Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Fluorogenic Labeling of 5-Formylpyrimidine Nucleotides in DNA and RNA

*Biswajit Samanta, Jan Seikowski, and Claudia Höbartner**

anie_201508893_sm_miscellaneous_information.pdf

Table of Contents

Index of Supporting Tables

Index of Supporting Figures

1. General Information

1.1. Materials

All standard chemicals and solvents were purchased from commercial suppliers and used without further purification. 5-Formyldeoxyuridine was obtained from AkPharm, 1-Ethyl-2,3,3-trimethylindolenium-5-sulfonate (**1**) was from Toronto Research Chemicals, 2,3,3-Trimethyl-3H-benz(e)indol-6-sulfonic acid (precursor to **2**) and 1-(4-Sulfobutyl)-2,3,3-trimethylbenz[e]indolium betaine (**3**) were from Organica Feinchemie.

Silica gel plates precoated with fluorescent indicator were used for thin layer chromatography (TLC) and the plates were visualized with UV light. Silica gel 60, 0.032-0.063 mm (230-450 mesh) was used for column chromatography. ^{1}H NMR and ^{13}C NMR were recorded on a 400 or 600 MHz instrument. The chemical shifts (in ppm) are reported downfield from TMS (0 ppm) or $CDCl_3$ (7.26 ppm for ^{1}H and 77.16 ppm for ^{13}C). Spin multiplicities in ^{1}H NMR are reported as singlet (s), doublet (d), triplet (t), triplet of triplet (tt) or multiplet (m).

Unmodified DNA oligonucleotides were purchased from Sigma Aldrich and purified by denaturing PAGE (20 % polyacrylamide). Standard DMT-protected 2'-deoxyribonucleotide 3'-β-cyanoethyl N,N-diisopropylphosphoramidites were from SAFC, 5'-DMT-5-(1,2-diacetoxyethyl)-2'-deoxyuridine 3'-β-cyanoethyl N,N-diisopropylphosphoramidite was from ChemGenes, benzylthiotetrazole (BTT) was from Carbosynth. Standard ribonucleotide triphosphates (NTPs) for in vitro transcription reactions were purchased from Jena Bioscience, 5fCTP, 4fUTP and 5fdUTP were from Trilink. γ-^{32}P-ATP (3000 Ci/mmol) was obtained from Perkin Elmer. T4 DNA ligase, T4 polynucleotide kinase (PNK) and Calf-intestine alkaline phosphatase (CIAP) were purchased from Fermentas. *Taq* DNA polymerase was purchased from New England BioLabs.

1.2. Chromatography

Anion exchange HPLC experiments were performed on an ÄKTApurifier 10 from GE Healthcare using a Dionex DNAPac PA200 (4 x 250 mm) column at 80°C. Eluent A: 25 mM Tris HCl, pH 8.0, 6 M urea; Eluent B: 25 mM Tris HCl, pH 8.0, 0.5 M $NaClO_4$, 6 M urea.

Reversed phase HPLC experiments were performed on a Shimadzu HPLC system with a diode array detector using a Nucleosil 100-5 C18 HD (4 x 250 mm) column at 40 °C. Eluent A: 100 mM TEAA in H_2O; Eluent B: 20 mM TEAA in H_2O:MeCN 1:4.

1.3. Gel electrophoresis

Denaturing polyacrylamide gels (0.4 or 0.7 mm thick, 20x30 cm, 20% acrylamide, 7 M urea, 1x TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3) were run in 1x TBE buffer at 35 W. Bromophenol blue and xylene cyanol were used as dye markers. On preparative gels, oligonucleotides were visualized by UV shadowing. Gels of radioactive samples were exposed to phosphor storage screens and visualized with a Storm 820 PhosphorImager, GE Healthcare.

1.4. Mass spectrometry

The masses of all the oligonucleotides (crude, purified and also after different chemoenzymatic reactions) were analyzed by electrospray ionization (ESI-MS) technique using a Q-Tof Ultima (Waters) instrument. The data collection and processing was performed by Uwe Pleßmann and Monika Raabe in the Bioanalytical Mass Spectrometry lab (Prof. Dr. Henning Urlaub, Max Planck Institute for biophysical Chemistry). Around 100-300 pmol of oligonucleotides were used for each measurement.

1.5. UV-VIS absorbance measurements

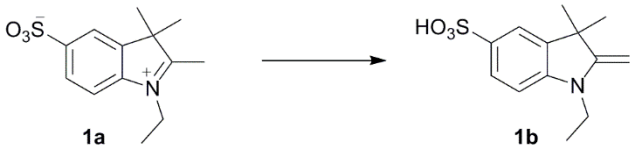
The absorbance measurements were performed in quartz cuvettes (path length 10 mm, 0.5 or 1 ml) in a Cary 100 UV-VIS spectrophotometer (Varian) equipped with a peltier controlled temperature device. Data were recorded at 20°C in absorbance mode with averaging time = 0.1 s, data interval = 1.00 nm, scan rate = 600 nm/min and source change at 380 nm.

1.6. Fluorescence measurements

The fluorescence was measured in Cary Eclipse Fluorescence spectrophotometer (Varian) equipped with a Xenon flash lamp. All the excitation and emission measurements were performed at 20 °C in fluorescence mode with scan rate = medium (600 nm/min), averaging time = 0.1 s, data interval = 1.0 nm, excitation and emission slit widths = 10 nm, PMT = 800 V. Data is presented as an average of three scans. For the kinetic measurements, the in-built kinetic program in the instrument was used. The data was recorded for fixed pair of excitation and emission wavelengths at every 5 min up to 180 min with excitation and emission slit widths 10 nm, PMT 800 V. The measurements were performed in quartz cuvettes (120 µl solution with 10 x 2 mm light path and additionally 1 ml solution with 10 x 4 mm light path for nucleoside).

2. Synthetic procedures

2.1. Synthesis of 1b [1-Ethyl-3,3-dimethyl-2-methylene-5-sulfonyl-indol]



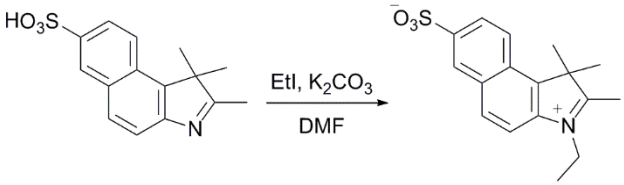
1a 1b

Commercially available **1a** (1-ethyl-2,3,3-trimethyliodoleninium-5-sulfonate) (50 mg, 0.19 mmol) was treated with 1N NaOH (0.8 ml, 4 ml/mmol) for 1.5 h. To the reaction mixture Dowex-H [300 mg; the commercial Dowex-H (50W X8 20-50 mesh, strongly acidic) was washed with MeOH followed by water, each multiple times and air dried] was added and stirred for 0.5 h. The supernatant was taken out, the Dowex-H beads were washed with water and the combined liquid layers were lyophilized to obtain deep brown solid **1b** (47.5 mg).

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.37 (dd, J = 8.0, 1.7 Hz, 1H), 7.34 (dd, J = 1.7, 0.5 Hz, 1H), 6.55 (d, J = 8.1Hz, 1H), **3.90 (dd, J = 16.8, 1.8 Hz, 2H)**, 3.56 (q, J = 7.0 Hz, 2H), 1.26 (s, 6H), 1.06 (t, J = 7.0 Hz, 3H). ^{13}C NMR (100 MHz, DMSO-d_6) δ 160.1, 145.1, 139.0, 136.1, 125.6, 119.7, 103.8, 74.3, 43.5,

35.9, 29.7, 29.6, 10.4. [For In2: ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.02 (d, J = 1.5 Hz, 1H), 7.90 (d, J = 8.4 Hz, 1H), 7.83 (dd, J = 8.3, 1.6 Hz, 1H), 4.47 (q, J = 7.3 Hz, 2H), **2.81 (s, 3H)**, 1.54 (s, 7H), 1.43 (t, J = 7.3 Hz, 3H); the highlighted singlet signal corresponding to exocyclic methyl group in **1a** was converted to a doublet of doublet in **1b**, which confirmed formation of the exocyclic methylene group.

2.2. Synthesis of 2a [1-Ethyl-2,3,3-trimethyl-3H-benzo(e)indol-6-sulfonic acid]



Compound **2a** was synthesized following reported procedures with slight modifications.[1] Commercially available 2,3,3-Trimethyl-3H-benzo(e)indol-6-sulfonic acid (6-sulfo-TMBI, 150 mg, 0.52 mmol) and ethyl iodide (833 µl, 10.4 mmol) were heated in DMF (1 ml, 1.9 ml/mmol) in presence of K_2CO_3 (108 mg, 0.8 mmol) at 80°C for 12 h. The product **2a** was purified by Lobar 310-25 (LiChroprep, RP-18, 40-63 µM, Merck column) with water-ACN in ÄKTA prime. The product was lyophilized to obtain off-white amorphous solid (95 mg, 58%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.39 (d, J = 1.7 Hz, 1 H), 8.32 (m, 2 H), 8.13 (d, J = 8.9 Hz , 1 H), 7.94 (dd, J = 8.7 Hz, 1.7 Hz, 1 H), 4.6 (q, J = 7.3 Hz 6 H), 2.91 (s, 3 H), 1.75 (s, 6 H), 1.50 (t, J = 7.3 Hz, 3 H); ^{13}C NMR (101 MHz, $CdCl_3$) δ 195.9, 146.7, 138.6, 136.9, 132.5, 131.3, 127.00, 126.4, 125.6, 123.1, 113.4, 55.4, 43.3, 21.4, 13.4, 12.8; ESI-MS calcd. for $C_{17}H_{19}NO_3S$ 317.40, found 316.3.

2.3. Synthesis of Hcy1U nucleoside [5-[2-(1-Ethyl-3,3-dimethyl-5-sulfonyl-indolenin-2-yl)-ethenyl]-2'-deoxyuridine]

Commercially available 5-formyldeoxyuridine (25.0 mg, 0.098 mmol) and **1a** (1-ethyl-2,3,3-trimethyliodoleninium-5-sulfonate) (31.3 mg, 0.12 mmol) were heated in EtOH (600 µl) at 70°C for 4 h. Then another 600 µl EtOH was added and the mixture was stirred for further 2 h at 70 °C. All the ethanol was evaporated and the residue was dissolved in water. The pure product was isolated by RP-C18 chromatography on Lobar 310-25 (LiChroprep, 40-63 µM, Merck column) with water-ACN in ÄKTA prime. The product was lyophilized to obtain yellow amorphous solid (27.1 mg, 69%). ^{1}H NMR (600 MHz, DMSO-d_6) δ 11.97 (s, 1H-N1), 8.92 (s, 1H-h), 8.10 (d, J = 15.8 Hz, 1H-g), 8.04 (m, 2H-a, g'), 7.84 (d, J = 8.3 Hz, 1H-c), 7.81 (dd, J = 8.3, 1.4 Hz, 1H-b), 6.12 (t, J = 6.1 Hz, 1H-1'), 5.29 & 5.19 (two br s, 2' & 3'-OH), 4.42 (q, J = 7.3 Hz, 2H-e), 4.34 – 4.27 (m, 1H-3'), 3.89 (q, J = 4.0 Hz, 1H-4'), 3.75 – 3.58 (m, 2H-5'), 2.38 – 2.27 (m, 2H-2'), 1.76 (d, J = 3.9 Hz, 6H-f,f'), 1.44 (t, J = 7.3 Hz, 3H-d). ^{13}C NMR (126 MHz, DMSO-d_6) δ 181.8, 161.5, 150.3, 149.2, 148.9, 148.7, 142.9, 140.3, 126.5, 120.2, 113.9, 109.6, 108.5, 88.0, 86.2, 69.3, 60.5, 51.7, 42.0, 25.8, 25.8, 12.8. HRMS calcd. for $C_{23}H_{27}N_3O_8S$ 505.1519, found 504.1444, 505.1457, 506.1374.

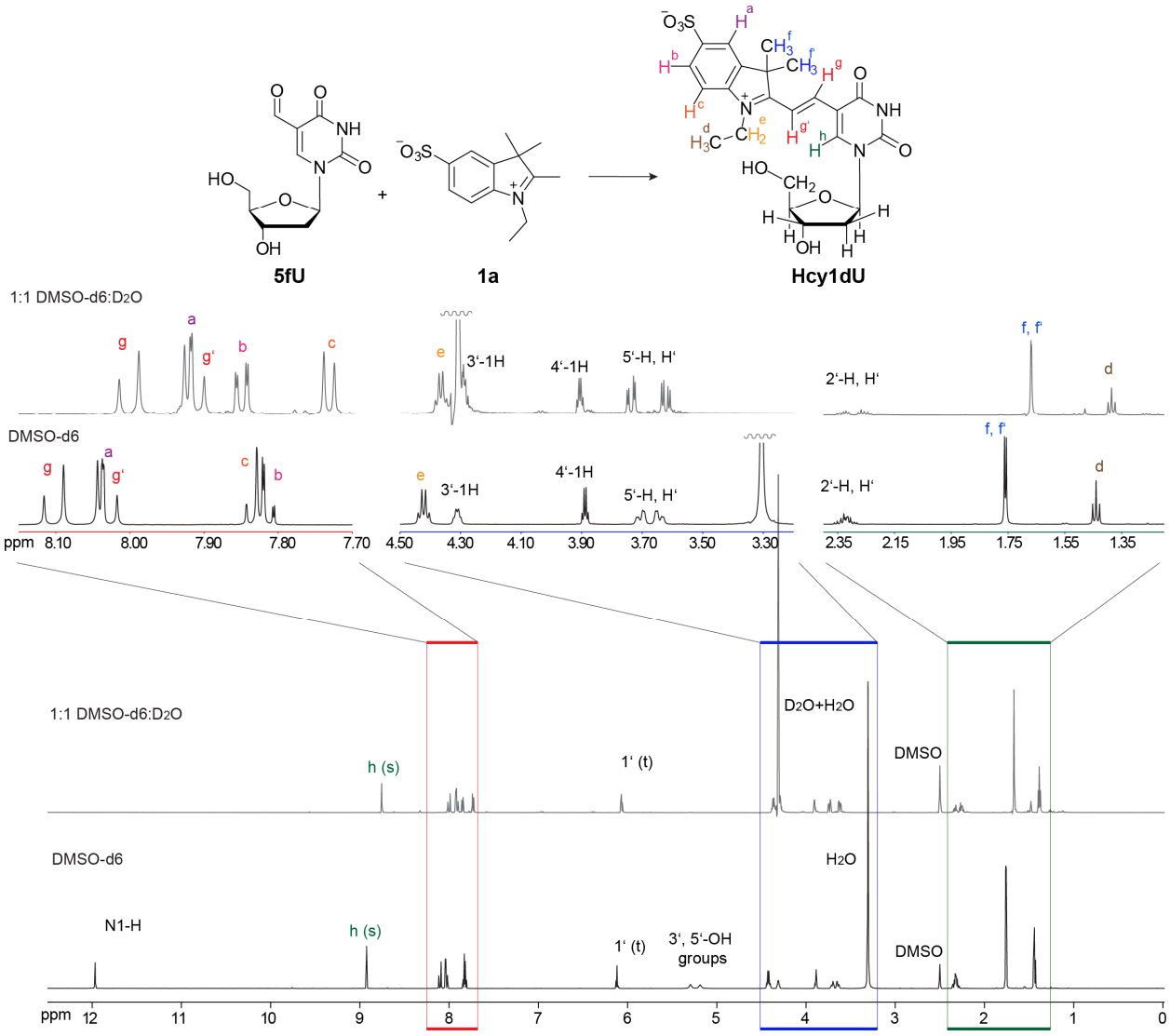


Figure S1a. Synthesis and NMR characterization of Hcy1dU. 1H-NMR (600 MHz) in DMSO-d6 and 50% D_2O in DMSO-d6.

Extinction coefficient and Quantum yield determination of Hcy1dU

The extinction coefficient of Hcy1dU nucleoside at 445 nm is 54900 $M^{-1}cm^{-1}$ in Tris buffer at pH 9, 45000 $M^{-1}cm^{-1}$ in MeOH, and 32900 $M^{-1}cm^{-1}$ in DMSO.

The quantum yield of the Hcy1dU deoxynucleoside was measured with respect to fluorescein as well as coumarin153 following the equation: $\Phi_F(x) = (A_s/A_x)(F_x/F_s)(n_x/n_s)^2\Phi_F(s)$

where 's' is the standard (fluorescein in 100 mM NaOH or coumarin 153 in EtOH), 'x' is the nucleoside (in 10 mM Tris pH 9, 150 mM NaCl), 'A' is the absorbance at excitation wavelength (445 nm), 'F' is the area under the emission curve, 'n' is the refractive index of the solvent, and Φ_F is the fluorescence quantum yield. The quantum yield of the standard fluorescein solution was taken as $\Phi_F(s)$ as 0.92 and that of coumarin 153 in EtOH was 0.38, and in water 0.11.[2]

The fluorescence quantum yield of Hcy1dU upon excitation at 445 nm is 0.01.

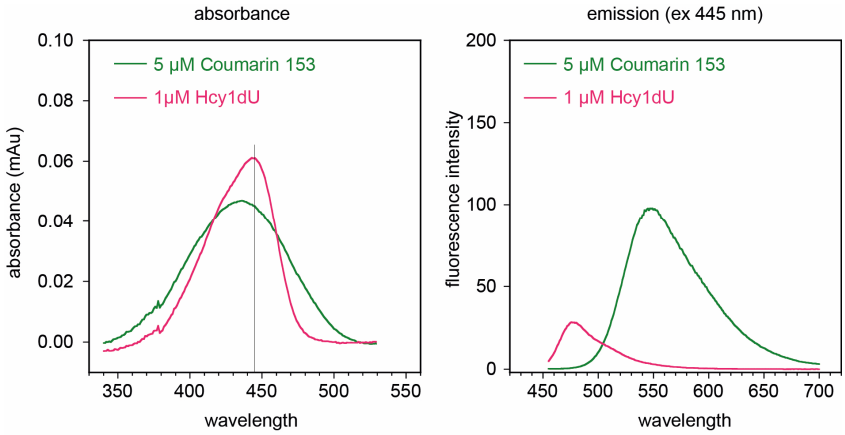
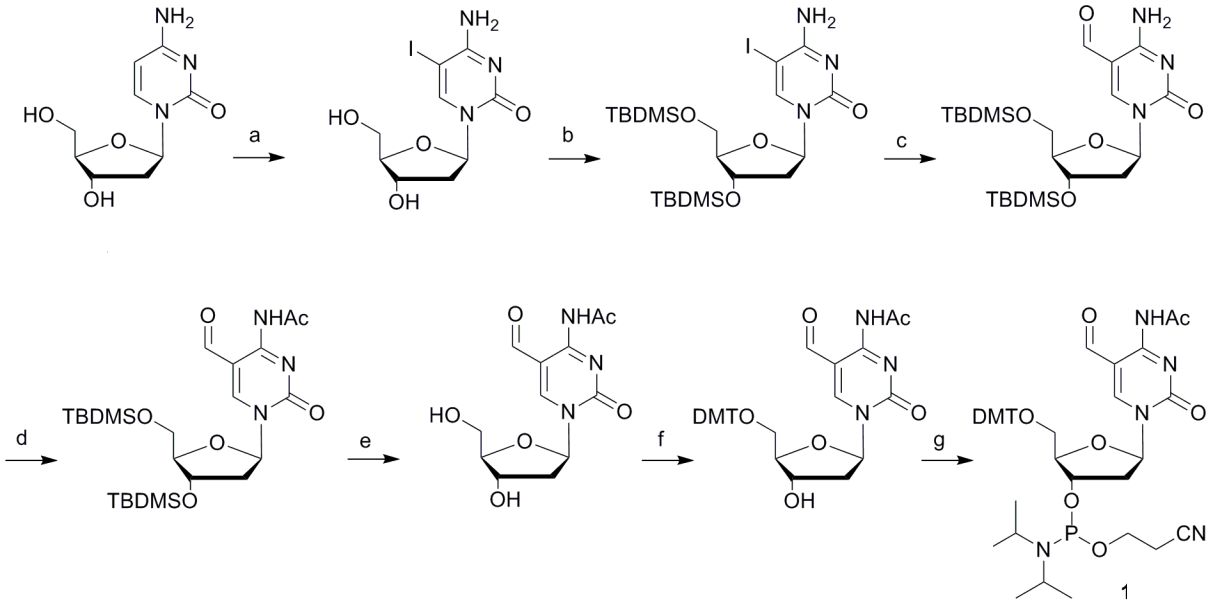


Figure S1b. Comparison of absorbance and emission of Hcy1dU nucleoside and coumarin 153 in aqueous solution. 1 µM Hcy1dU in 10 mM Tris buffer pH 9, 5 µM coumarin 153 in water (with 5% EtOH).

2.4. Synthesis of N^4-Acetyl-5'-DMT-5-formyl-2'-deoxycytidine 3'-β-cyanoethyl N,N-diisopropyl phosphoramidite

The 5fC phosphoroamidite building block was synthesized in 7 steps starting from 2'-deoxycytidine, according to published procedures with only minor modifications.[3] The synthetic scheme is depicted below. The intermediates and the final product were characterized by NMR spectroscopy and mass spectrometry, and the data were identical to reported values.



Scheme S1. Synthesis of 5fdC phosphoramidite. a) I_2, mCPBA, DMF, RT, 2 h, 45 %; b) tBDMS-Cl, imidazole, DMF, RT, 18 h, 93 %; c) CO, Bu_3SnH, $Pd_2(dba)_3$, toluene, 80 °C, 3 h, 65 %; d) AC_2O, DMF, RT, 3 d, 87 %; e) 1M TBAF, 0.5 M AcOH, THF, RT, 4.5 h, 74 %; f) DMT-Cl, pyridine, RT, 23 h, 57 %; g) CEP-Cl, iPr_2NEt, CH_2Cl_2, RT, 2 h, 68 %.

2.5. Synthesis of 5fU and 5fC-modified oligonucleotides

Solid-phase synthesis of DNA oligonucleotides

Solid-phase synthesis was carried out using standard phosphoramidite chemistry, at 0.6 µmol scale, using 100 mM solutions of phosphoramidites in dry acetonitrile, with benzylthiotetrazol as coupling

reagent, as regularly performed in our group.[4] Deprotection was performed under standard conditions with 25% NH_4OH/EtOH 4:1 at 55°C for 12-16h. Oligonucleotides were purified by PAGE, and characterized by anion exchange HPLC and ESI-MS.

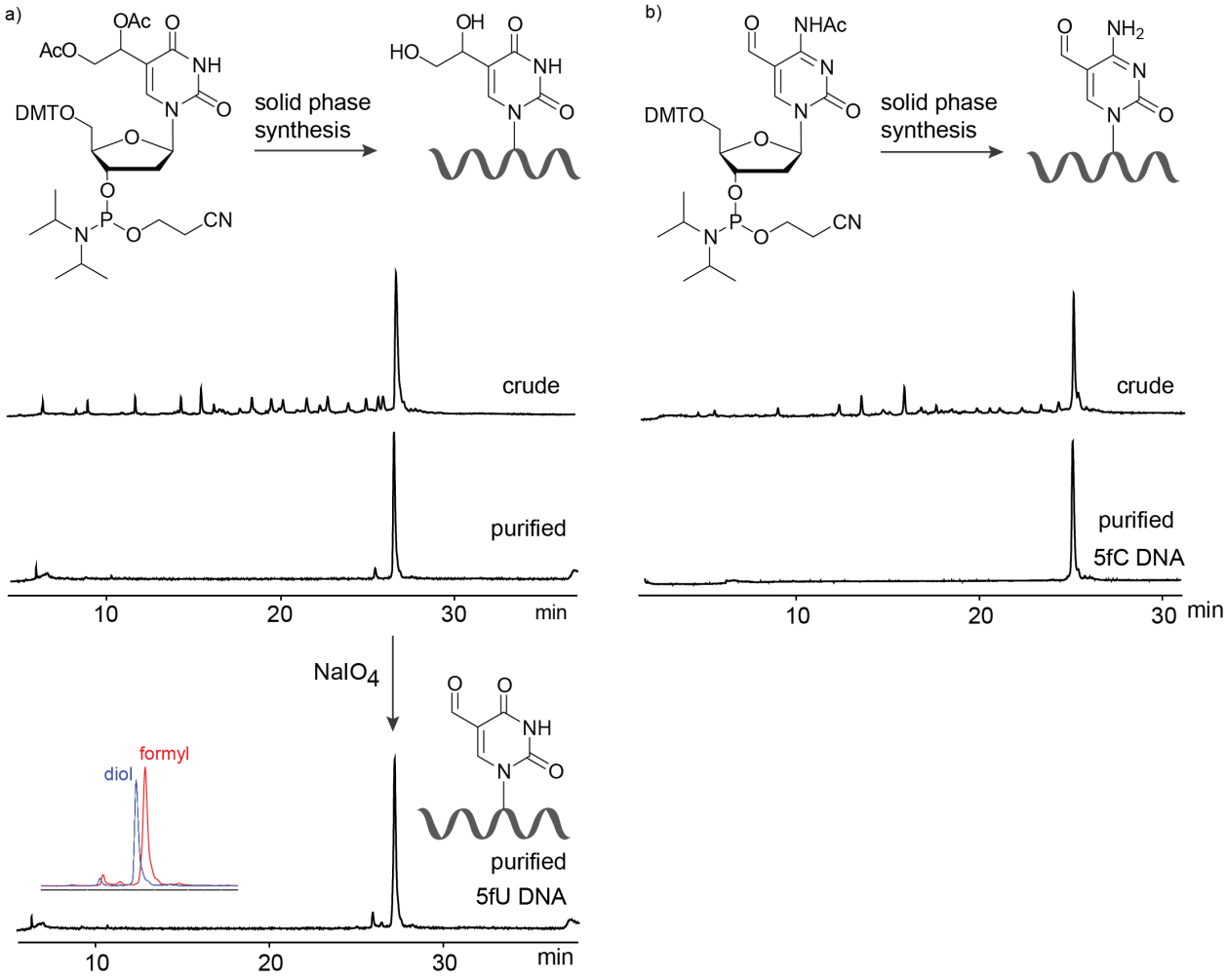
5fU-containing DNA oligonucleotides were prepared from 5-(1,2-dihydroxyethyl)-2'-deoxyuridine modified DNAs by oxidation with $NaIO_4$, according to reported procedures.[5] Briefly, an ice-cooled solution of 5-(1,2-dihydroxyethyl)-2'-deoxyuridine modified DNA (7.1 nmol in a final volume of 100 µl) was treated with 50 equivalent $NaIO_4$ (final concentration 3.5 mM) for 30 min at 4°C. Excess periodate was quenched by addition of glycerol (3 µL of 1 M) and the oxidized product was isolated by ethanol precipitation.

Enzymatic synthesis of RNA oligonucleotides by in vitro transcription

The in vitro transcription reaction with homemade T7 RNA polymerase was performed in 100 µL final volume using 1 µM template strand T1 / T2 and 1 µM T7 promoter strand in 40 mM Tris-HCl, pH 8.0, 30 mM $MgCl_2$, 10 mM DTT, 4 mM each unmodified NTP, 0.5 mM 5-formyl-modified NTP (5fCTP/5fUTP) and 2 mM spermidine at 37°C for 5 h. The reaction was quenched by adding 50 mM EDTA pH 8.0. The RNA transcript was purified by gel electrophoresis (20 % polyacrylamide, 35 W, 0.7 mm, 2 hr) yielding formyl modified 5'-triphosphorylated RNAs, which were characterized by ESI-MS. Yield: 6.0 nmol 5fU-RNA pppR1; 3.7 nmol 5fC-RNA pppR2 (+3.2 nmole with one extra nucleotide addition).

5'-Dephosphorylation by alkaline phosphatase to fC & fU-RNA

1 nmol of fC/fU-modified transcript was incubated in 1x CIAP buffer (Fermentas; 50 mM Tris-HCl, pH 8.5, 0.5 mM EDTA) and 1 u of Calf-intestine alkaline phosphatase in 25 µl at 37 °C for 0.5 h. Extraction with phenol-chloroform-isoamyl alcohol (PCI) followed by ethanol precipitation of aqueous layer yielded fC/fU-modifed RNA (yield ~70 %) which was directly used without further purification.



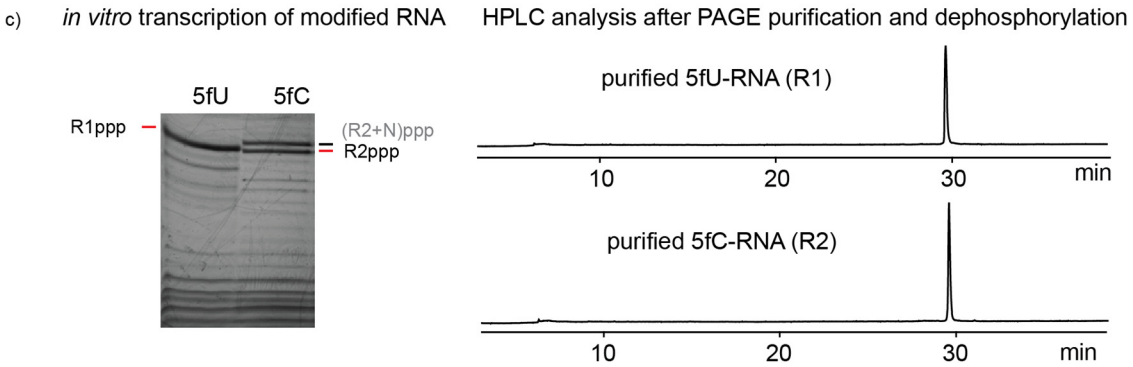


Figure S2. Synthesis and characterization of 5fU and 5fC modified oligonucleotides. a) Solid-phase synthesis provided 5-(1,2-dihydroxethyl)2'-deoxyruridine-modified DNA, which was converted to 5fU-DNA upon NaIO4 oxidation. Anion exchange HPLC traces for D1 are shown. b) Solid-phase synthesis provided 5fC DNA. Anion exchange HPLC traces for crude and purified D6 are shown. c) In vitro transcription with T7 RNA polymerase provided 5fU- and 5fC-modified RNA. Image of denaturing PAGE is shown, and anion exchange HPLC ttraces of purified RNAs after dephosphorylation.

3. Oligonucleotide sequences

Table S1. Sequences and ESI-MS of 5-formylpyrimidine modified oligonucleotides

name	sequence	calc. mw.	obs. m.w.
D1	CTCTTGA**G(fU)G**TTATG	4602.9	4604.0
D2	CTCTTGA**A(fU)A**TTATG	4570.9	4570.8
D3	CTCTTGA**C(fU)C**TTATG	4522.9	4522.4
D4	CTCTTGA**T(fU)T**TTATG	4552.9	4552.1
D5	GACTCA**A(fU)A**GCCGTG	4590.9	4591.4
D6	CTCTTGA**G(fC)G**TTATG	4601.9	4601.7
D7	CTCTTGA**A(fC)A**TTATG	4569.5	4569.5
D8	CTCTTGA**C(fC)C**TTATG	4521.9	4522.6
D9	CTCTTGA**T(fC)T**TTATG	4551.9	4551.6
D10	CATA**G(fC)G**CTCAAGAGAAATCTCGATGG	8345.4	8345.1
D11	C**G(fC)G**GAGCTCGCTTGTCG	5535.6	5535.0
R1	r(GGAAGAG**A(fU)G**GCGACGG)	5635.5	5633.4
R2	r(GGAAGAG**A(fC)G**GUGAUGG)	5636.5	5635.0

unmodified analogue DNAs for quantification experiments (all 5'-sequence-3')

analogue of D5: GACTCAATAGCCGTG
analogue of D10: CATAGCGCTCAAGAGAAATCTCGATGG

Complementary oligonucleotides to D6-D10

complement of D6: CATAACGCTCAAGAG
complement of D7: CATAATGTTCAAGAG
complement of D8: CTCTTGACCCTTATG
complement of D9: CATAAAGATCAAGAG

Templates for transcription of 5fC/5fU modified RNA

T1 CCGTCGC**CAT**CTCTTCCTATAGTGAGTCGTATTACAG
T2 CCATCAC**CGT**CTCTTCCTATAGTGAGTCGTATTACAG

Templates for enzymatic incorporation of multiple 5fU nucleotides

T3 CTATTGTCGAGAAATCTCGATGG
T4 CTATTGTCTATTGTCTATTGTCCTATAGTGAGTCGTATTACAG
T5 CTATTGTCTATTGTCTATTGTCTATTGTCTATTGTCTATTGTCCTATAGTGAGTCGTATTACAG

Other unmodified DNAs

T7 promoter CTGTAATACGACTCACTATA (for *in vitro* transcription and for primer extension in Fig S6)
Primer P1 CCATCGAGATTTCTC (used in polymerase stop assays with D10, Fig S12, S13)
Primer P2 CCATCGAGATTTCTCGAGTT (see Fig S12)

4. Fluorogenic labeling of 5fU- and 5fC-modified oligonucleotides

Protocol for labeling of 5fU-oligonucleotides

with 1a. The 5fU-oligonucleotides (1 nmol) were mixed with MES buffer stock, pH 6.0 (prepared by mixing 0.5 M MES, pH 6.0 and 2 M KCl stocks in 2:1 (v/v) ratio). To the DNA solution in buffer, 8 µl of 125 mM **1a** was added. The final reaction volume was 10 µl, which contained 50 mM MES, pH

6.0 and 100 mM KCl. The reactions were performed at 45°C for 6-7 h. The reaction mixture was diluted to 100 µL with water, and excess reagents were removed by ethanol precipitation and ultrafiltration using Vivaspin 500 columns (Sartorious, 3000 MWCO PES). The products were used without PAGE purification.

with 2a & 3a. The 5fdU-DNA (2 nmole) were incubated with **2a** (40 mM) or **3a** (80 mM) in 50 mM MES, pH 6.0, 100 mM KCl in a final volume of 40 µl at 45°C for 6 h. The excess reagents were removed by ethanol precipitation and purification in a Sephadex size exclusion column by elution with water, followed by lyophilization. The products were used without PAGE purification.

Table S2. ESI-MS of HcyxU-labeled DNA and RNA oligonucleotides

product	expected m.w.	observed m.w.
D1-Hcy1U	4852.3	4851.5
D2-Hcy1U	4820.3	4819.1
D3-Hcy1U	4772.2	4771.2
D4-Hcy1U	4802.3	4801.2
D5-Hcy2U	4891.4	4889.3
D5-Hcy3U	4916.4	4917.8
R1-Hcy1U	5884.8	5884.8

Protocol for labeling of 5fC-oligonucleotides

The 5fC-oligonucleotides (1 nmole) and **1b** (100 mM final conc, added from a 1 M stock in water, deep brown solution) were incubated in 50 mM KPi, pH 7.4, 100 mM KCl in a final volume of 10 µl at 45°C for 6-7 h. The excess reagents were removed by ethanol precipitation and the reaction mixtures were dissolved in 10 µl water. Glacial acetic acid (0.3 -0.5 µl) was added to solutions and incubated at 45°C for 13 min. To the light yellow colored solutions, TEN buffer, pH 8.0 (190 µl) was added and precipitated with ethanol. The products were purified by denaturing PAGE (20% polyacrylamide, 35 W, 2.5 h). Alternatively, after the aldol-addition reaction in KPi buffer, the reaction mixture was subjected to ethanol precipitation and PAGE purification. The purified addition intermediate was subjected to dehydration (1-2 nmole in 10 µl) with glacial acetic acid (0.3-0.5µl) as above. Neutralization with TEN buffer, pH 8.0 (190 µl) followed by ethanol precipitation yielded the desired product as confirmed by ESI-MS.

Table S3. ESI-MS of Hcy1C-labeled DNA and RNA oligonucleotides

product	expected m.w.	observed m.w.
D6-Hcy1C	4851.3	4850.1
D7-Hcy1C	4818.8	4818.5
D8-Hcy1C	4771.2	4771.3
D9-Hcy1C	4800.0	4801.2
D10-Hcy1C	8594.8	8594.5
R2-Hcy1C	5885.8	5885.4

5. Enzymatic incorporation of multiple 5fU in DNA and reaction with 1a

Primer extension and reaction with 1a: 400 pmoles of template (T3/T4/T5) and 400 pmoles of primer (P1) were annealed in 43.5 µl of water by heating at 95 °C for 3 min and cooling at RT for 30

min. The primer extensions were performed at 37 °C for 1 h in a final volume of 50 µl in 1x extension buffer (10 mM Tris, pH 7.5, 5 mM $MgCl_2$, 7.5 mM DTT) using 200 µM of dNTP (fdUTP instead of dTTP) each and 0.1 u/µl Klenow exo- (from Jena Bioscience). After the extension reaction, the polymerase was removed by PCI extraction and the excess dNTPs by ultrafiltration using Vivaspin 500 followed by ethanol precipitation. The primer extension mixtures were dissolved in 2 µl water containing 250 mM MES, pH 6.0 and 500 mM KCl. To each of these three mixtures, 8 µl of 125 mM **1a** solution was added and the reactions were incubated at 45°C for 6.5 h. The excess reagents were removed by ethanol precipitation and the reaction mixtures were dissolved in 20 µl water. These stock solutions were diluted to 1µM (calculated with respect to the starting amount of template/primer).

Fluorescence measurement: 13 µl of 1 µM stock solutions were dissolved in 10 mM Tris, pH 9.0 and 150 mM NaCl in a final volume of 130 µl (final conc. ~0.1 µM). Fluorescence was measured with excitation at 445 nm after incubation in the buffer for 90 min at 20°C. Results are depicted in Figure S6.

6. Polymerase stop assay

The 5fC-DNA (D10) and its unmodified analogue were mixed in various ratios (containing 0, 5, 10, 20, 50, 80, and 100 % 5fC) for a final amount of 400 pmol DNA. The mixtures were reacted with 100 mM **1b** in 50 mM KPi pH 7.4, 100 mM KCl in final volume of 10 µl at 45 °C for ~11 h. The excess reagents were removed by ethanol precipitation. The reaction mixtures were then dissolved in 10 µl water and treated with 0.3 µl glacial AcOH at 45 °c for 13 min. The acid was neutralized by addition of TEN buffer pH 8.0 (190 µl) and the product was isolated by ethanol precipitation and dissolved to a concentration of 10 µM.

Primer extension conditions: The 5'-^{32}P-labeled primer (~6 pmol P1) and products from step ii (~20 pmol) was annealed by heating at 95 °C for 2 min and cooling at RT for 15 min in a total volume of 7 µl. The extension reaction was performed at 37 °C for 10 min in 1x Klenow reaction buffer (10 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$ and 7.5 mM DTT) using 200 µM of dNTP each and either 0.02 u/µl or 0.5 u/µl Klenow exo$^-$ polymerase. The primer extensions were analyzed by denaturing PAGE. Band intensities were quantified by ImageQuant software. The summation of band intensities for 6 and 7 divided by the summation of all bands in a lane yielded the amount of stop. The results are presented in Figure S12.

7. Additional Supporting Information Figures

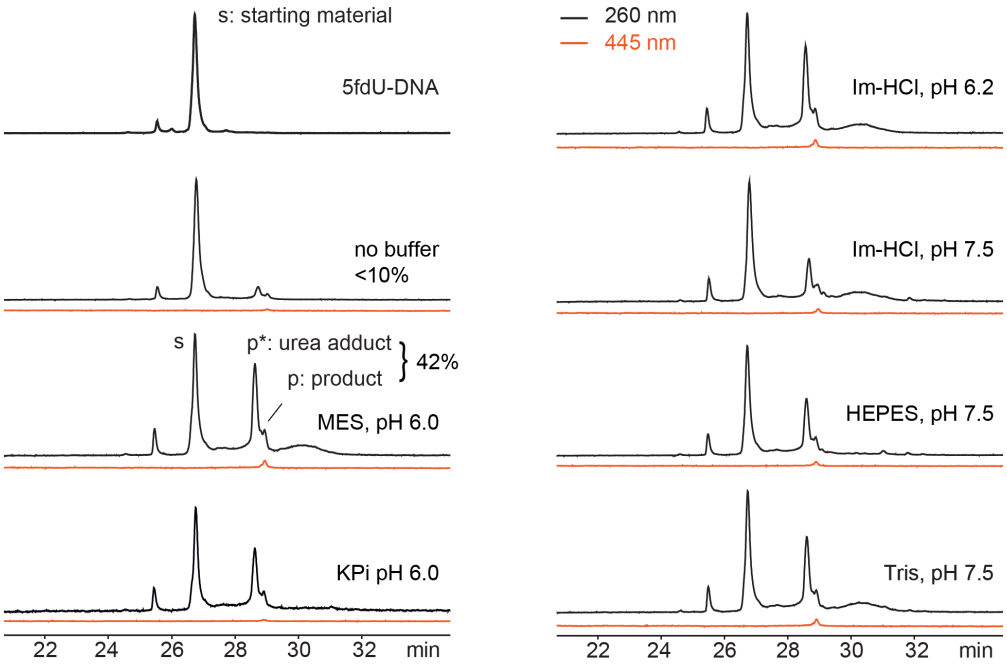


Figure S3. Anion exchange HPLC traces for reaction optimization of 5fU-DNA (D1) with **1a**.

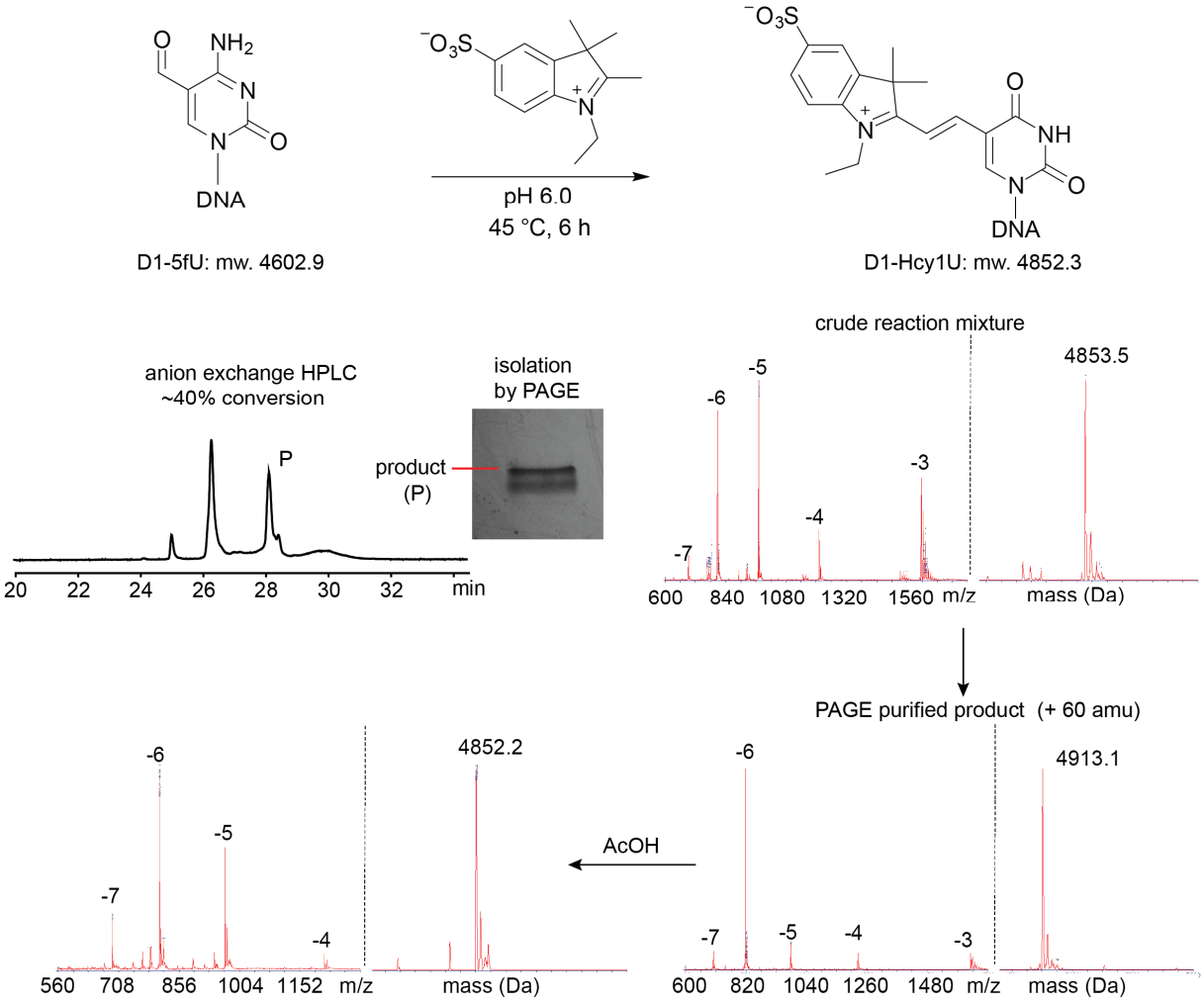


Figure S4. Analysis of reaction of D1-5fU with **1a** by ESI-MS analysis before and after PAGE purification shows formation of urea adduct with the product D1-Hcy1U, which reverted back to the desired Hcy1U product by acid treatment.

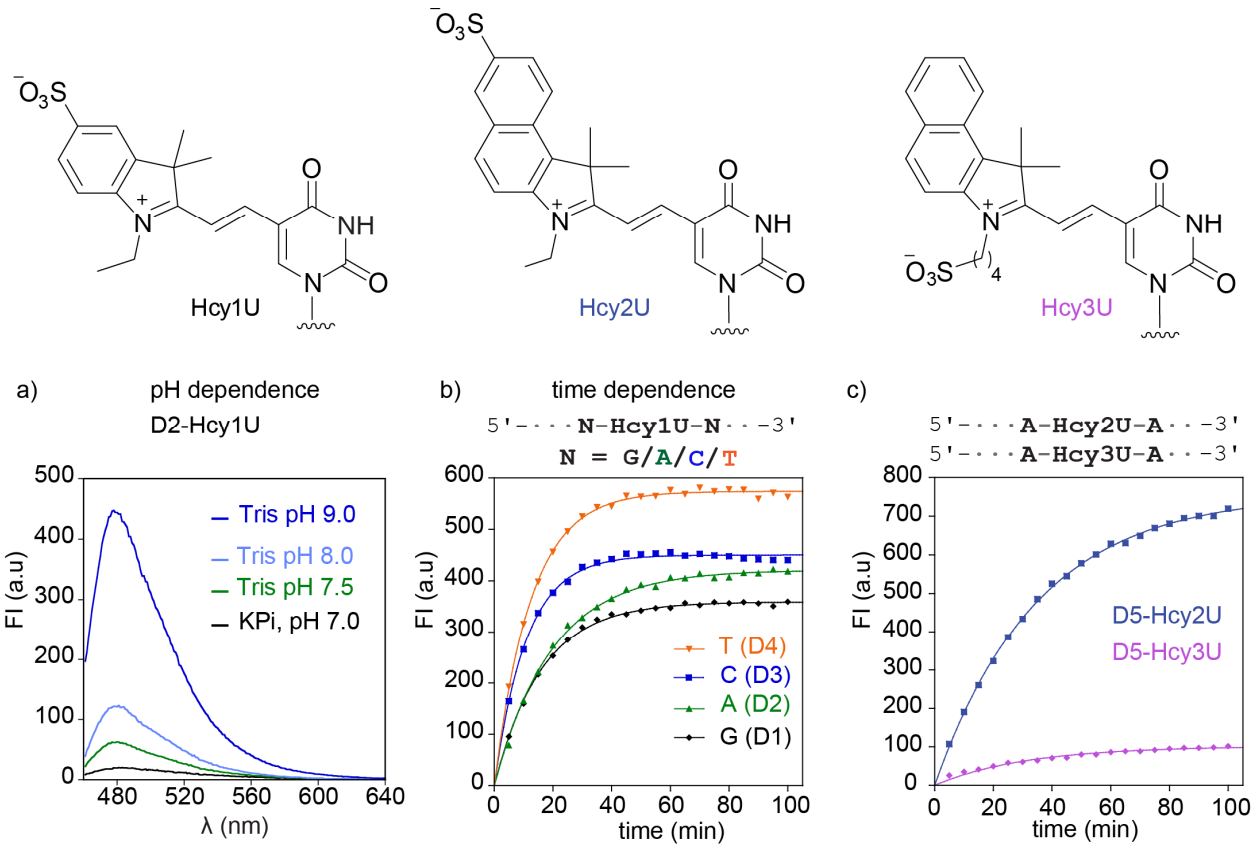


Figure S5. a) pH-Dependence of Hcy1U-DNA (D2) after 90 min of sample preparation. b) Time-dependent fluorescence intensity enhancement for Hcy1U-labeled DNAs with various adjacent nucleobases (D1-D4). c) Time-dependence of fluorescence emission of Hcy2U- and Hcy3U-labeled D5.

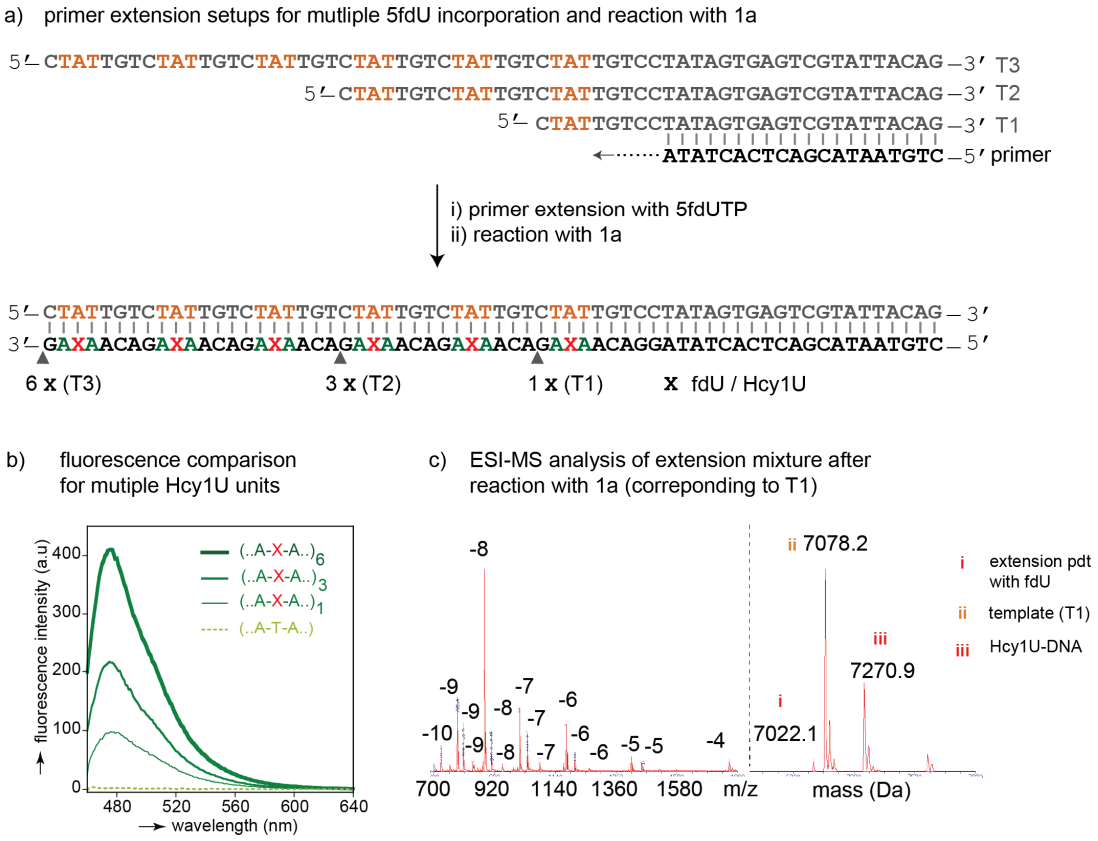


Figure S6. a) Incorporation of multiple 5fU units into DNA strand by primer extension and labeling reaction with **1a**. b) Comparison of the fluorescence intensities after labeling shows linearity with the number of 5fU units present in the DNAs. c) ESI-MS analysis of the reaction mixture after primer extension and labeling for one 5fU nucleotide.

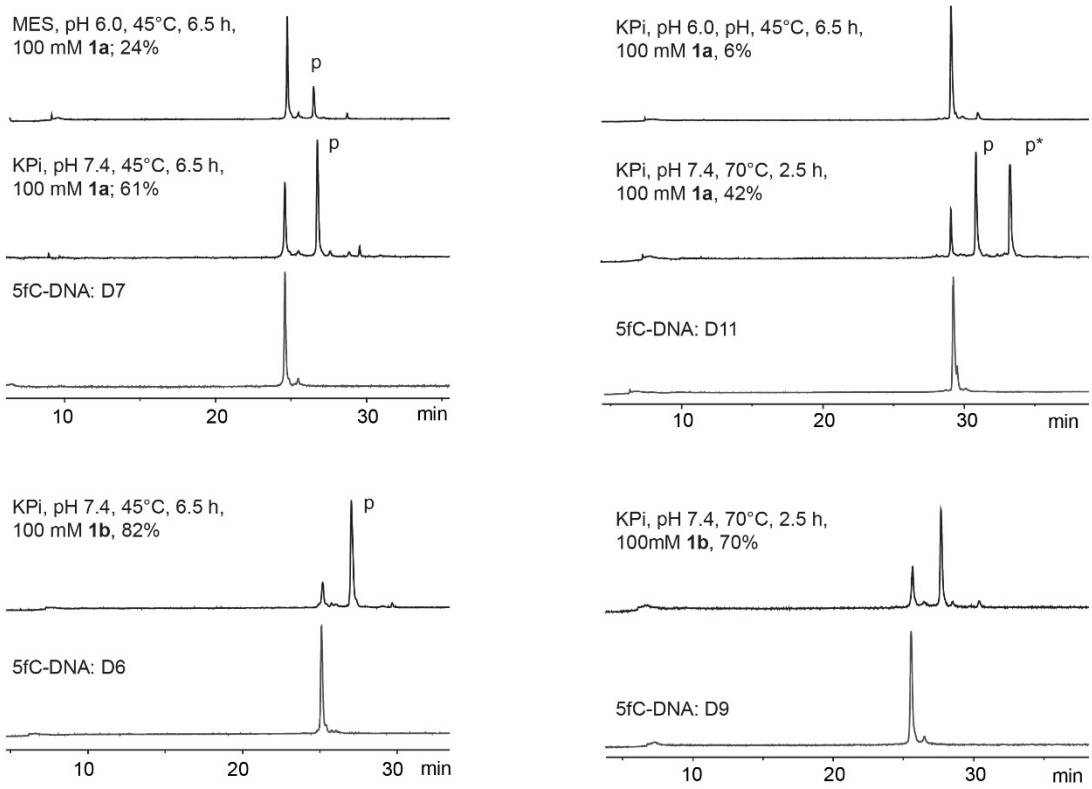


Figure S7. Anion exchange HPLC analysis of labeling reactions of 5fC-DNAs with **1a** and **1b**. p = aldol addition product, p* = double addition product (see main text).

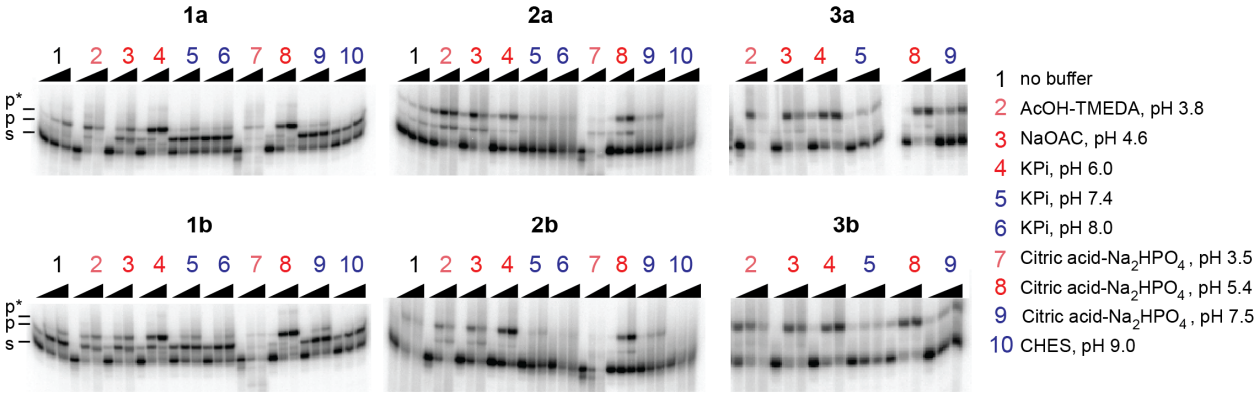


Figure S8. Reactions of 5'-^{32}P-labeled 5fC-DNA (D6, labeled "s" for starting material) with indoleninium derivatives under various conditions 1-10. The acidic pHs are marked with numbers in red color while the basic pHs in blue color. The reactions were performed with 10 mM (**1**, **2**) or 35 mM (**3**) reagent in 10 µl volume at 70°C. The time points are 1, 3, 6 h. All the reactions contained 100 mM NaCl. The reactions were analyzed by 20% PAGE. Single and double addition products are labeled as p and p* (see also Fig S7 and main text).

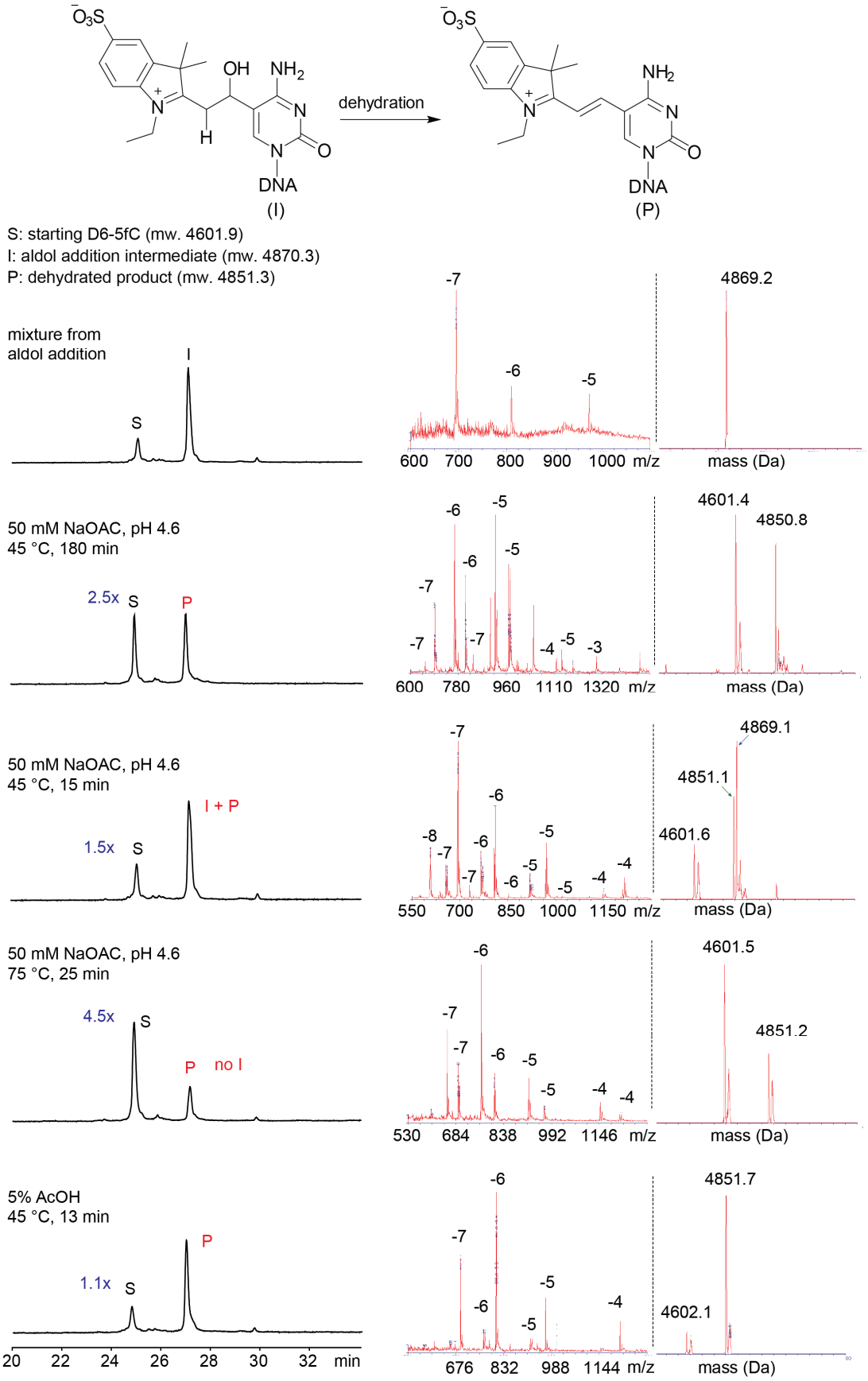


Figure S9. Optimizations for dehydration reaction after aldol addition of 5fC-DNA (D6) with **1b**. The relative peak intensities in anion exchange HPLC (left panel) of the non-consumed starting 5fC-DNA before and after elimination reaction were compared to determine the extent of retro-aldol reaction. As the intermediate and the

dehydrated product could not be distinguished based on their retention times, formation of the product was confirmed by ESI-MS analysis (right panel).

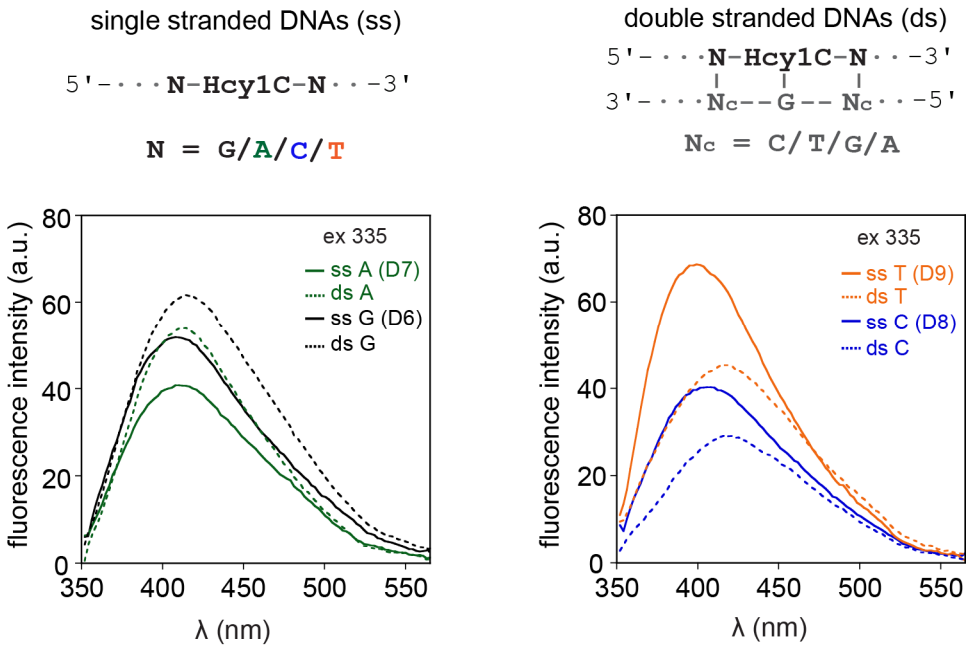


Figure S10. Effect of duplex formation on the fluorescence emission of Hcy1C in DNA. The labeled DNAs were PAGE-purified after the aldol addition reaction, followed by dehydration and precipitation. Duplex samples were prepared by annealing with complementary strands in 10 mM KPi buffer pH 7.0, 150 mM NaCl.

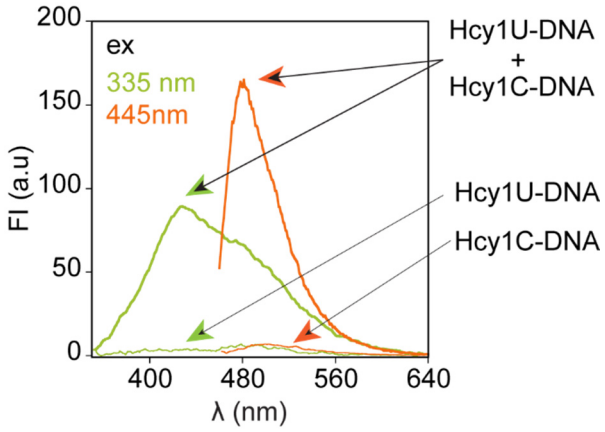


Figure S11. Mutually exclusive emission behavior of Hcy1U and Hcy1C labeled DNAs (D5-5fU and D7-5fC). The sample containing both modifications is represented by thick lines (0.1 µM Hcy1U and 3 µM Hcy1C), showing emission of Hcy1C and Hcy1U, but only upon individual excitation. The thin lines show samples containing either Hcy1U or Hcy1C, each excited at the wavelengths specific for the other nucleoside, respectively, demonstrating only background emission. The data for the cognate excitation wavelengths for Hcy1U and Hcy1C in single-labeled samples are not shown. All the samples were measured in 10 mM Tris pH 9, 150 mM NaCl.

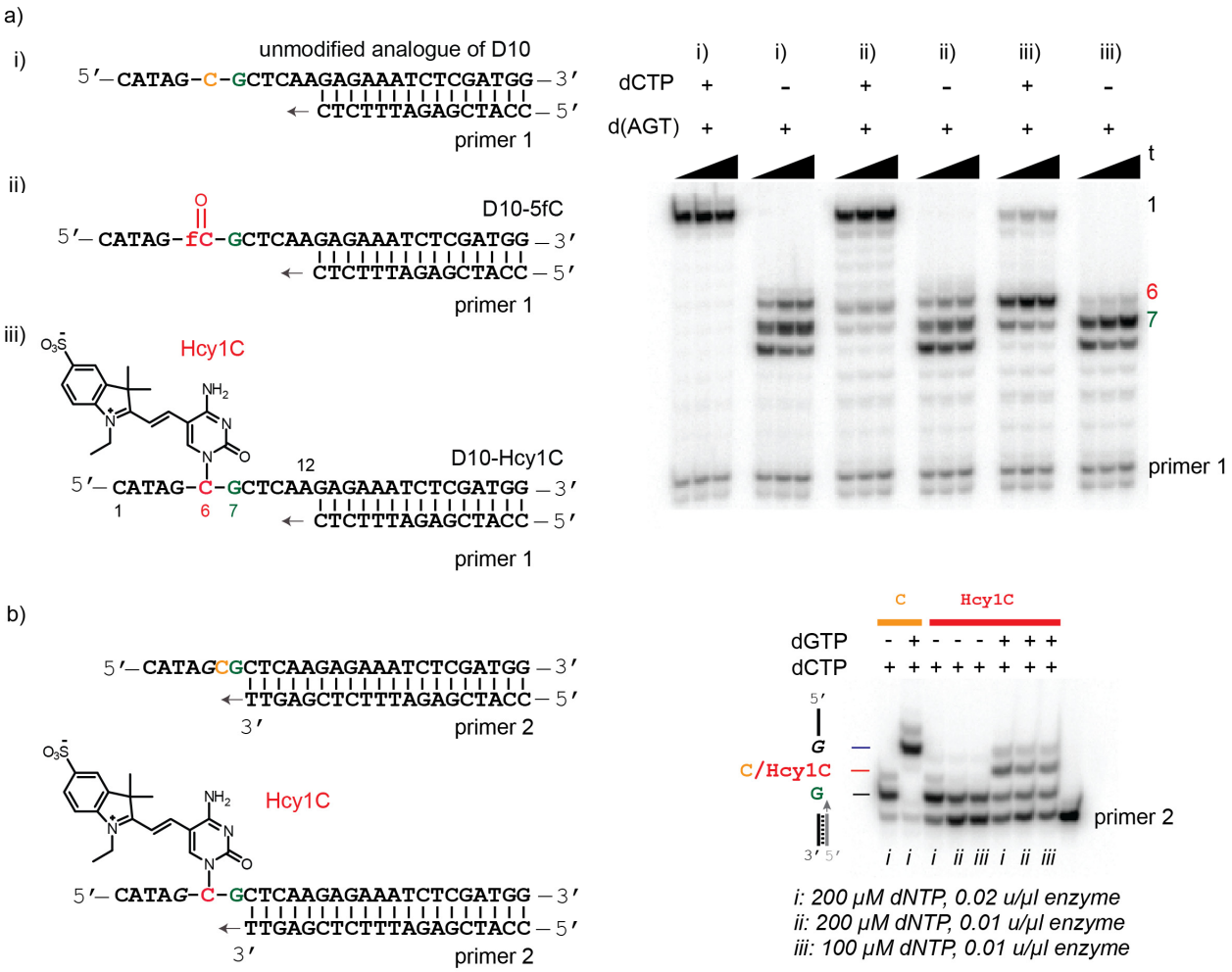


Figure S12. a) Comparing primer extension reactions with DNA polymerase on templates containing (i) canonical C, (ii) 5fC and (iii) Hcy1C in an identical sequence context. The time points are 10, 25, 45 min. Primer extension conditions as described on page S11. b) Deciphering the specificity of the polymerase stop.

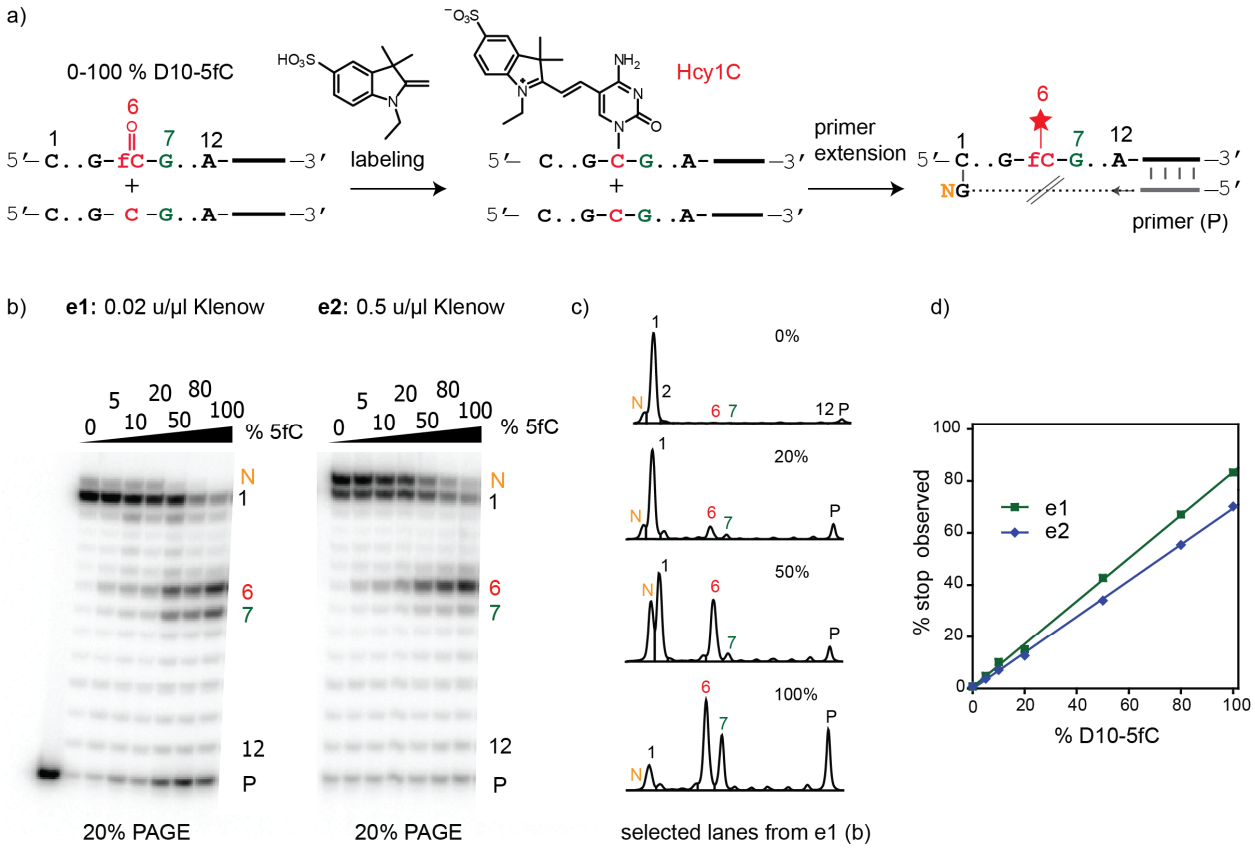


Figure S13. Deciphering % 5fC content from polymerase stop assay. a) Schematic depiction of the experimental setup. b) PAGE analysis of the primer extension reactions using different amount of enzymes. c) Exemplary line graph of selected lanes from gels in b) generated by ImageQuant software. d) Linear correlation between the observed polymerase stop (calculated as % intensity in bands 6+7) and the content of 5fC-DNA in the sample.

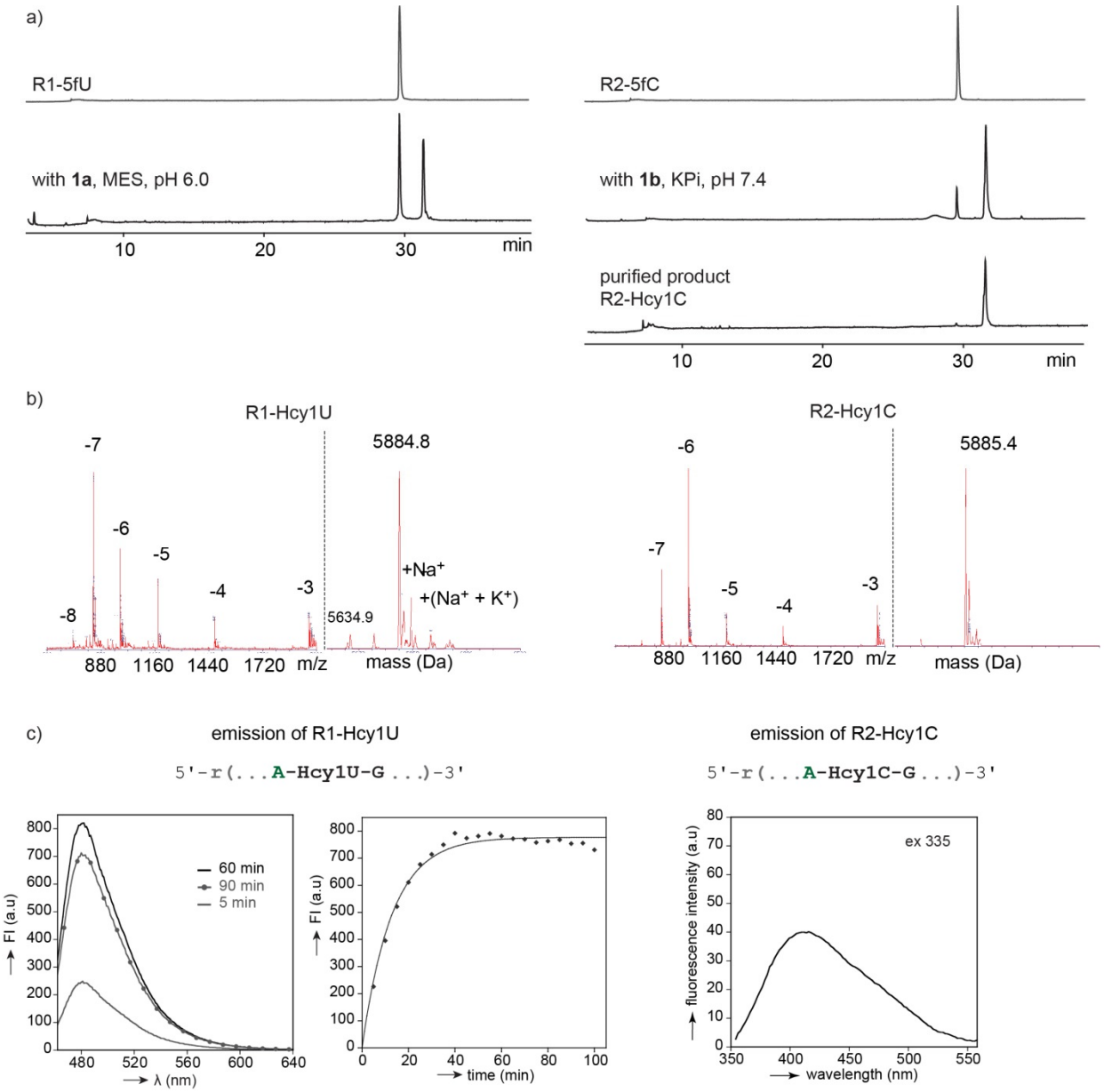


Figure S14. Labeling of 5fU and 5fC containing RNAs with **1a** and **1b** respectively, analyzed by a) anion exchange HPLC and b) ESI-MS. c) Fluorescence emission properties of the labeled RNAs (0.25 µM Hcy1U-RNA in 10 mM Tris pH 9.0; 5 µM Hcy1C-RNA in 10 mM KPi, pH 7.0).

References for supporting information

[1] a) F. Shao, H. Yuan, L. Josephson, R. Weissleder, S. A. Hilderbrand, *Dyes and pigments* **2011**, *90*, 119-122; b) R. B. Mujumdar, L. A. Ernst, S. R. Mujumdar, C. J. Lewis, A. S. Waggoner, *Bioconjug. Chem.* **1993**, *4*, 105-111.

[2] D. Magde, R. Wong, P. G. Seybold, *Photochem. Photobiol.***2002**, *75*, 327-334.

[3] a) M. Munzel, U. Lischke, D. Stathis, T. Pfaffeneder, F. A. Gnerlich, C. A. Deiml, S. C. Koch, K. Karaghiosoff, T. Carell, *Chemistry* **2011**, *17*, 13782-13788; b) Q. Dai, C. He, *Org. Lett.* **2011**, *13*, 3446-3449.

[4] a) F. Wachowius, F. Javadi-Zarnaghi, C. Höbartner, *Angew. Chem. Int. Ed.* **2010**, *49*, 8504-8508; b) B. Samanta, C. Höbartner, *Angew. Chem. Int. Ed.* **2013**, *52*, 2995-2999.

[5] a) K. Sato, W. Hirose, A. Matsuda, *Curr. Protoc. Nucleic Acid Chem.* **2008**, *Chapter 1*, Unit 1 21; b) H. Sugiyama, S. Matsuda, K. Kino, Q. M. Zhang, S. Yonei, I. Saito, *Tetrahedron Lett.* **1996**, *37*, 9067-9070.