

## SUPPORTING INFORMATION

### Reaction of *S*-geranylated 2-thiouracil modified oligonucleotides with alkyl amines leads to the *N2*-alkyl *isocytosine* derivatives

Grazyna Leszczynska,<sup>a\*</sup> Klaudia Sadowska,<sup>a</sup> Malgorzata Sierant,<sup>b</sup> Milena Sobczak,<sup>b</sup> Barbara Nawrot,<sup>b</sup> and Elzbieta Sochacka<sup>a</sup>

<sup>a</sup> Institute of Organic Chemistry, Lodz University of Technology, Zeromskiego 116, 90-924 Lodz, Poland

<sup>b</sup> Department of Bioorganic Chemistry, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112, 90-363 Lodz, Poland

E-mail: [grazyna.leszczynska@p.lodz.pl](mailto:grazyna.leszczynska@p.lodz.pl)

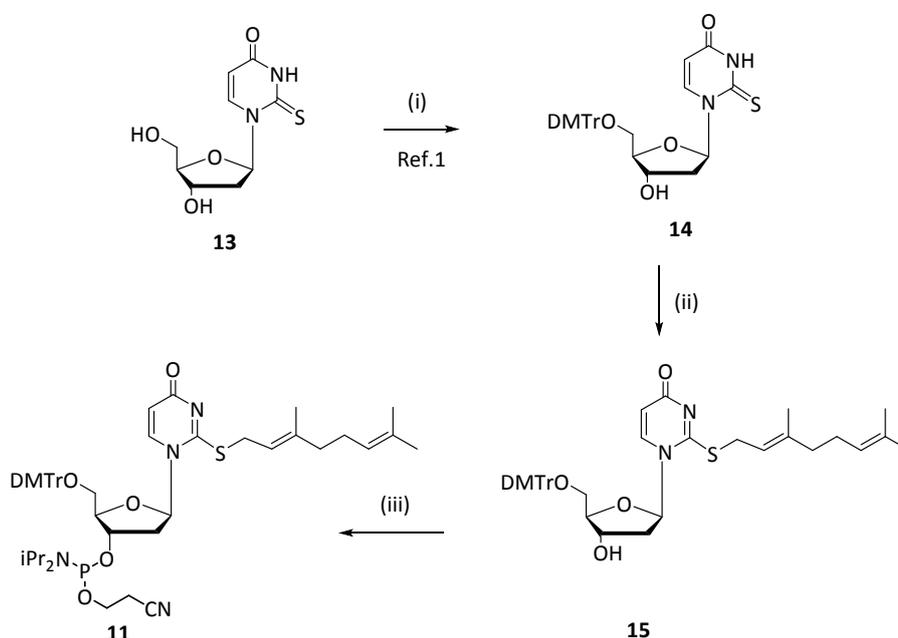
#### Table of contents:

I. General methods.....	S2
II. Synthesis of <i>S</i> -geranyl-2-thio-2'-deoxyuridine phosphoramidite.....	S2
III. <sup>1</sup> H, <sup>13</sup> C and <sup>31</sup> P NMR spectra of <b>11</b> and <b>15</b> .....	S4
IV. Synthesis of geS2dU-DNA via direct incorporation of the geS2dU phosphoramidite <b>11</b> (method I).....	S7
V. Deprotection and support cleavage of CPG-linked geS2dU-DNA with aq. ammonia.....	S8
VI. Synthesis and characterization of geS2dU-DNA via post-synthetic geranylation of fully deprotected S2dU-DNA (method II).....	S11
VII. Post-synthetic conversion geS2dU-DNA → <i>m2isodC</i> -DNA.....	S13
VIII. Post-synthetic conversion geS2dU-DNA → <i>b2isodC</i> -DNA.....	S14
IX. Post-synthetic conversion geS2U-RNA → <i>m2isoC</i> -RNA.....	S15
X. Deprotection and support cleavage of CPG-linked geS2U-RNA with aq. ammonia .....	S17
XI. UV melting temperature measurements.....	S19
XII. Thermodynamic data for DNA and RNA duplexes containing geS2Ura and <i>m2isoCyt</i> .....	S20
XIII. References.....	S20

## I. General Methods

Anhydrous solvents were prepared using standard procedures. All solid reagents were dried under a high vacuum line prior to use. Thin layer chromatography was done on silica gel coated plates (60F254, Merck), and silica gel 60 (mesh 230–400, Merck) was used for column chromatography. HPLC was performed with a Waters chromatograph equipped with a 996 spectral diode array detector. NMR spectra were recorded at a 700 MHz (for  $^1\text{H}$ ) instrument (176 MHz for  $^{13}\text{C}$  and 101 MHz for  $^{31}\text{P}$ ). Chemical shifts ( $\delta$ ) are reported in ppm relative to TMS (an internal standard) for  $^1\text{H}$  and  $^{13}\text{C}$ , and external  $\text{H}_3\text{PO}_4$  for  $^{31}\text{P}$ . The signal multiplicities are described as s (singlet), d (doublet), dd (doublet of doublets), t (triplet), q (quartet), and m (multiplet). High-resolution mass spectra were obtained from a Finnigan MAT 95 spectrometer by a peak matching technique in FAB positive or negative ion detection mode or from MALDI SYNAPT G2-S HDMS (ESI ionization). MALDI-TOF spectra were recorded on Applied Biosystems Voyager-Elite mass spectrometer.

## II. Synthesis of *S*-geranyl-2-thio-2'-deoxyuridine phosphoramidite (**11**)



**Scheme S1.** Synthetic route to *S*-geranyl-2-thio-2'-deoxyuridine (**11**). Reagents and conditions: (i) DMTr-Cl, pyridine, 24 h, rt; (ii) geBr, DIPEA, MeOH, 15 min., rt; (iii) iPr<sub>2</sub>NP(O)Cl, DIPEA, DCM, 3.5 h, rt.

### Synthesis of 5'-*O*-dimethoxytrityl-2-thio-2'-deoxy-2-thiouridine (**14**)

To a solution of S2dU (**13**) [1] (494 mg, 2.02 mmol) in anhydrous pyridine (9 ml) 4,4'-dimethoxytrityl chloride (820 mg, 2.43 mmol, 1.2 equiv) was added. The reaction mixture was stirred at room temperature (rt) for 3.5 h. The reaction was quenched with water (25 ml) and the mixture was extracted with  $\text{CH}_2\text{Cl}_2$  (3 x 30 ml). The combined organic layers were dried over anhydrous  $\text{MgSO}_4$ , filtered and concentrated in vacuo. The residue was subjected to silica gel chromatography ( $\text{CHCl}_3$ ) to yield 5'-*O*-DMTr-S2dU (**14**) as a yellow foam (794 mg, 72%). TLC  $R_f$  = 0.54 ( $\text{CHCl}_3/\text{MeOH}$  95:5(v/v)) Analytical data was consistent with the literature [1].

### Synthesis of 5'-O-(4,4'-dimethoxytrityl)-S-geranyl-2-thio-2'-deoxyuridine (15)

To a solution of 5'-O-DMTr-S2dU (**14**) (794 mg, 1.45 mmol) in MeOH (15 ml), *N,N'*-diisopropylethylamine (DIPEA) (380  $\mu$ l, 2.18 mmol) and geranyl bromide (432  $\mu$ l, 2.18 mmol) were added. The reaction mixture was stirred at room temperature for 15 min. The reaction was quenched with water (25 ml) and the mixture was extracted with AcOEt (30 ml). The combined organic layers were dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated in vacuo. The residue was subjected to silica gel chromatography (AcOEt) to yield 5'-O-DMTr-geS2dU (**15**) as a white foam (902 mg, 91%). TLC R<sub>f</sub> = 0.60 (CHCl<sub>3</sub>/MeOH 95:5(v/v)); <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>)  $\delta$  [ppm]: 1.58 (s, 3H, Hj), 1.67 (s, 3H, Hi), 1.69 (s, 3H, Hd), 2.00-2.03 (m, 2H, He), 2.05-2.08 (m, 2H, Hf), 2.24-2.28 (m, 1H, H2'), 2.51-2.55 (m, 1H, H2''), 3.45-3.49 (m, 2H, H5', H5''), 3.75 (s, 6H, OCH<sub>3</sub>), 3.84-3.91 (m, 2H, Ha), 4.11-4.12 (m, 1H, H4'), 4.33 (s, 1H, -OH), 4.65-4.66 (m, 1H, H3'), 5.04-5.06 (m, 1H, Hg), 5.28-5.31 (m, 1H, Hb), 5.71 (d, *J*=7.7 Hz, 1H, H5), 6.18 (t, *J*=6.3 Hz, 1H, H1'), 6.81 (d, *J*=8.4 Hz, 4H, H<sub>Ar</sub>), 7.18-7.22 (m, 1H, H<sub>Ar</sub>), 7.25-7.27 (m, 6H, H<sub>Ar</sub>), 7.36-7.37 (m, 2H, H<sub>Ar</sub>), 7.99 (d, *J*=7.7 Hz, 1H, H6). <sup>13</sup>C NMR (700 MHz, CDCl<sub>3</sub>)  $\delta$  [ppm]: 16.45 (Cd), 17.78 (Ci), 25.75 (Cj), 26.42 (Cf), 30.76 (Ca), 39.62 (Ce), 41.95 (C2'), 55.28 (OCH<sub>3</sub>), 62.78 (C5'), 70.78 (C3'), 86.88 (C4'), 87.12 (Ph<sub>3</sub>C), 88.23 (C1'), 109.33 (C5), 113.36 (C<sub>Ar</sub>), 116.28 (Cb), 123.77 (Cg), 127.18 (C<sub>Ar</sub>), 128.16 (C<sub>Ar</sub>), 130.13 (C<sub>Ar</sub>), 131.86 (Ch), 135.19 (C<sub>Ar</sub>), 135.39 (C<sub>Ar</sub>), 139.04 (C6), 142.98 (Cc), 144.34 (C<sub>Ar</sub>), 158.75 (C<sub>Ar</sub>), 162.16 (C4), 169.23 (C2). HRMS calcd for C<sub>40</sub>H<sub>46</sub>N<sub>2</sub>O<sub>6</sub>S [M+H]<sup>+</sup> 683.3154, found 683.3149.

### Synthesis of 5'-O-(4,4'-dimethoxytrityl)-S-geranyl-2-thio-2'-deoxyuridine-3'-O-( $\beta$ -cyanoethyl-*N,N*-diisopropyl)phosphoramidite (11)

To a solution of 5'-O-DMTr-geS2dU (**15**) (250 mg, 0.35 mmol, 1 equiv) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (2.6 ml) DIPEA (243  $\mu$ l, 1.40 mmol, 4 equiv) and 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (157  $\mu$ l, 0.70 mmol, 2 equiv) were added. The reaction mixture was stirred at rt for 3.5 h under argon. Then, the mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> (10 ml), treated with 5% aq. NaHCO<sub>3</sub> (3 ml) and washed with water (3 ml). Organic layer was dried over MgSO<sub>4</sub>, filtered and concentrated in vacuo. The residue was co-evaporated with anhydrous toluene and the resulting oil was subjected to silica gel flash chromatography (33-50% AcOEt in petroleum ether) using argon pressure. The product **11** was obtained as a white solid in 61% yield (190 mg). TLC R<sub>f</sub> = 0.60 (C<sub>6</sub>H<sub>6</sub>/CH<sub>2</sub>Cl<sub>2</sub>/TEA, 7:2:1, v/v); <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>)  $\delta$  [ppm]: 1.14-1.26 (m, 12H, 2xNCH(CH<sub>3</sub>)<sub>2</sub>), 1.65 (s, 3H, Hj), 1.73 (s, 3H, Hi), 1.78 (s, 3H, Hd), 2.04-2.17 (m, 4H, He, Hf), 2.30-2.46 (m, 1H, H2''), 2.49-2.58 (m, 1H, NCH(CH<sub>3</sub>)<sub>2</sub>), 2.60-2.70 (m, 2H, H2', NCH(CH<sub>3</sub>)<sub>2</sub>), 3.44-3.80 (m, 6H, H5', H5'', CH<sub>2</sub>CH<sub>2</sub>CN), 3.84 (s, 6H, 2xOCH<sub>3</sub>), 3.94-4.02 (m, 2H, Ha), 4.17-4.25 (m, 1H, H4'), 4.69-4.80 (m, 1H, H3'), 5.09-5.16 (m, 1H, Hg), 5.34-5.43 (m, 1H, Hb), 5.76-5.80 (m, 1H, H5), 6.19-6.25 (m, 1H, H1'), 6.85-6.93 (m, 4H, H<sub>Ar</sub>), 7.23-7.37 (m, 9H, H<sub>Ar</sub>), 7.93-8.01 (m, 1H, H6). <sup>31</sup>P NMR (250 MHz, C<sub>6</sub>D<sub>6</sub>)  $\delta$  [ppm]: 149.38, 149.73. HRMS calcd for C<sub>49</sub>H<sub>63</sub>N<sub>4</sub>O<sub>7</sub>PS [M+H]<sup>+</sup> 883.4233, found 883.4232.

III.  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{31}\text{P}$  NMR spectra of 15 and 11

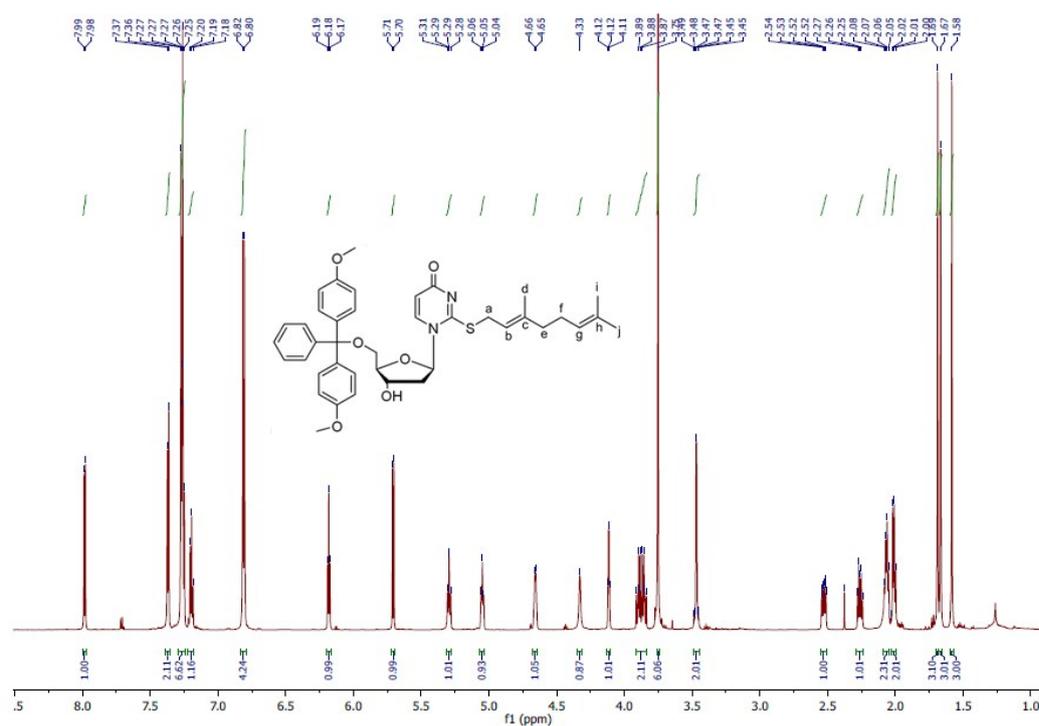


Figure S1.  $^1\text{H}$  NMR (700 MHz,  $\text{CDCl}_3$ ) spectrum of 15.

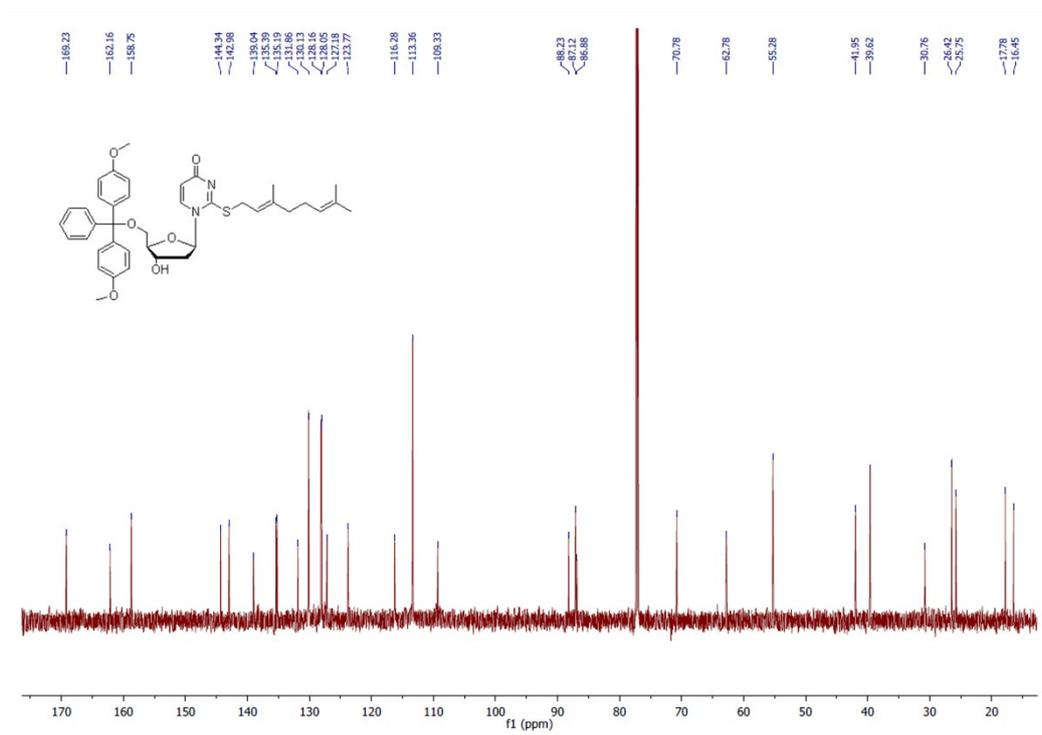
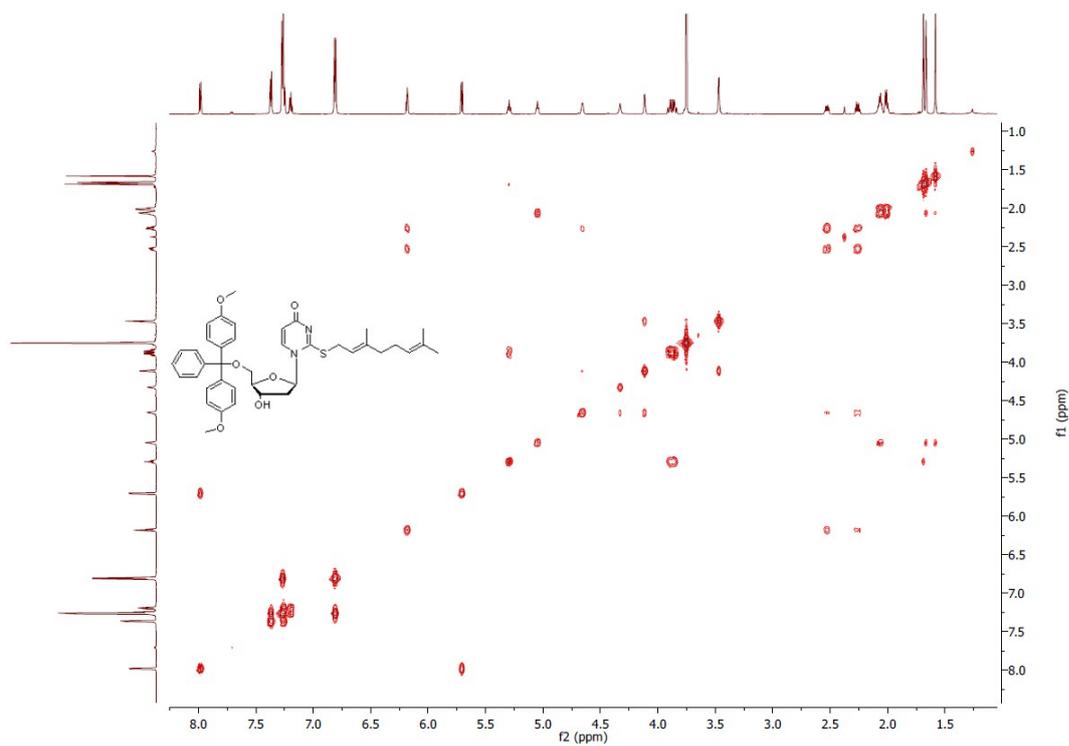
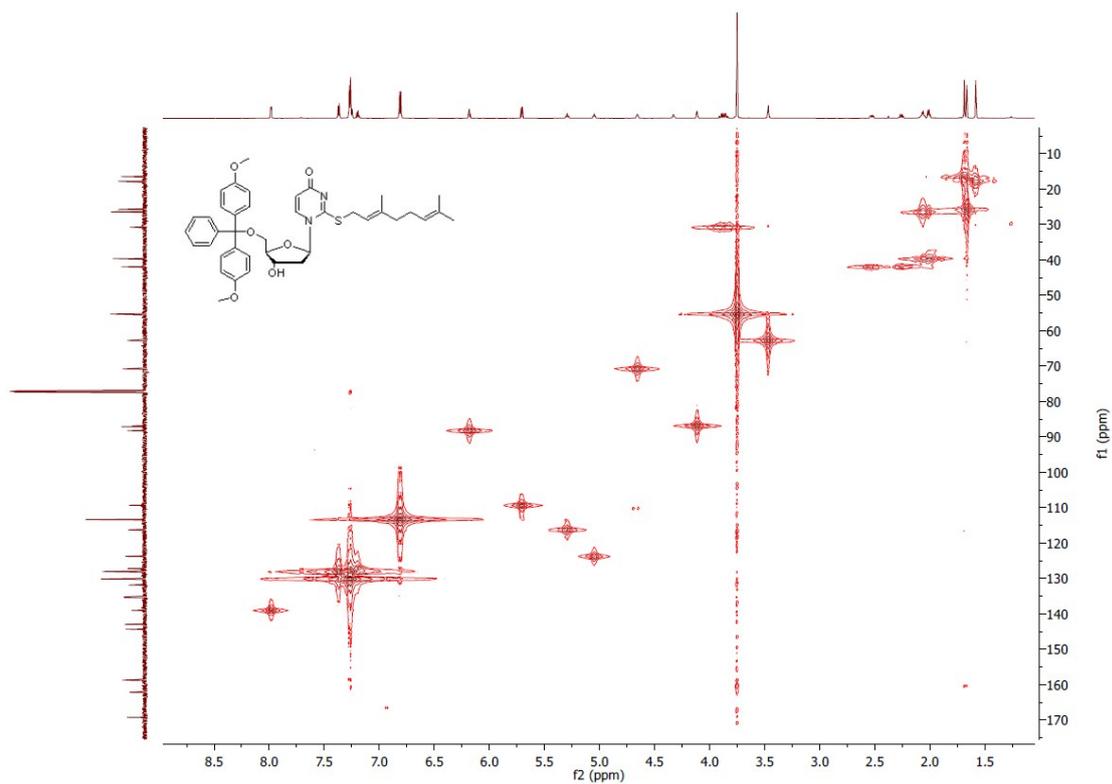


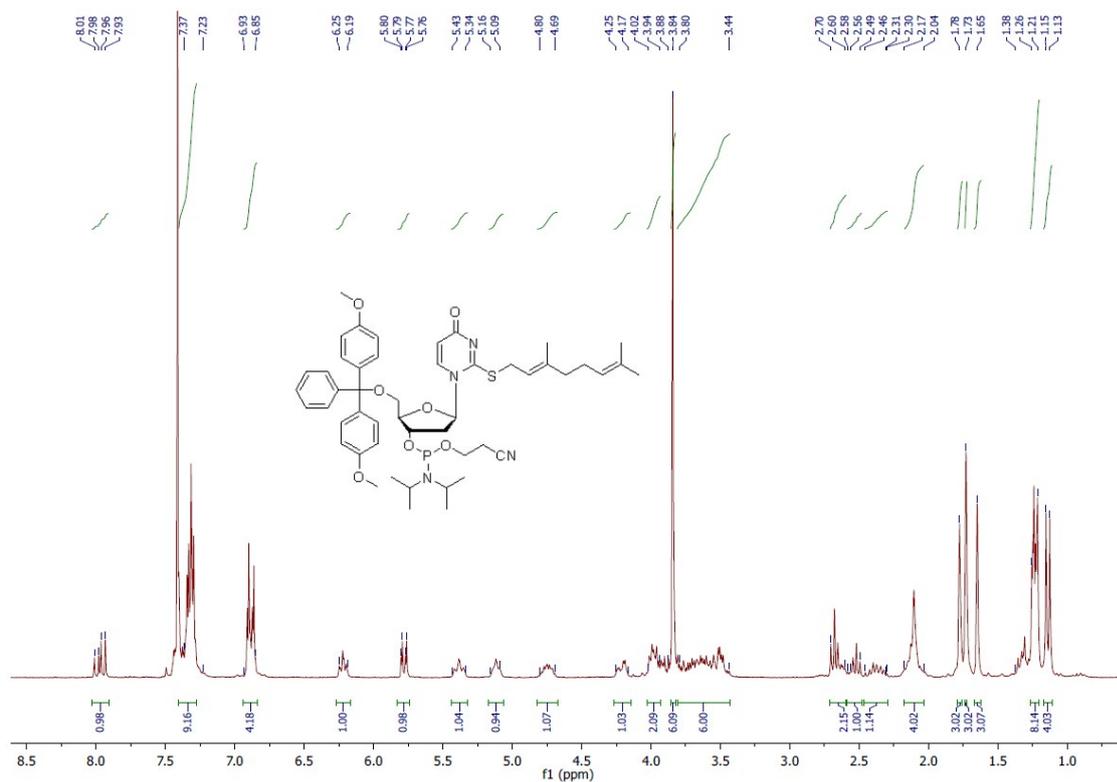
Figure S2.  $^{13}\text{C}$  NMR (700 MHz,  $\text{CDCl}_3$ ) spectrum of 15.



**Figure S3.** COSY (700 MHz, CDCl<sub>3</sub>) spectrum of **15**.



**Figure S4.** HSQC (700 MHz, CDCl<sub>3</sub>) spectrum of **15**.



Figur

e S5.  $^1\text{H}$  NMR (250 MHz,  $\text{CDCl}_3$ ) spectrum of **11**.

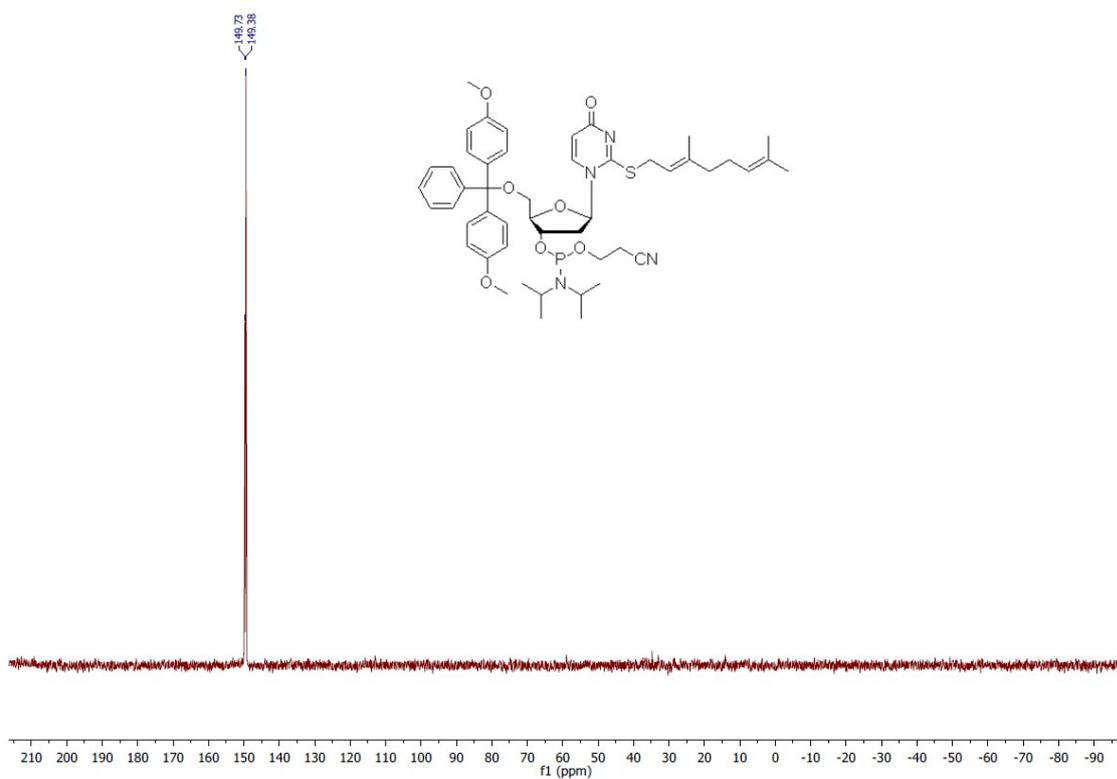


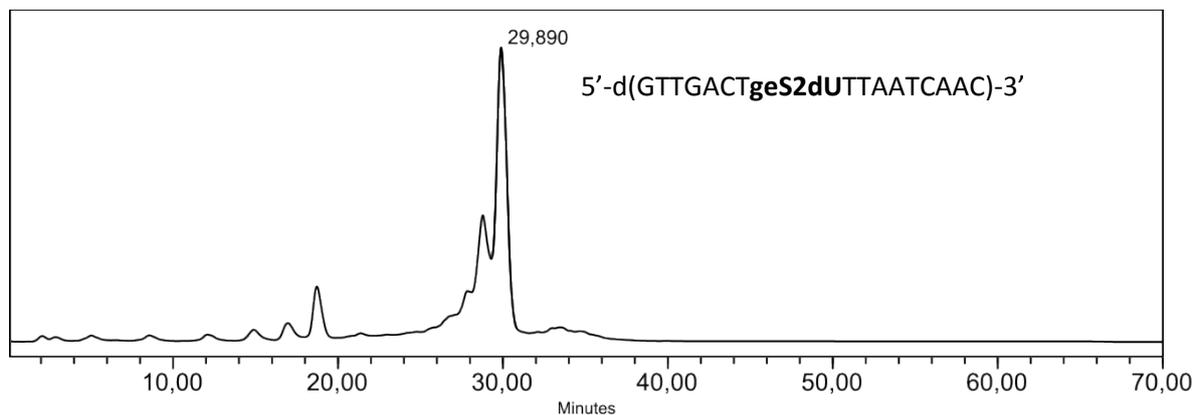
Figure S6.  $^{31}\text{P}$  NMR (250 MHz,  $\text{CDCl}_3$ ) spectrum of **11**.

#### IV. Synthesis of geS2dU-DNA *via* direct incorporation of the geS2dU phosphoramidite **11** (method I)

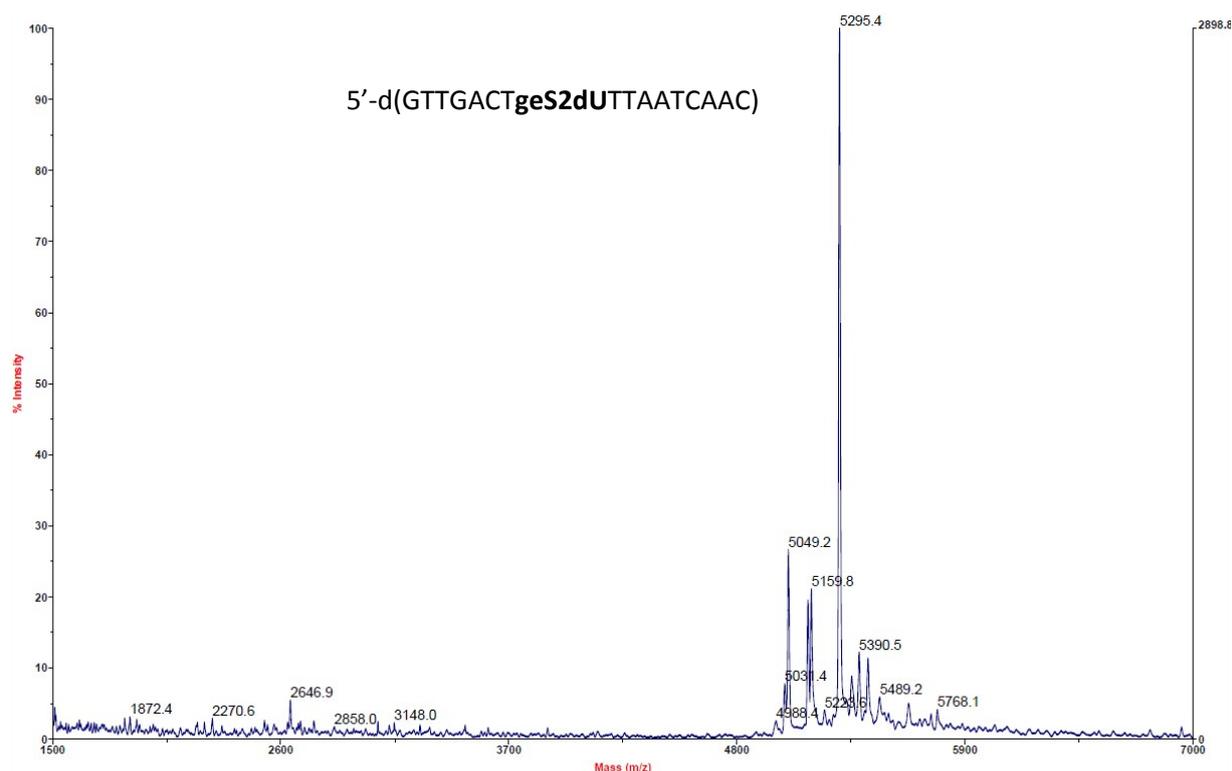
**Synthesis.** The geS2dU-DNA of the sequence 5'-d(GTTGACTgeS2dUTTAATCAAC)-3' was synthesized automatically (three runs at a 0.2  $\mu$ mol scale) with the use of an H6 Gene World DNA/RNA automated synthesizer (K&A, Laborgeraete GbR, Schaafheim, Germany). The commercially available monomeric units A, C, U, and G were protected with DMTr on the 5'-OH functions, whereas the exocyclic amine functions of A, G were masked with phenoxyacetyl (pac) and of C with acetyl (Ac) (Proligo<sup>®</sup>). The rC(tac)-succinyl-1000-CPG (Proligo<sup>®</sup>) support and 0.07 M acetonitrile (ACN) solution of monomeric units were used. All amidites were delivered at a 10-fold molar excess and the coupling lasted for 10 min. in the presence of a BMT activator (0.25 M solution in ACN). Capping was performed with tac<sub>2</sub>O in THF using a mixture of Fast protection Cap A (Proligo<sup>®</sup>) and Cap B (Proligo<sup>®</sup>) (Cap A : Cap B, 1.1/1 v/v) for 2 min. For oxidation step, the solution of 0.02 M I<sub>2</sub> in THF-H<sub>2</sub>O-pyridine (90.54:9.05:0.41 v/v/v) was used for 1 min. After the last coupling the DMTr group was removed and the support was transferred to a screw cap glass vial.

**Deprotection and support cleavage.** The CPG-linked geS2dU-DNA (0.2  $\mu$ mol) was stirred with 260  $\mu$ l of TEA/ACN (1/1, v/v) at rt for 30 min. The supernatant was removed and the resin was washed with ACN (3  $\times$  200  $\mu$ l) and dried in vacuo for 30 min. The support-bound geS2dU-DNA was treated with 340  $\mu$ l of 8 M ethanolic ammonia at rt for 8h. The solution was removed and the support was washed with anhydrous ethanol (3  $\times$  200  $\mu$ l). To release the oligomer from the support the resin was suspended with 300  $\mu$ l of 30% aq. ammonia at rt for 1 h. The supernatant was transferred to the vial and the support was washed with water (3  $\times$  300  $\mu$ l). The combined fractions were evaporated on a Speed-Vac concentrator.

**Purification.** The fully deprotected geS2dU-DNA was purified by anion-exchange (IE) HPLC (Source 15Q 4.6/100PE<sup>®</sup>) (Fig. S7). The column was eluted with a linear gradient 50–650 mM NaBr in 20 mM Na<sub>2</sub>HPO<sub>4</sub>-NaH<sub>2</sub>PO<sub>4</sub> buffer solution (pH 7.5), containing EDTA (50  $\mu$ M) and 10% ACN (flow 1 mL/min). Fractions containing the desired product (Rt = 29.89 min) were combined, concentrated on a Speed-Vac, and desalted with a C-18 cartridge (Sep-Pak<sup>®</sup>, Waters). The geS2dU-DNA afforded in 70% yield was lyophilized and analyzed by MALDI-TOF mass spectrometry (*m/z* 5295, MW 5298) (Fig. S8).



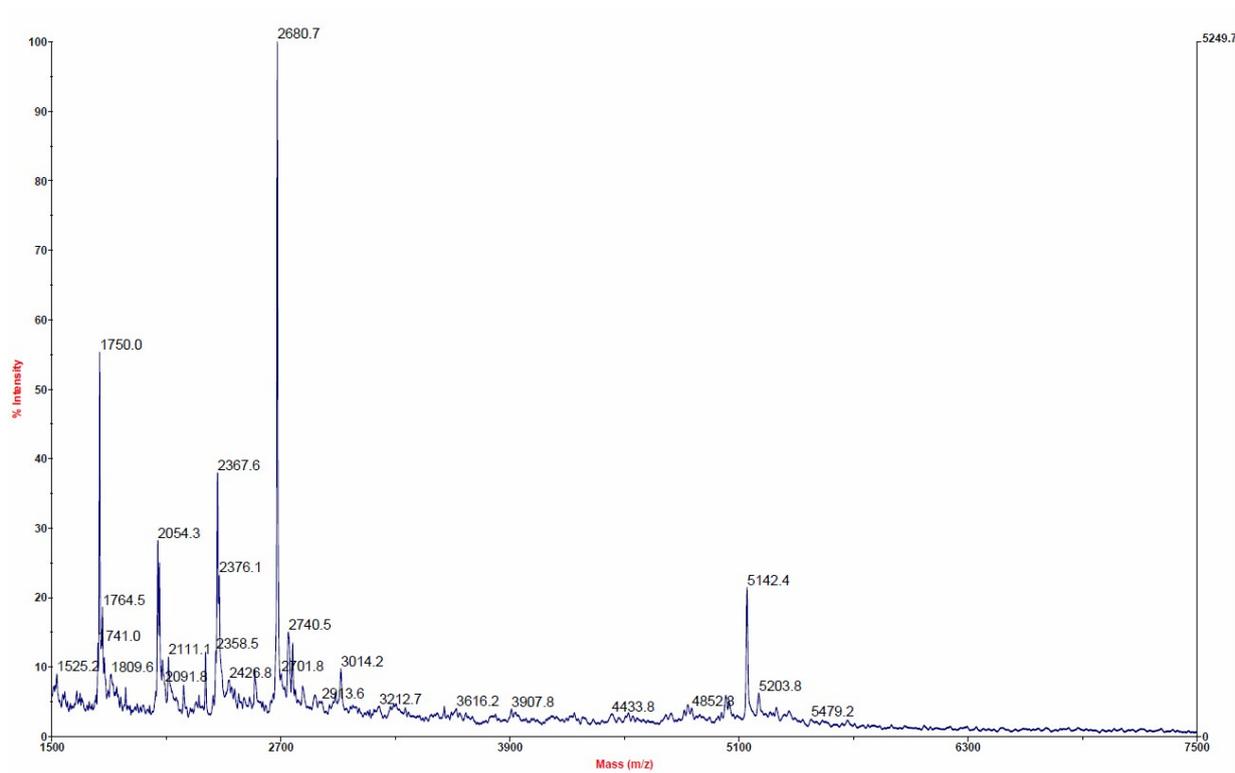
**Figure S7.** IE-HPLC analysis of the crude geS2dU-DNA obtained by direct incorporation of geS2dU phosphoramidite **11**.



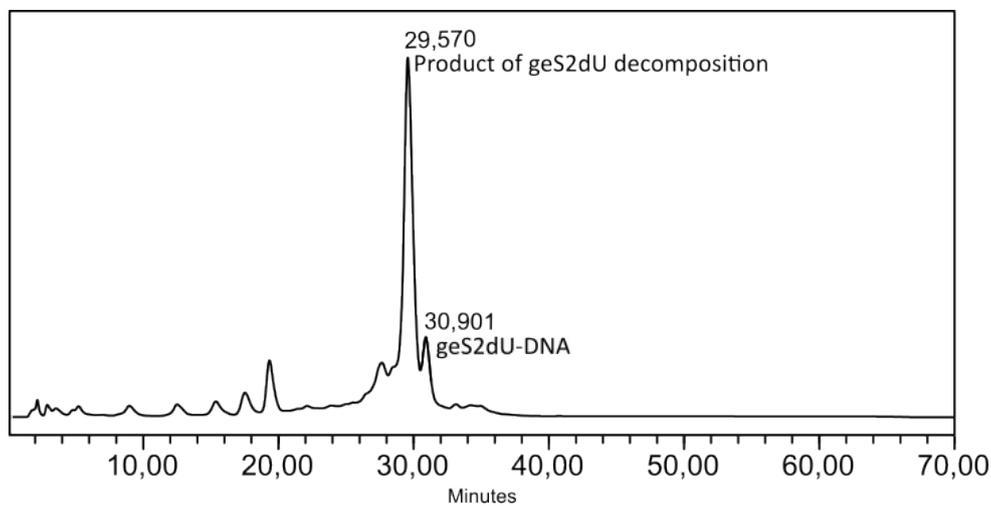
**Figure. S8.** The MALDI-TOF spectrum of 5'-d(GTTGACTgeS2dUTTAATCAAC)-3' (geS2dU-DNA);  $m/z$  5295, MW 5298.

#### V. Deprotection and support cleavage of CPG-linked geS2dU-DNA with aq. ammonia

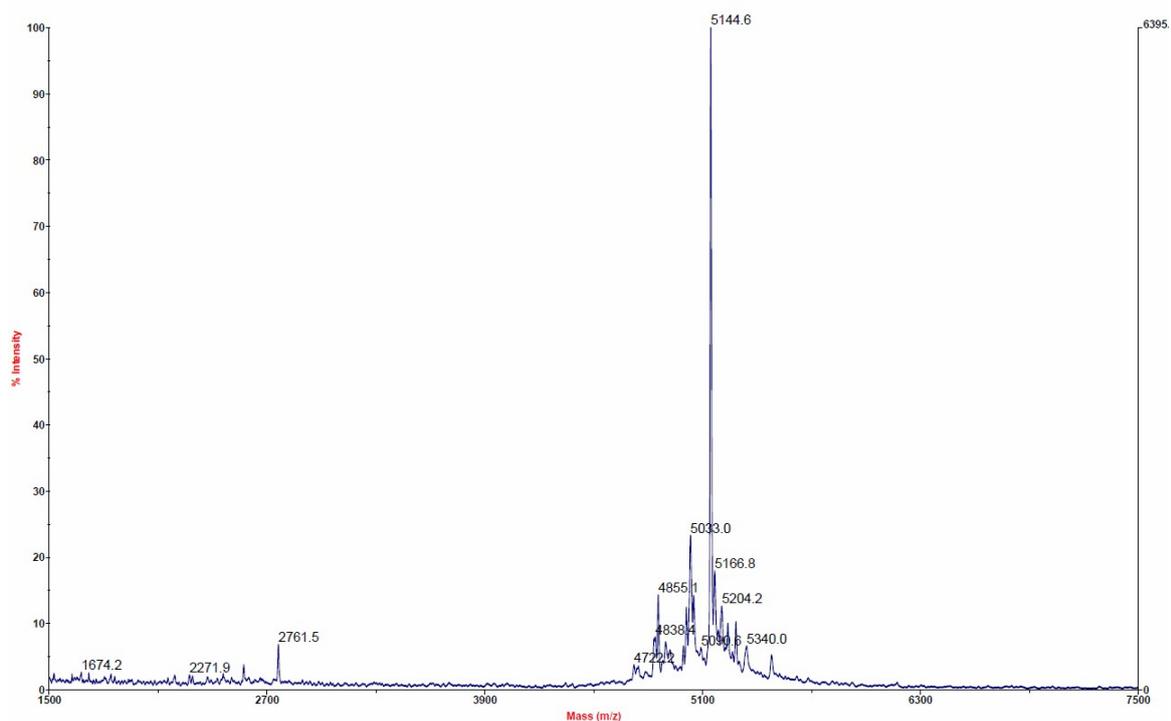
The CPG-linked geS2dU-DNA (0.025  $\mu\text{mol}$ ) was stirred with 33  $\mu\text{l}$  of TEA/ACN (1/1, v/v) at rt for 20 min. The supernatant was removed and the resin was washed with ACN ( $3 \times 100 \mu\text{l}$ ) and dried in vacuo for 30 min. The support-bound geS2dU-DNA was treated with 37  $\mu\text{l}$  of 30% aq. ammonia at rt for 14 h. The supernatant was transferred to the vial and the support was washed with water ( $3 \times 100 \mu\text{l}$ ). The combined fractions were evaporated on a Speed-Vac concentrator. The MALDI-TOF analysis of crude mixture (Fig. S9) indicated decomposition of geS2dU-DNA ( $m/z$  5142.4 in the place of 5298 calculated for geS2dU-DNA). The crude mixture of oligomers was analyzed by IE-HPLC (Fig. S 10). Two fractions: Rt = 29.57 min and 30.90 min were desalted with a C-18 cartridge (Sep-Pak<sup>®</sup>, Waters) and analyzed by MALDI-TOF mass spectrometry. The pick at Rt = 29.57 min derives from *isodC*-DNA or dU-DNA ( $m/z$  5144.6, MW 5141, Fig. S11) whereas the pick at Rt = 30.9 min derives from geS2dU-DNA ( $m/z$  5295, MW 5298, Fig. S12). On the base of isolated amounts of both oligomers the loss of -Sge residue was estimated on 94%.



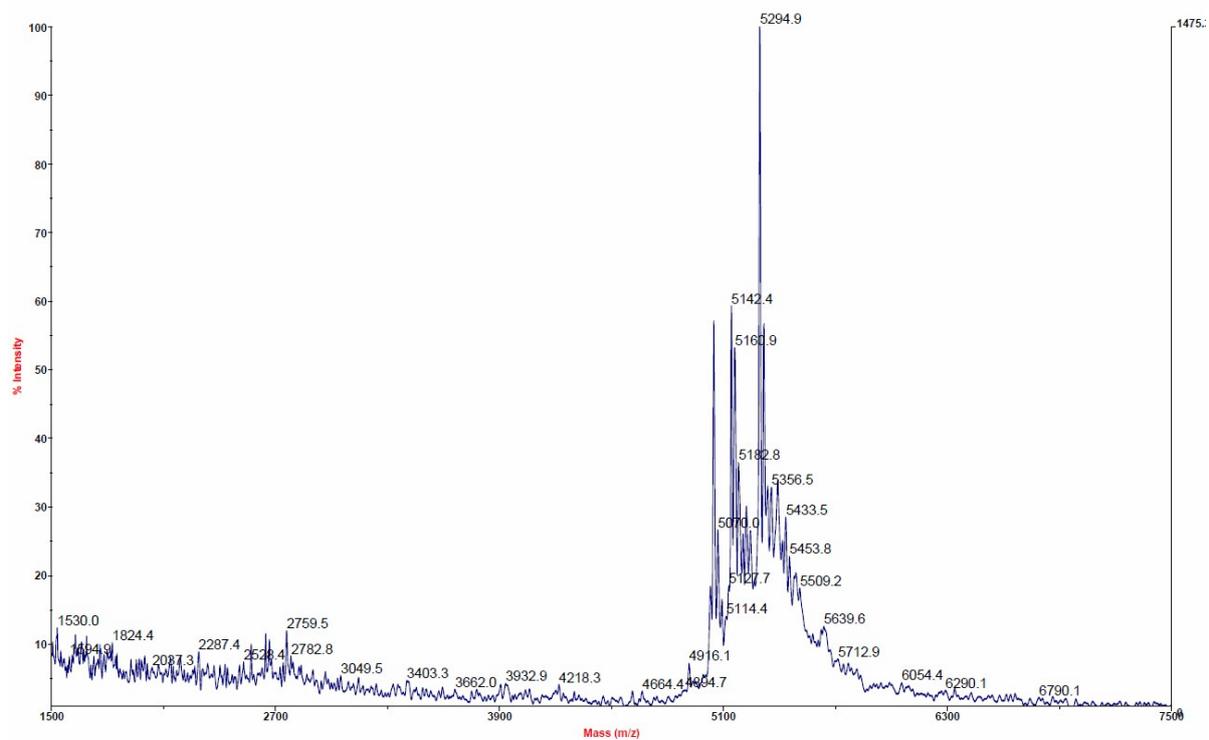
**Figure. S9.** The MALDI-TOF spectrum of the crude mixture obtained after treatment of CPG-linked geS2dU-DNA with 30% aq. ammonia for 14 h at rt.



**Figure S10.** IE-HPLC analysis of the crude mixture of oligomers obtained after treatment of CPG-linked geS2dU-DNA with 30% aq. ammonia for 14 h at rt.



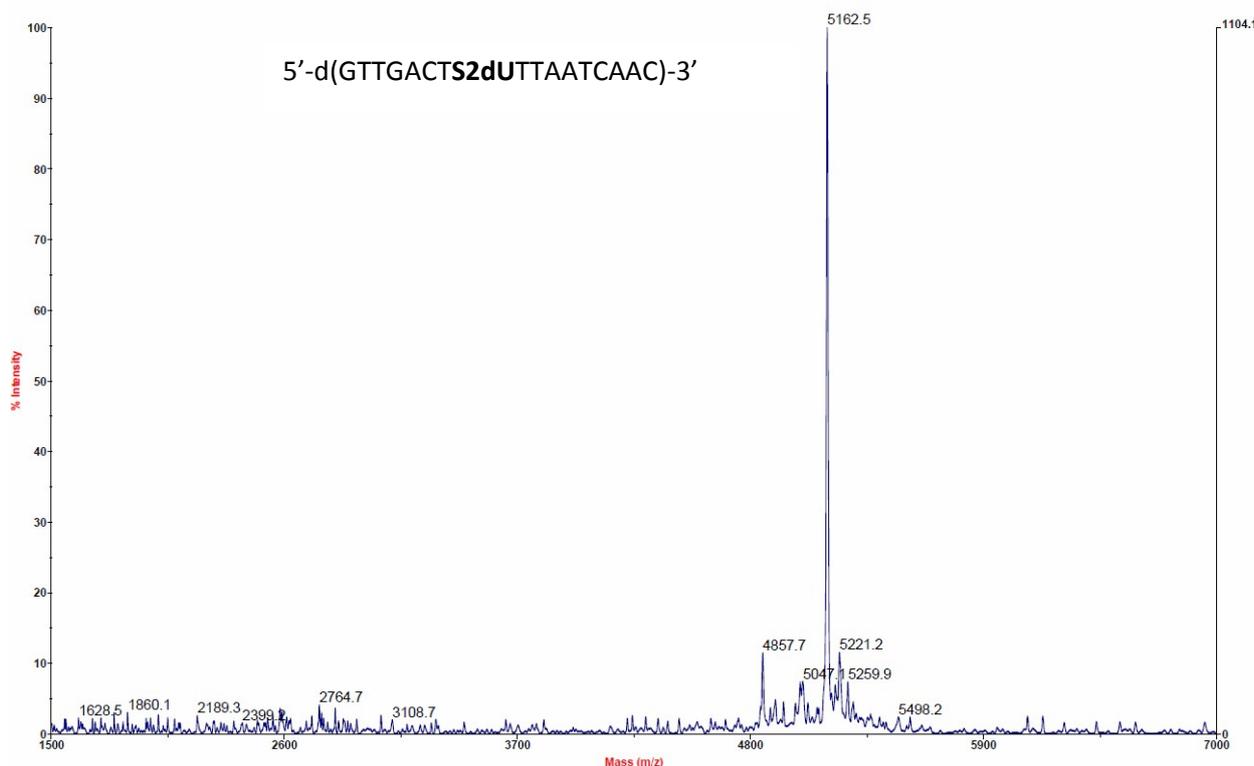
**Figure. S11.** The MALDI-TOF spectrum of oligomer at  $R_t=29.57$  min (dU-DNA or *isodC*-DNA,  $m/z$  5144.6, MW 5141).



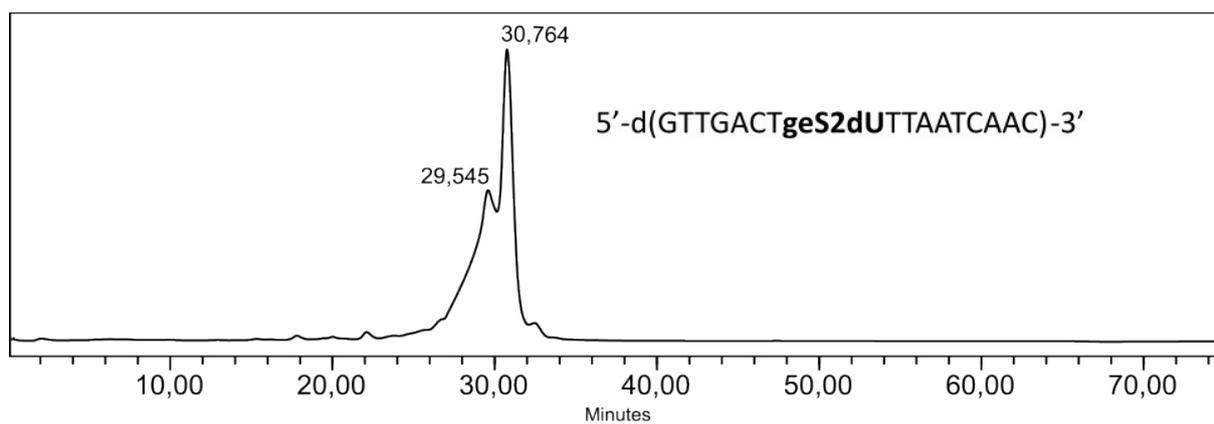
**Figure. S12.** The MALDI-TOF spectrum of oligomer at  $R_t=30.9$  min (*geS2dU*-DNA;  $m/z$  5295, MW 5298).

## VI. Synthesis of geS2dU-DNA via post-synthetic geranylation of fully deprotected S2dU-DNA (method II)

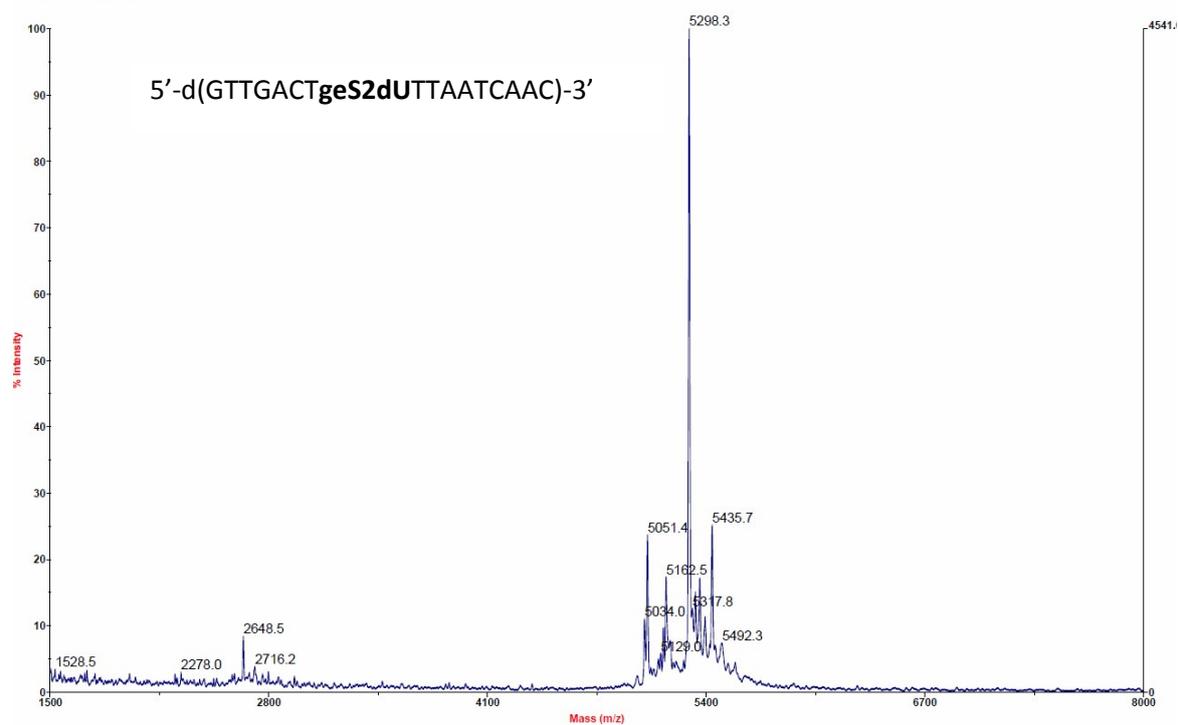
The S2dU-DNA of the sequence 5'-d(GTTGACTS2dUTTAATCAAC)-3' was synthesized automatically on a 0.2  $\mu$ mole analogously to the synthesis of geS2dU-DNA (ESI, chapter IV) with the exception of the use of S2dU-phosphoramidite, and conditions for the oxidation steps, where a solution of 0.25 M *tert*-butylhydroperoxide (*t*BuOOH) in anhydrous toluene was applied for 2 min [2]. The "DMTr-off" support-linked S2dU-DNA (0.2  $\mu$ mole scale) was treated with 300  $\mu$ L of aq. 30% ammonia / EtOH (3/1, v/v, 300  $\mu$ L) at rt for 16 h. The supernatant was collected and the support was washed with the ethanol/H<sub>2</sub>O (1/1, v/v, 3  $\times$  200  $\mu$ L). The combined fractions were evaporated on a Speed-Vac concentrator, purified and analyzed by mass spectrometry (MALDI TOF, *m/z* 5162, MW 5159, Fig. S13). For the geranylation, lyophilized S2dU-DNA (24 OD) was treated with the solution of geBr (11.8  $\mu$ L, 350 equiv), TEA (6.9  $\mu$ L, 350 equiv) in EtOH/H<sub>2</sub>O (1/1, v/v, 130  $\mu$ L). The mixture was shaken vigorously for 3 h at rt. The resultant solution was diluted with 20% aq. EtOH and the DNA oligomer was isolated using a NAP-25 column<sup>®</sup> (GE Healthcare). The eluate was evaporated on a Speed-Vac concentrator and purified by IE HPLC (Rt = 30.76 min, Fig. S14) according to the procedure described in chapter IV. Fractions containing the geS2dU-DNA were concentrated, and desalted with a C-18 cartridge (Sep-Pak<sup>®</sup>, Waters). The desalted DNA was lyophilized and analyzed by MALDI-TOF mass spectrometry (*m/z* 5298, MW 5298, Fig. S15) and enzymatic digestion [3] (Fig. S16). The yield of conversion was calculated as 80%. The pick at Rt = 29.54 min was attributed to dU-DNA.



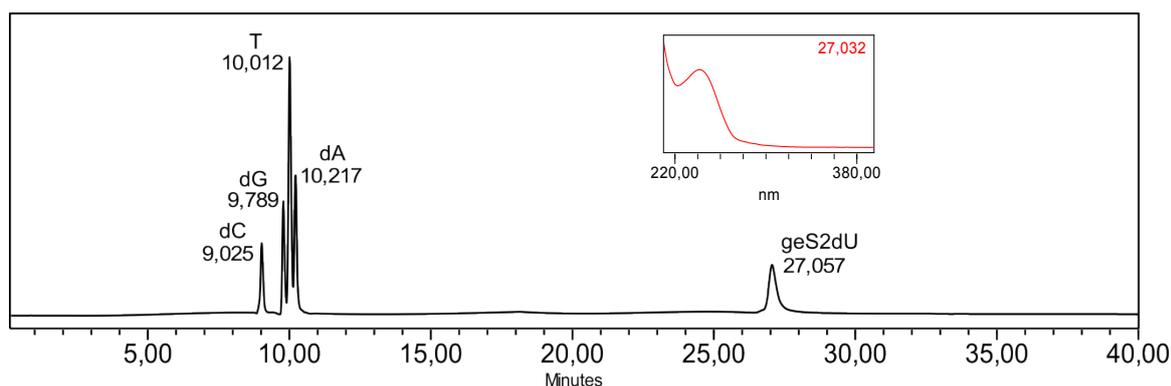
**Figure S13.** The MALDI-TOF spectrum of 5'-d(GTTGACTS2dUTTAATCAAC)-3' (S2dU-DNA); *m/z* 5162, MW 5159.



**Figure S14.** IE-HPLC analysis of geS2dU-DNA ( $R_t = 30.76$  min) obtained by post-synthetic geranylation of fully deprotected S2dU-DNA.



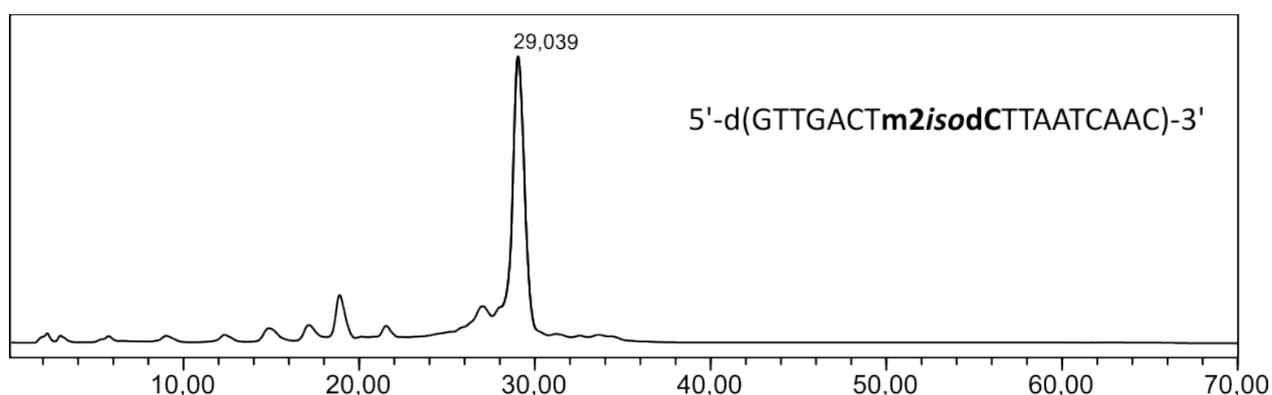
**Figure S15.** The MALDI-TOF spectrum of 5'-d(GTTGACTgeS2dUTTAATCAAC)-3' (geS2dU-DNA;  $m/z$  5298, MW 5298).



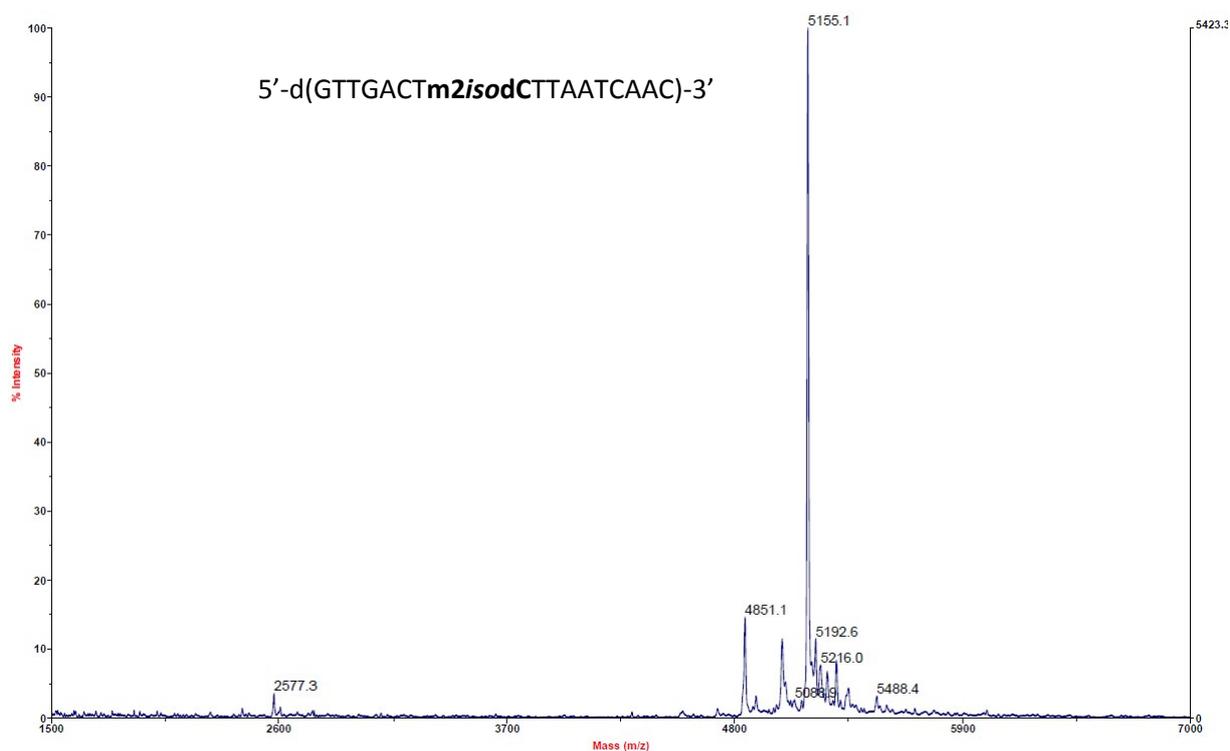
**Figure S16.** RP-HPLC analysis of nucleoside composition of 5'-d(GTTGACTgeS2dUTTAATCAAC)-3' (obtained via post-synthetic geranylation of S2dU-DNA) by a C18 column (Kinetex, 4.6 x 250 mm, 5  $\mu$ m). The mobile phase composition was as follows: linear gradient from 100% A (0.01 % aq. formic acid) to 100% B (50% acetonitrile in 0.01% aq. formic acid, pH 3.5) for 20 min.; maintain at 100% B for 15 min. The inset indicates the UV spectrum of geS2dU which shows a UV absorption peak at 240 nm.

#### VII. Post-synthetic conversion geS2dU-DNA $\rightarrow$ m2isodC-DNA

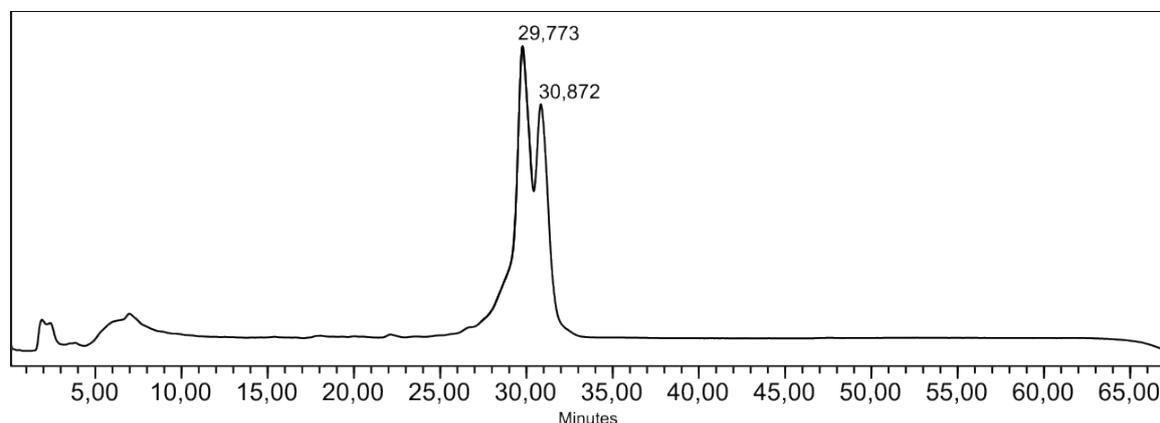
The support-bound geS2dU-DNA (0.2  $\mu$ mol scale) synthesized according to the procedure described in chapter IV was treated with 300  $\mu$ L of AMA (aq 30% ammonia/ aq 40% MeNH<sub>2</sub>, 1/1, v/v) at 65  $^{\circ}$ C for 30 min. The supernatant was transferred to the vial and the resin was washed with water (3  $\times$  300  $\mu$ l). The combined fractions were evaporated on a Speed-Vac concentrator and purified by IE HPLC (Fig. S17) according to the procedure described in chapter IV. Fractions containing the m2isodC-DNA (Rt = 29.04 min) were concentrated and desalted with a C-18 cartridge (Sep-Pak<sup>®</sup>, Waters) affording product in 70% yield. The desalted DNA was lyophilized and analyzed by MALDI-TOF mass spectrometry ( $m/z$  5155, MW 5155, Fig. S18). Due to the similarity of m2isodC-DNA and S2dU-DNA molecular weights IE-HPLC co-injection of both pure oligomers was performed (Fig. S19).



**Figure S17.** IE-HPLC analysis of the crude m2isodC-DNA.



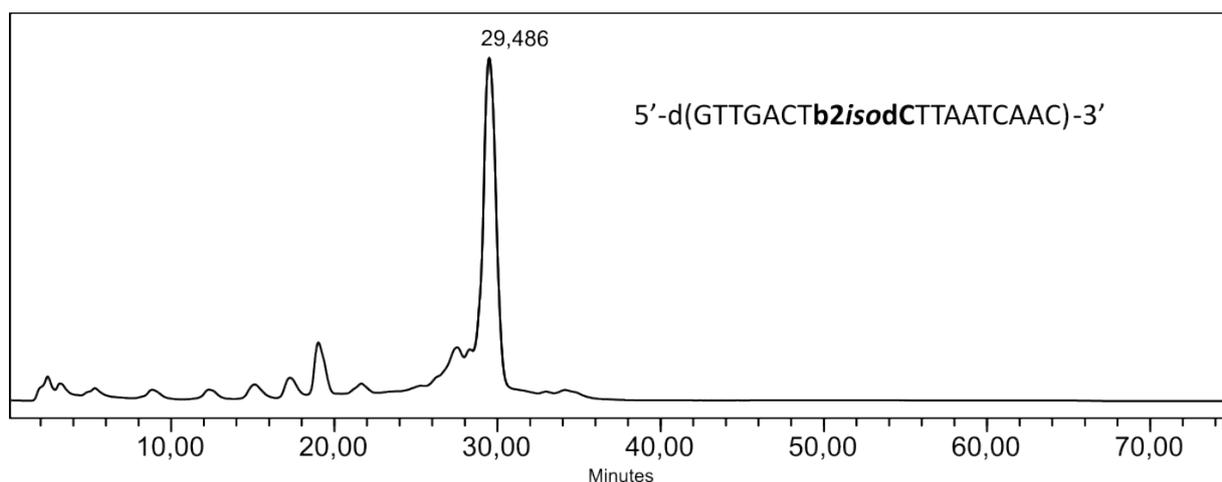
**Figure S18.** The MALDI-TOF spectrum of 5'-d(GTTGACTm2isodCTTAATCAAC)-3' (m2isodC-DNA);  $m/z$  5155, MW 5155.



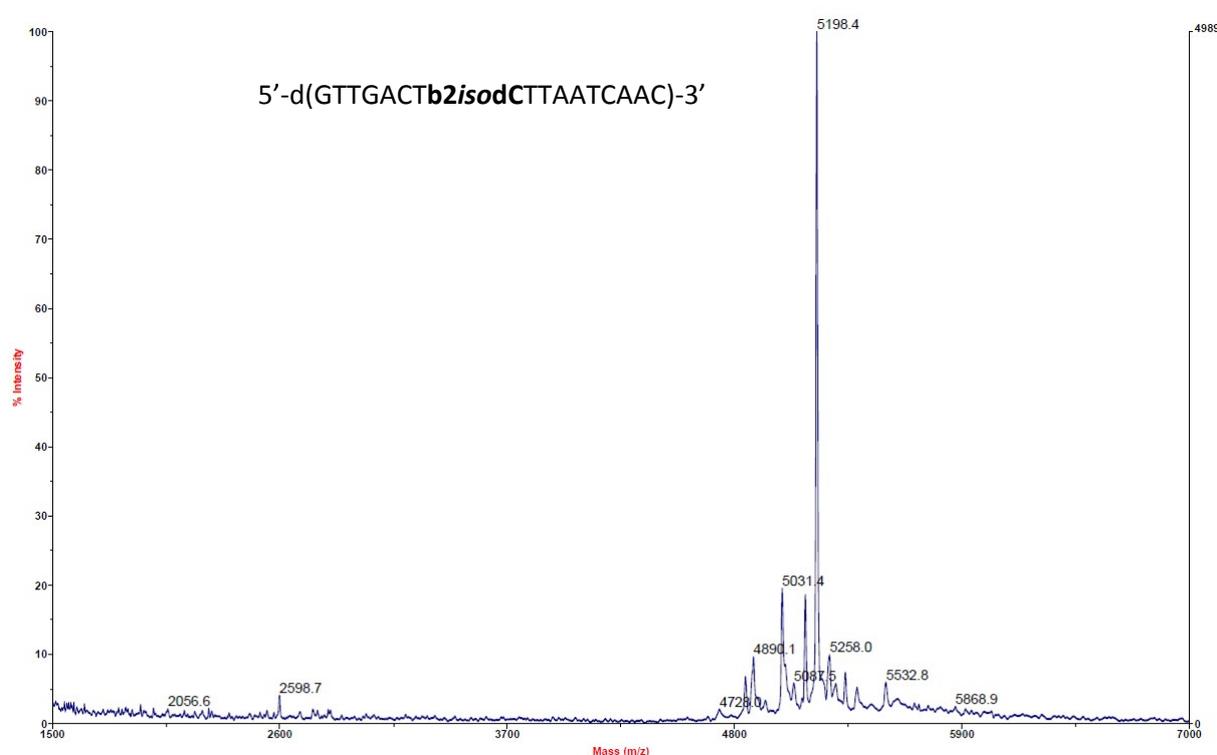
**Figure S19.** IE-HPLC co-injection of pure 5'-d(GTTGACTm2isodCTTAATCAAC)-3' ( $R_t$  = 29.77 min) and 5'-d(GTTGACTS2dUTTAATCAAC)-3' ( $R_t$  = 30.87 min)

#### VIII. Post-synthetic conversion geS2dU-DNA → b2isodC-DNA

The support-bound geS2dU-DNA (0.2  $\mu$ mol scale) synthesized according to the procedure described in chapter IV was treated with the solution of water and n-butylamine (800  $\mu$ L, 1/1, v/v) at 37  $^{\circ}$ C for 24 h. The supernatant was transferred to the vial and the resin was washed with water (3  $\times$  300  $\mu$ l). The combined fractions were evaporated on a Speed-Vac concentrator and purified by IE HPLC (Fig. S20) according to the procedure described in chapter IV. Fractions containing the b2isodC-DNA ( $R_t$  = 29.49 min) were concentrated and desalted with a C-18 cartridge (Sep-Pak<sup>®</sup>, Waters) affording product in 70% yield. The desalted DNA was lyophilized and analyzed by MALDI-TOF mass spectrometry ( $m/z$  5198, MW 5198) (Fig. S21).



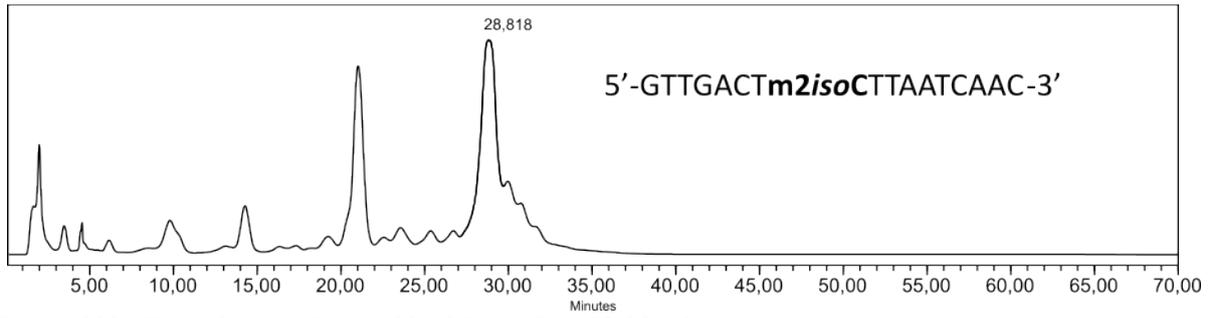
**Figure S20.** IE-HPLC analysis of the crude *b2isodC*-DNA.



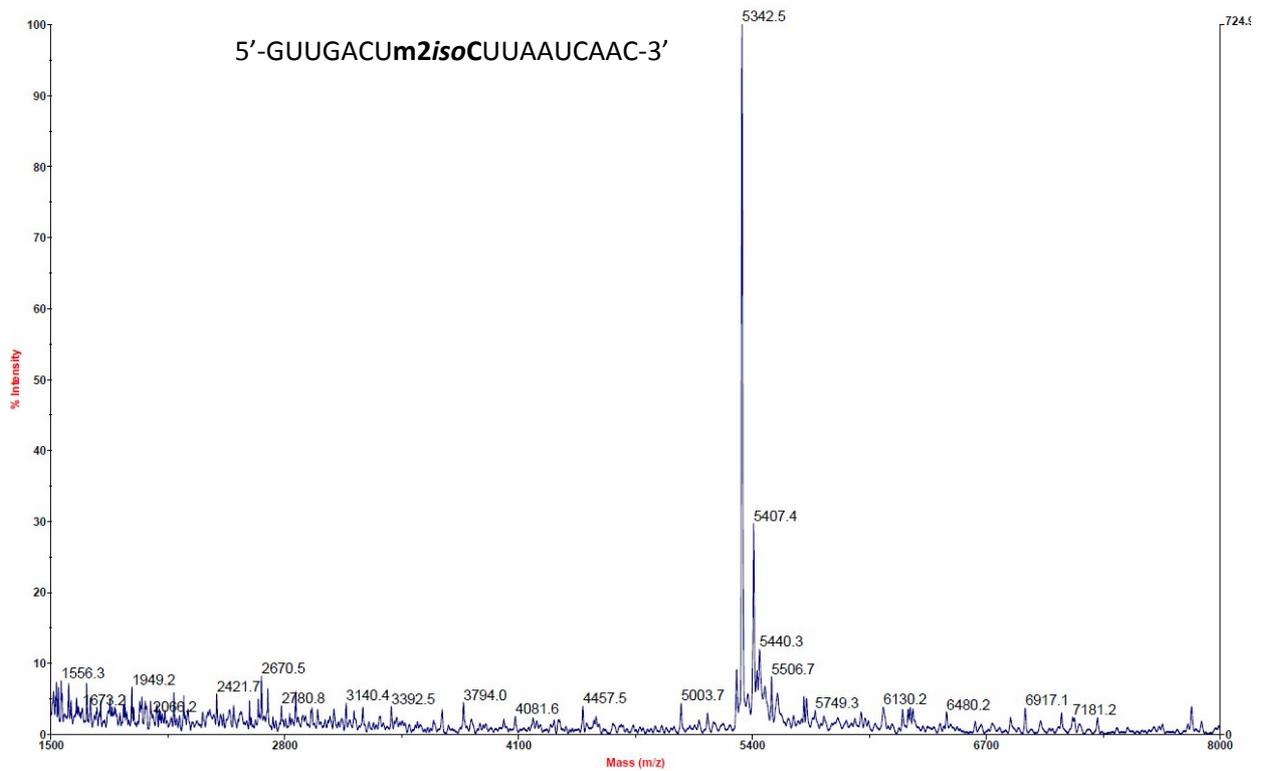
**Figure S21.** The MALDI-TOF spectrum of 5'-d(GTTGACT***b2isod***CTTAATCAAC)-3' (*b2isodC*-DNA);  $m/z$  5198, found 5198.

### IX. Post-synthetic conversion *ges2U*-RNA → *m2isoC*-RNA

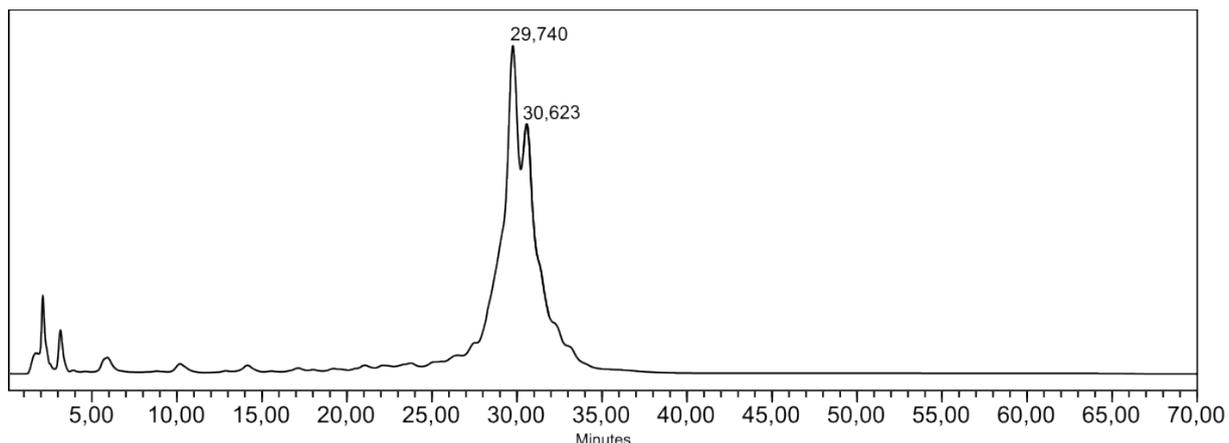
The support-bound *ges2U*-RNA (0.2  $\mu$ mol scale) synthesized according to the procedure described in ref. 4. was treated with 300  $\mu$ L of AMA (aq 30% ammonia/ aq 40% MeNH<sub>2</sub>, 1/1, v/v) at 65 °C for 30 min. The supernatant was transferred to the vial and the resin was washed with water (3  $\times$  300  $\mu$ l). The combined fractions were evaporated on a Speed-Vac concentrator and purified by IE HPLC (Fig. S22) according to the procedure described in chapter IV. Fractions containing the *m2isoC*-RNA (Rt = 28.82 min) were concentrated and desalted with a C-18 cartridge (Sep-Pak<sup>®</sup>, Waters). The desalted RNA was lyophilized and analyzed by MALDI-TOF mass spectrometry ( $m/z$  5342, MW 5343) (Fig. S23). Due to the similarity of *m2isoC*-RNA and *S2U*-RNA molecular weights, IE-HPLC co-injection of both pure oligomers was performed (Fig. S24).



**Figure S22.** IE-HPLC analysis of *m2isoC*-RNA (Rt = 28.82 min).



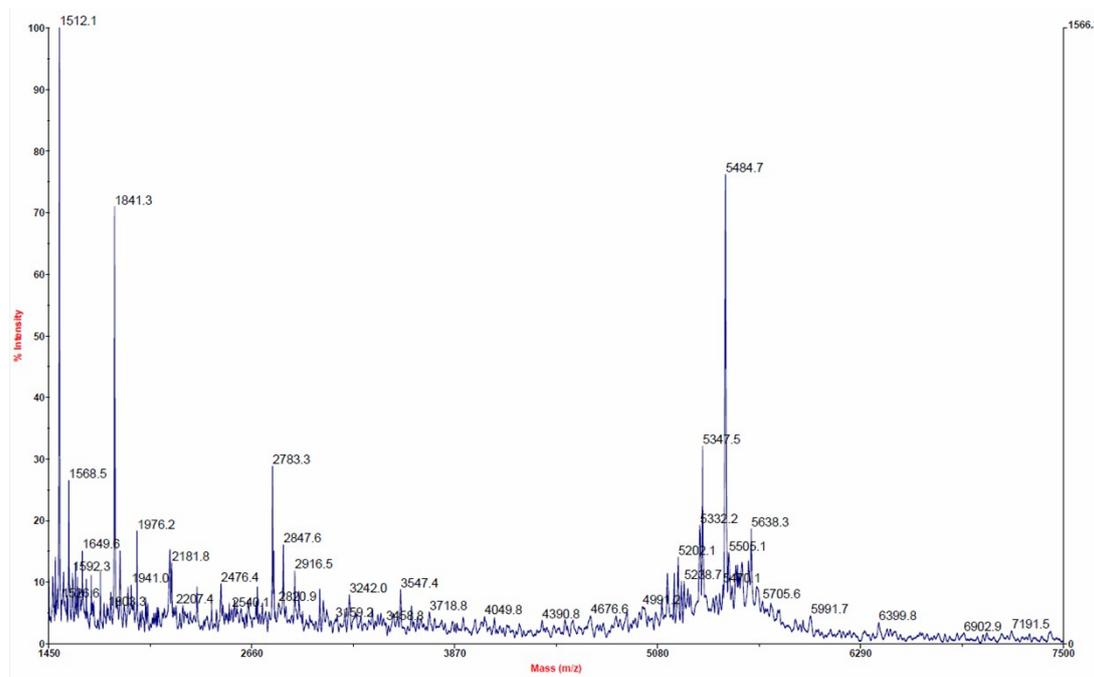
**Figure S23.** The MALDI-TOF spectrum of 5'-GUUGACUm2isoCUUAAUCAAC-3' (*m2isoC*-RNA); *m/z* 5342, MW 5343.



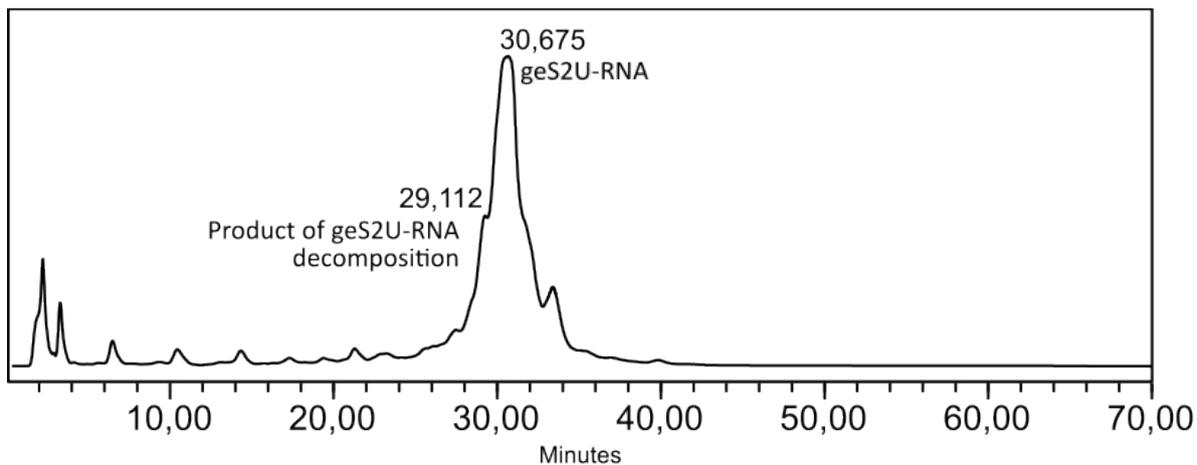
**Figure 24.** IE-HPLC co-injection of pure 5'-GUUGACUm2isoCUUAAUCAAC-3' (Rt = 29.74 min) and 5'-GUUGACUS2UUUAAUCAAC-3' (Rt = 30.62 min).

#### X. Deprotection and support cleavage of CPG-linked geS2U-RNA with aq. ammonia

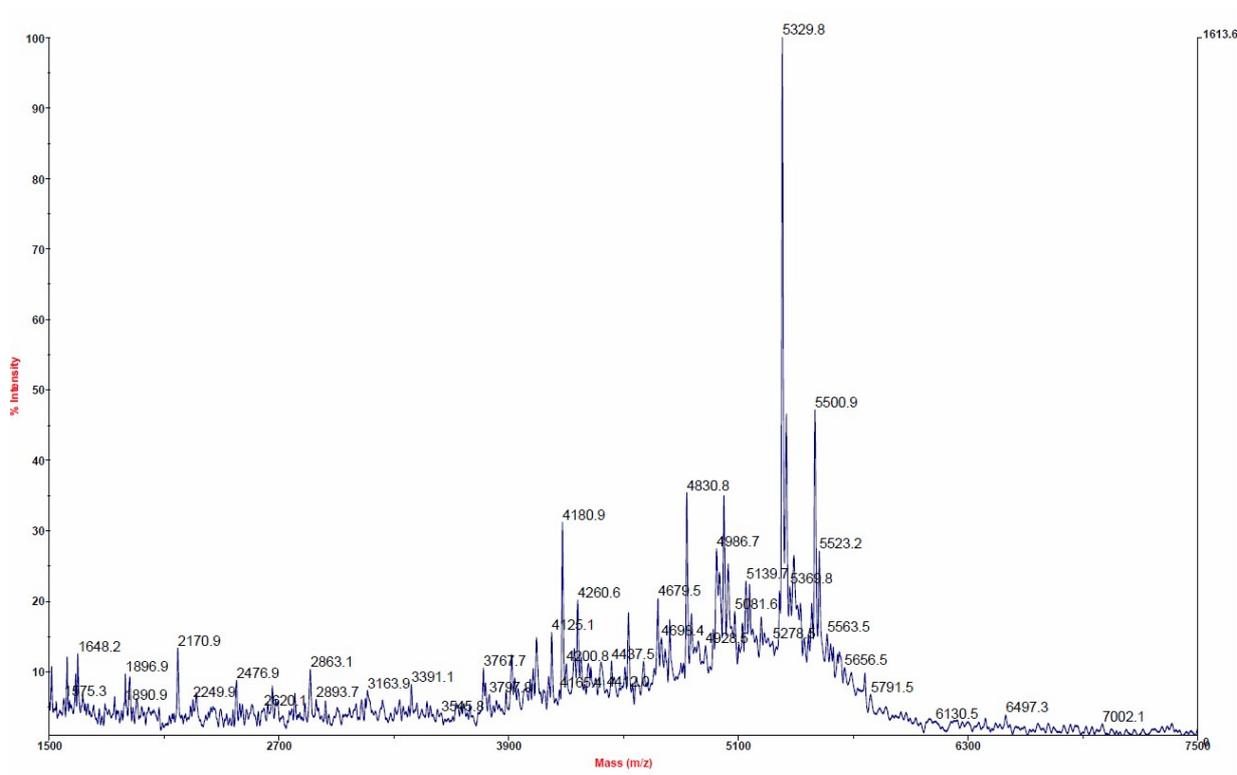
The CPG-linked geS2U-RNA (0.05  $\mu$ mol) was stirred with 66  $\mu$ l of TEA/ACN (1/1, v/v) at rt for 20 min. The supernatant was removed and the resin was washed with ACN (3  $\times$  100  $\mu$ l) and dried in vacuo for 30 min. The support-bound geS2U-RNA was treated with 75  $\mu$ l of 30% aq. ammonia at rt for 14 h. The supernatant was transferred to the vial and the support was washed with water (3  $\times$  100  $\mu$ l). The combined fractions were evaporated on a Speed-Vac concentrator. The MALDI-TOF analysis of crude mixture (Fig. S25) indicated slight decomposition of geS2U-RNA ( $m/z$  5332.2 derives from U-RNA or isoC-RNA;  $m/z$  5347 derives from S2U-RNA). The crude mixture of oligomers was analyzed by IE-HPLC (Fig. S26). Two fractions (picks at Rt = 29.1 min and 30.6 min) were desalted with a C-18 cartridge (Sep-Pak<sup>®</sup>, Waters) and analyzed by MALDI-TOF mass spectrometry. The pick at Rt = 29.1 min derives from U-RNA or isoC-RNA (Fig. S27) whereas the pick at Rt = 30.6 min derives from geS2U-RNA (Fig. S28). On the base of isolated amounts of both oligomers the loss of -Sge residue was estimated on 18%.



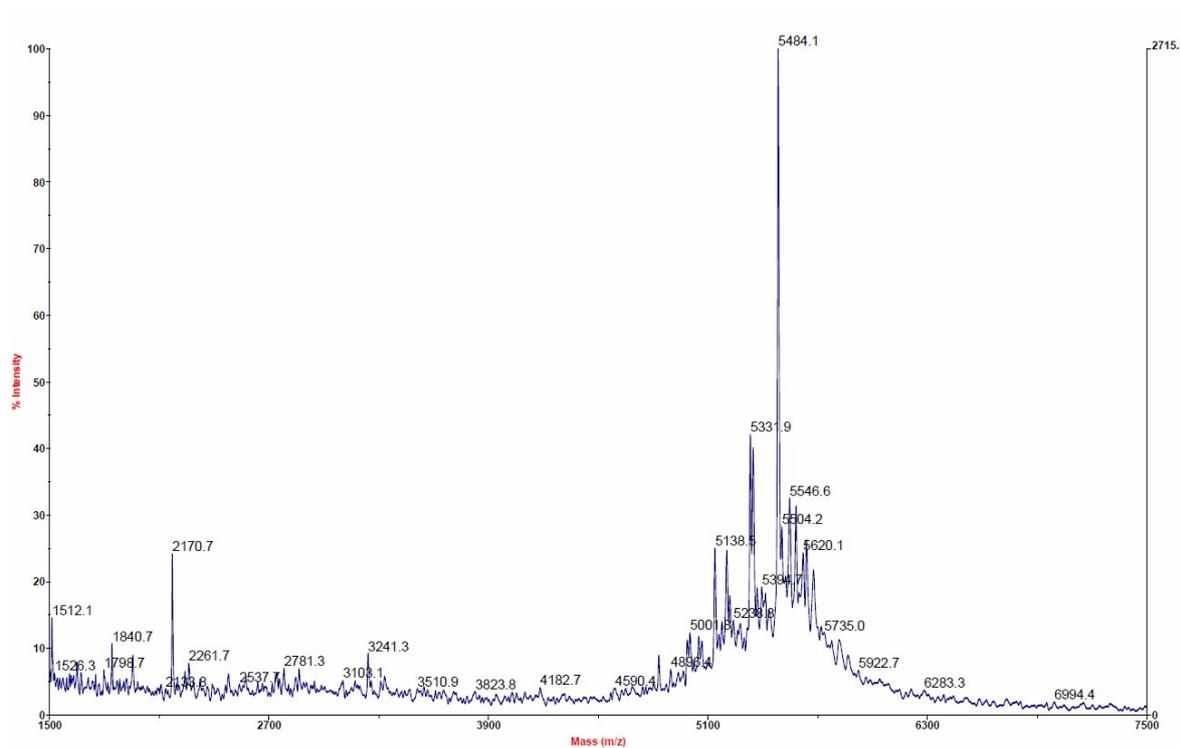
**Figure. S25.** The MALDI-TOF spectrum of crude mixture obtained after treatment of CPG-linked geS2U-RNA with 30% aq. ammonia.



**Figure S26.** IE-HPLC analysis of the crude mixture of oligomers obtained after treatment of CPG-linked geS2U-RNA with 30% aq. ammonia (14 h, rt).



**Figure. S27.** The MALDI-TOF spectrum of oligomer at Rt=29.1 min (U-RNA or *iso*C-RNA,  $m/z$  5329.8, MW 5330.6).



**Figure. S28.** The MALDI-TOF spectrum of oligomer at  $R_t = 30.67$  min (geS2U-RNA;  $m/z$  5484.1, MW 5482.7).

## XI. UV melting temperature measurements

UV absorption measurements were carried out in a 1-cm path length cells using a Cintra 4040 spectrophotometer, equipped with a Peltier Thermocell (GBC, Dandenong, Australia), with a detector set at 260 nm. Complementary DNA/DNA oligonucleotide strands were mixed in a phosphate buffer (10 mM sodium phosphate pH 7.4 containing 0.1 M NaCl) at the final 2  $\mu$ M concentration. Then, the samples were heated to 85  $^{\circ}$ C and cooled to 15  $^{\circ}$ C with a temperature gradient of 1.5  $^{\circ}$ C/min. The melting profiles were recorded from 15 to 85  $^{\circ}$ C, with the temperature gradient of 0.5  $^{\circ}$ C/min. The melting temperatures were calculated using the first derivative method implemented in the Cintra 4040 software. The calculation of thermodynamic parameters ( $\Delta G^{\circ}$ ,  $\Delta H^{\circ}$ ,  $\Delta S^{\circ}$ ) was done by numerical fitting of a given melting curve using a two-state model algorithm provided by a MeltWin v.3.5 software. Each result was taken as an averaged one from three independent experiments.

## XII. Thermodynamic data for DNA/DNA and RNA/RNA duplexes containing geS2Ura and m2isoCyt.

**Table S1.** Base pairing properties of DNA duplex: 5'-d(GTTGACTXTTAATCAAC)-3'·3'-d(CAACTGAYAATTAGTTG)-5' or RNA duplex: 5'-GUUGACUXUUAUCAAC-3' 3'-CAACUGAYAAUUA GUUG-5' (X pairs with Y). Thermodynamic parameters of the investigated duplexes were determined by MeltWin. v. 3.5 melting curves fitting software.

No	X	Y	T <sub>m</sub> (°C)	ΔS°(cal/K mol)	ΔH°(kcal/mol)	ΔG°(kcal/mol)
<b>DNA duplex</b>						
1.	T	dA	51.2 ± 0,4	-278,02±1,3	-98,58±0,4	-12.4 ± 0,03
2.	dU	dA	50.7 ± 0.3	-271,26±22,4	-96,04±0,4	-11.9 ± 7.4
3.	S2dU	dA	50.4 ± 0.3	-256,69±3,2	-91,24±0,1	-11.6 ± 1.1
4.	ges2du	dA	41.5 ± 1.0	-130,04±2,3	-50,16±0,8	-9.8 ± 1.5
5.	m2isodC	dA	38.3 ± 1.1	-115,86±0,9	-44,25±0,3	-8.3 ± 0.0
6.	T	dG	43.8 ± 0,3	-168,30±2,1	-61,50±0,7	-9.3 ± 0,1
7.	dU	dG	42.6 ± 0.4	-169,10±1,1	-61,46±0,04	-9.03 ± 0.4
8.	S2dU	dG	43.1 ± 0.2	-128,87±1,2	-48,66±0,4	-8.7 ± 0.02
9.	geS2dU	dG	45.2 ± 0.6	-232,11±1,1	-82,25±0,4	-10.3 ± 0.06
10.	m2isodC	dG	45.7 ± 0.5	-241,23±9,3	-85,10±2,9	-10.3 ± 0.1
<b>RNA duplex</b>						
11.	U	A	55.9 ± 0.3	-313.7 ± 0.8	-112.4 ± 0.2	-15.1 ± 0.1
12.	S2U	A	59.5 ± 0.2	-292.4 ± 2.8	-106.4 ± 0.9	-15.8 ± 0.1
13.	geS2U	A	46.9 ± 1.0	-138.1 ± 3.9	-53.5 ± 1.3	-10.7 ± 0.2
14.	m2isoC	A	47.1 ± 0.7	-101,9±2,8	-40,8±0,9	-9.2 ± 0.04
15.	U	G	53.5 ± 0.3	-295.9 ± 11.5	-105.8 ± 3.9	-13.9 ± 0.3
16.	S2U	G	51.6 ± 0.5	-253.9 ± 5.9	-91.7 ± 1.9	-12.9 ± 0.2
17.	geS2U	G	50.1 ± 0.2	-218.3 ± 11.5	-79.9 ± 3.8	-12.2 ± 0.2
18.	m2isoC	G	48.9 ± 0.5	-218,8±9	-78,9±2,9	-11.0 ± 0.1

## XIII. References

1. R. G. Kuimelis, H. Hope and K. P. Nambiar, *Nucleosides & Nucleotides*, 1993, **12**, 737.
2. G. Leszczynska, J. Pieta, P. Leonczak, A. Tomaszewska and A. Malkiewicz, *Tetrahedron Lett.* 2012, **53**, 1214.
3. (A) C. W. Gehrke and K. C. Kuo, *J. Chromatogr.* 1989, **471**, 3; (B) C. W. Gehrke, K. C. Kuo, R. A. McCune and K. O. Gerhardt, *J Chromatogr Biomed Sci Appl* 1982, **230**, 297.
4. M. Sierant, G. Leszczynska, K. Sadowska, A. Dziergowska, M. Rozanski, E. Sochacka and B. Nawrot, *Nucleic Acids Res.*, 2016, **44**, 10986.