Supporting Information

Selective chemical labeling of natural T modifications in DNA

Robyn E. Hardisty, Fumiko Kawasaki, Aleksandr B. Sahakyan, Shankar Balasubramanian

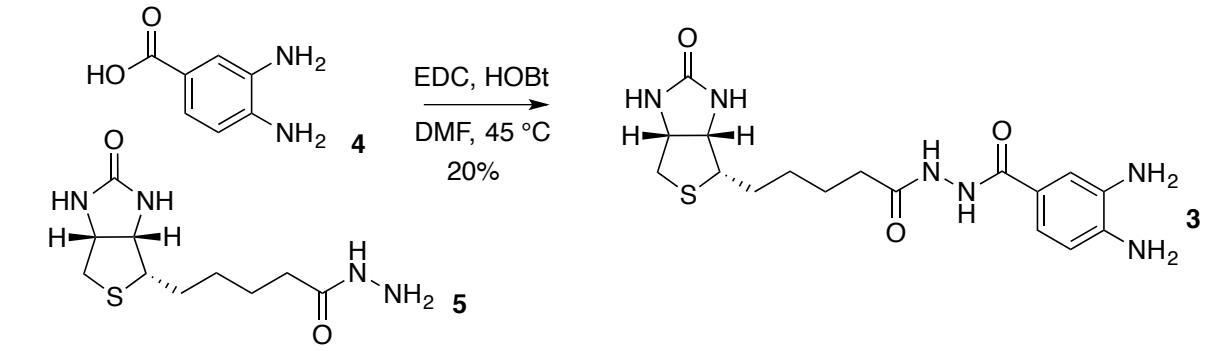
Table of Contents

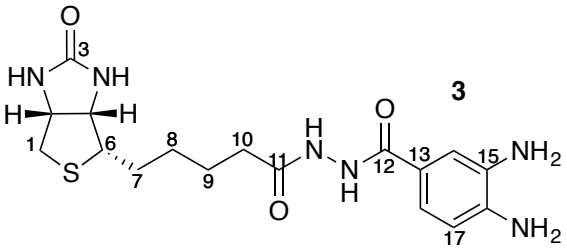
Section 1: General remarks

All solvents and reagents were purified by standard techniques reported in Armarego, W. L. F., Chai, C. L. L., Purification of Laboratory Chemicals, 5th edition, Elsevier, 2003; or used as supplied from commercial sources (Sigma Aldrich® unless stated otherwise). Thin layer chromatography (TLC) was performed on Merck Kieselgel 60 F254 plates, and spots were visualized under UV light. LC-MS was performed on an Agilent LC-MS (Agilent Technologies, Santa Clara, CA). Flash chromatography was carried out using CombiFlash Rf (Teledyne Isco) with puriFlash columns (Interchim). NMR spectra were acquired on a Bruker® DRX-500 instrument using deuterated solvents as indicated and at ambient probe temperature (300 K). Notation for the ^{1}H NMR spectral splitting patterns includes: doublet (d), double doublet (dd), double double doublet (ddd), multiplet/overlapping peaks (m). Signals are quoted as δ values in ppm, coupling constants (*J*) are quoted in Hertz and approximated to the nearest 0.1. Data analysis for the NMR spectra was performed using MestReNova® software. Accurate mass spectra were recorded on a Waters LCT Premier (ESI) spectrometer. DNA concentration was measured using a ND-1000 Spectrophotometer (Nanodrop).

Section 2: Experimental procedure and characterisation of Probe 3

Scheme S1: Synthesis of Probe 3.





To a solution of 3,4-diaminobenzoic acid **4** (70 mg, 0.46 mmol) in DMF (20 mL) was added *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (96 mg, 0.50 mmol), 1-hydroxybenzotriazole (84 mg, 0.55 mmol), NEt_3 (0.12 ml, 0.86 mmol) and (+)-biotin hydrazide **5** (107 mg, 0.41 mmol). The reaction was stirred at 45 °C for 18 h, then the solvent was removed *in vacuo*. The reaction mixture was subjected to flash chromatography (CH_2Cl_2:MeOH=19:1 to 3:2, v/v) and size exclusion chromatography using Sephadex LH-20 to afford **3** as a white solid. (35 mg, 20%); ^{1}H NMR (500 MHz, methanol-*d*4) δ 7.23 (d, *J* = 2.1 Hz, 1H, *H*14), 7.19 (dd, *J* = 8.2, 2.1 Hz, 1H, *H*18), 6.67 (1H, d, *J* = 8.2 Hz, *H*17), 4.49 (ddd, *J* = 7.9, 5.0, 0.9 Hz, 1H, *H*2), 4.32 (dd, *J* = 7.9, 4.5 Hz, 1H, *H*5), 3.22 (ddd, *J* = 8.4, 6.2, 4.5 Hz, 1H, *H*6), 2.93 (dd, *J* = 12.7, 5.0 Hz, 1H, *H*1), 2.70 (d, *J* = 12.7 Hz, 1H, *H*1), 2.32 (m, 2H, *H*10), 1.75 (m, 2H, *H*9), 1.67–1.45 (m, 4H, *H*7, *H*8); ^{13}C NMR (100 MHz, methanol-*d*4) δ 173.9, 168.2, 164.8, 140.0, 133.3, 121.2, 119.5, 115.2, 114.1, 61.8, 60.3, 55.5, 39.7, 33.2, 28.2, 28.0, 25.0; HRMS $C_{17}H_{25}N_6O_3S^+$ $[M+H]^+$ calculated 393.1709, found 393.1701.

Section 3: ODN sequences, reactions and characterisation

Source of fU-ODN, fC-ODN and hmU-ODN. fU-ODN (see Table S1) was synthesized using a protected 5-formyldeoxyuridine phosphoramidite. The identity of the product was confirmed by LC-MS analysis. fC-ODN (see Table S1) was obtained from Eurogentec and subjected to further HPLC purification using a Agilent Technologies 1200 series HPLC to remove impurities. A Pursuit C18 column (5 μM , 150 x 10.0 mm, Agilent) was used, (solvent A = 50 mM NH_4OAc, solvent B = MeCN, flow-rate 4 mL/min, 3% B for 5 min, and a gradient of 3-10% B for 25 min). hmU-ODN (See Table S1) was purchased from ATD Bio.

Synthesis of fU-DNA, fC-DNA and GCAT-DNA by polymerase chain reaction (PCR).

fU-DNA (See Table S1) was synthesized using **template 1** and **forward primer 1 and reverse primer 1** (See Table S8) in the presence of dATP, dCTP, dGTP and dfUTP.

fC-DNA (See Table S1) was synthesized using **template 2, forward primer 2** and **reverse primer 2** (See Table S8) in the presence of dATP, dfCTP, dGTP, and dTTP.

GCAT-DNA was synthesized using **template 3**, **forward primer 3** and **reverse primer 3** (See Table S8) in the presence of dATP, dCTP, dGTP and dTTP.

PCRs were made up with 4 x dNTP (1 mM), forward primer 1, 2, or 3 (1 μM), reverse primer 1, 2 or 3 (1 μM), template 1, 2, or 3 (0.01 μM), 10 × DreamTaq Buffer (2.5 μL) and DreamTaq Polymerase (Thermo Scientific, 0.25 μL) to give a final volume of 25 μL. Primers and templates were obtained from Sigma Aldrich or Invitrogen. dGTP, dCTP, dATP and dTTP were obtained from ThermoFischer, while dfCTP and dfUTP were obtained from Trilink Biotechnologies. The mixture was then subjected to the following thermal cycle: 95 °C for 3 min, 40 cycles of (95 °C

for 30 s, 60 °C for 60 s, 72 °C for 90 s), 72 °C for 5 min using a T100 Thermal Cycler (BioRad). The PCR products were purified using a GeneJET PCR Purification Kit (Thermo Scientific) according to the manufacturer's instructions. Formation of 80mer ODNs was confirmed by Tapestation 2200 (Agilent Technologies) using D1000 screen-tape. Confirmation of incorporated bases was confirmed by digestion of synthetic DNA and injection onto a QExactive quadrupole-orbitrap hybrid tandem MS spectrometer (Thermo Fischer), where the mass of constituent nucleosides $[M+H]^+$ was extracted.

Digestion of DNA. DNA (500 ng) in the presence of Degradase Plus (0.5 μL) and 10 × Degradase Plus reaction buffer (4 μL) (Zymo Research) in a final volume of 40 μL was digested to its constituent nucleosides *via* incubation at 37 °C for 3 h. The nucleoside mixture was subsequently cleaned-up by filtration using Amicon Ultra-0.5 mL Centrifugal Filters 10K (Millipore).

Table S1: Sequence of ODNs used for this study (modifications are highlighted in bold).

ODNs	**Sequences**
fU-ODN	5'-ATCGCA**fU**GTA-3'
fC-ODN	5'-TAATTATC**fC**GGACTCATAAG-3'
fU-DNA	5'- TTCTTGGCTGTGGCTCTGCGTCCTTGTCCTGCCCAC**fU**GCC**fU**GACGGGCGGAGG CACAACAGAGAGCAACACCGCCGAGGA-3' Complement sequence 3'- AAGAACCGACACCGAGACGCAGGAACAGGACGGG**fU**GACGGACGCCCGCC**fU**C CGTGTTGTCTCTCGTTGTGGCGGCTCCT-5'
fC-DNA	5'- CTAAATCTACTAAATCCTCTAAATCTATTCTATA**fC**ATGAAT**fC**TTA GTTAAAGGTAGTAGTAGTAGATATAAGATGATAGG-3 Complement sequence 3'- GATTTAGATGATTTAGGAGATTTAGATAA**fC**ATATGTACTTA**fC**AATCAATTTCC ATCATCATCATCTATATTCTACTATCC-5'
GCAT-DNA	5'- GCTCGCTTTGTTGGTTTCCTTGTTCTCTGTGCCCACTGCCTGACGGGCGGAAAG CAGCGCGAGCAAGCGAGACAGGACAC-3' Complement sequence 3'- CGAGCGAAACAACCAAAGGAACAAGAGACACGGGTGACGGACTGCCCGCCTT TCGTCGCGCTCGTTCGCTCTGTCCTGTG -5'
hmU-ODN	5Phos-ATCGCA**hmU**GTA

Chemical tagging reactions with fU-ODN and fC-ODN. Biotinylated probes were made up as a stock solution in DMSO; **1** (100 mM), **2** (100 mM), **3** (50 mM). *p*-anisidine was made up as a stock solution in MeOH (1 M). fU-ODN (2 μL, 100 uM) and/or fC-ODN (2 μL, 100 mM) were subjected to reaction conditions (Table 1**,** main text) in the presence of **1** (0.4 mM), **2** (10 mM), or **3** (5 mM) in the presence of absence of *p*-anisidine (100 mM). pH buffer solution (sodium phosphate or NH_4OAc) had a final concentration of 40 mM. Reactions on ODNs were purified by mini quick spin oligo columns (Roche), which were pre-washed with water (2 × 300 μL). Analysis was performed by LC-MS analysis. Both O-(Biotinylcarbazoylmethyl)Hydroxylamine **1** (Cayman Chemicals) and (+)-Biotinamidohexanoic acid hydrazide **2** (Sigma Aldrich) were commercially sourced.

Chemical oxidation of hmU-ODN. Using a protocol used for 5-hmC oxidation[1], hmU-ODN (2 μL, 100 uM) (Table S1) was incubated with NaOH (1.25 μL, 1M) and $KRuO_4$ (1 μL, 15 mM in 0.05 M NaOH) (Alfa Aesar) on ice for 1 h. The reaction was purified by mini quick spin oligo column (Roche), which was pre-washed with water (2 × 300 μL). Analysis was performed by LC-MS analysis, and nucleobase composition of digested DNA was analyzed by a Q-exactive (Thermo Fischer) quadrupole-orbitrap hybrid tandem MS spectrometer in positive ion mode, where ions of masses 228.1, 243.1, 252.1, 268.1, 259.1, 257.1 were fragmented in a positive ion mode at 10% collision energy. Spectra was processed using Xcalibur software (ThermoFischer). Extracted ion chromatograms of base fragments 112.05054, 127.05020, 136.06177, 152.05669, 143.04512, 141.02947 were used corresponding to C, T, A, G, 5-hmU and 5-fU respectively. Gaussian smoothing (7 points) was applied.

Selective chemical labeling of hmU-ODN over fU-ODN. To a mixture of fU-ODN (3 μL, 100 μM) and hmU-ODN (3 μL, 100 μM) was added *N*-methylhydroxylamine hydrochloride (5 μL, 50 mM) and pH = 6 Sodium Phosphate Buffer (40 mM) in a final reaction volume of 50 μL. The reaction was left for 3 h at RT, before being purified using two Bio-Spin P-6 Gel Columns, SSC Buffer (Bio-Rad), which had been pre-washed with water (6 × 500 μL). The resultant solution was incubated with NaOH (2.5 μL, 1M) and $KRuO_4$ (2 μL, 15 mM in 0.05 M NaOH) (Alfa Aesar) on ice for 1 h. The mixture was purified by mini quick spin oligo column (Roche), which was pre-washed with water (2 × 300 μL), before the addition of **2** (10 μL, 100 mM) and pH = 7 Sodium Phosphate Buffer to give a final concentration of 40 mM. Purification was completed *via* mini quick spin oligo column (Roche) which was pre-washed with water (2 × 300 μL). Analysis was performed by LC-MS analysis.

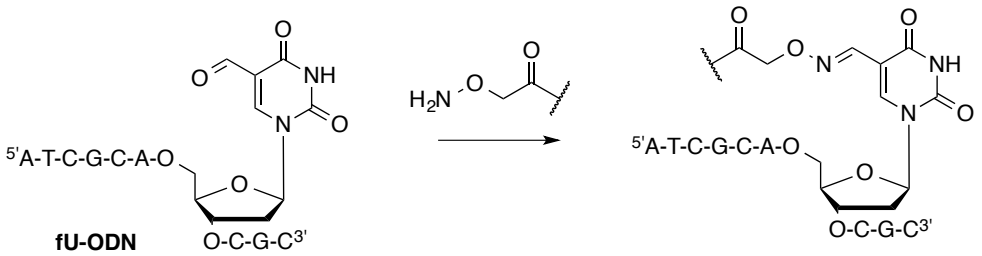
LC-MS analysis of ODNs. LC-MS was performed on a Bruker amaZon system, using an XTerra MS C18 column (2.5 μM, 2.1 x 50 mm), using solvents A (100 mM 1,1,1,3,3,3-hexafluoro-2-propanol, 10 mM NEt_3) and B (MeOH), at a flow-rate of 0.2 mL/min, with a gradient 5%–30% B increasing at 1% per minute. Reaction conversion was calculated by integration of UV signals of the starting material and product(s) at 260 nm. Identity of products was confirmed by ESI-MS. Observed ESI^- signals and their retention times are listed in Table S2.

Table S2: LC-MS data for ODNs used in this study and their reaction products.

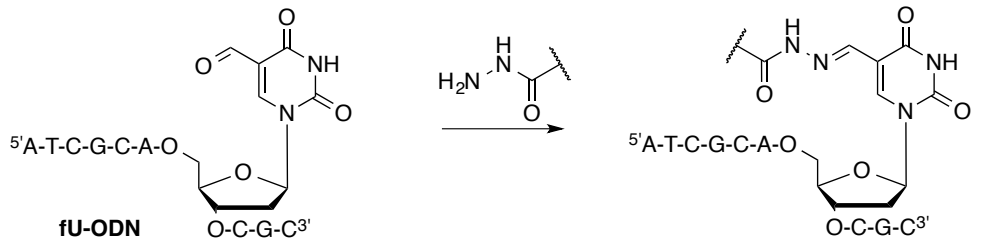
ODN	MW	ESI-MS	Retention Time
fU-ODN	3041.02	M^{-2} = 1519	13.7
fC-ODN	6128.04	M^{-3} = 2041 M^{-4} = 1530	17.5
hmU-ODN	3121.51	M^{-2} = 1560	14.4

fU-ODN + **1**	3354.44	M^{-2} = 1676	15.8
fU-ODN + **2**	3393.51	M^{-2} = 1696	17.5
fU-ODN + **3**	3413.46	M^{-2} = 1705	16.3
fU-ODN + NH_2OMe	3068.56	M^{-2} = 1534	15.0
fC-ODN + **1**	6441.42	M^{-3} = 2145 M^{-4} = 1609	18.2
fC-ODN + **2**	6481.53	M^{-3} = 2159 M^{-4} = 1619	19.5
fC-ODN + **3**	6500.49	M^{-3} = 2166 M^{-4} = 1624	18.8
hmU-ODN + $KRuO_4$	3120.20	M^{-2} = 1559	14.6
hmU-ODN + $KRuO_4$ + **2**	3472.69	M^{-2} = 1736	17.7

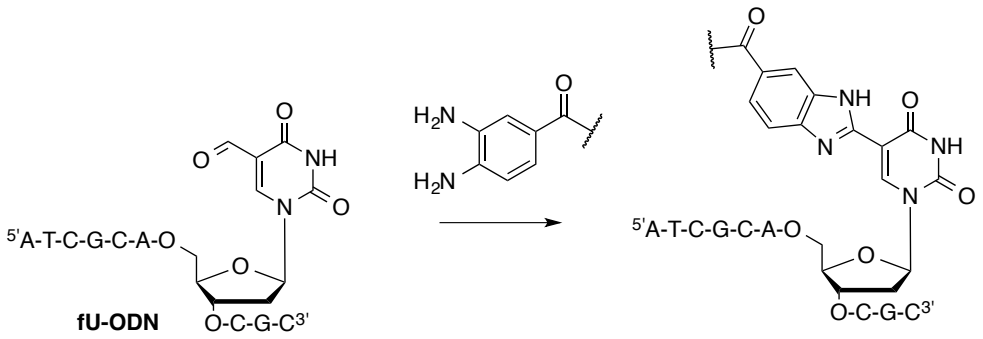
Scheme S2: Reaction of fU-ODN with biotinylated oxyamine **1** gave an ESI^- corresponding to the formation of an oxime.



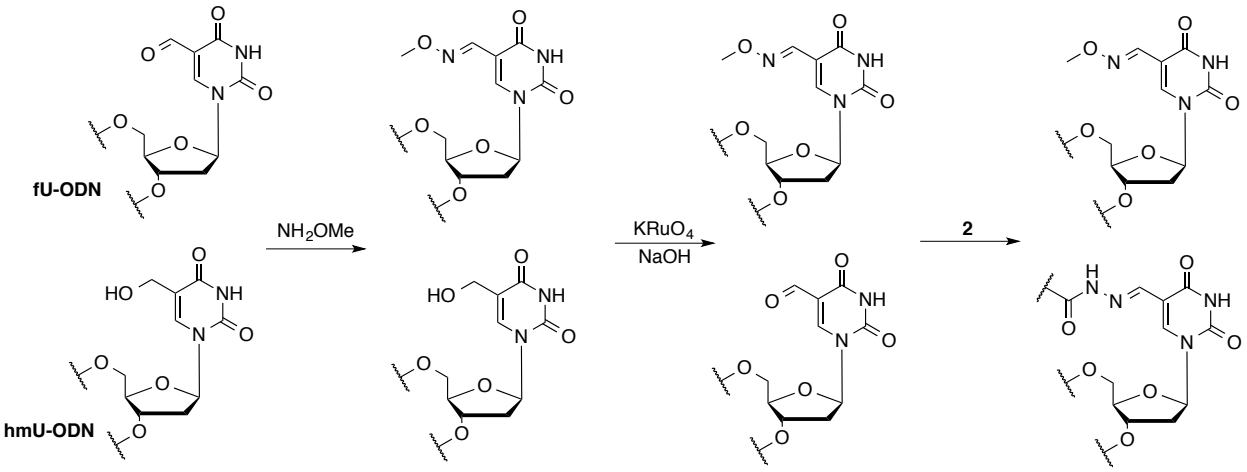
Scheme S3: Reaction of fU-ODN with biotinylated hydrazide **2** gave an ESI^- corresponding to the formation of a hydrazone.



Scheme S4: Reaction of fU-ODN with the phenylenediamine probe 3 gave an ESI^- corresponding to a stable benzimidazole, as shown previously for the reaction of 5-fU mononucleotide with an unsubstituted phenylenediamine.[2]



Scheme S5: Selective biotinylation of hmU-ODN with **2**. ESI^- signals correspond to fU-ODN functionalized with NH_2OMe and hmU-ODN that has been oxidized and subsequently functionalized with **2** to form a hydrazone.



Section 4: Experimental procedure for qPCR enrichment studies

Reaction procedure for chemical enrichment studies. DNA (500 ng) was subjected to reaction conditions a, b, c, d or e. The reactions were purified using mini quick spin oligo columns (Roche) pre-washed with water (2 × 300 μL). The resulting purified mixture was diluted 100 fold to give an approximate concentration of 100 pg/μL per ODN. Reactions were done in triplicate:

a) fU-DNA (500 ng) and fC-DNA (500 ng) was incubated with NH_4OAc buffer pH = 5 (40 mM), **1** (2 M) and *p*-anisidine (100 mM), to make a final reaction volume of 50 μL, at RT for 24 h;

b) fU-DNA (500 ng) and fC-DNA (500 ng) was incubated with sodium phosphate buffer pH = 6 (40 mM) and **1** (0.4 mM) to make a final reaction volume of 50 μL, at RT for 4 h;

c) fU-DNA (500 ng) and fC-DNA (500 ng) was incubated with sodium phosphate buffer pH = 7 (40 mM) and **2** (10 mM) to make a final reaction volume of 50 μL, at RT for 4 h;

d) fU-DNA (500 ng) and fC-DNA (500 ng) was incubated with sodium phosphate buffer pH = 7 (40 mM) and **3** (5 mM) to make a final reaction volume of 50 μL, at RT for 4 h;

e) fU-DNA (500ng) and GCAT-DNA (500 ng) was incubated with sodium phosphate buffer pH = 7 (40 mM) and **2** (10 mM) to make a final reaction volume of 50 μL, at RT for 4 h.

DNA enrichment procedure. A reported DNA enrichment protocol was used with some modifications.[3] MagneSphere streptavidin magnetic beads (50 μg, Promega), were washed with 1 × binding buffer (5 mM Tris pH 7.5, 0.5 mM EDTA, 1M NaCl, 0.05% Tween 20) (3 x 500 μL) and then resuspended in 50 μL 2 × binding buffer (10 mM Tris pH 7.5, 1 mM EDTA, 2M NaCl,

0.1% Tween 20). Input DNA (10 μL, 1000 pg/ODN) and Salmon sperm DNA (10 μg, Invitrogen) were mixed and made up to a final volume of 50 μL, and then added to the magnetic beads, before incubation for 15 minutes at RT. Beads were washed with 1 × binding buffer (6 × 500 μL), and the beads were then resuspended in 100 μL elution buffer (95% formamide, 10 mM EDTA) and were heated to 95 °C for 5 min. The eluent was then removed from the beads and placed on ice. The step was then repeated using 50 μL elution buffer to remove residual DNA from the magnetic beads. The eluent was diluted with water (350 μL), and purified by filtration using Amicon Ultra-0.5 mL Centrifugal Filters 10K (Millipore), following a wash by water (450 μL) and centrifugation for 15 min. The Amicon filters were washed with water and centrifuged for a further 15 min (2 × 450 μL). DNA was then recovered from the Amicon filter (25 μL). Enrichments were carried out in either duplicate or triplicate from each reaction.

qPCR analysis for chemical enrichment studies. qPCRs were performed using a CFX96 Real-Time System (BioRad), and data was processed using the CFX Software manager (BioRad). Enriched DNA (1 μL) was added to a mixture of Brilliant III Ultra-Fast SYBR Green qPCR Master Mix (5 μL) (Agilent Technologies), forward primer 1, 2 or 3 (1 μM) (See Table S8), reverse primer 1, 2 or 3 (1 μM) (See Table S8) and diluted with water to give a final volume of 10 μL. The mixture was subject to qPCR according to the protocol outlined by the manufacturer. DNA concentration was quantified by comparison with calibration lines of known concentration of input ODNs. (See Figures S1, S2 and S3).

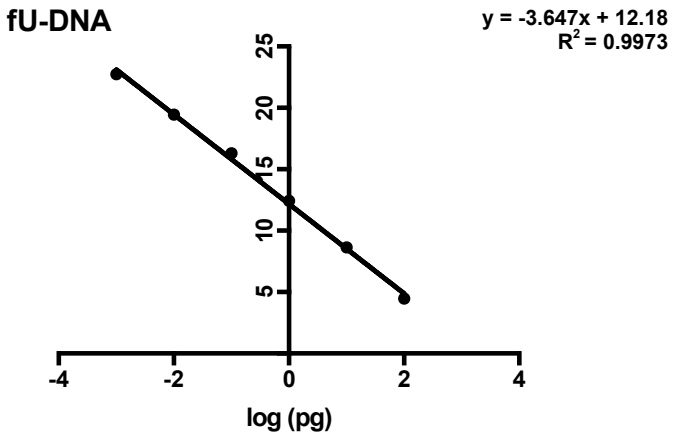


Figure S1: Example calibration line for fU-DNA.

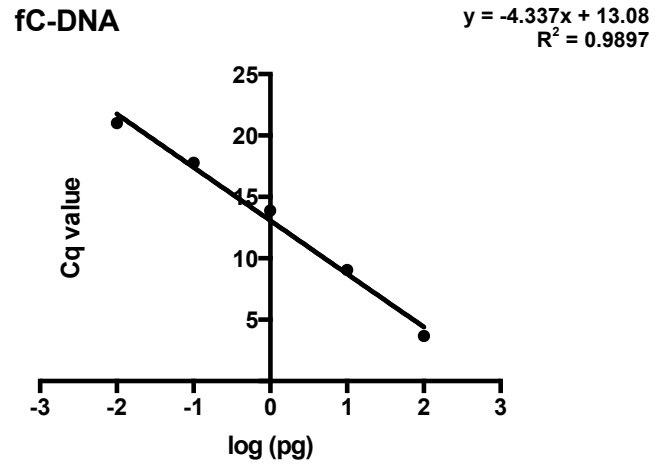


Figure S2: Example calibration line for fC-DNA.

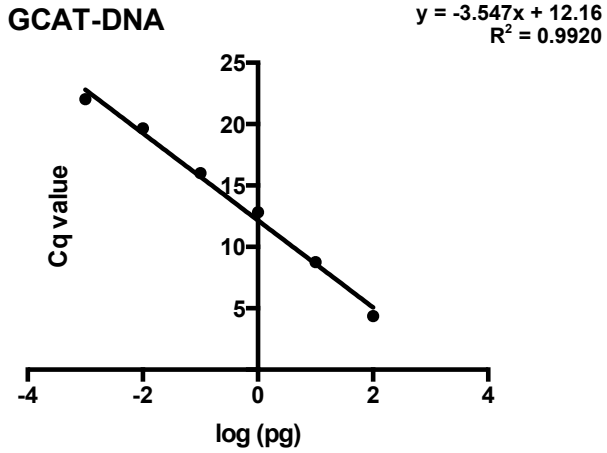


Figure S3: Example calibration line for GCAT-DNA.

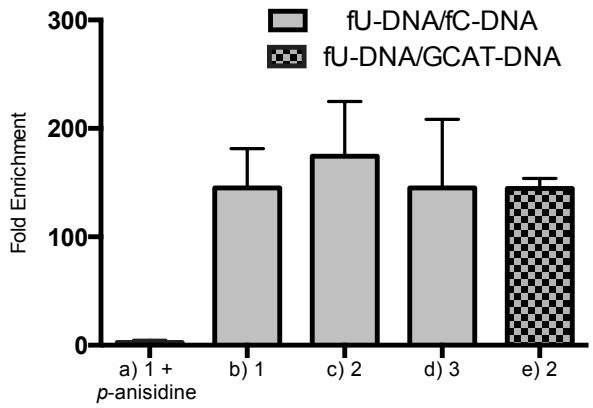


Figure S4: Extent of enrichment of fU-DNA over fC-DNA or GCAT-DNA under different conditions: a) enrichment under the previously reported 5-fC tagging conditions[3]; b) enrichment using biotinylated oxyamine **1**; c) enrichment using biotinylated acyl hydrazide **2**; d) enrichment using biotinylated o-phenylenediamine **3**; e) enrichment of fU-DNA using biotinylated acyl hydrazide **2**.

Table S3: Enrichment data for Probe **1** + *p*-anisidine for fU-DNA and fC-DNA.

Probe 1 + *p*-anisidine	**fU-DNA Cq value**	**fU-DNA (pg)**	**fC-DNA Cq value**	**fC-DNA (pg)**	**Selectivity**	**Mean Selectivity**
Bio 1, Tech 1	11.19	1.809	13.03	1.024	1.8	1.5
Bio 1, Tech 2	11.55	1.435	12.57	1.307	1.1	
Bio 2, Tech 1	10.60	2.625	12.31	1.507	1.7	1.4
Bio 2, Tech 2	10.44	2.899	11.35	2.510	1.2	
Bio 3, Tech 1	9.40	5.781	12.56	1.315	4.4	5.0
Bio 3, Tech 2	9.10	6.954	12.36	1.467	4.8	
Bio 3, Tech 3	9.66	4.879	13.45	0.821	5.9	

Table S4: Enrichment data for Probe **1** for fU-DNA and fC-DNA.

Probe 1	**fU-DNA Cq value**	**fU-DNA (pg)**	**fC-DNA Cq value**	**fC-DNA (pg)**	**Selectivity**	**Mean Selectivity**
Bio 1, Tech 1	8.34	10.943	18.52	0.056	197	184
Bio 1, Tech 2	8.44	10.266	18.98	0.043	236	
Bio 1, Tech 3	9.26	6.1134	18.67	0.051	119	
Bio 2, Tech 1	9.63	3.295	19.09	0.023	143	112

Bio 2, Tech 2	9.07	4.762	16.67	0.089	54	
Bio 3, Tech 3	9.31	4.071	18.65	0.030	138	
Bio 3, Tech 1	8.47	7.080	17.97	0.043	164	139
Bio 3, Tech 2	8.43	7.263	17.26	0.063	114	

Table S5: Enrichment data for Probe **2** for fU-DNA and fC-DNA

Probe 2	**fU-DNA Cq value**	**fU-DNA (pg)**	**fC-DNA Cq value**	**fC-DNA (pg)**	**Selectivity**	**Mean Selectivity**
Bio 1, Tech 1	8.05	13.126	17.82	0.080	163	176
Bio 1, Tech 2	7.92	14.220	18.13	0.068	208	
Bio 1, Tech 3	8.35	10.857	18.10	0.070	156	
Bio 2, Tech 1	9.09	6.811	18.79	0.048	141	123
Bio 2, Tech 2	9.12	6.69	18.25	0.064	104	
Bio 3, Tech 1	8.58	9.372	19.15	0.040	236	224
Bio 3, Tech 2	8.73	8.559	19.11	0.041	211	

Table S6: Enrichment data for Probe **3** for fU-DNA and fC-DNA.

Probe 3	**fU-DNA Cq value**	**fU-DNA (pg)**	**fC-DNA Cq value**	**fC-DNA (pg)**	**Selectivity**	**Mean Selectivity**
Bio 1, Tech 1	10.18	3.420	19.70	0.030	115	120
Bio 1, Tech 2	10.53	2.743	19.56	0.032	86	
Bio 1, Tech 3	9.97	3.899	20.07	0.024	160	
Bio 2, Tech 1	10.18	3.421	21.05	0.014	236	217
Bio 2, Tech 2	10.28	3.221	20.83	0.016	197	
Bio 3, Tech 1	10.49	0.450	19.68	0.030	94	98
Bio 3, Tech 2	10.48	0.451	19.83	0.028	102	

Table S7: Enrichment data for Probe **2** for fU-DNA and GCAT-DNA.

Probe 2	**fU-DNA (Cq value)**	**fU-DNA (pg)**	**GCAT-DNA Cq value**	**GCAT-DNA (pg)**	**Selectivity**	**Mean Selectivity**
Bio 1, Tech 1	11.34	3.565	18.05	0.022	162	151
Bio 1, Tech 2	11.04	4.054	17.63	0.029	141	
Bio 1, Tech 3	10.82	4.692	17.52	0.031	151	
Bio 2, Tech 1	11.36	3.289	17.70	0.028	120	134
Bio 2, Tech 2	10.92	4.386	17.52	0.031	142	
Bio 2, Tech 3	10.43	6.093	16.98	0.044	139	
Bio 3, Tech 1	10.48	5.887	17.12	0.040	147	149
Bio 3, Tech 2	10.84	4.627	17.70	0.028	169	
Bio 3, Tech 3	11.40	3.208	17.88	0.024	131	

Table S8: Sequence of templates and primers used for PCR synthesis of fU-DNA, fC-DNA and GCAT-DNA.

	Sequence
Template 1	5’-TTCTTGGCTGTGGCTCTGCGTCCTTGTCCTGCCCACTGCCTGACGGGCGGAGGCACAACAGAGAGCAACACCGCCGAGGA-3’
Template 2	5’-CTAAATCTACTAAATCCTCTAAATCTATTCTATACATGAATCTTAGTTAAAGGTAGTAGTAGTAGATATAAGATGATAGG-3’
Template 3	5’-GCTCGCTTTGTTGGTTTCCTTGTTCTCTGTGCCCACTGCCTGACGGGCGGAAAGCAGCGCGAGCAAGCGAGACAGGACAC-3’
Forward Primer 1	5’-TTCTTGGCTGTGGCTCTGCGTCCTTGTCCT-3’
Reverse Primer 2	5’-TCCTCGGCGGTGTTGCTCTCTGTTGTGCCT-3’
Forward Primer 2	5’-CTAAATCTACTAAATCCTCTAAATCTATTC-3’
Reverse Primer 2	5’-CCTATCATCTTATATCTACTACTACTACCT-3’

Forward Primer 3	5’-GCTCGCTTTGTTGGTTTCCTTGTTCTCTGT-3’
Reverse Primer 3	5’-GTGTCCTGTCTCGCTTGCTCGCGCTGCTTT-3’

All templates and primers were sourced from either Invitrogen or Sigma Aldrich.

Section 5: Ab initio study on 5-fU and 5-fC reactivity

To obtain a theoretical insight on what might facilitate the increased reactivity of 5-fU, as compared to 5-fC, in the Schiff base, oxime or hydrazone formation reactions described in this work, we have performed *ab initio* quantum mechanical calculations on reduced model systems.

The models and the level of theory used in the computational study. The models used in the computational study are shown in Figure S5a, where the N-glycosidic bonds in the 5-fU and 5-fC nucleotides are reduced via capping methyl groups. Taking into account that all the reactive ends of the studied tagging reagents have -NH_2 groups, methylamine is taken as the model reactant. The rate-determining step in our experimental condition and used reactants is most likely to be the nucleophilic addition to the carbonyl, hence, we considered this stage of the reaction (Figure S5b) to reveal the possible stationary points along the pathway.

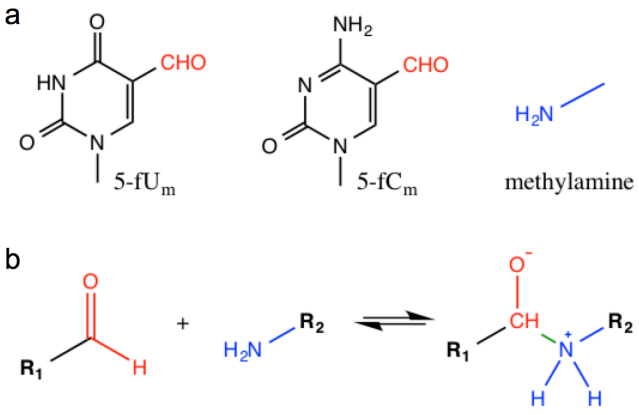


Figure S5. a) The used reduced molecules for the computational study. b) The general first stage of the addition reaction, expected to be the rate-limiting one.

Closed-shell restricted Hartree-Fock (RHF) calculations have been done with the Møller-Plesset correlation energy correction truncated at second-order (MP2)[4,5] and with the double-zeta cc-

pVDZ Dunning's correlation consistent basis set.[6,7] All the calculations have been done using the Gaussian 03 suite of programs.[8]

The aldehyde group rotation barriers in 5-fU_m and 5-fC_m. The aldehyde group in the 5-fU and 5-fC can exist in two different rotameric states. To this end, we have first performed a characterisation of the stationary points in the rotation process. The energy minima were found through a fully relaxed geometry optimisation of two (*syn* and *anti*) rotameric structures constructed for both 5-fU_m and 5-fC_m. The transition states were located via synchronous transit-guided quasi-Newton search[9] in between the two minima. All the found stationary points were verified to be either energy minima or transition states (first order saddle points) *via* an additional vibrational frequency calculation to find out the number (or the absence) of the imaginary force constants. The relative energies and structures for the rotation stationary points, referenced against the minimum energy structures (MES) for 5-fU_m and 5-fC_m correspondingly, are shown in Figure S6.

As can be seen from the characterised structures and energies, the conformations of MESs differ for 5-fU_m and 5-fC_m. The former prefers the *anti* rotameric state (11.58 kcal/mol rotation barrier) for the aldehyde group, where the oxygen is far from the 4-O of uracil, while the possibility of a hydrogen bond in between the cytosine 4-NH_2 and the aldehyde oxygen favours the *syn* conformation in 5-fC_m (rotation barrier 15.43 kcal/mol). The rotation barriers and the *syn*/*anti* energy differences are substantial enough for the molecules to prevalently occupy their minimum potential energy states, which are also the states observed in the existing X-ray structures for 5-fU and 5-fC containing small molecules (Cambridge Structure Database accession IDs AFURID,

AYUNOY, YAGSAC, RAKLOG), and for 5-fC bearing nucleic acids (Protein Data Bank IDs 1VE8, 4R2D, 4R2Q).

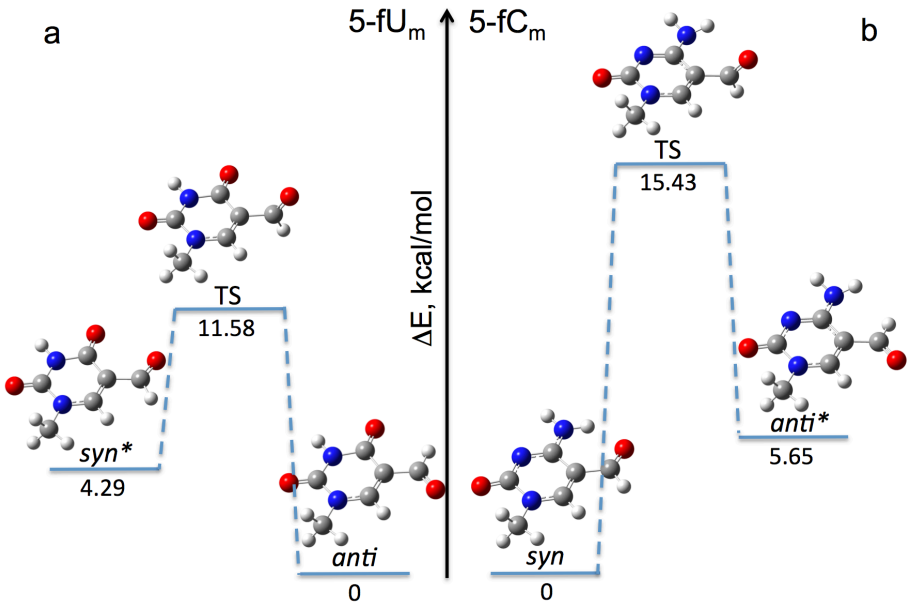


Figure S6. The energy minima (syn and anti) and the transition states (TS) along the aldehyde group rotation pathway in a) 5-fU$_m$ and b) 5-fC$_m$. For both model molecules, the higher energy conformations are additionally marked with asterisks.

In general, to avoid steric hindrances, the products of the reaction with the formylated nucleobases would prefer the *anti* conformation of the C=N group. While the *anti* arrangement of the prior C=O is the lowest energy native rotameric state in 5-fU, that state requires a crossing of a rotation barrier to be reached by 5-fC, starting from its preferred *syn* conformation. The noted difference between the 5-fU and 5-fC conformational preference can be among the facilitating factors in the observed high reactivity of 5-fU with the proposed tagging reagents.

The partial charges at the aldehyde carbons of 5-fU$_m$ and 5-fC$_m$. Next, we have calculated the atomic charges in the studied model molecules (Figure S7), to see whether the variation of the partial positive charge at the aldehyde carbonyl in 5-fU$_m$ and 5-fC$_m$ can contribute to the

difference in reactivity. Besides the conventional Mulliken charges,[10] the more realistic electrostatic potential fitted charges calculated via the Merz-Singh-Kollman (MSK) scheme[11,12] were considered.

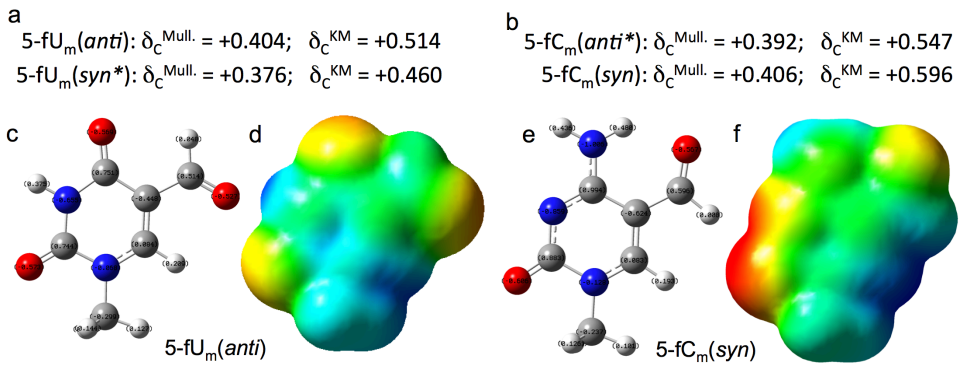


Figure S7. The Mulliken and Merz-Singh-Kolmann (MSK) partial charges at the aldehyde carbon calculated for both syn and anti conformers of a) 5-fU$_m$ and b) 5-fC$_m$. The minimum energy conformers along with the MSK charges are shown in c) and e) for 5-fU$_m$ and 5-fC$_m$ respectively, with the charge distribution (colour pallet from red to blue for -0.08 to +0.08 charge range) onto the electron density surface (iso-value 0.001) plotted in d) and f).

The calculated charges vary in different conformers, with the tendencies also inverting while comparing 5-fU$_m$ and 5-fC$_m$ using the Mulliken versus MSK charges. To this end, the results are inconclusive. However, interestingly, the minimum energy conformers (*anti* for 5-fU$_m$ and *syn* for 5-fC$_m$) attribute greater partial positive charges to the aldehyde carbon.

Natural bond orbital analysis of 5-fU$_m$ and 5-fC$_m$. The formation of a transient C-N bond is among the key stages in our reactions (see Figure S5b). In this process, the aldehyde carbon in 5-fU or 5-fC should become more tetrahedral, which is expected to occur more easily if the conjugation of the nucleobase ring extends less to the aldehyde group. In order to find out whether such a connection may be behind the core electronic cause of the increased reactivity of 5-fU as compared to 5-fC, we have performed a natural bond orbital (NBO) analysis[13,14] of the 5-

fU_m and 5-fC_m models, localising and focusing our attention to the properties of the C_{ring}-$C_{aldehyde}$ bonding orbital. The calculated orbital energies are listed below (Table S9).

Table S9: Orbital energies (in atomic units) of the C_{ring}-$C_{aldehyde}$ bonding orbitals in the studied model molecules, as calculated via NBO analysis.

Model Molecule	Ground State	Orbital Energy (a. u.)
5-fC_m *anti*	no	-0.91041
5-fC_m *syn*	yes	-0.92741
5-fU_m *anti*	yes	-0.89813
5-fU_m *syn*	no	-0.90472

The data show, that regardless the conformational state of the nucleobases, 5-fU features a weaker C_{ring}-$C_{aldehyde}$ bond (the bonding orbital energies are higher), owing to which the aldehyde carbon can gain its pyramidality and form the transient C-N bond much easier as compared to 5-fC. If we consider only the *anti* conformations, the C_{ring}-$C_{aldehyde}$ bonding orbital energy in 5-fC_m is more stable by 7.71 kcal/mol, then that in 5-fU_m (1 a.u. = 627.509 kcal/mol). For the ground state conformation (*syn* for 5-fC_m and *anti* for 5-fU_m), the stability difference is 18.37 kcal/mol (Table S9). Such difference can reflect the core electronic reason for the increased reactivity of 5-fU over 5-fC. The same trend is observed while we consider the effects of the ring on the aldehyde C=O bonding orbital (Table S10). However, the influence is less pronounced owing to the relatively increased distance of the C=O moiety from the ring and the much lower basal energy of the C=O bonding orbital. For the ground state conformations, the C=O bonding orbital energy in 5-fC_m is more stable by 1.71 kcal/mol compared to that in 5-fU_m (Table S10).

Table S10: Orbital energies (in atomic units) of the C=O bonding orbitals in the aldehyde moieties of the studied model molecules, as calculated via NBO analysis.

Model Molecule	Ground State	Orbital Energy (a. u.)
5-fC$_m$ *anti*	no	-1.36378
5-fC$_m$ *syn*	yes	-1.36138
5-fU$_m$ *anti*	yes	-1.35866
5-fU$_m$ *syn*	no	-1.36961

Intermediates along the pathway of the hemiaminal formation. We have attempted to characterise the stationary points in the energy landscape of the hemiaminal formation reaction of 5-fU$_m$ and 5-fC$_m$ with methylamine (Figure S5a), where the first stage (Figure S5b) is expected to be the rate limiting one defining the reactivity of 5-fU$_m$ and 5-fC$_m$ with our tagging reagents. The pathways were initially explored via four techniques; a) we constructed multiple initial geometries (with methylamine and both *anti* and *syn* conformers of the modified bases) that would be close to the expected intermediate state (the product of the scheme in Figure S5b) and performed a geometry optimisation to a transition state; b) we constructed two geometries in two directions from the expected transition states, and performed a synchronous transit-guided quasi-Newton search[9] in between the two structures; c) and d) we performed the above mentioned a and b techniques but without the constraints on the number of imaginary force constants, hence enabling a convergence to any stationary point.

All the above attempts, while using only the modified base and methylamine as reactants, did not locate a transition state, or a local minimum differing from just a set of initial molecules with no or little interaction. This might be an indication of the necessity of having water molecules in the system to stabilise the transition/intermediate states and facilitate the hydrogen transfers via a 6-membered transient ring formation.[15]

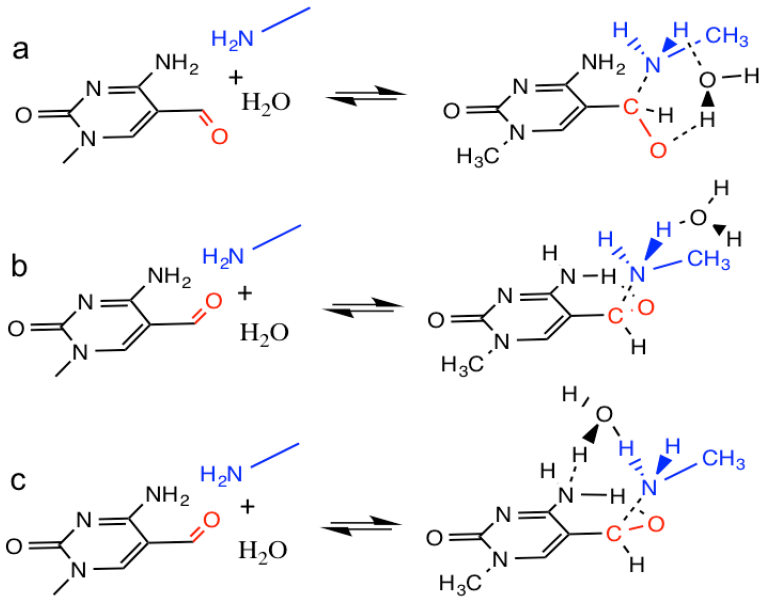


Figure S8: The considered transient initial interactions involving anti and syn variants of 5-fC_m, along the pathway of the hemiaminal formation.

We have thus considered a system, with one additional water molecule, placed in the positions to facilitate the hydrogen transfers in the hemiaminal formation pathway. The schematic representation of the transient reactions considered for 5-fC_m in its *anti* and most stable *syn* conformations are shown in Figure S8. We have applied the above-described a-d techniques, which unfortunately were only breaking the system into the individual molecules with little interaction in between. In the reference [15], the Schiff base formation reaction was studied on different model molecules via DFT, and the whole water cage was necessary to stabilise the transition states.

However, while considering similar interactions involving 5-fU_m, we quickly located a stationary point upon the interaction of its minimum-energy *anti* conformer with methylamine and just one water molecule (Figure S9a). The further vibrational analysis verified the stationary point to be

an energy minimum which appeared to be more stable (ΔE=-13.79 kcal/mol) than the system of non-interacting 5-fU$_m$, water and methylamine molecules.

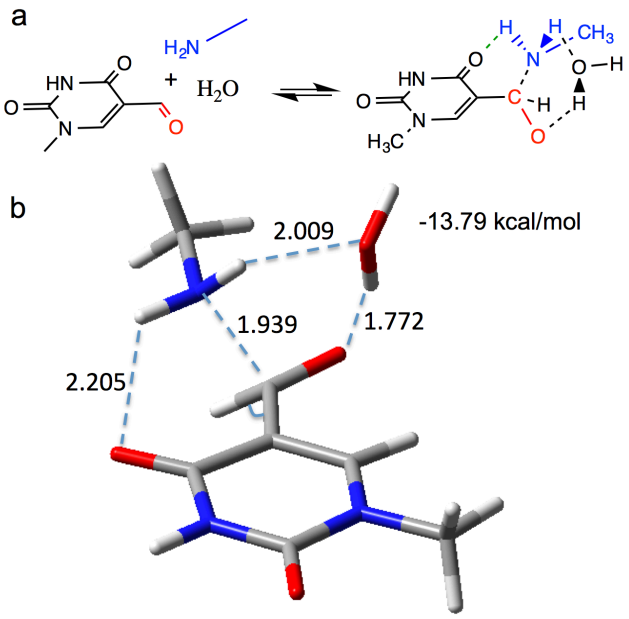


Figure S9. The found intermediate state (ΔE=-13.79 kcal/mol) along the pathway of hemiaminal formation with 5-fU$_m$ (a). The structure is stabilised via the extra hydrogen bond in between the amino group and the 4-O of 5-fU$_m$. The aldehyde carbon has partially gained its tetrahedricity (b) upon the C-N bond formation, and rotated a bit to facilitate the formation of the above mentioned hydrogen bond. All the outlined distances are in Å.

The found intermediate structure is shown in Figure S9b, where the spatial approach of methylamine to 5-fU$_m$ is further stabilised by the hydrogen bond in between the methylamine -NH$_2$ and 4-O of 5-fU$_m$. The 6-membered hydrogen transfer transient ring in also formed, with the carbon of the aldehyde group partially gaining its pyramidality upon the C-N bond formation. This stable intermediate can surely be one of the factors behind the high reactivity of 5-fU, as compared to 5-fC with the tagging reagents used in this study.

The stationary structures reported in this work. All the calculations were done at the RMP2/cc-pVDZ level of theory as described above. All the stationary points, including the *IS* intermediate state, had 0 imaginary frequencies, except the *TS* transition states that had 1. The absolute MP2 energies in atomic units (a.u., the same Hartrees) are written at the top of the Cartesian coordinates.

5-fC$_m$ *anti* (E_{MP2}=-546.0437708 a.u.)

N	0.156301	0.692521	-0.178386
C	0.056302	-0.654981	0.100300
C	1.340445	1.257489	-0.331829
N	1.292050	-1.387652	0.266573
O	4.977894	0.473128	-0.313797
O	-0.989752	-1.271605	0.235124
N	1.343749	2.594412	-0.668421
C	2.592211	0.542498	-0.222060
C	2.486315	-0.797208	0.098362
C	3.936190	1.102367	-0.463047
C	1.162795	-2.806654	0.594188
H	2.114805	3.152180	-0.315727
H	0.430582	3.023728	-0.532417
H	3.386252	-1.411206	0.210241
H	3.961310	2.161251	-0.819266
H	2.169168	-3.237561	0.690176
H	0.598720	-3.319396	-0.198348
H	0.605991	-2.918966	1.535710

5-fC$_m$ *syn* (E_{MP2}=-546.0527682 a.u.)

N	0.148429	0.688404	-0.140450
C	0.054513	-0.651018	0.142920
C	1.336110	1.256250	-0.301761
N	1.296086	-1.404225	0.255773
O	4.147807	2.312094	-0.622635
O	-0.983308	-1.274370	0.310292
N	1.384598	2.574342	-0.578946
C	2.595548	0.533539	-0.195616
C	2.489868	-0.814705	0.089540
C	3.918402	1.128568	-0.367859
C	1.161146	-2.827462	0.557708
H	2.290457	3.017935	-0.703293
H	0.507666	3.076411	-0.651171
H	3.379465	-1.449019	0.190106
H	4.763389	0.403534	-0.247735
H	2.164881	-3.272647	0.614658
H	0.569091	-3.319215	-0.227691
H	0.631590	-2.955874	1.512846

5-fC$_m$ *TS* (E_{MP2}=-546.0281768 a.u.)

N	0.516449	1.532433	-0.043732
C	1.611631	0.684612	-0.037833
C	-0.701486	1.037668	0.050310
N	1.344634	-0.722427	0.006554
O	-3.145156	-1.184547	-0.548061
O	2.777759	1.051480	-0.090223
N	-1.757555	1.929972	0.096090
C	-0.999441	-0.370128	0.167152
C	0.083810	-1.212766	0.122429
C	-2.401819	-0.866070	0.367982
C	2.512436	-1.598949	-0.039110
H	-2.604942	1.621482	-0.375285
H	-1.470645	2.868860	-0.176302
H	-0.014375	-2.301873	0.186201
H	-2.752423	-0.907202	1.428968
H	2.171787	-2.643606	-0.005695
H	3.174224	-1.385223	0.813331
H	3.080074	-1.413946	-0.962918

5-fU$_m$ *anti* (E_{MP2}=-565.9111727 a.u.)

N	0.114526	0.627170	-0.181797
C	0.035133	-0.702540	0.201330
C	1.258319	1.413211	-0.436237
N	1.289834	-1.327846	0.349872
O	4.889650	0.724565	-0.328581
O	-1.009498	-1.300646	0.398619
O	1.154640	2.586114	-0.770694
C	2.503129	0.663632	-0.251869
C	2.451658	-0.651653	0.126961
C	3.820693	1.307339	-0.469367
C	1.265975	-2.734445	0.754003
H	-0.784996	1.094813	-0.294047
H	3.383777	-1.209826	0.264108
H	3.765253	2.375909	-0.774905
H	2.303312	-3.087052	0.828562
H	0.717467	-3.331224	0.010651
H	0.761449	-2.839359	1.725635

5-fU$_m$ *syn* (E_{MP2}=-565.9043332 a.u.)

N	0.094960	0.621352	-0.179900
C	0.029031	-0.706321	0.202583
C	1.225427	1.435534	-0.441950
N	1.291864	-1.315921	0.346224
O	4.046656	2.454974	-0.804397
O	-1.006140	-1.319623	0.404189
O	1.082290	2.599806	-0.772894
C	2.487902	0.697371	-0.261242
C	2.442242	-0.620803	0.118255
C	3.821687	1.305127	-0.468707
C	1.287430	-2.722379	0.750004
H	-0.809677	1.079893	-0.289263
H	3.370174	-1.187809	0.258041

H	4.657732	0.576458	-0.280156
H	2.328854	-3.064003	0.821869
H	0.744348	-3.326544	0.008554
H	0.787441	-2.834459	1.723227

5-fU$_m$ *TS* (E_{MP2}=-565.892721 a.u.)

N	0.517610	1.444637	-0.029286
C	1.627073	0.610055	-0.071909
C	-0.842039	1.107119	0.096731
N	1.314452	-0.749882	0.044556
O	-3.196265	-1.005821	-0.654783
O	2.775177	1.007871	-0.195165
O	-1.714960	1.965201	0.122188
C	-1.057355	-0.340041	0.198631
C	0.014685	-1.183948	0.174839
C	-2.472458	-0.845230	0.311413
C	2.444520	-1.677002	-0.000047
H	0.726976	2.439572	-0.109693
H	-0.105716	-2.269602	0.252316
H	-2.830047	-1.062306	1.347834
H	2.057880	-2.698990	0.112478
H	3.151267	-1.452817	0.812579
H	2.977030	-1.582858	-0.958272

5-fU$_m^{anti}$•CH_3NH_2•H_2O *IS* (E_{MP2}=-737.7093785 a.u.)

N	1.684193	-1.517560	0.191369
C	2.591968	-0.463932	0.225728
C	0.320294	-1.495043	-0.135480
N	2.039679	0.762930	-0.153132
O	-1.969865	1.230686	-1.234650
O	3.762767	-0.588000	0.553717
O	-0.360112	-2.520370	-0.053180
C	-0.151139	-0.173899	-0.536794
C	0.723551	0.874412	-0.526246
C	-1.558487	0.059075	-1.028713
C	2.944208	1.911707	-0.139565
C	-3.913740	-0.520402	0.288043
N	-2.477032	-0.610029	0.542085
H	2.063787	-2.418596	0.481069
H	0.380664	1.867375	-0.830095
H	-1.929650	-0.784247	-1.656707
H	2.381445	2.794060	-0.472147
H	3.796118	1.732006	-0.812162
H	3.332521	2.076863	0.876511
H	-4.210004	-1.312936	-0.416337
H	-4.103169	0.456800	-0.184696
H	-4.519986	-0.611593	1.205052
H	-2.153945	-1.550687	0.785598
H	-2.187451	0.087278	1.242846
O	-1.801830	2.055727	1.350333
H	-1.838773	1.975148	0.367875
H	-2.527045	2.665285	1.539342

Section 6: NMR, LC-MS and HPLC-MS spectra

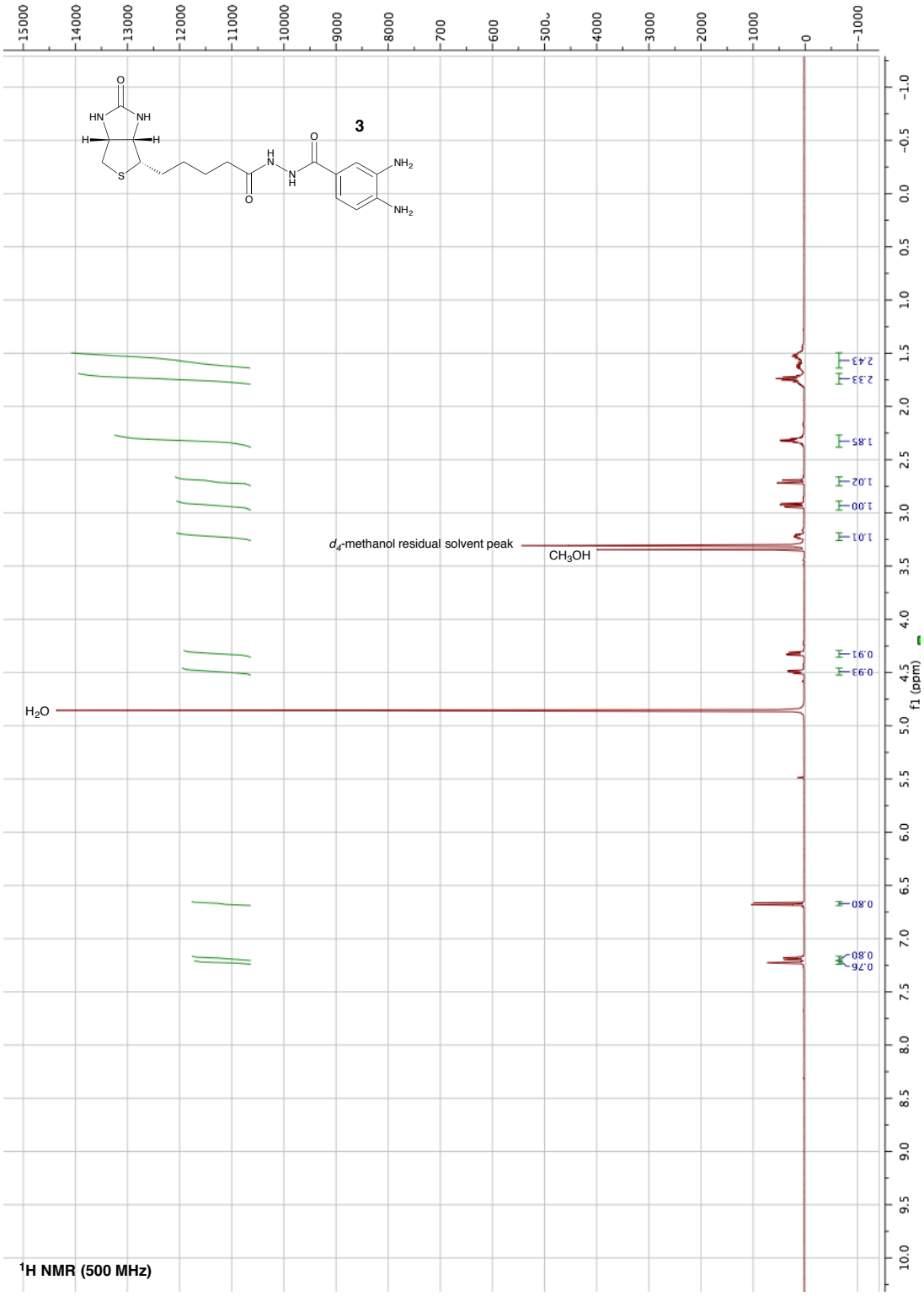


Figure S10: ^{1}H NMR spectra for **2** in d_4-methanol and labeled solvent residual peaks.[16]

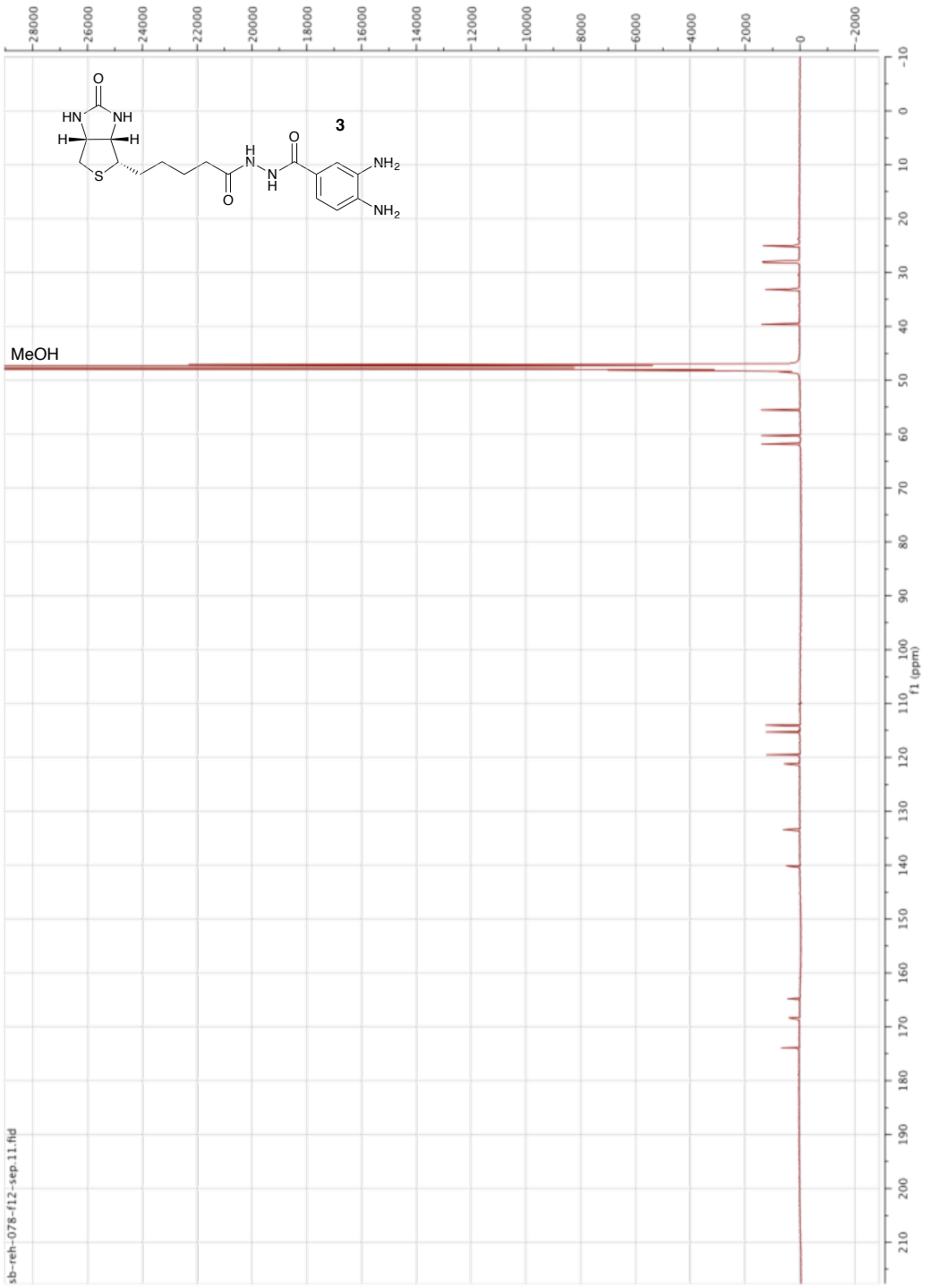


Figure S11: ^{13}C NMR spectra for Probe **2** in d_4-methanol.

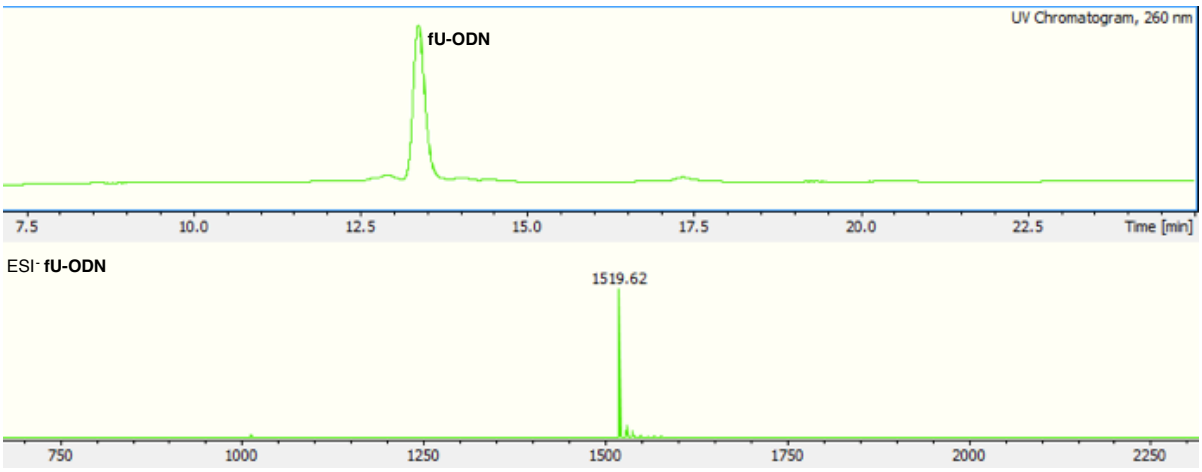


Figure S12: LCMS profile for fU-ODN.

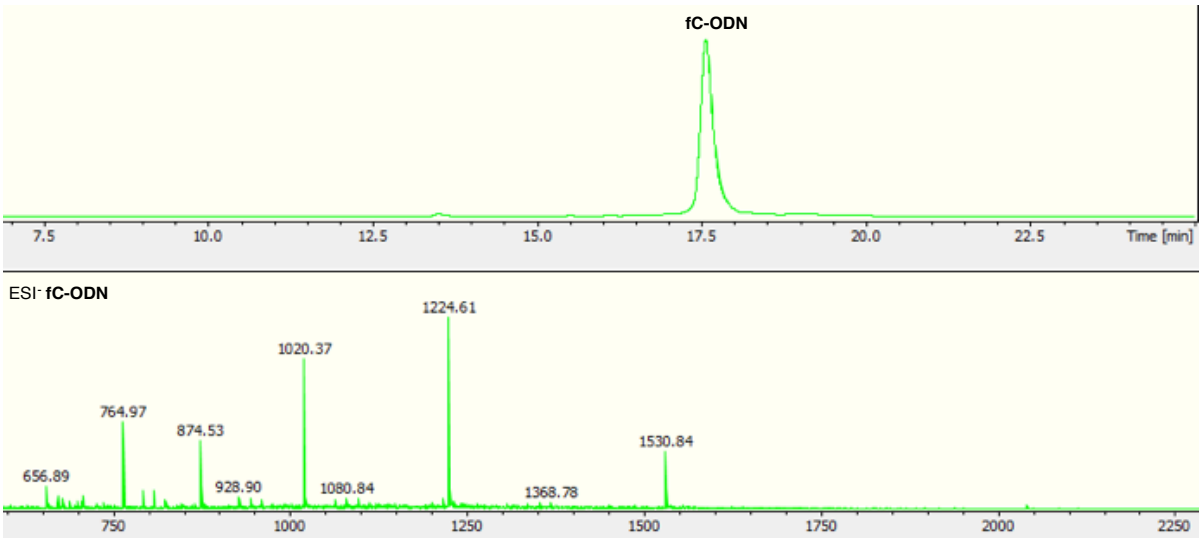


Figure S13: LC-MS profile for fC-ODN.

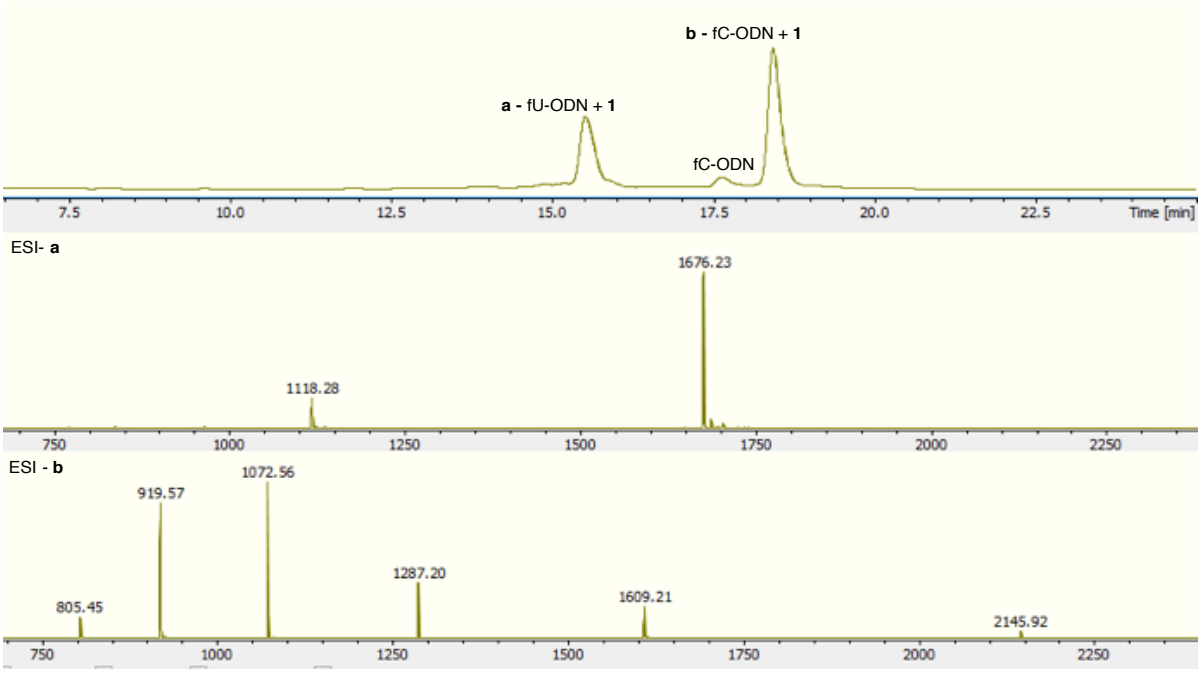


Figure S14: LC-MS profile for reaction of fU-ODN and fC-ODN with **1** at pH = 5, 24 h, in the presence of *p*-anisidine (Table 1, Entry 1, main text).

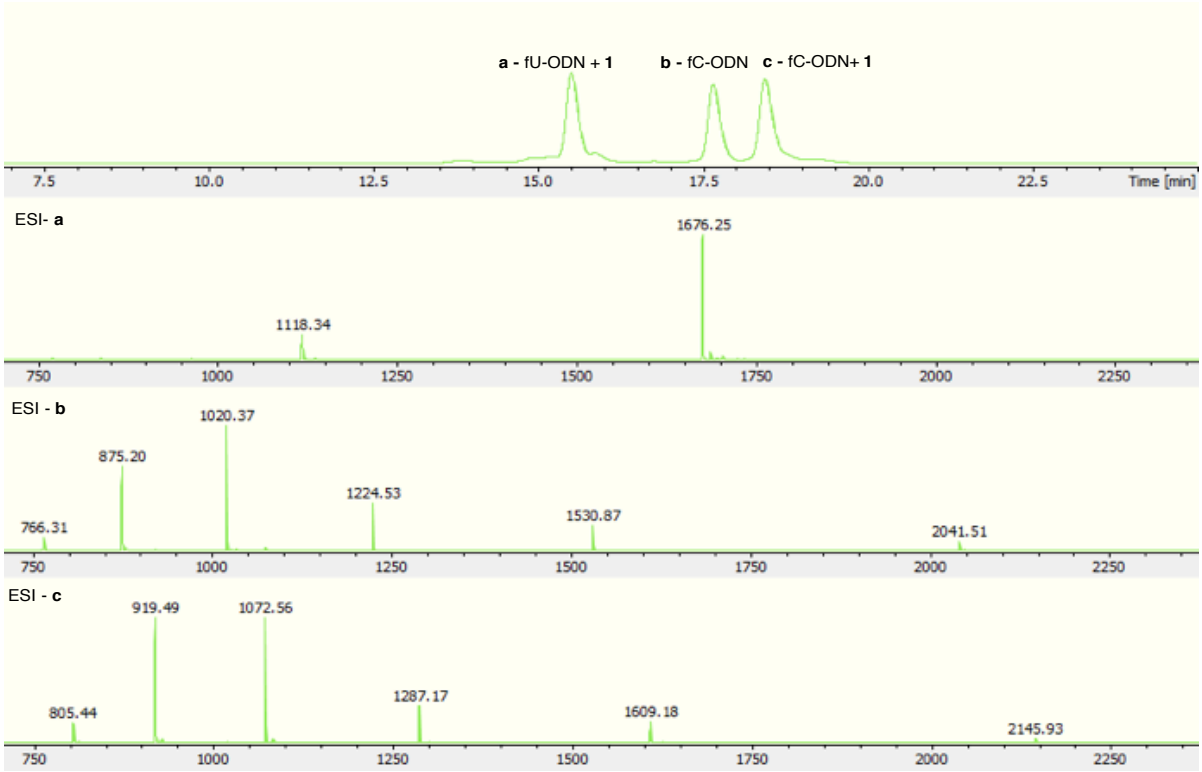


Figure S15: LC-MS profile for reaction of fU-ODN and fC-ODN with **1** at pH = 5, 24 h (Table 1, Entry 2, main text).

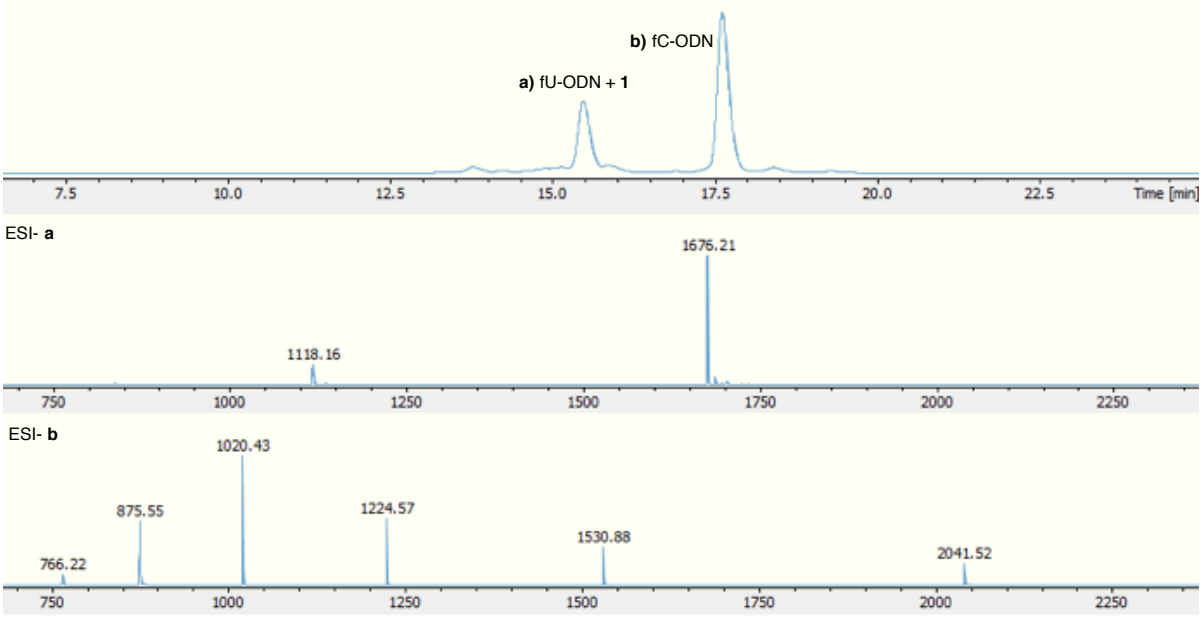


Figure S16: LC-MS profile for reaction of fU-ODN and fC-ODN with **1** at pH = 6, 24 h (Table 1- Entry 3, main text).

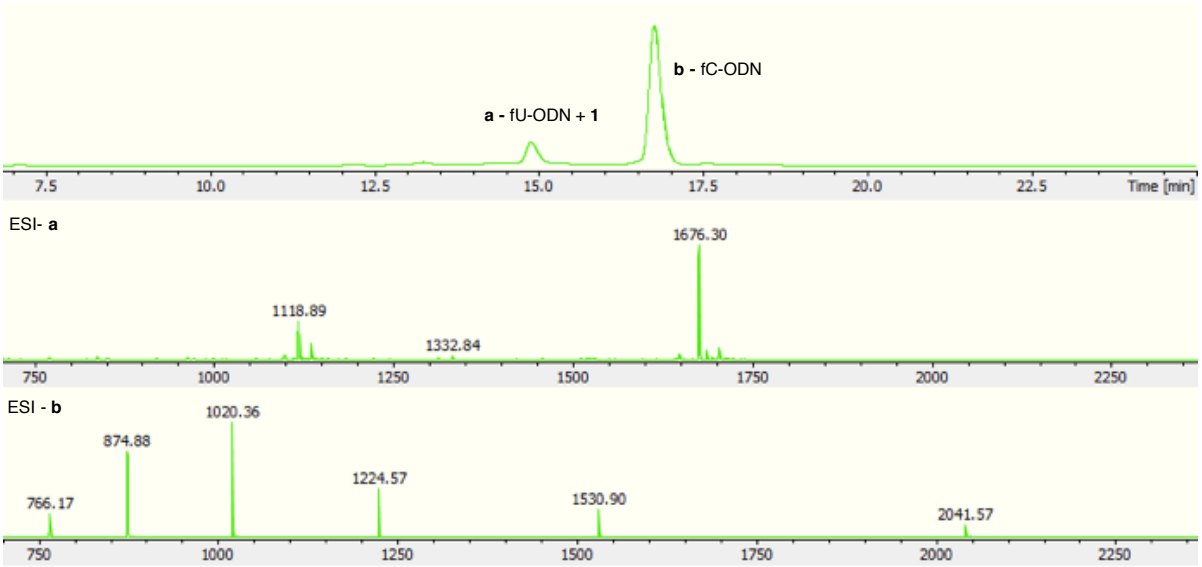


Figure S17: LC-MS profile for reaction of fU-ODN and fC-ODN with **1** at pH = 6, 4 h (Table 1 – Entry 4, main text).

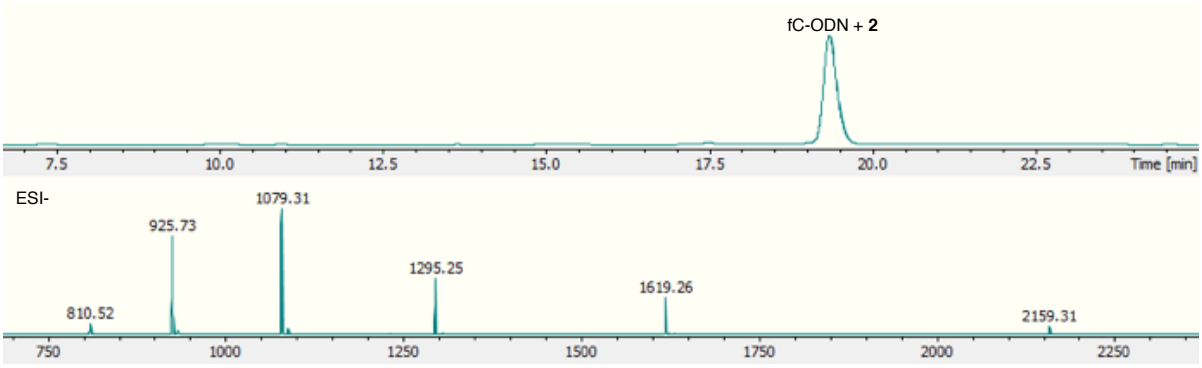


Figure S18: LC-MS profile for reaction of fC-ODN with **2** at pH = 5, 24 h, in the presence of *p*-anisidine (Table 1, entry 5, main text).

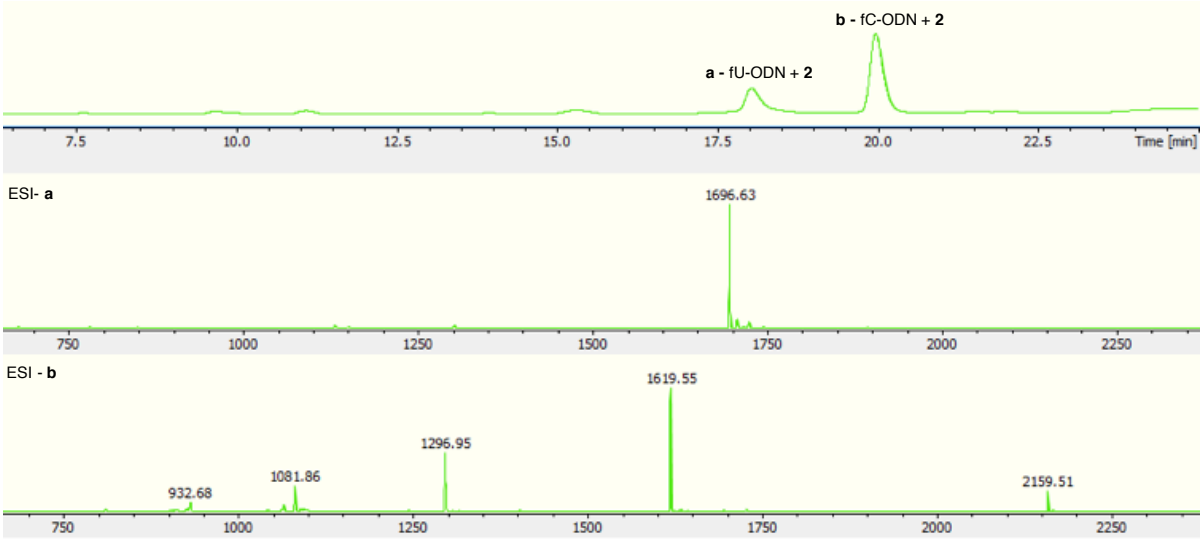


Figure S19: LC-MS profile for reaction of fU-ODN and fC-ODN with **2** at pH = 5, 24 h, in the presence of *p*-anisidine. (Table 1, entry 5, main text)

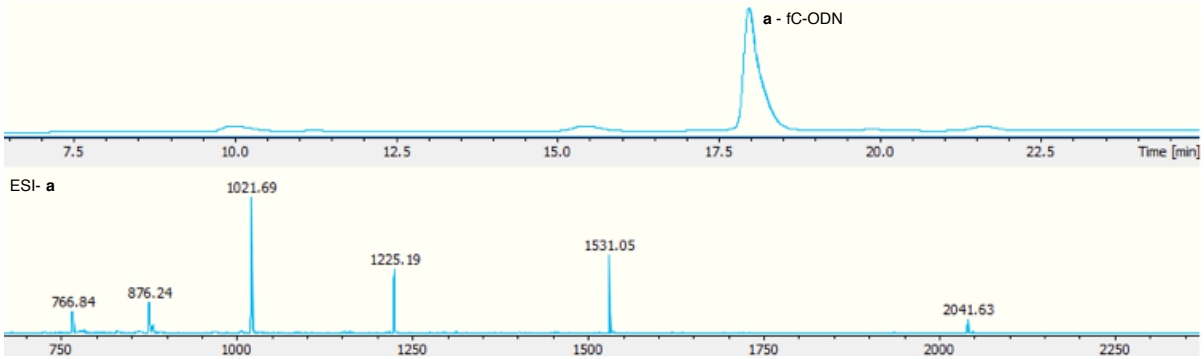


Figure S20: LC-MS profile for reaction of fC-ODN with **2** at pH = 7, 24 h. (Table 1, entry 6, main text)

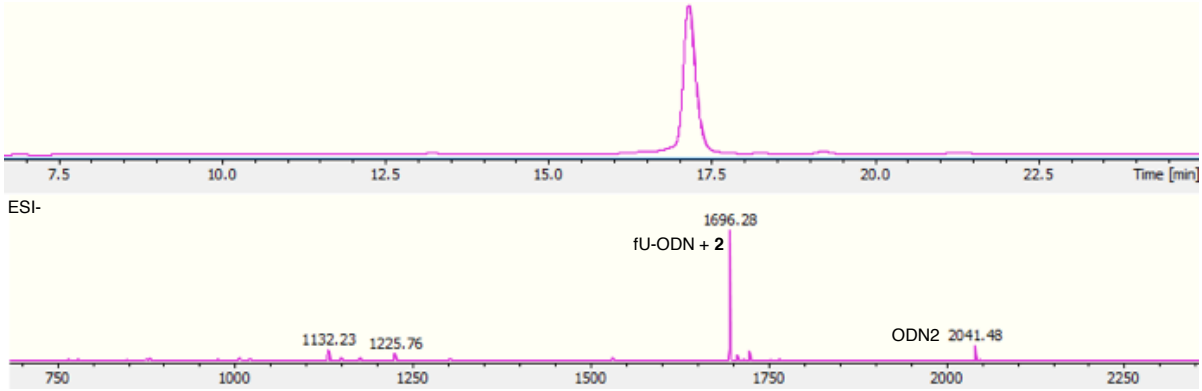


Figure S21: LC-MS profile for reaction of fU-ODN and fC-ODN with **2** at pH = 7, 24 h. Retention times for fC-ODN and fU-ODN + **2** co-elute. Their respective masses can be observed in the ESI-spectrum. (Table 1, entry 6, main text)

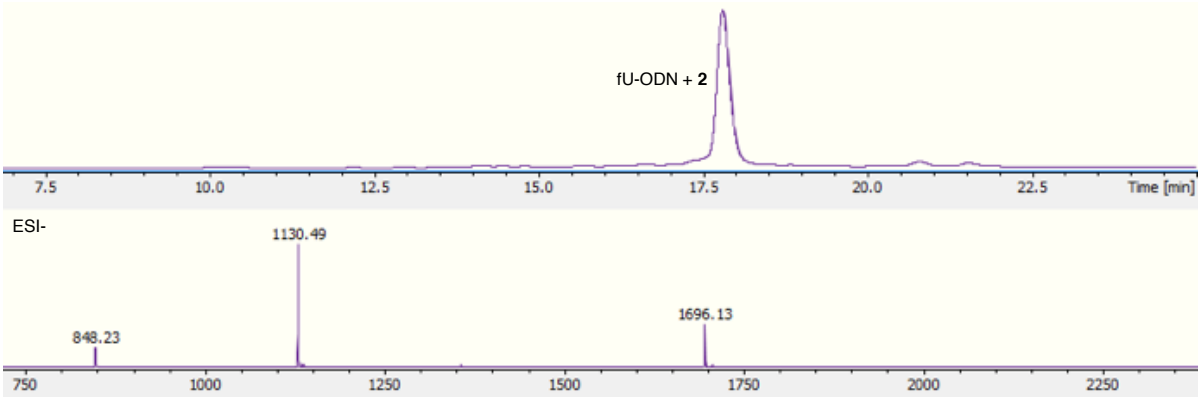


Figure S22: LC-MS profile for reaction of fU-ODN with **2** at pH = 7, 4 h. (Table 1, entry 7, main text)

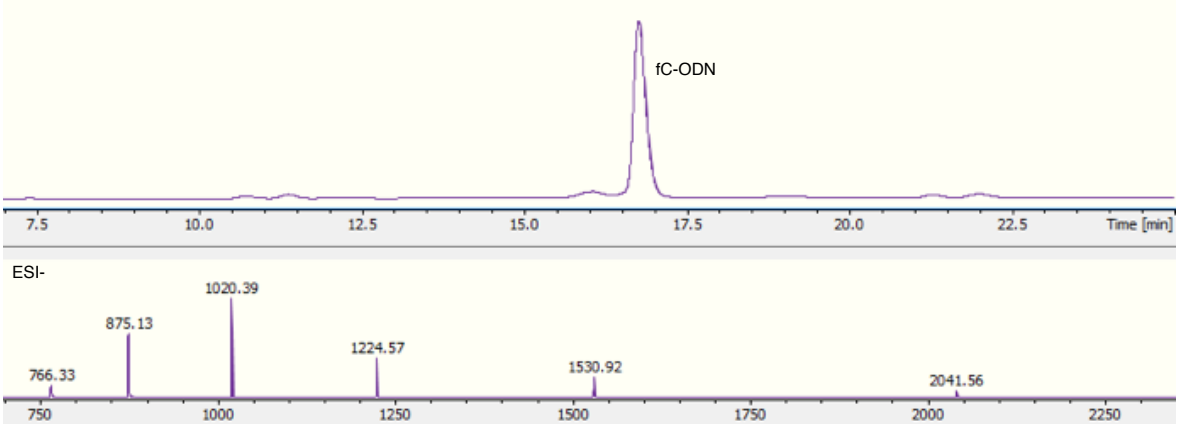


Figure S23: LC-MS profile for reaction of fC-ODN with **2** at pH = 7, 4 h (Table 1, entry 7, main text).

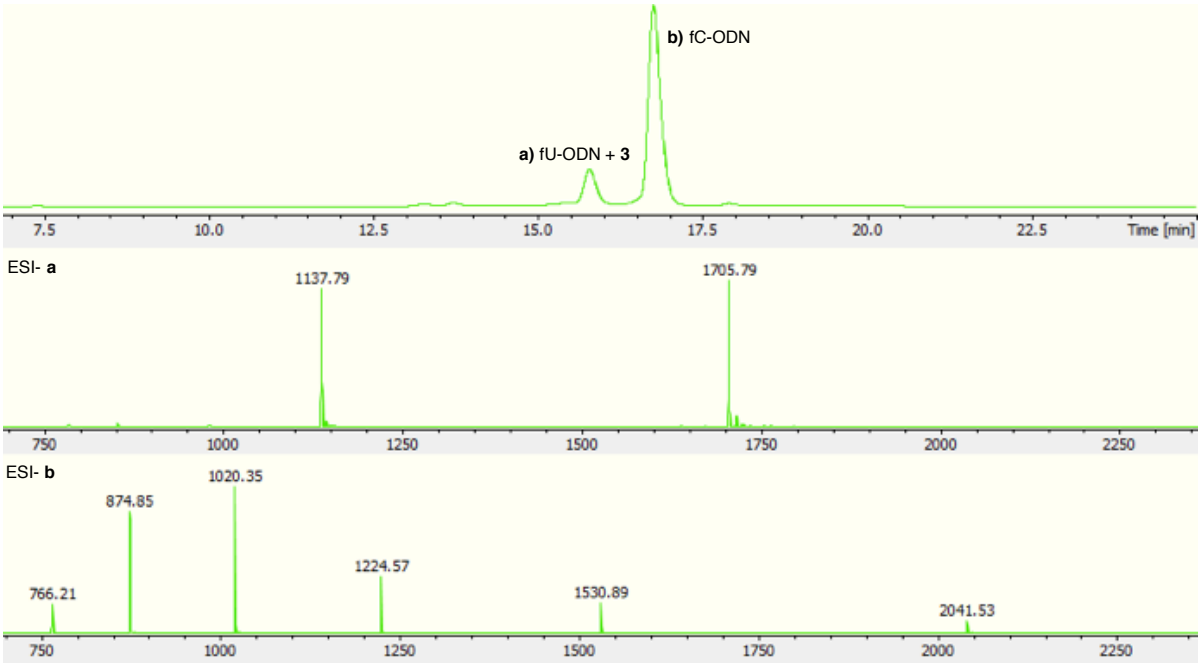


Figure S24: LC-MS profile for reaction of fU-ODN and fC-ODN with **3** at pH = 7, 4 h (Table 1 – Entry 8, main text).

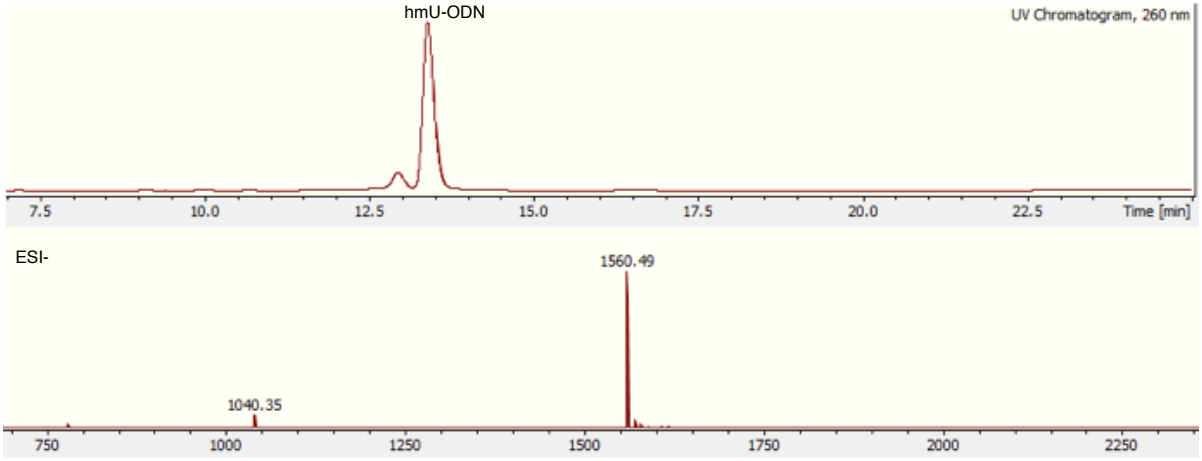


Figure S25: LC-MS profile for hmU-ODN.

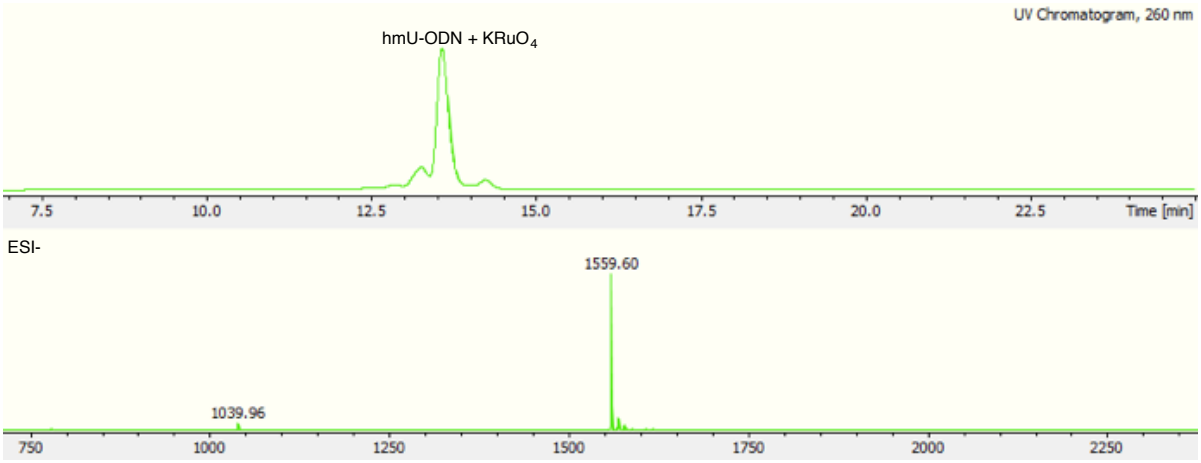


Figure S26: LC-MS profile for oxidation of hmU-ODN with $KRuO_4$.

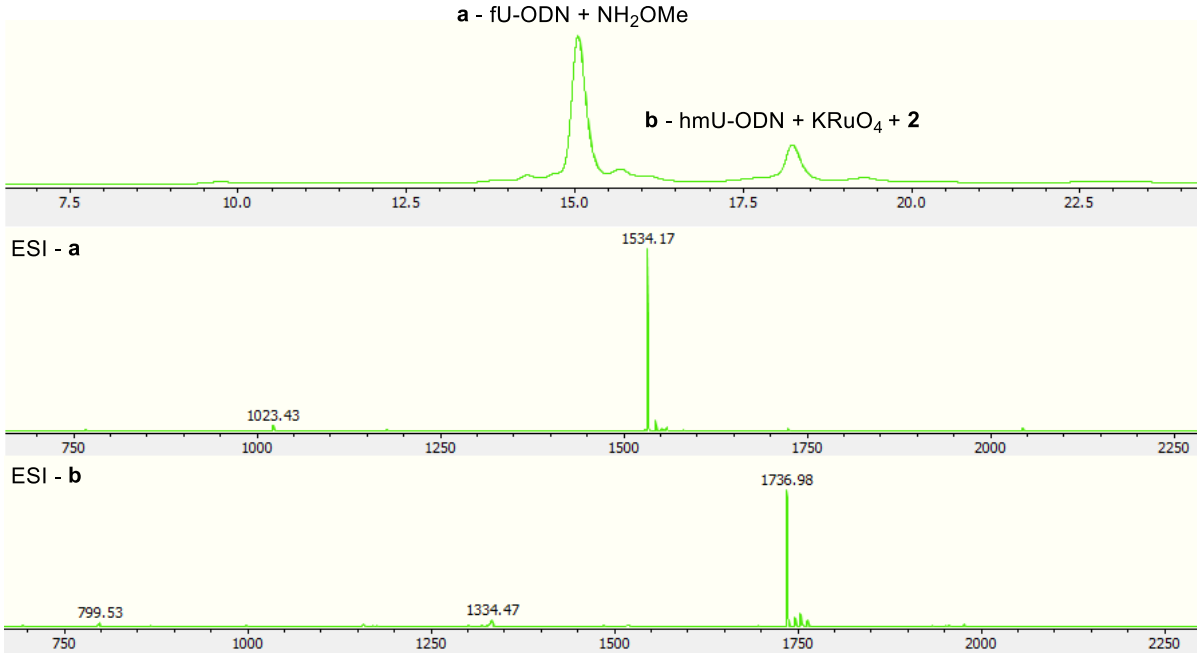


Figure S27: LC-MS profile for reaction of fU-ODN and hmu-ODN with NH_2OMe (pH = 6, 3 h), followed by oxidation with $KRuO_4$ and reaction with **2** (pH = 7, 4 h). No peak corresponds to fU-ODN + **2**, indicating that 5-hmU can be tagged selectively in the presence of 5-fU.

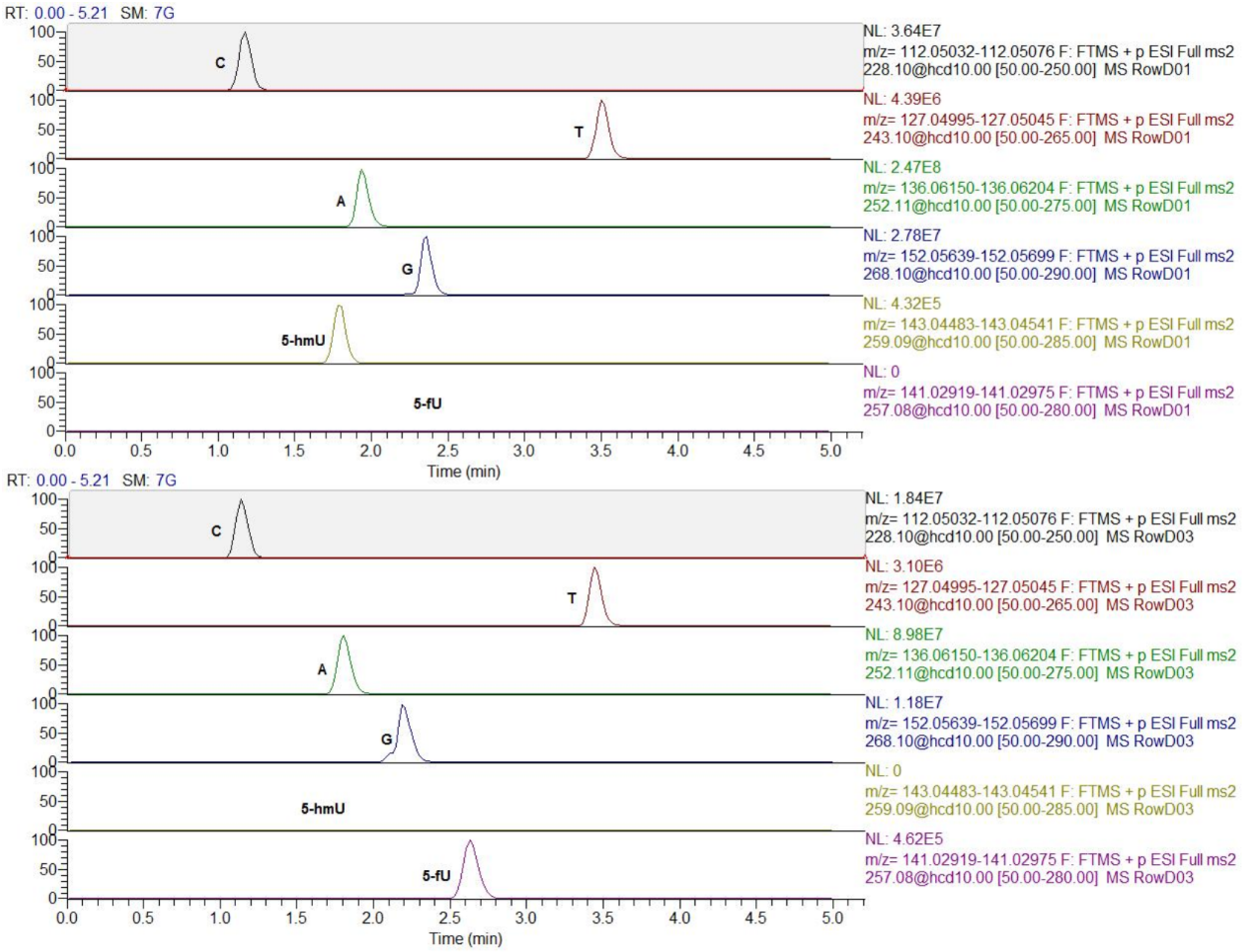


Figure S28: HPLC-MS extracted $[M+H]^+$ ion count for C, T, A, G, 5-hmU and 5-fU deoxynucleosides after digestion of a) hmU-ODN, b) hmU-ODN after treatment with $KRuO_4$.

Section 7: References

(1) Booth, M. J.; Branco, M. R.; Ficz, G.; Oxley, D.; Krueger, F.; Reik, W.; Balasubramanian, S. *Science* **2012**, *336*, 934.
(2) Guo, P.; Xu, X.; Qiu, X.; Zhou, Y.; Yan, S.; Wang, C.; Lu, C.; Ma, W.; Weng, X.; Zhang, X.; Zhou, X. *Organic & Biomolecular Chemistry* **2013**, *11*, 1610.
(3) Raiber, E.-A.; Beraldi, D.; Ficz, G.; Burgess, H. E.; Branco, M. R.; Murat, P.; Oxley, D.; Booth, M. J.; Reik, W.; Balasubramanian, S. *Genome Biology* **2012**, *13*, R69.
(4) Møller, C.; Plesset, M. S. *Physical Review* **1934**, *46*, 618.
(5) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chemical Physics Letters* **1988**, *153*, 503.
(6) Dunning, T. H. *The Journal of Chemical Physics* **1989**, *90*, 1007.
(7) Peterson, K. A.; Woon, D. E.; Dunning, T. H. *The Journal of Chemical Physics* **1994**, *100*, 7410.
(8) Gaussian 03, Revision E 01, M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G; M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
(9) Peng, C.; Bernhard Schlegel, H. *Israel Journal of Chemistry* **1993**, *33*, 449.
(10) Mulliken, R. S. *The Journal of Chemical Physics* **1955**, *23*, 1833.
(11) Singh, U. C.; Kollman, P. A. *Journal of Computational Chemistry* **1984**, *5*, 129.
(12) Besler, B. H.; Merz, K. M.; Kollman, P. A. *Journal of Computational Chemistry* **1990**, *11*, 431.
(13) Reed, A. E.; Weinhold, F. *The Journal of Chemical Physics* **1983**, *78*, 4066.
(14) Foster, J. P.; Weinhold, F. *Journal of the American Chemical Society* **1980**, *102*, 7211.
(15) Solís-Calero, C.; Ortega-Castro, J.; Hernández-Laguna, A.; Muñoz, F. *Theor Chem Acc* **2012**, *131*, 1.
(16) Gottlieb, H. E.; Kotlyar, V.; Nudelham, A. *J. Org. Chem*. 1997, *62*, 7512-7515