI, M]+ m/z: m/6: found: 1779; calculated: 1778.

DNA2-2-G/C after 0 min Irradiation

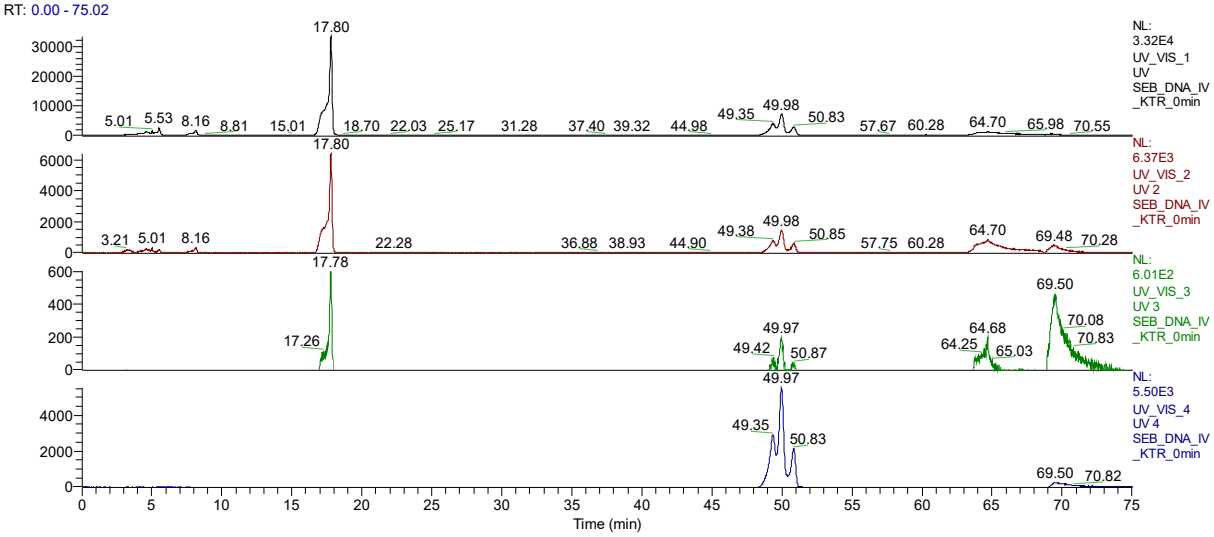


Figure S70: Overview of the LC-MS chromatogram at different wavelengths (260 nm, 290 nm, 365 nm, and 554 nm) after 0 minutes of exposure time.

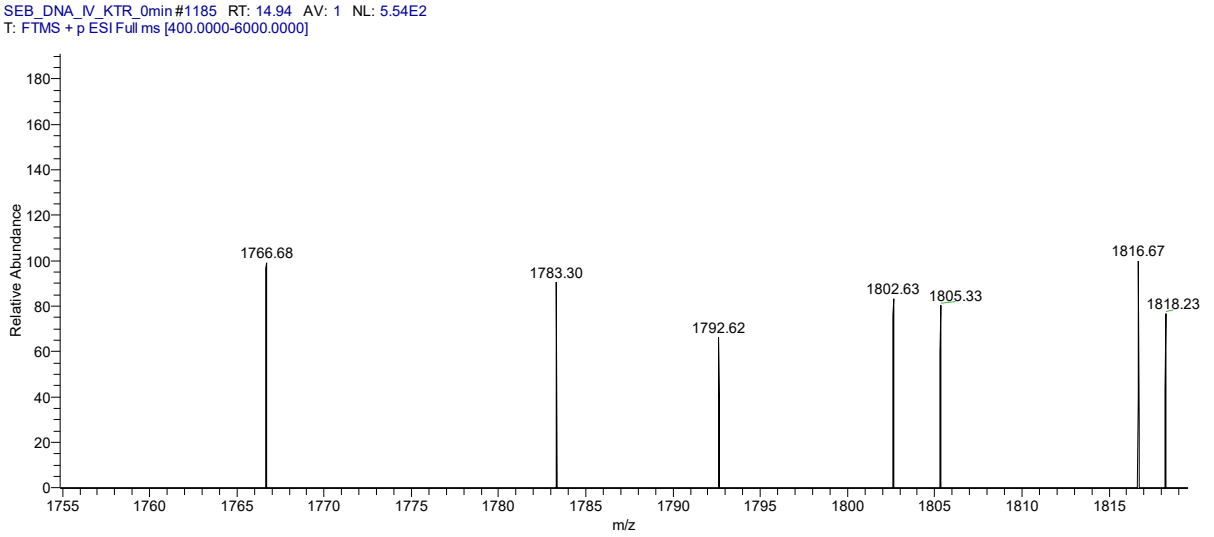


Figure S71: LC-MS analysis of the complementary strand (GS). RT = 14.94 min: [ESI, M]+ m/z: observed: 1783; calculated: 1781.

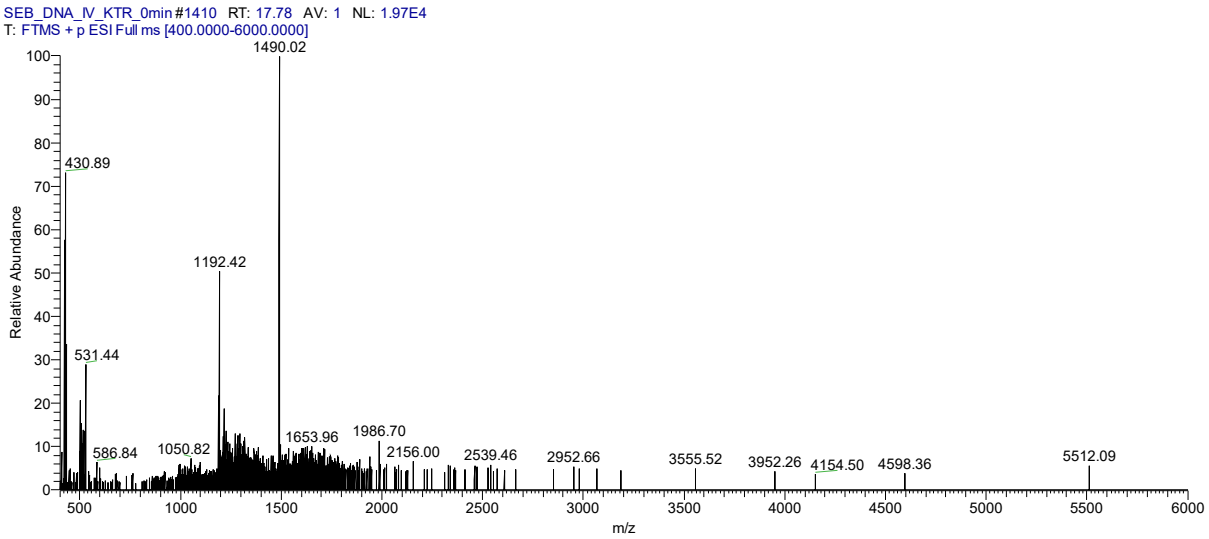


Figure S72: LC-MS analysis of the template strand (TS). RT = 17.78 min: [ESI, M]+ m/z: m/3: observed: 1987; calculated: 1985; m/4: observed: 1490; calculated: 1489.

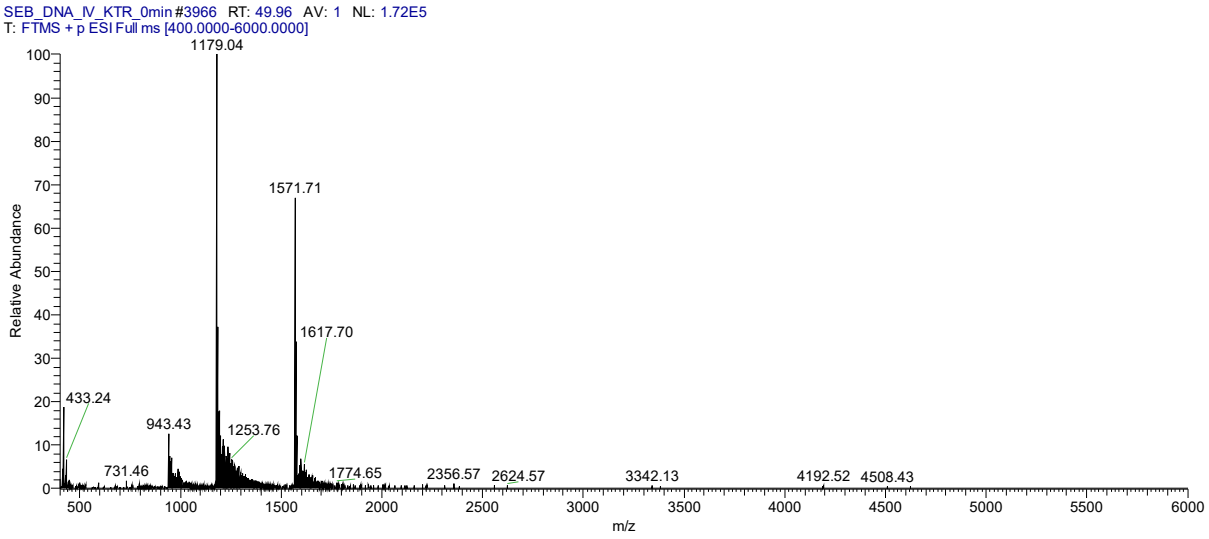


Figure S73: LC-MS analysis of the ATTO-strand (AS). RT = 49.96 min: [ESI, M]+ m/z: m/2: observed: 2357; calculated: 2356; m/3: observed: 1572; calculated: 1571; m/4: observed: 1179; calculated: 1178.

DNA2-2-G-C after 1 h Irradiation

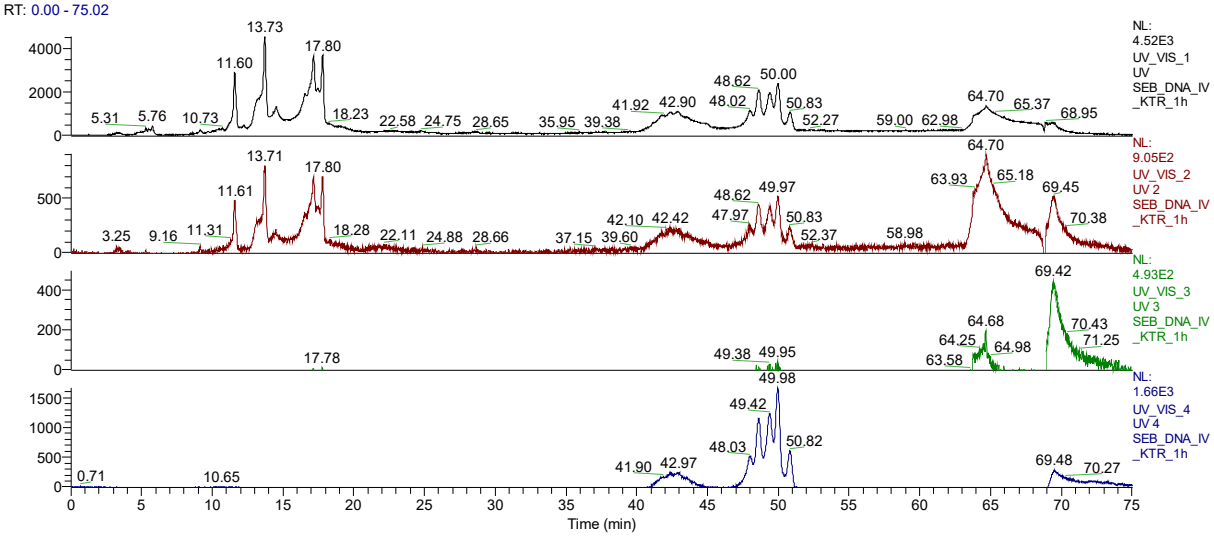


Figure S74: Overview of the LC-MS chromatogram at different wavelengths (260 nm, 290 nm, 365 nm, and 554 nm) after 1 hour of exposure time.

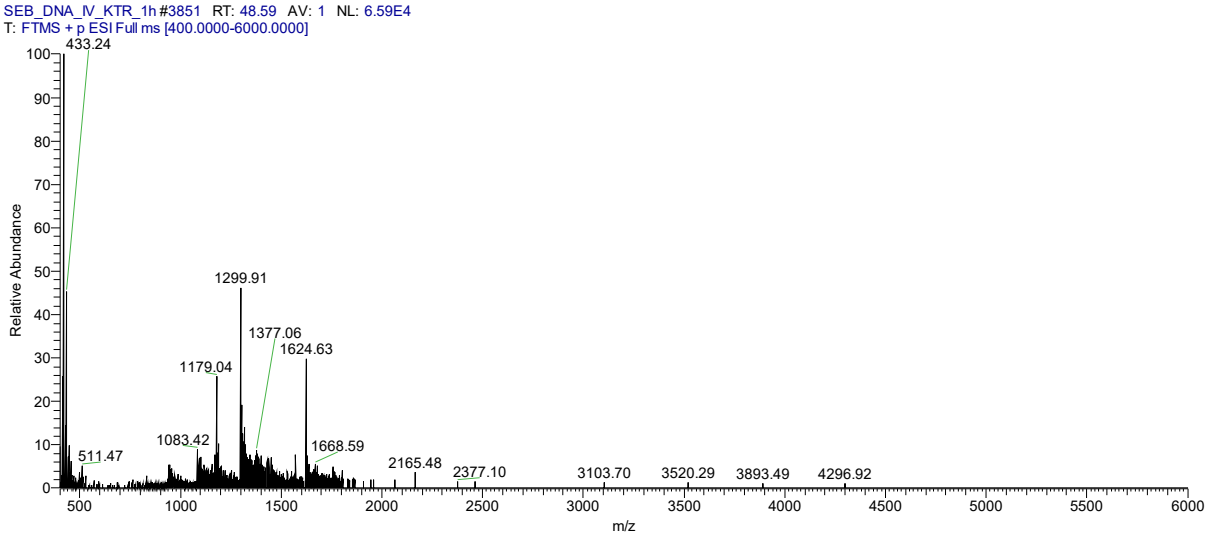


Figure S75: LC-MS analysis of the CPD damage. RT = 48.59 min: [ESI, M]+ m/z: m/3: observed: 2165; calculated: 2164; m/4: observed: 1624; calculated: 1623.

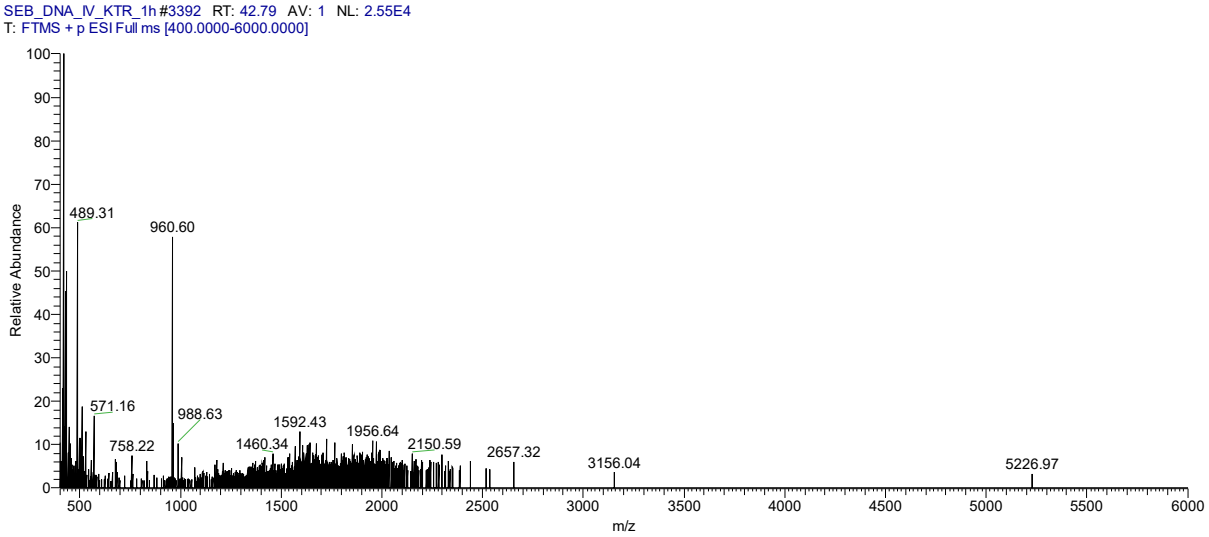


Figure S76: LC-MS analysis of the ICL damage. RT = 42.79 min: [ESI, M]+ m/z: m/6: observed: 1777; calculated: 1778; m/7: observed: 1524; calculated: 1524; m/8: observed: 1333; calculated: 1333.

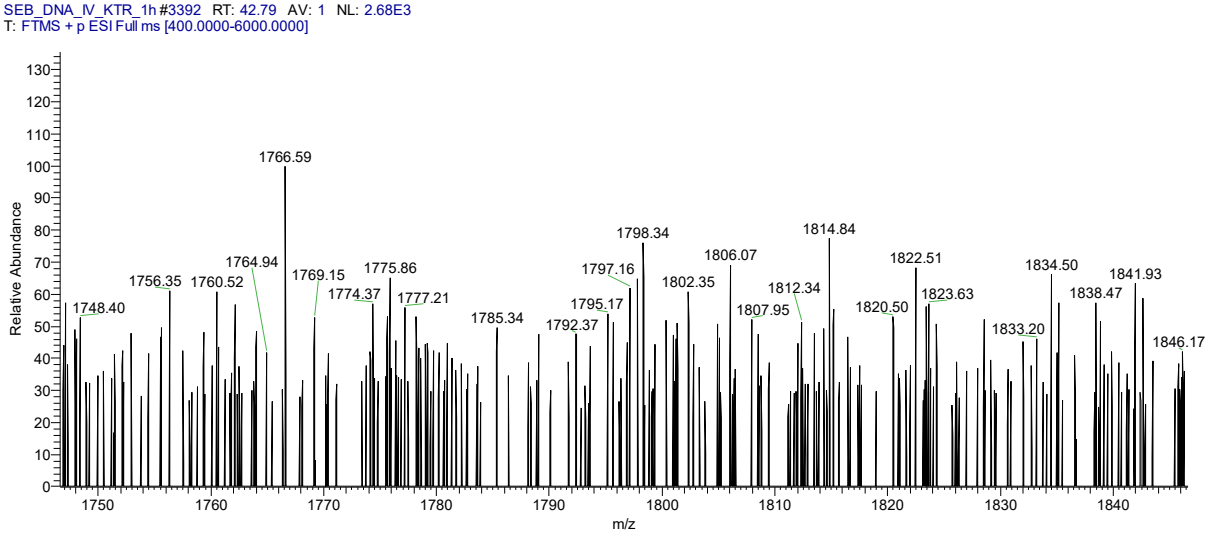


Figure S77: Zoom: LC-MS analysis of the ICL damage. RT = 42.79 min: [ESI, M]+ m/z: m/6: observed: 1777; calculated: 1778.

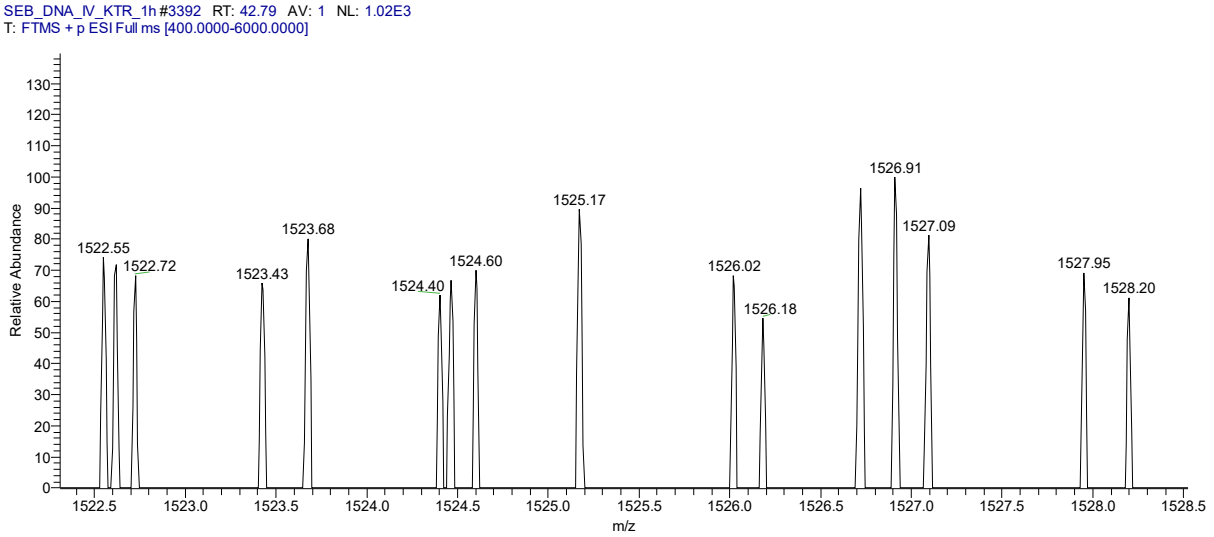


Figure S78: Zoom: LC-MS analysis of the ICL damage. RT = 42.79 min: [ESI, M]+ m/z: m/7: observed: 1524; calculated: 1524.

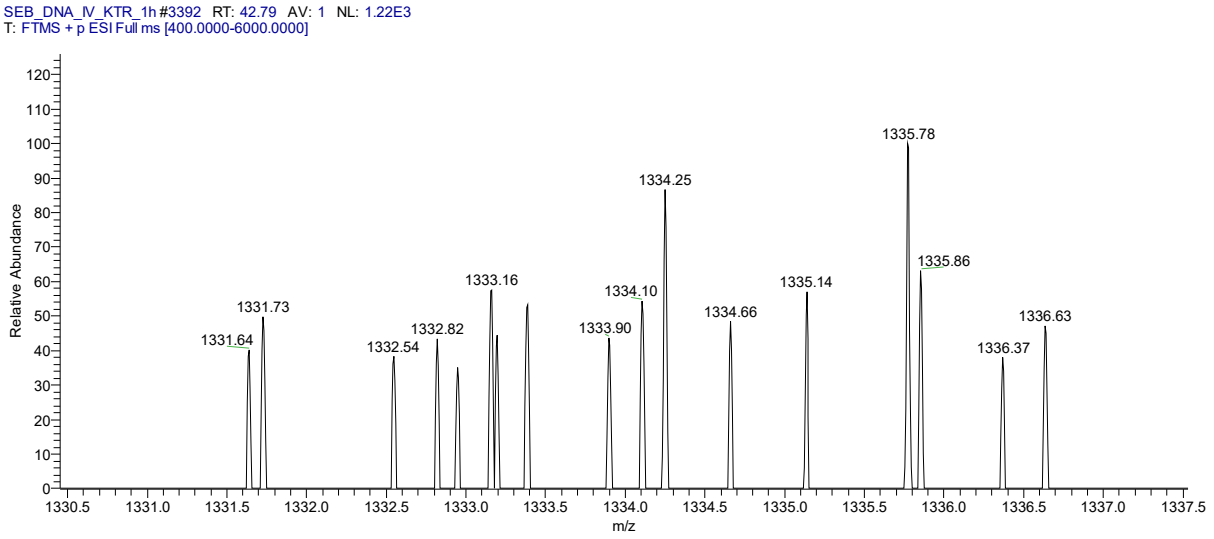


Figure S79: Zoom: LC-MS analysis of the ICL damage. RT = 42.79 min: [ESI, M]+ m/z: m/8: observed: 1333; calculated: 1333.

8. Phosphorescence spectroscopy

For low temperature phosphorescence measurements, a liquid nitrogen dewar was integrated into the LP980KS setup from Edinburgh Instruments equipped with a Nd:YAG laser from Quantel (Q-smart 450). The frequency-tripled (355 nm) output served as the excitation source. The laser pulse duration was ~ 10 ns and the pulse frequency was 10 Hz. The typical pulse energy used emission studies was ~20 mJ. The samples in EPR tubes were placed in the dewar and cooled to 77 K using liquid nitrogen. For detection, an iCCD camera from Andor was employed. The concentration of the samples was chosen to have an optical density (OD) of < 0.3 at the excitation wavelength to avoid inner filter effects.

Single-stranded **DNA1-0** ($E_T(1)$) and **DNA2-0** ($E_T(2)$) were used. The sample labeled with TX_{DNA} is a thioxanthone-containing single stranded DNA from our previous study.[1]

To determine the triplet energy, we used the wavelength at the high-energy edge of the 77 K phosphorescence spectrum, where the intensity amounts to 10 % compared to the phosphorescence maximum. This is a widely-used approach for analyzing broad and unstructured phosphorescence spectra.

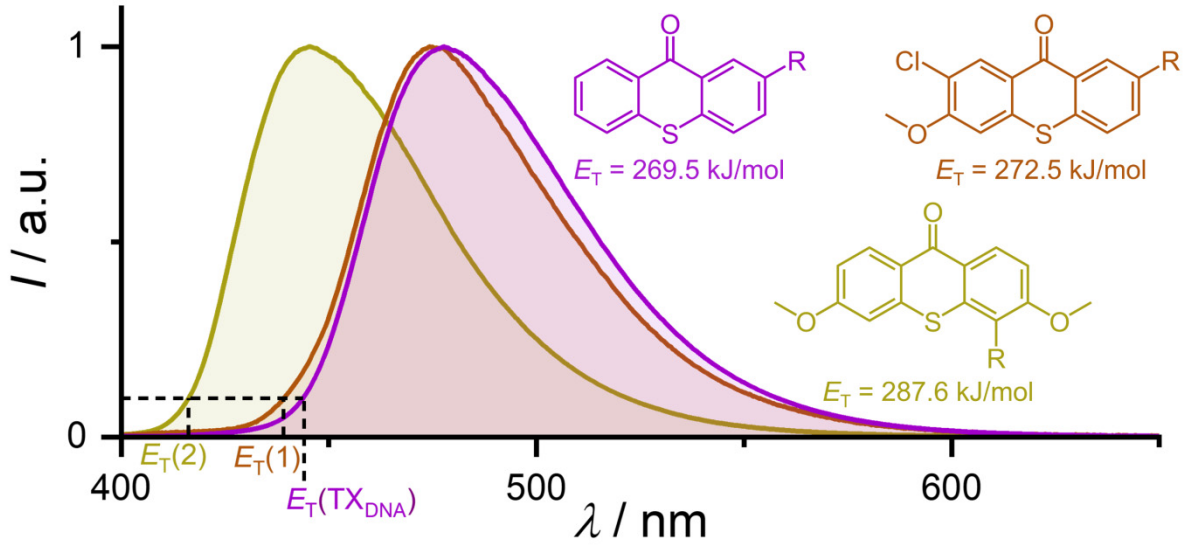


Figure S80: Phosphorescence measurements at 77 K of single strand DNA, modified with different thioxanthones in water. The samples were measured using a 355 nm laser pulse and the spectra were recorded with an iCCD-camera. The spectra were integrated over 10 ms with an initial delay of 1 ms (10 ms for TX_{DNA}) from the laser pulse to avoid residual fluorescence emission.

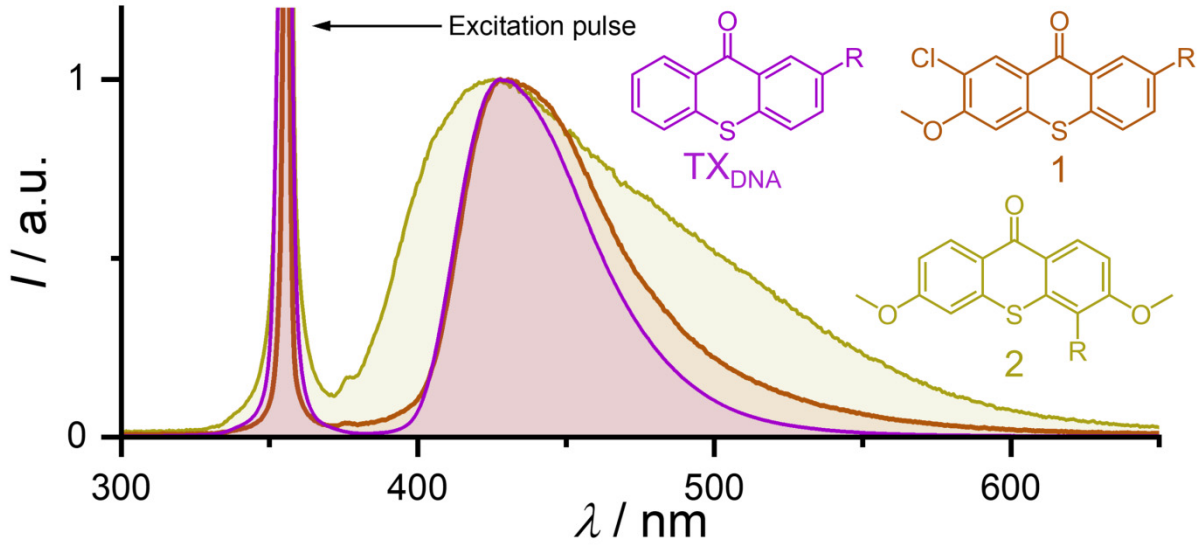


Figure S81: Fluorescence measurements at 77 K of single strand DNA, modified with different thioxanthones in water. The samples were measured using a 355 nm laser pulse and the spectra were recorded with an iCCD-camera. The spectra were integrated over 50 ns (10 ns for TX_{DNA}) with the laser pulse as starting point. The fluorescence emission spectra of **1** and **2** are broadened due to partial overlap with phosphorescence emission.

9. References

[1] S. Häcker T. J. B. Zähringer, H.-A. Wagenknecht, C. Kerzig C. Direct Observation of Triplet-Triplet Energy Transfer Between C-Nucleotides of Thioxanthone and Naphthalene in DNA. *ChemRxiv.* **2025**; doi:10.26434/chemrxiv-2025-nrzhp.