Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Analysis of an Active Deformylation Mechanism of 5-Formyldeoxycytidine (fdC) in Stem Cells

Alexander Schön[+], *Ewelina Kaminska*[+], *Florian Schelter*[+], *Eveliina Ponkkonen, Eva Korytiaková, Sarah Schiffers, and Thomas Carell**

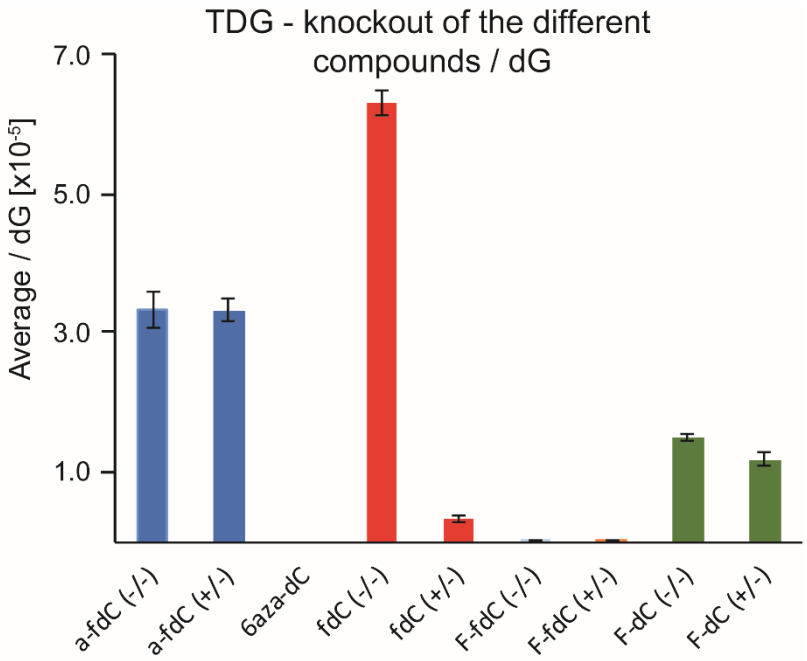
anie_202000414_sm_miscellaneous_information.pdf

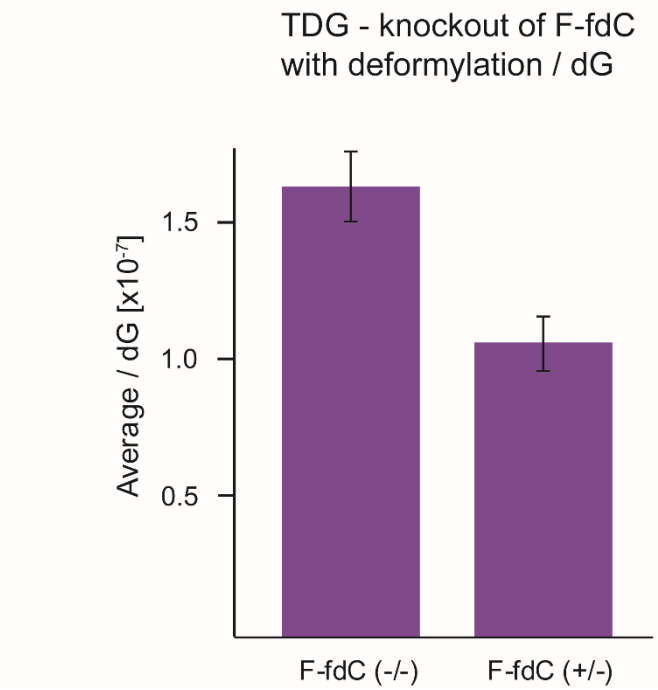
Supporting information

Table of contents

Supplementary figures

Table 1: TDG knockout experiments of the fed nucleosides





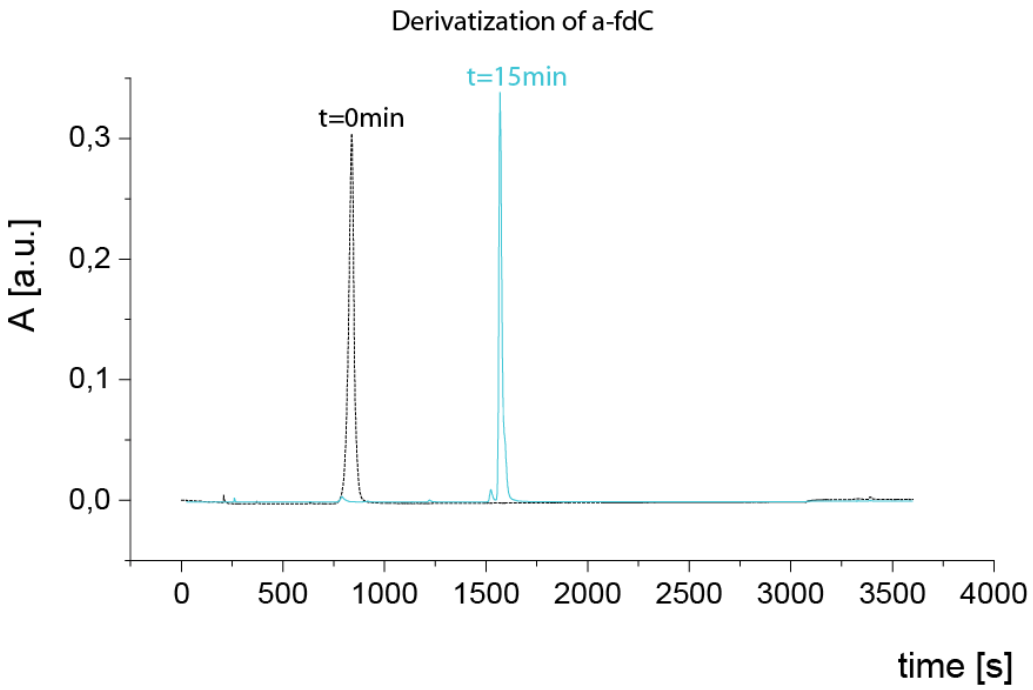


Figure 1: Quantitative derivatization of a-fdC with $MeONH_2$ after 15 min.

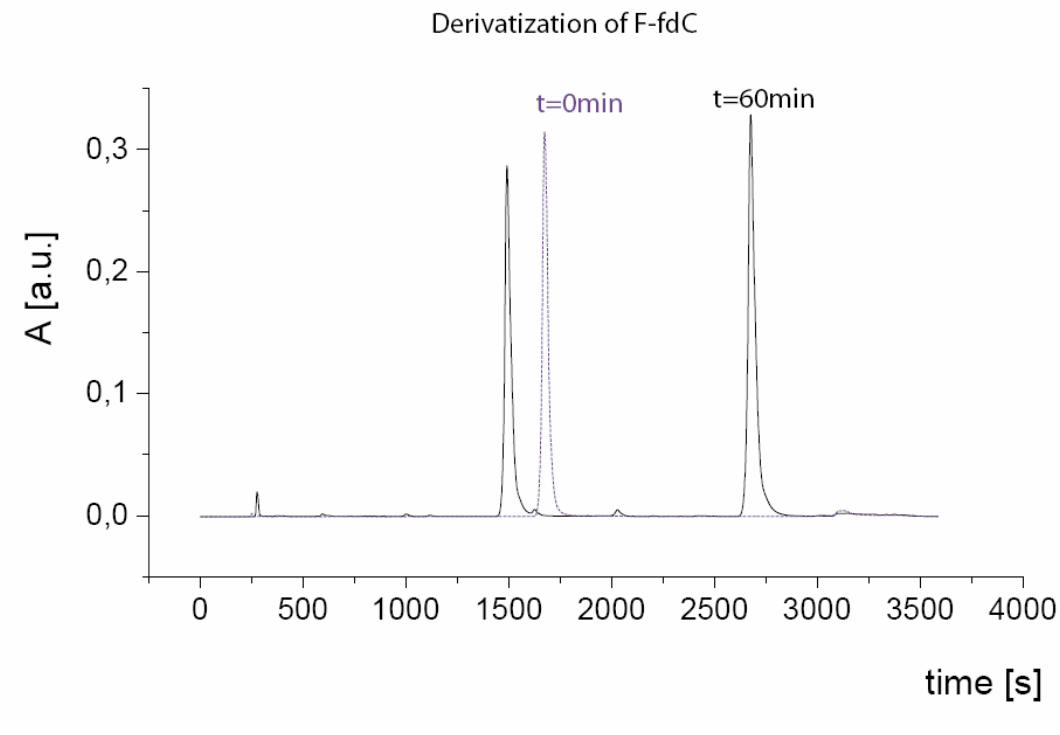


Figure 2: Derivatisation of F-fdC after 60 min, mixture of product and starting material.

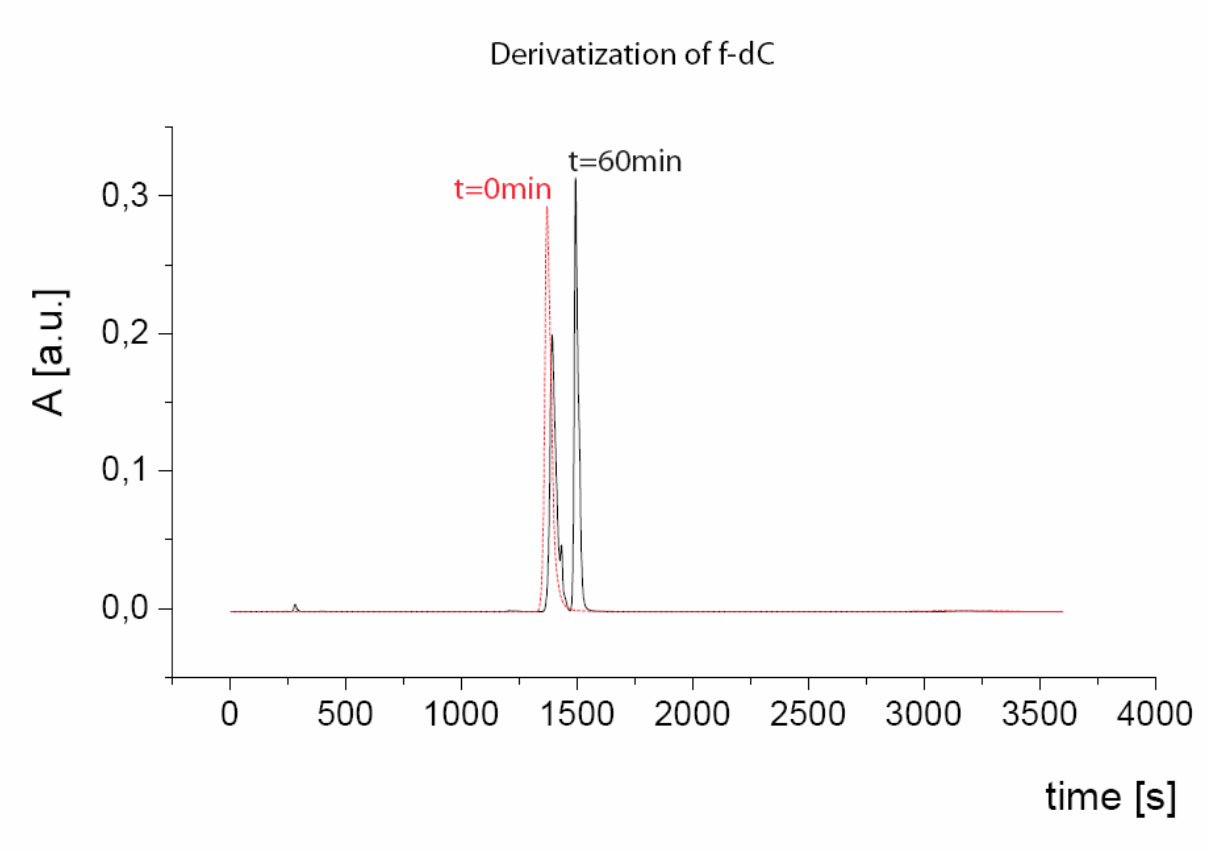


Figure 3: Derivatisation of fdC after 60 min, mixture of product and starting material.

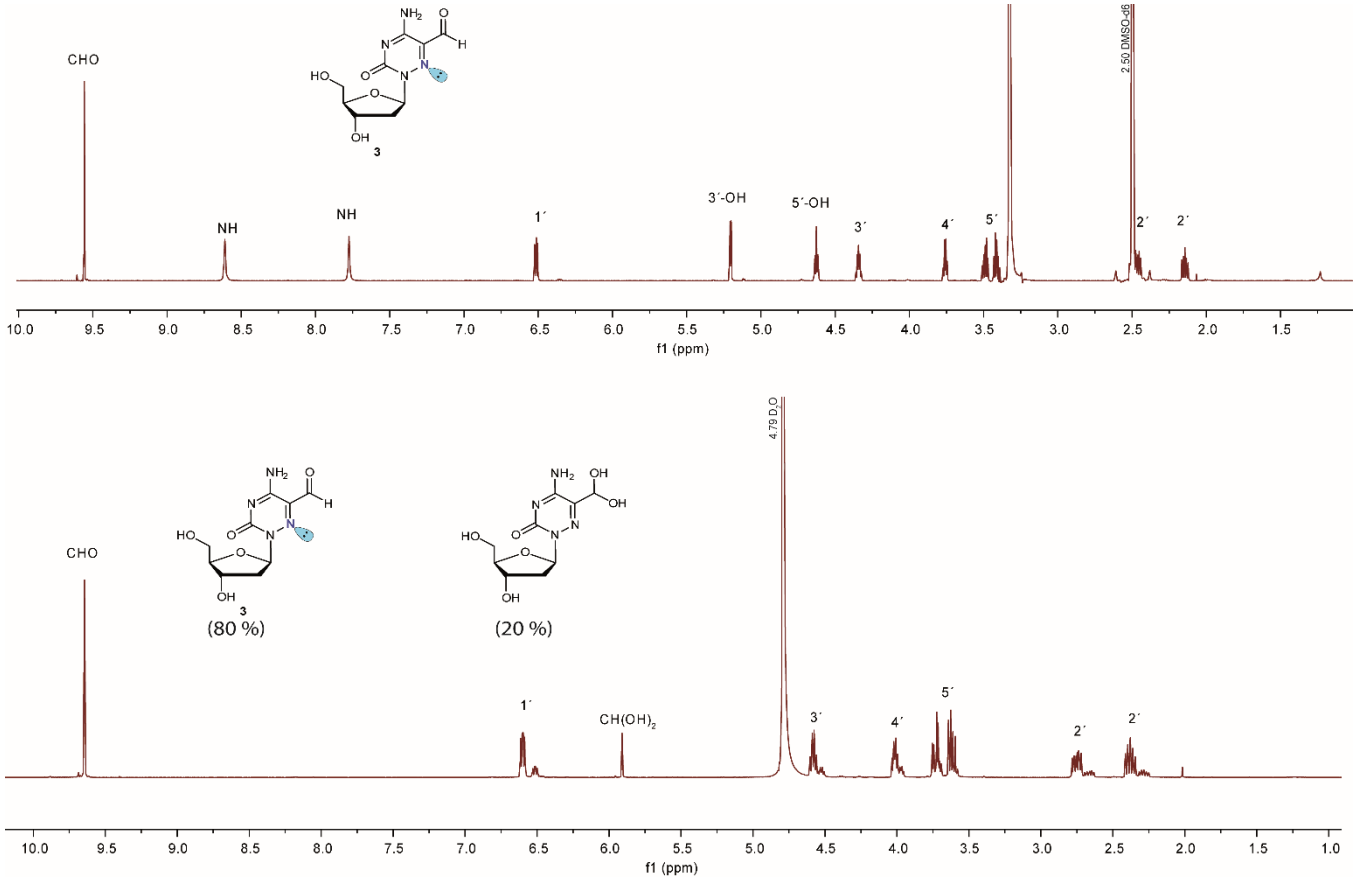


Figure 4: NMR spectra of a-fdC showing the hydrate formation of a-fdC in D_2O. Upper panel: NMR in DMSO-d6. Lower panel: NMR in D_2O.

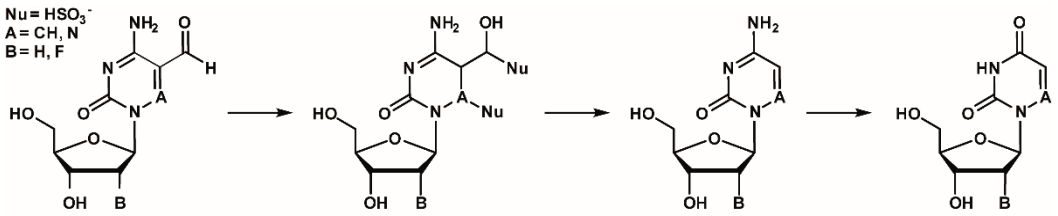


Figure 5: Proposed reaction mechanism with bisulfite

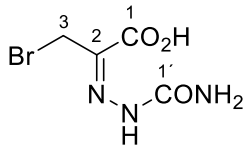
General Methods

Chemicals were purchased from *Sigma-Aldrich*, *TCI*, *Fluka*, *ABCR*, *Carbosynth* or *Acros Organics* and used without further purification. Solvents of reagent grade were purified by distillation. Reactions and column chromatography fractions were monitored by thin-layer chromatography (TLC) on silica gel F254TLC plates from *Merck KGaA*. Flash column chromatography was performed on *Geduran*® Si60 (40-63 µm) silica gel from *Merck KGaA* applying slight nitrogen pressure. Chemical transformations were conducted under nitrogen or argon atmosphere in oven-dried glassware unless otherwise specified. NMR spectra were recorded on *Bruker AVIIIHD 400* (400 MHz) or *Varian NMR-System600* (600 MHz) spectrometers. ^{1}H-NMR shifts were calibrated to the residual solvent resonances: $CDCl_3$ (7.26 ppm), DMSO-d6 (2.50 ppm) and D_2O (4.79 ppm). ^{13}C-NMR shifts were calibrated to the residual solvent: $CDCl_3$ (77.16 ppm), DMSO-d6 (39.52 ppm). All NMR spectra were analysed using the program *MestReNova* 10.0.1 from *Mestrelab Research S. L*. Low resolution mass spectra were measured on a LT Q FT-ICR by *Thermo Finnigan GmbH*. High resolution mass spectra were measured by the analytical section of the Department of Chemistry of the Ludwigs-Maximilians-Universität München on a MAT 90 (ESI) from *Thermo Finnigan GmbH*. IR spectra were recorded on a *PerkinElmer Spectrum BX II FT-IR* system. Substances were applied as a film or directly as solids on the ATR unit. Analytical RP-HPLC was performed on an analytical HPLC *Waters Alliance* (2695 Separation Module, 2996 Photodiode Array Detector) equipped with the column *Nucleosil* 120-2 C18 from *Macherey Nagel* applying eluent flow of 0.5 mL/min. Preparative RP-HPLC was performed on a HPLC *Waters Breeze* (2487 Dual λArray Detector, 1525 Binary HPLC Pump) equipped with the column VP 250/32 C18 from *Macherey Nagel*. A flowrate of 5 mL/min was applied.

Synthetic Procedures

Synthesis of 6-Aza-5-formyl-2´-deoxycytidine 3

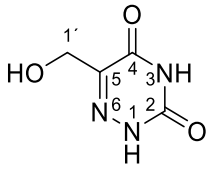
3-Bromopyruvic acid semicarbazone **5**



A solution of semicarbazide hydrochloride (3.34 g, 30.0 mmol, 1.0 eq.) and NaOAc (3.20 g, 38.6 mmol, 1.3 eq.) in 25.0 mL ddH_2O was added to a solution of 5.00 g 3-bromopyruvic acid (30.0 mmol, 1.0 eq.) in 15.0 mL glacial HOAc and 5.0 mL ddH_2O at 0 °C. The mixture was allowed to stand at room temperature for 2.5 h and at 0 °C for 19 h. The resulting white precipitate was filtered, washed with ice cold ddH_2O and Et_2O and dried subsequently under high vacuum to yield 3.26 g of the semicarbazone **5** as a mixture of (E)- and (Z)-isomers (14.6 mmol, 49 %) as a colorless solid.

^{1}H-NMR (400 MHz, DMSO-*d6*): δ/ppm = 10.58 (s, 0.6H, NN**H**) , 10.52 (s, 0.4H, NN**H**), 7.39 (br s, 0.8H, CON**H_2**), 6.92 (br s, 1.3H, CON**H_2**), 4.57 (s, 1.3H, 3-H), 4.46 (s, 0.8H, 3-H). **^{13}C-NMR** (101 MHz, DMSO-*d6*): δ/ppm = 163.51 (1-C), 155.63 (1´-C), 132.48 (2-C), 32.48 (3-C). **IR (ATR):** ν (cm^{-1}) = 3462 (w), 3245 (w), 2394 (br w), 1894 (br w), 1693 (m), 1445 (m), 1421 (s), 1364 (m), 1297 (m), 1234 (w), 1201 (s), 1164 (m), 1143 (m), 992 (m), 821 (s), 695 (m), 658 (s).

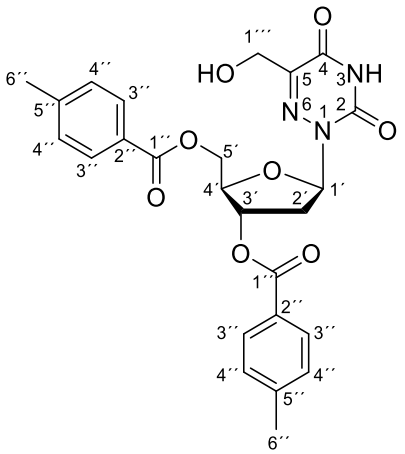
5-Hydroxymethyl-6-azauracil **6**



Compound **6** was synthesized according to a modified procedure of *Alekseeva et al.*[1] Under Ar atmosphere 1.40 g semicarbazone **5** (6.2 mmol, 1.0 eq.) were suspended in 50.0 mL $SOCl_2$ and two drops of pyridine were added. The reaction mixture was refluxed at 80 °C for 75 min, allowed to cool to room temperature, filtered through Celite and concentrated *in vacuo* to approximately 15 mL. Crystallization at 0 °C for 4 d resulted in a yellow precipitate which was filtered, dissolved in 30 ml dry CH_2Cl_2 and concentrated to dryness. After drying under high vacuum the obtained yellow solid was suspended in 20.0 mL ddH_2O and refluxed at 110 °C for 17 h. Subsequently, the solution was concentrated *in vacuo*, the residue was dissolved in 15.0 mL ddH_2O and lyophillized to yield 657 mg 5-hydroxy-6-azauracil **6** (4.60 mmol, 74 %) as a beige solid.

^{1}H-NMR (400 MHz, DMSO-*d6*): δ/ppm = 12.16 (s, 1H), 11.93 (s, 1H), 4.25 (s, 2H, 1´-H). **^{13}C-NMR** (101 MHz, DMSO-*d6*): δ/ppm = 156.75 (4-C), 149.55 (2-C), 144.33 (5-C), 58.02 (1´-C). **HRMS (ESI$^-$)**: calculated for $C_4H_4N_3O_3^-$ [M-H]$^-$ 142.0258; found: 142.0258. **IR (ATR):** ν (cm^{-1}) = 3429 (w), 3197 (w), 3034 (w), 2846 (w), 1682 (s), 1476 (m), 1446 (m), 1415 (m), 1244 (m), 1207 (m), 1061 (m), 1030 (m), 869 (m), 811 (m), 742 (s).

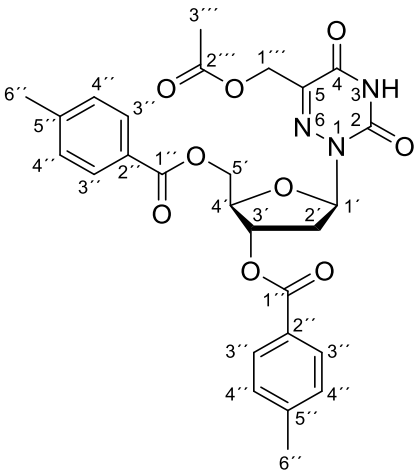
5-Hydroxymethyl-(3´,5´-di-*O*-p-toluoyl)-6-aza-2´-deoxyuridine **7**



In a heat-dried round bottom flask and under argon atmosphere 290 mg 5-hydroxymethyl-6-azauracil **6** (2.03 mmol, 1.0 eq.) were suspended in 10.64 mL hexamethyldisilazane and 351 µL TMSCl were added. The mixture was refluxed at 135 °C for 75 min, the resulting brown solution was cooled to room temperature and concentrated *in vacuo* at 30 °C. The oily residue was dried under high vacuum for 1 h and subsequently dissolved in 6.38 mL dry $CHCl_3$. 867 mg Hoffer´s chlorosugar (2.23 mmol, 1.1 eq.) were added and the reaction mixture was stirred at room temperature for 17 h. The reaction mixture was poured into 250 mL 0.2 M aq. HCl and extracted with EtOAc (3 x 200 mL). Combined organic layers were dried over Na_2SO_4 and concentrated to dryness. Purification via silica gel column chromatography (iHex:EtOAc 3:1 → iHex:EtOAc 1:1 → EtOAc) yielded a 9:1 mixture of diastereomeric β- and α- nucleosidation products (760 mg, 1.53 mmol, 75 %) as a colorless solid. Recrystallization from absolute EtOH yielded 558 mg of the pure β-nucleoside **7** (1.13 mmol, 56%) as a colorless solid.

^{1}H-NMR (600 MHz, $CDCl_3$): δ/ppm = 9.52 (s, 1H, 3-H), 7.96 – 7.88 (m, 4H, 3´´-H), 7.26 – 7.19 (m, 4H, 4´´-H), 6.73 (dd, *J* = 6.7, 6.6 Hz, 1H, 1´-H), 5.65 (ddd, *J* = 6.3, 3.1, 3.0 Hz, 1H, 3´-H), 4.73 (dd, *J* = 11.5, 7.0 Hz, 1H, 5´-H), 4.61 – 4.46 (m, 4H, 4´-H, 5´-H, 1´´´-H), 2.99 (ddd, *J* = 14.3, 6.7, 6.5 Hz, 1H, 2´-H), 2.50 (ddd, *J* = 14.2, 6.8, 3.3 Hz, 1H, 2´-H), 2.42 (s, 3H, 6´´-H), 2.38 (s, 3H, 6´´-H). **^{13}C-NMR** (151 MHz, $CDCl_3$): δ/ppm = 166.70 (1´´-C), 166.19 (1´´-C), 155.59 (4-C), 148.47 (2-C), 145.58 (5-C), 144.54 (5´´-C), 144.29 (5´´-C), 129.94 (3´´-C), 129.37 (4´´-C), 129.30 (4´´-C), 126.85 (2´´-C), 126.59 (2´´-C), 86.27 (1´-C), 82.74 (4´-C), 75.20 (3´-C), 64.36 (5´-C), 60.29 (1´´´-C), 35.04 (2´-C), 21.87 (6´´-C), 21.82 (6´´-C). **HRMS (ESI$^+$)**: calculated for $C_{25}H_{26}N_3O_8^+$ $[M+H]^+$ 496.1714, found: 496.1715; calculated for $C_{25}H_{29}N_4O_8^+$ $[M+NH_4]^+$ 513.1980, found: 513.1975; calculated for $C_{25}H_{25}N_3O_8Na^+$ $[M+Na]^+$ 518.1534, found: 518.1529. **HRMS (ESI$^-$)**: calculated for $C_{25}H_{24}N_3O_8^-$ $[M-H]^-$ 494.1566; found: 494.1572. **IR (ATR):** ν (cm^{-1}) = 1714 (s), 1611 (m), 1450 (w), 1273 (s), 1178 (m), 1105 (s), 1020 (w), 753 (m). **R_f** (iHex: EtOAc 1:1): 0.27.

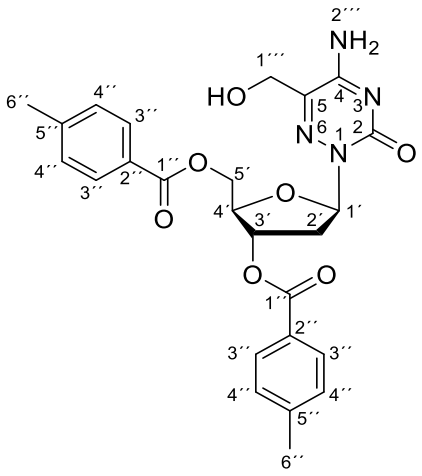
5-Methylacetate-(3´,5´-di-*O*-p-toluoyl)-6-aza-2´-deoxyuridine **8**



273 mg hydroxymethyl nucleoside **7** (0.55 mmol, 1.0 eq.) were dissolved in 3.30 mL dry pyridine and 220 µL acetic anhydride (2.33 mmol, 4.2 eq.) were added under argon atmosphere. The reaction mixture was stirred at room temperature for 22 h and stopped by addition of 5.0 mL absolute EtOH. The solution was concentrated *in vacuo*, coevaporated twice from dry toluene (2 x 15 mL) and the residue was purified by silica gel column chromatography (CH_2Cl_2 → CH_2Cl_2:MeOH 20:1) to obtain 283 mg of acetyl protected nucleoside **8** (0.53 mmol, 96 %) as a colorless foam.

^{1}H-NMR (600 MHz, $CDCl_3$): δ/ppm = 8.54 (s, 1H, 3-H), 7.94 – 7.88 (m, 4H, 3´´-H), 7.26 – 7.18 (m, 4H, 4´´-H), 6.67 (dd, *J* = 6.9, 5.4 Hz, 1H, 1´-H), 5.69 (ddd, *J* = 7.4, 4.5, 3.2 Hz, 1H, 3´-H), 5.04 (d, *J* = 14.0 Hz, 1H, 1´´´-H), 4.91 (d, *J* = 14.0 Hz, 1H, 1´´´-H), 4.58 – 4.49 (m, 3H, 4´-H, 5´-H), 2.96 (ddd, *J* = 14.1, 6.8, 5.4 Hz, 1H, 2´-H), 2.51 (ddd, *J* = 14.1, 7.0, 4.5 Hz, 1H, 2´-H), 2.43 (s, 3H, 6´´-H), 2.39 (s, 3H, 6´´-H), 2.15 (s, 3H, 3´´´-H). **^{13}C-NMR** (151 MHz, $CDCl_3$): δ/ppm = 170.33 (2´´´-C), 166.33 (1´´-C), 166.07 (1´´-C), 154.46 (4-C), 147.79 (2-C), 144.52 (5´´-C), 144.13 (5´´-C), 141.45 (5-C), 129.94 (3´´-C), 129.89 (3´´-C), 129.38 (4´´-C), 129.27 (4´´-C), 127.05 (2´´-C), 126.65 (2´´-C), 86.29 (1´-C), 82.54 (4´-C), 74.86 (3´-C), 64.33 (5´-C), 59.91 (1´´´-C), 35.38 (2´-C), 21.88 (6´´-C), 21.82 (6´´-C), 20.73 (3´´´-C). **HRMS (ESI$^+$):** calculated for $C_{27}H_{31}N_4O_9^+$ $[M+NH_4]^+$ 555.2086, found: 555.2019. **HRMS (ESI$^-$)**: calculated for $C_{27}H_{26}N_3O_9^-$ $[M-H]^-$ 536.1675; found: 536.1677. **IR (ATR)**: ν (cm^{-1}) = 1952 (w), 1711 (s), 1610 (m), 1442 (m), 1380 (m), 1269 (s), 1177 (m), 1100 (s), 1019 (m), 839 (w), 751 (s). **R_f** (CH_2Cl_2:MeOH 20:1): 0.52.

5-Hydroxymethyl-(3´,5´-di-*O*-p-toluoyl)-6-aza-2´-deoxycytidine **9**

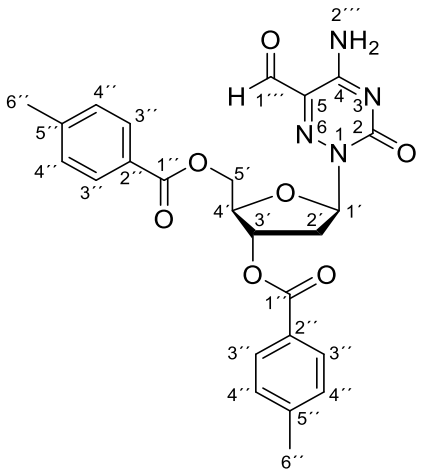


In a heat dried Schlenk-flask 487 mg 1,2,4-triazole (7.05 mmol, 9.0 eq.) were dissolved in 19.50 mL dry MeCN and cooled to 0 °C under Ar atmosphere. 146 µL $POCl_3$ (1.57 mmol, 2.0 eq.) were added dropwise and the mixture was stirred at 0 °C for 10 min. Subsequently, 966 µL NEt_3 (6.97 mmol, 8.9 eq.) were added and the mixture was stirred for another 20 min at 0 °C before 421 mg 2´-deoxyuridine derivative **8** (0.78 mmol, 1.0 eq.) were added. The reaction mixture was allowed to warm to room temperature and stirred for 18 h. After complete conversion the mixture was poured into 100 mL saturated aq. $NaHCO_3$ solution and extracted with CH_2Cl_2 (3 x 150 mL). Combined organic layers were dried over Na_2SO_4 and concentrated to dryness *in vacuo*.

The residue was dissolved in 15.60 mL 1,4-dioxane, 4.91 mL conc. NH_4OH were added and the mixture was stirred at 40 °C for 5 h. After cooling to room temperature the solution was poured into saturated aq. NH_4Cl solution (100 mL) and extracted with CH_2Cl_2 (3 x 100 mL). Combined organic layers were dried over Na_2SO_4, concentrated to dryness *in vacuo* and purified by silica gel column chromatography (CH_2Cl_2:MeOH 80:1 → CH_2Cl_2:MeOH 60:1 → CH_2Cl_2:MeOH 50:1 → CH_2Cl_2:MeOH 30:1 → CH_2Cl_2:MeOH 20:1 → CH_2Cl_2:MeOH 15:1) to obtain 324 mg of the 2´-deoxycytidine derivative **9** (0.66 mmol, 84 %) as an off-white solid.

^{1}H-NMR (600 MHz, $CDCl_3$): δ/ppm = 9.96 (s, 1H, 2´´´-H), 9.21 (s, 1H, 2´´´-H), 7.91 – 7.84 (m, 4H, 3´´-H), 7.24 – 7.16 (m, 4H, 4´´-H), 6.67 (d, *J* = 9.6 Hz, 1H, 1´-H), 5.64 (dd, *J* = 6.5, 3.4 Hz, 1H, 3´-H), 4.66 – 4.48 (m, 6H, 4´-H, 5´-H, 1´´´-H), 3.02 – 2.94 (m, 1H, 2´-H), 2.56 – 2.48 (m, 1H, 2´-H), 2.39 (s, 3H, 6´´-H), 2.36 (s, 3H, 6´´-H). **^{13}C-NMR** (151 MHz, $CDCl_3$): δ/ppm = 166.45 (1´´-C), 166.15 (1´´-C), 154.50 (4-C), 144.47 (5´´-C), 144.30 (5´´-C), 137.91 (5-C), 129.94 (3´´-C), 129.87 (3´´-C), 129.42 (4´´-C), 129.36 (4´´-C), 126.91 (2´´-C), 126.63 (2´´-C), 87.22 (1´-C), 83.01 (4´-C), 75.07 (3´-C), 64.27 (5´-C), 60.98 (1´´´-C), 35.38 (6´´-C), 21.85 (6´´-C). **HRMS (ESI$^+$)**: calculated for $C_{25}H_{27}N_4O_7^+$ $[M+H]^+$ 495.1874, found: 495.1872; calculated for $C_{25}H_{26}N_4O_7Na^+$ $[M+Na]^+$ 517.1694, found: 517.1691. **HRMS (ESI$^-$)**: calculated for $C_{25}H_{25}N_4O_7^-$ $[M-H]^-$ 493.1729; found: 493.1734. **IR (ATR)**: ν (cm^{-1}) = 3229 (br w), 1716 (s), 1611 (m), 1450 (w), 1272 (s), 1178 (m), 1104 (s), 1020 (m), 752 (s). **R_f** (CH_2Cl_2:MeOH 20:1): 0.26.

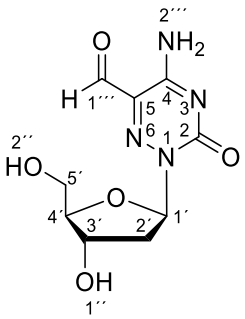
5-Formyl-(3´,5´-di-*O*-p-toluoyl)-6-aza-2´-deoxycytidine **10**



In a heat dried Schlenk flask 306 mg of 5-hydroxymethyl-2´-deoxycytdidine derivative **9** (0.62 mmol, 1.0 eq.) were dissolved in 18.6 mL dry CH_2Cl_2 and cooled to -15 °C in a NaCl/ice cooling bath. Subsequently, 289 mg Dess-Martin periodinane (0.68 mmol, 1.1 eq.) were added, the mixture was allowed to warm to room temperature and stirred for 1 h at room temperature. After complete conversion the reaction was stopped by addition of a solution of 294 mg $Na_2S_2O_3$ (1.86 mmol, 3.0 eq.) in 100 mL saturated aq. $NaHCO_3$ and extracted with with EtOAc (3 x 150 mL). Combined organic layers were dried over Na_2SO_4, concentrated to dryness *in vacuo* and purified by silica gel column chromatography (CH_2Cl_2 → CH_2Cl_2:MeOH 40:1) to yield 272 mg of aldehyde **10** (0.55 mmol, 89 %) as a yellow foam.

^{1}H-NMR (600 MHz, $CDCl_3$): δ/ppm = 9.40 (s, 1H, 1´´´-H), 8.31 (s, 1H, 2´´´-H), 7.94 – 7.86 (m, 4H, 3´´-H), 7.28 – 7.20 (m, 4H, 4´´-H), 6.87 (dd, *J* = 5.9 Hz, 1H, 1´-H), 5.74 (dd, *J* = 5.9, 2.9 Hz, 1H, 3´-H), 4.72 – 4.52 (m, 3H, 4´-H, 5´-H), 3.01 – 2.94 (m, 1H, 2´-H), 2.66 – 2.59 (m, 1H, 2´-H), 2.43 (s, 3H, 6´´-H), 2.41 (s, 3H, 6´´-H). **^{13}C-NMR** (151 MHz, $CDCl_3$): δ/ppm = 188.81 (1´´´-C), 166.26 (1´´-C), 166.17 (1´´-C), 155,17 (4-C), 144.69 (5´´-C), 144.43 (5´´-C), 129.97 (3´´-C), 129.83 (3´´-C), 129.44 (4´´-C), 129.37 (5-C), 126.93 (2´´-C), 126.48 (2´´-C), 88.08 (1´-C), 83.34 (4´-C), 74.73 (3´-C), 63.87 (5´-C), 36.14 (2´-C), 21.89 (6´´-C), 21.84 (6´´-C). **HRMS (ESI$^+$)**: calculated for $C_{25}H_{25}N_4O_7^+$ $[M+H]^+$ 493.1718, found: 493.1717; calculated for $C_{25}H_{24}N_4O_7Na^+$ $[M+Na]^+$ 515.1537, found: 515.1534. **HRMS (ESI$^-$)**: calculated for $C_{25}H_{23}N_4O_7^-$ $[M-H]^-$ 491.1572; found: 491.1577. **IR (ATR)**: ν (cm^{-1}) = 3237 (br w), 1720 (s), 1612 (m), 1272 (s), 1178 (w), 1104 (m), 753 (m). **R_f** (CH_2Cl_2:MeOH 20:1): 0.34.

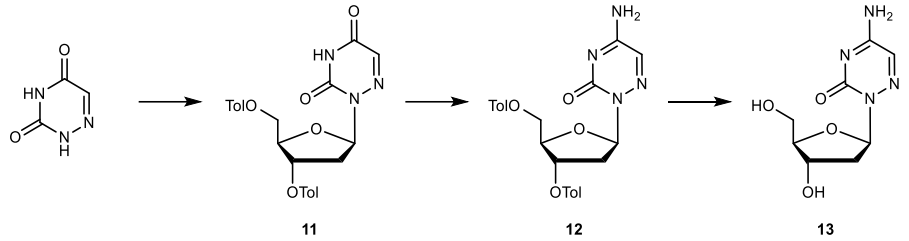
6-Aza-5-formyl-2´-deoxycytidine **3**



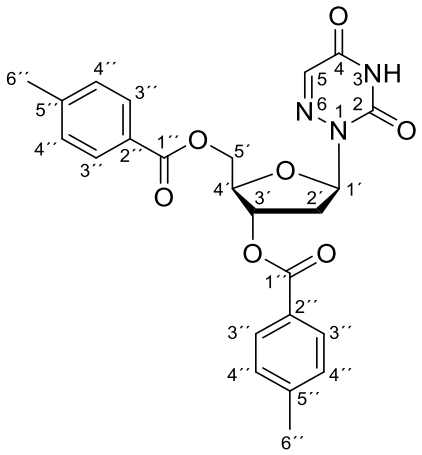
The deprotcetion was performed according to a modified procedure of *Mitchell et al.*[2] 36 mg of toluoyl protected 6-Aza-nucleoside **10** (73 µmol, 1.0 eq.) were dissolved in 0.73 mL dry MeOH and 0.73 dry benzene. To the solution 540 µL 0.5 M NaOMe in dry MeOH (270 µmol, 3.7 eq.) were added and the reaction mixture was stirred for 1.5 h at room temperature. After complete conversion the reaction mixture was neutralized with 2 M aq. HCl and evaporated to dryness. The resulting residue was dissolved in 10 mL ddH_2O and extracted with CH_2Cl_2 (5 x 10 mL). The aqueous layer was lyophyllized and the residue was purified by *reversed phase* HPLC (0 % → 3 % MeCN in H_2O in 45 min) to yield 10.1 mg of the desired product **3** (39 µmol, 54 %) as a colorless solid.

^{1}H-NMR (600 MHz, DMSO-*d6*): δ/ppm = 9.56 (s, 1H, 1´´´-H), 8.61 (s, 1H, 2´´´-H), 7.78 (s, 1H, 2´´´-H), 6.52 (dd, *J* = 7.0, 5.0 Hz, 1H, 1´-H), 5.21 (d, *J* = 4.7 Hz, 1H, 1´´-H), 4.63 (t, *J* = 5.9 Hz, 1H, 2´´-H), 4.35 (ddd, *J* = 10.2, 5.3, 5.1 Hz, 1H, 3´-H), 3.76 (dd, *J* = 5.2, 5.1 Hz, 1H, 4´-H), 3.49 (dt, *J* = 11.2, 5.5 Hz, 1H, 5´-H), 3.41 (dt, *J* = 11.8, 6.0 Hz, 1H, 5´-H), 2.47 (ddd, *J* = 13.3, 6.3, 4.9 Hz, 1H, 2´-H), 2.15 (ddd, *J* = 13.1, 7.0, 5.3 Hz, 1H, 2´-H). **^{13}C-NMR** (151 MHz, DMSO-*d6*): δ/ppm = 189.77 (1´´´-C), 156.56 (4-C), 151.66 (2-C), 129.61 (5-C), 87.87 (4´-C), 86.48 (1´-C), 70.50 (3´-C), 62.07 (5´-C), 37.74 (2´-C). **HRMS (ESI$^+$)**: calculated for $C_9H_{13}N_4O_5^+$ $[M+H]^+$ 257.0881, found: 257.0880; calculated for $C_9H_{12}N_4O_5Na^+$ $[M+Na]^+$ 279.0700, found: 279.0698. **HRMS (ESI$^-$)**: calculated for $C_9H_{11}N_4O_5^-$ $[M-H]^-$ 255.0735; found: 255.0735.

Synthesis of 6-Aza-2´-deoxycytidine



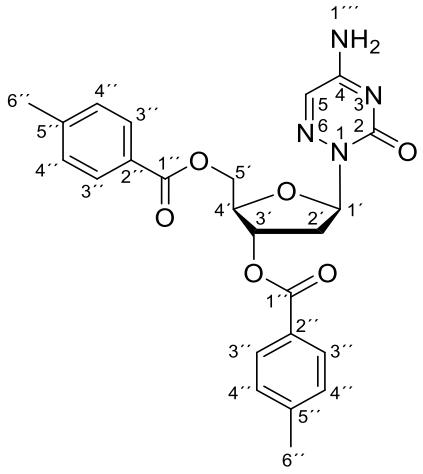
(3´,5´-Di-*O*-p-toluoyl)-6-aza-2´-deoxyuridine **11**



In a heat-dried round bottom flask and under argon atmosphere 300 mg 6-azauracil (2.65 mmol, 1.0 eq.) were suspended in 2.88 mL hexamethyldisilazane and 100 µL TMSCl were added. The mixture was refluxed at 145 °C for 1 h and concentrated *in vacuo* at 30 °C after cooling to room temperature. The obtained dark oil was dried under high vacuum for 1 h and subsequently dissolved in 8.33 mL dry $CHCl_3$. 1.13 g Hoffer´s chlorosugar (2.92 mmol, 1.1 eq.) were added and the reaction mixture was stirred at room temperature for 16 h. The reaction mixture was poured into 250 mL saturated aq. $NaHCO_3$ solution and extracted with EtOAc (3 x 200 mL). Combined organic layers were dried over Na_2SO_4 and concentrated to dryness. Purification via silica gel column chromatography (iHex:EtOAc 3:1 → iHex:EtOAc 1:1 → iHex:EtOAc 1:2) yielded a mixture of diastereomeric β- and α- nucleosidation products (868 mg, 1.86 mmol, 70 %) as a colorless solid. Recrystallization from absolute EtOH yielded 742 mg of the pure β-nucleoside **11** (1.59 mmol, 60%) as a colorless solid.

^{1}H-NMR (400 MHz, $CDCl_3$): δ/ppm = 8.82 (s, 1H,3-H), 7.97 – 7.91 (m, 4H, 3´´-H), 7.29 – 7.20 (m, 5H, 5-H, 4´´-H), 6.68 (dd, *J* = 6.9, 6.7 Hz, 1H, 1´-H), 5.72 (ddd, *J* = 6.8, 3.9, 3.5 Hz, 1H, 3´-H), 4.66 (dd, *J* = 11.6, 4.3 Hz, 1H, 5´-H), 4.53 (dd, *J* = 9.7, 4.4 Hz, 1H, 4´-H), 4.46 (dd, *J* = 11.6, 5.0 Hz, 1H, 5´-H), 2.98 (dt, *J* = 13.4, 6.5 Hz, 1H, 2´-H), 2.48 (ddd, *J* = 14.1, 6.8, 3.9 Hz, 1H, 2´-H), 2.43 (s, 3H, 6´´-H), 2.41 (s, 3H, 6´´-H). **^{13}C-NMR** (101 MHz, $CDCl_3$): δ/ppm = 166.36 (1´´-C), 166.16 (1´´-C), 155.30 (4-C), 147.77 (2-C), 144.56 (5´´-C), 144.11 (5´´-C), 136.10 (5-C), 129.94 (3´´-C), 129.90 (3´´-C), 129.39 (4´´-C), 129.27 (4´´-C), 127.11 (2´´-C), 126.61 (2´´-C), 86.05 (1´-C), 82.73 (4´-C), 75.03 (3´-C), 64.06 (5´-C), 35.17 (2´-C), 21.90 (6´´-C), 21.86 (6´´-C). **HRMS (ESI$^+$)**: calculated for $C_{24}H_{24}N_3O_7^+$ $[M+H]^+$ 466.1609, found: 466.1610; calculated for $C_{24}H_{27}N_4O_7^+$ $[M+NH_4]^+$ 483.1874, found: 483.1874. **HRMS (ESI$^-$)**: calculated for $C_{24}H_{22}N_3O_7^-$ $[M-H]^-$ 464.1463; found: 464.1469. **IR (ATR)**: ν (cm^{-1}) = 3200 (w), 1727 (s), 1699 (s), 1610 (w), 1437 (w), 1392 w), 1374 (w), 1329 (m), 1265 (s), 1173 (w), 1100 (s), 1080 (s), 954 (m), 808 (m), 752 (s). **R_f** (iHex: EtOAc 1:1): 0.44.

(3´,5´-Di-*O*-p-toluoyl)-6-aza-2´-deoxycytidine **12**

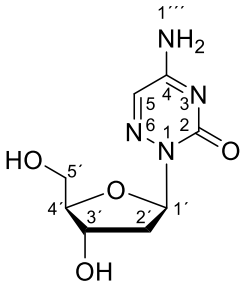


In a heat dried Schlenk-flask 668 mg 1,2,4-triazole (9.67 mmol, 9.0 eq.) were dissolved in 26.80 mL dry MeCN and cooled to 0 °C under Ar atmosphere. 201 µL $POCl_3$ (2.15 mmol, 2.0 eq.) were added dropwise and the mixture was stirred at 0 °C for 10 min. Subsequently, 1.33 mL NEt_3 (9.56 mmol, 8.9 eq.) were added and the mixture was stirred for another 20 min at 0 °C before 500 mg of protected 6-aza-2´-deoxyuridine **11** (1.07 mmol, 1.0 eq.) were added. The reaction mixture was allowed to warm to ambient temperature and stirred for 17 h at room temperature. After complete conversion the mixture was poured into 250 mL saturated aq. $NaHCO_3$ solution and extracted with CH_2Cl_2 (3 x 200 mL). Combined organic layers were washed with brine (150 mL), dried over Na_2SO_4 and concentrated to dryness *in vacuo*.

The residue was dissolved in 21.40 mL 1,4-dioxane and 6.74 mL conc. NH_4OH were added. After stirring the mixture at 45 °C for 3 h the solution was poured into saturated aq. NH_4Cl solution (150 mL) and extracted with EtOAc (3 x 150 mL). Combined organic layers were washed with brine (150 mL), dried over Na_2SO_4 and concentrated to dryness *in vacuo*. The crude product was purified by silica gel column chromatography (CH_2Cl_2:MeOH 40:1 → CH_2Cl_2:MeOH 20:1 → CH_2Cl_2:MeOH 10:1) to obtain 460 mg of the amination product **12** (0.99 mmol, 93 %) as a slightly yellowish solid.

^{1}H-NMR (600 MHz, $CDCl_3$): δ/ppm = 8.57 (s, 1H, 1´´´-H)), 7.92 (m, 4H, 3´´-H), 7.71 (s, 1H, 5-H), 7.32 (s, 1H, 1´´´-H), 7.25 – 7.16 (m, 4H, 4´´-H), 6.72 (dd, *J* = 6.6, 6.3 Hz, 1H, 1´-H), 5.73 (ddd, *J* = 7.2, 3.9, 3.1 Hz, 1H, 3´-H), 4.58 (ddd, *J* = 11.1, 5.4, 1.4 Hz, 1H, 5´-H), 4.55 – 4.52 (m, 1H, 4´-H), 4.47 (ddd, *J* = 11.2, 5.2, 1.2 Hz, 1H, 5´-H), 3.01 (ddd, *J* = 13.6, 6.8, 6.4 Hz, 1H, 2´-H), 2.45 (ddd, *J* = 13.9, 6.9, 4.2 Hz, 1H, 2´-H), 2.41 (s, 3H, 6´´-H), 2.36 (s, 3H, 6´´-H). **^{13}C-NMR** (151 MHz, $CDCl_3$): δ/ppm = 166.60 (1´-C), 166.20 (1´´-C), 159.07 (4-C), 154.24 (2-C), 144.34 (5´´-C), 143.96 (5´´-C), 129.94 (3´´-C), 129.94 (3´´-C), 129.32 (4´´-C), 129.21 (4´´-C), 128.61 (5-C), 127.10 (2´´-C), 126.84 (2´´-C), 87.76 (1´-C), 82.23 (4´-C), 75.41 (3´-C), 64.55 (5´-C), 35.17 (2´-C), 21.85 (6´´-C), 21.79 (6´´-C). **HRMS (ESI$^+$)**: calculated for $C_{24}H_{25}N_4O_6^+$ $[M+H]^+$ 465.1769, found: 465.1770. **HRMS (ESI$^-$)**: calculated for $C_{24}H_{23}N_4O_6^-$ $[M-H]^-$ 463.1623; found: 463.1630**. IR (ATR)**: ν (cm^{-1}) = 3305 (br w), 3059 (br w), 1715 (m), 1643 (m), 1610 (m), 1536 (w), 1466 (w), 1334 (w), 1265 (s), 1178 (m), 1102 (s) 1020 (m), 961 (w), 839 (w), 730 (s). **R_f** (CH_2Cl_2:MeOH 20:1): 0.30.

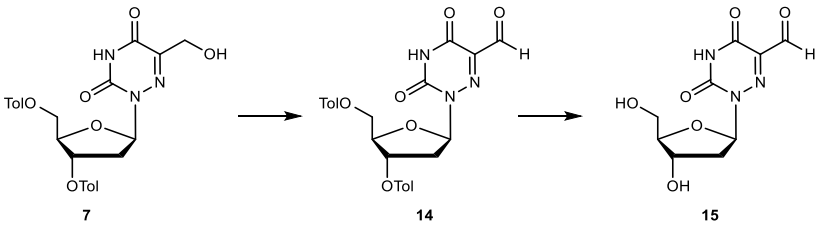
6-Aza-2´-deoxycytidine **13**



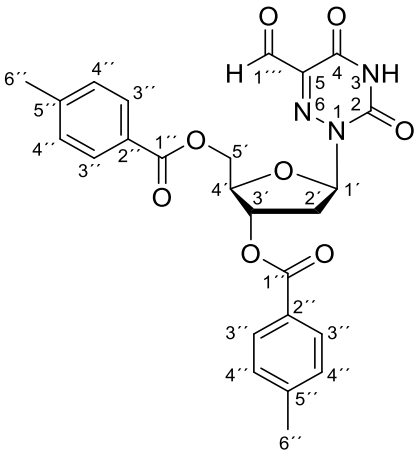
In a round bottom flask 10.0 mL of a solution of $MeNH_2$ in EtOH (33%) were added to 200 mg of toluoyl protected compound **12** (0.43 mmol, 1.0 eq.) and the mixture was stirred at room temperature for 14 h. As reaction control still showed protected starting material another 5.0 mL 33 % $MeNH_2$ in EtOH were added and the mixture was heated to 45 °C for 1 h. The solution was evaporated to dryness, dissolved in 10 mL ddH_2O and 10 mL CH_2Cl_2. Phases were separated, the aqueous layer was extracted with CH_2Cl_2 (4 x 10 mL) and subsequently concentrated *in vacuo*. The crude product was dissolved in ddH_2O (24 mL) and purified by reversed phase HPLC (0 % → 5 % MeCN in H_2O in 45 min) to yield 57.6 mg of the deprotected 6-Aza-nucleoside **13** as a white solid (0.25 mmol, 58 %).

^{1}H-NMR (600 MHz, D_2O): δ/ppm = 7.75 (s, 1H, 5-H), 6.57 (dd, *J* = 7.2, 4.8 Hz, 1H, 1´-H), 4.57 (ddd, *J* = 6.9, 4.4, 1.7 Hz, 1H, 3´-H), 4.05 – 4.01 (m, 1H, 4´-H), 3.85 – 3.69 (m, 1H, 5´-H), 3.71 – 3.61 (m, 1H, 5´-H), 2.72 (ddd, *J* = 13.9, 7.1, 4.8, 2.0 Hz, 1H, 2´-H), 2.35 (ddd, *J* = 14.7, 7.4, 4.0, 1.9 Hz, 1H, 2´-H). **^{13}C-NMR** (151 MHz, D_2O): δ/ppm = 158.97 (4-C), 155.80 (2-C), 129.37 (5-C), 86.52 (1´-C), 86.44 (4´-C), 70.93 (3´-C), 61.80 (5´-C), 36.67 (2´-C). **HRMS (ESI$^-$)**: calculated for $C_8H_{11}N_4O_4^-$ $[M-H]^-$ 227.0786; found: 227.0787.

Synthesis of 6-Aza-5-formyl-2´-deoxyuridine



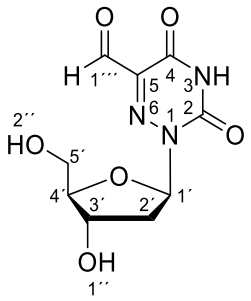
5-Formyl-(3´,5´-di-*O*-p-toluoyl)-6-aza-2´-deoxyuridine **14**



In a Schlenk flask 158 mg of 5-hydroxymethyl-2´-deoxyuridine derivative **7** (0.32 mmol, 1.0 eq.) were dissolved in 9.6 mL dry CH_2Cl_2 and cooled to -15 °C. At this temperature 162 mg Dess-Martin periodinane (0.38 mmol, 1.2 eq.) were added, the mixture was allowed to warm to room temperature slowly and was furthe stirred for 2.5 h. The reaction was stopped by addition of a solution of 182 mg $Na_2S_2O_3$ (1.15 mmol, 3.6 eq.) in 50 mL saturated aq. $NaHCO_3$ and the mixture was extracted with with EtOAc (3 x 100 mL). Combined organic layers were dried over Na_2SO_4, concentrated to dryness *in vacuo* and purified by silica gel column chromatography (iHex:EtOAc 1:1 → EtOAc) to yield 152 mg of aldehyde **14** (0.31 mmol, 96 %) as an off-white foam.

^{1}H-NMR (400 MHz, $CDCl_3$): δ/ppm = 9.58 (s, 1H, 1´´´-H), 9.02 (s, 1H, 3H), 8.01 – 7.79 (m, 4H, 3´´-H), 7.29 – 7.18 (m, 4H, 4´´-H), 6.70 (t, *J* = 6.3 Hz, 1H, 1´-H), 5.72 (dd, *J* = 6.7, 3.4 Hz, 1H, 3´-H), 4.70 – 4.49 (m, 3H, 4´-H, 5´-H), 3.07 (dt, *J* = 14.2, 6.4 Hz, 1H, 2´-H), 2.58 (ddd, *J* = 14.2, 6.7, 4.0 Hz, 1H, 2´-H), 2.43 (s, 3H, 6´´-H), 2.40 (s, 3H, 6´´-H). **^{13}C-NMR** (101 MHz, $CDCl_3$): δ/ppm = 184.16 (1´´´-C), 166.33 (1´´-C), 166.14 (1´´-C), 152.79 (4-C), 147.15 (2-C), 144.68 (5´´-C) , 144.40 (5´´-C), 137.22 (5-C), 129.94 (3´´-C), 129.81 (3´´-C), 129.42 (4´´-C), 129.40 (4´´-C), 126.81 (2´´-C), 126.41 (2´´-C), 87.16 (1´-C), 83.20 (4´-C), 74.59 (3´-C), 63.74 (5´-C), 35.40 (2´-C), 21.91 (6´´-C), 21.85 (6´´-C). **HRMS (ESI$^+$)**: calculated for $C_{25}H_{27}N_4O_8^+$ $[M+NH_4]^+$ 511.1823, found:511.1823. **HRMS (ESI$^-$)**: calculated for $C_{25}H_{22}N_3O_8^-$ $[M-H]^-$ 492.1412; found: 492.1413**. IR (ATR)**: ν (cm^{-1}) = 2963 (w), 1712 (s), 1611 (m), 1439 (w), 1398 (w), 1309 (w), 1260 (s), 1178 (m), 10963 (s), 1020 (s), 909 (m), 800 (s), 752 (s), 732 (s). **R_f** (iHex: EtOAc 1:1): 0.17.

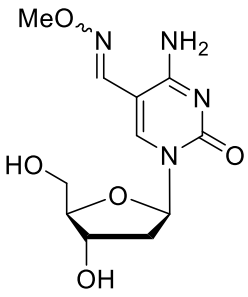
5-Formyl-6-aza-2´-deoxyuridine **15**



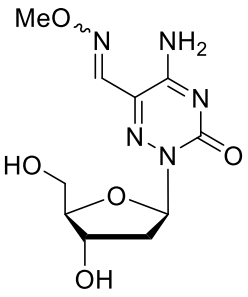
97 mg of toluoyl protected 6-Aza-nucleoside **14** (0.20 mmol, 1.0 eq.) were dissolved in 2.0 mL dry MeOH and 2.0 dry benzene. To the solution freshly prepared 0.5 M NaOMe in dry MeOH (1.46 mL, 0.73 mmol, 3.7 eq.) was added and the reaction mixture was stirred for 2 h at room temperature. Subsequently, the reaction mixture was neutralized with 2 M aq. HCl, evaporated to dryness and redissolved in 10 mL ddH_2O. The mixture was extracted with CH_2Cl_2 (4 x 10 mL). The aqueous layer was lyophyllized and the residue was purified by *reversed phase* HPLC (0 % → 3 % MeCN in H_2O in 45 min) to yield 16.3 mg of the product **15** (63 µmol, 32 %) as a colorless solid.

^{1}H-NMR (400 MHz, D_2O): δ/ppm = 9.62 (s, 0.2H, 1´´´-H), 6.55 (dd, *J* = 7.2, 4.5 Hz, 0.2H, 1´-H), 6.48 (dd, *J* = 7.2, 5.2 Hz, 0.8H, 1´´-H), 5.87 (s, 0.7H, 1´´´-H), 4.62 – 4.48 (m, 1H, 3´-H), 3.96 (m, 1H, 4´-H), 3.75 – 3.55 (m, 2H, 5´-H), 2.71 (m, 1H, 2´-H), 2.28 (m, 1H, 2´-H). **^{13}C-NMR** (101 MHz, D_2O): δ/ppm = 190.31 (1´´´-C), 165.12 (4-C), 162.87 (4-C)), 156.31 (2-C), 155.12 (2-C), 142.78, 137.09 (5-C), 87.81 (4´-C), 86.93 (4´-C), 86.48 (1´-C), 86.42 (1´-C), 86.14 (1´´´-C), 71.05 (3´-C), 70.83 (3´-C), 61.81 (5´-C), 61.70 (5´-C), 36.89 (2´-C), 36.57 (2´-C). **HRMS (ESI$^-$)**: calculated for $C_9H_{10}N_3O_6^-$ $[M-H]^-$ 256.0575; found: 256.0576.

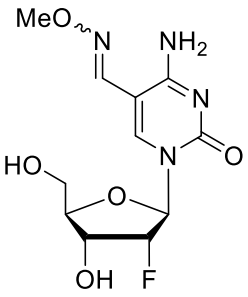
Reactions of **1**, **2**, **3** and **15** with Methoxyamine



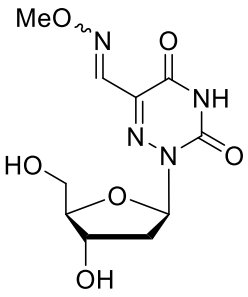
16



17



18



19

To a solution of 100 nmol nucleoside (1.0 eq.) in 35.0 µL aq. NaOH (pH = 10) was added 33.33 µL aq. $MeONH_2$ (150 mM, 50.0 eq.) and the mixture was incubated at 25 °C for 1 h. The reaction was stopped by the addition of 35.0 µL of aq. HCOOH (pH = 3) and analyzed by reversed phase HPLC. The resulting reaction products were collected and analyzed via HRMS.

	16	**17**	**18**	**19**
Conversion of SM	50 %	quant.	57 %	quant.
HPLC gradient [MeCN in H_2O in 45 min]	0 % → 15 %	0 % → 30 %	0 % → 13 %	0 % → 30 %
HRMS (calcd.) $ESI^{(+)}$	285.1194 $C_{11}H_{17}O_5N_4^+$ $[M+H]^+$	286.1146 $C_{10}H_{16}O_5N_5^+$ $[M+H]^+$	303.1099 $C_{11}H_{16}O_5N_4F^+$ $[M+H]^+$	
$ESI^{(-)}$	283.1048 $C_{11}H_{15}O_5N_4^-$ $[M-H]^-$	284.1000 $C_{10}H_{14}O_5N_5^-$ $[M-H]^-$		285.0841 $C_{10}H_{13}O_6N_4^-$ $[M-H]^-$
HRMS (found) $ESI^{(+)}$	285.1193 $[M+H]^+$	286.1151 $[M+H]^+$	303.1098 $[M+H]^+$	
$ESI^{(-)}$	283.1048 $[M-H]^-$	284.1000 $[M-H]^-$		285.0840 $[M-H]^-$

Reaction of **1**, **2** and **3** with $NaHSO_3$

20 mM nucleoside solution in ddH_2O was shaken with 4.36 M aq. $NaHSO_3$ (435 eq, pH = 5) at 55 °C. Samples were taken at time points of 0, 5 min, 20 min, 30 min, 40 min, 1 h, 2 h, 3 h, 4 h, 5 h, 6 h and 15 h. Bisulfite was cleaved by preparing 1 mM solution of reaction mixture and 2 M NaOH (1700 eq), shaken for 10 min at rt and followed by immediate injection to HPLC (Gradients: **1** 0% to 13% MeCN in H_2O in 45 min, **2** 0% to 15% MeCN in H_2O in 45 min, **3** 0% to 3% MeCN in H_2O in 45 min).

Chromatograms of each time point were normalized to the integral of the reagent and corrected accordingly. Conversion was determined by the decrease of the corrected integrals of the starting materials.

Cell culture

DMEM high glucose (Sigma Aldrich) containing 10% heat-inactivated FBS (Gibco), 100 U/mL penicillin, 100 µg/mL streptomycin, 2 mM L-glutamine were used to culture Neuro-2a and RBL-2H3 cell lines. DMEM high glucose and Ham's Nutrient Mixture F12 (Sigma Aldrich) containing 10% heat-inactivated FBS (Gibco), 100 U/mL penicillin, 100 µg/mL streptomycin, 2 mM L-glutamine were used to culture CHO-K1 cells. Metabolic-labelling experiments were conducted by plating cells in their dedicated medium containing 350 µM of F-fdC or a-fdC (in case of the co-feeding study: 350 µM of each) for three days.

DMEM high glucose (Sigma Aldrich) containing 10% FBS (PAN Biotech), 2 mM L-glutamine, 1x MEM Non-essential Amino Acid Solution and 0.1 mM β-mercaptoethanol (Sigma Aldrich) were used as basal medium for E14 TDG +/- and -/- (obtained from Cortázar *et al.)* mESC cultures. The mESC lines were maintained in naïve state on gelatin coated plates by supplementing basal medium with 1000 U/mL LIF (ORF Genetics), 3.0 µM GSK3 inhibitor CHIR99021 and 1.0 µM MNK inhibitor CGP57380 (a2i medium). Metabolic labelling experiments with isotope-labeled nucleosides were performed by plating mESCs in priming conditions consisted of basal mESC medium supplemented with 3. µM GSK3 inhibitor CHIR99021 and 2.5 µM tankyrase inhibitor IWR1-endo. Labelled nucleosides were added at the concentration of 350 µM each to the primming medium over three days.

All samples were washed with PBS (Sigma Aldrich) once befor harvesting and lysed directly in the plates by adding RLT buffer (Qiagen) supplemented with 400 µM 2,6-di-tert-butyl-4-methylphenol (BHT) and desferoxamine mesylate (DM). Next, DNA was sheered in MM400 bead mill (Retsch) at 30 Hz for 1 min in 2 mL microcentrifuge tubes (Eppendorf) with 5 mm diamiter stainless steel beads (one per tube) and centrifuged at 21000rcf for 5 minutes. Genomic DNA (gDNA) was extrated using Zymo Quick gDNA mini-prep® kit according to the protocol with an addition of RNAse A treatment (Qiagen) at 0.2 mg/mL in Genomic Lysis Buffer – 400 µL of the solution transferred directly on the column and incubated for 15 minutes. All samples were eluted in DNAse-free ddH_2O with 20 µM BHT, the concentration of gDNA was measured on Nanodrop.

While investigating the deformylation levels using F-fdC as a metabollic label, the following conditions were used for the gDNA digestion procedure: Degradase (1.5 µL) per up to 10 µg of F-fdC labelled gDNA in 35 µL in H_2O. The digestion misture was incubated at 37 ^oC for 4 hours. Then, the sampes were filtered using AcroPrep Advance 0.2 µm 96-well filter plate (Pall Life Sciences) prior to LC-MS/MS analysis (39 µL injection volume at 4 °C). This method was compared to NEB enzyme mix digestion (described below) and showed no

difference in the QQQ measurement outcome, thus we concluded that both can be used at equal efficiency.

Chemical labelling and spiking

Chemical labelling of a-fdC was needed to receive a sharper signal during the MS analysis. Therefore, methoxyamine was used as a derivatization reagent. We optimized the reaction conditions including time and temperature.

The quantification of the the Fluoro compounds as well as the other known nucleosides was done by spiking of heavy labelled nucleosides which were synthesized by ourselves.

For the determination of a-fdC in the gDNA samples we did an external calibration.

Analysis of labelled and unlabelled products of F-fdC and a-fdC

The analysis was performed using an UHPLC-QQQ-MS/MS system consisting of a Triple Quad™ 6490 mass spectrometer (Agilent) with an ESI source and an Agilent Infinity 1290 UHPLC. The elution was monitored at 260 nm (Agilent InfinityLab Deuterium Lamp G1314). Data Acquisition and processing were performed using MassHunter Workstation Software Version B.07.01 (Agilent).
The UHPLC separation was performed for the Fluoro-fdC as well as mdC, hmdC, fdC on an InfinityLab Poroshell 120 SB-C8 column (2.1 mm x 150 mm, 2.7 µm, Agilent Technologies, USA) at 35 °C. Water containing 0.0085% FA (v/v, solvent A) and MeCN containing 0.0085% FA (v/v, solvent B) was used as the mobile phase. A gradient of 0 - 3.5% B for 4 min, 3.5 - 5% B for 2.9 min, 5 - 80% B for 0.3 min, 80% B for 3.3 min was used. The flow rate of the mobile phase was set to 0.35 mL min^{-1}.

The derivatized a-fdC, F-fdC and fdC were separated on an InfinityLab Poroshell 120 SB-C18 column (2.1 mm x 150 mm, 2.7 µm, Agilent Technologies, USA) at 35 °C. Water containing 0.0085% FA (v/v, solvent A) and MeCN containing 0.0085% FA (v/v, solvent B) was used as the mobile phase. A gradient of 0% B for 3 min, 0 - 15% B for 27 min, 15 - 100% B for 5 min was used. The flow rate of the mobile phase was set to 0.35 mL min^{-1}.

The mass spectrometry detection was performed under positive ESI mode. The nucleosides and labelled products were monitored using the multiple reaction monitoring (MRM) mode. The MRM parameters were optimized to achieve maximal detection sensitivity (Tables 2 and 3).

Table 2 MRM parameters for the detection of F-fdC and its derivatives

Name	Transition	Scan	Type	Precursor Ion	Product Ion	Ion Polarity
UV-dG	0,0 -> 0,0	MRM	Target	0,0	0,0	Positive
UV-dC	0,0 -> 0,0	MRM	Target	0,0	0,0	Positive
UV total	0,0 -> 0,0	MRM	Target	0,0	0,0	Positive
Fluoro-fdC-dN	274,1 -> 140,1	MRM	Target	274,1	140,1	Positive
Fluoro-fdC-dN-$^{15}N_2$	276,1 -> 142,0	MRM	ISTD	276,1	142,0	Positive
Fluoro-dU	245,1 -> 225,1	MRM	Target	245,1	225,1	Negative
Fluoro-dC	246,1 -> 112,1	MRM	Target	246,1	112,1	Positive
F-fdU	273,1 -> 253,1	MRM	Target	273,1	253,1	Negative
fdC-dN-$^{15}N_2$	258,1 -> 142,0	MRM	ISTD	258,1	142,0	Positive
fdC-dN	256,1 -> 140,1	MRM	Target	256,1	140,1	Positive
$^{15}N_2$-FdC	248,1 -> 114,0	MRM	ISTD	248,1	114,0	Positive

Table 3 MRM parameters for the derivatized a-fdC samples

Name	Transition	Scan	Type	Precursor Ion	Product Ion	Ion Polarity
MeON-a-fdC	286,1 -> 170,1	MRM	Target	286,1	170,1	Positive
a-fdC	257,1 -> 141,1	MRM	Target	257,1	141,1	Positive
a-dC	229,1 -> 188,1	MRM	Target	229,1	188,1	Positive
a-dC	229,1 -> 112,1	MRM	Target	229,1	112,1	Positive

Digestion of the DNA of N2a cell line

As shown in Table 4 we used the following chemicals for the analysis of our samples. Reaction buffer 10X and Enzyme mix was bought as a Nucleoside Digestion Mix (M0649S) kit (*New England BioLabs Inc.*). The nucleosides mix was prepared by ourselve and contains heavy labelled mdC, hmdC, fdC, cadC, 8oxodG and hmdU. Furthermore we spiked heavy labelled Fluoro-dC and Fluoro-fdC for later quantification of the Fluoro-compounds. The concentration of the DNA for the digestion was received from the Nanodrop. The total volume was then incubated at 37 °C for 1.5 hours.

Table 4: Digestion sheet of the gDNA samples

	Sample	c(ng/µL)	m (DNA) [ng]	V (H2O) [µL]	V (DNA in H_2O) [µL]	V (Reaction buffer 10X)	V (Enzyme mix)	V (Nucleosides mix)	Spiking heavy Fluoro-dC	Spiking heavy Fluoro-fdC	Total Volume [µL]	1 h 30 min 37 °C	Injection volume [µL]
Blank 1	1	0	0	38,4	0,0	5,0	3,0	3,0	0,35	0,25	50,0		39
Blank 2	2	0	0	38,4	0,0	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 10 µg	3	528	10000	19,5	18,94	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 10 µg	4	528	10000	19,5	18,94	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 5 µg	5	528	5000	28,9	9,47	5,0	3,0	3,0	0,35	0,25	50,0		39
Control N2a Cofeeding 5 µg	6	528	5000	28,9	9,47	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 10 µg	7	412	10000	14,1	24,27	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 10 µg	8	412	10000	14,1	24,27	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 5 µg	9	412	5000	26,3	12,14	5,0	3,0	3,0	0,35	0,25	50,0		39
Sample N2a Cofeeding 5 µg	10	412	5000	26,3	12,14	5,0	3,0	3,0	0,35	0,25	50,0		39

Derivatization of a-fdC in digested gDNA of N2a cell line

Derivatization of a-fdC from the digested DNA (5 µg) of biological samples by methoxyamine was performed under the following conditions. Briefly, the nucleoside mixture was derivatized in 50 µL H_2O with 40 µL aq. NaOH (pH 10) and 8 µL methoxyamine (163 mM in H_2O) for 45 min at 25 °C. Afterwards the solution was neutralized with 40 µL aq. FA (pH 3). The derivatized nucleosides were then lyophylized and resuspended in 50 µL H_2O.

After resuspension the samples were filtered utilizing a 0.2 µm Supor filtration plate (*Pall Corporation*) and subjected to UHPLC-QQQ-MS/MS.

The determination of F-fdC and some other nucleosides in gDNA (10 µg) was performed by using heavy labelled compounds. After the digestion the samples were filtered utilizing a 0.2 µm Supor filtration plate (*Pall Corporation*) and subjected to UHPLC-QQQ-MS/MS.

Calibration curve for external quantification of a-fdC in gDNA of N2a cell lines

The quantification of a-fdC was done by an external calibration.[3] The external calibration curve was done by serially diluting pure a-fdC (see Table 5) and measured in technical triplicates prior to each measurment. Linear regression was done by Microsoft Excel (Figure 6). The injection volume was 29 µL.

Table 5 Measured values for the external calibration curve of a-fdC

n [fmol]	Average area	standard deviation
0,11	1006	110
0,22	1822	98
0,44	3160	414
0,88	4754	428
1,77	8439	333
3,53	14692	346
7,07	28982	522
14,14	56445	1107
28,28	111343	2041
56,55	220616	2561
113,10	439962	2460

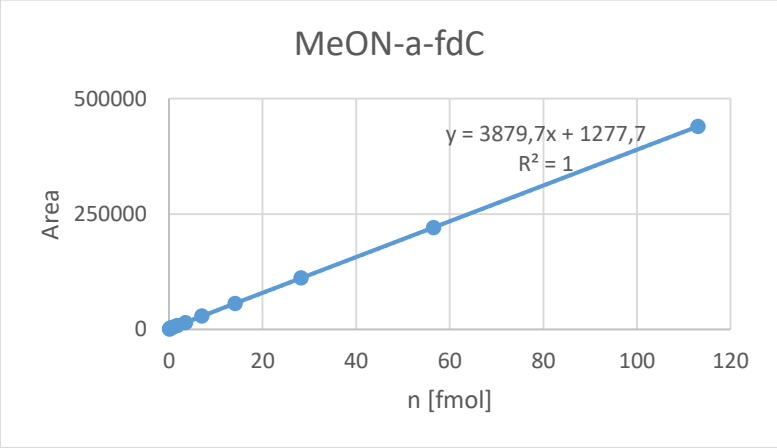


Figure 6: Calibration curve for the external quantification of a-fdC

Quantifiation of concentration of the fed compounds

The quantification of the fed compounds was performed in two technical replicates of a biological triplicate.

For the quantification of a-fdC one has to say, that there was a small background in the a-fdC peak, which was continously substracted from the a-fdC value. This value was determined by the control sample.

The quantification of the Fluoro-fdC and the deformylated Fluoro-dC was done over the ratio of unlabelled to labelled compound. Furthermore it was compared to the amount of deoxy-cytidine and deoxy-Guanosine and dN. The final results can be seen in Figures 7.

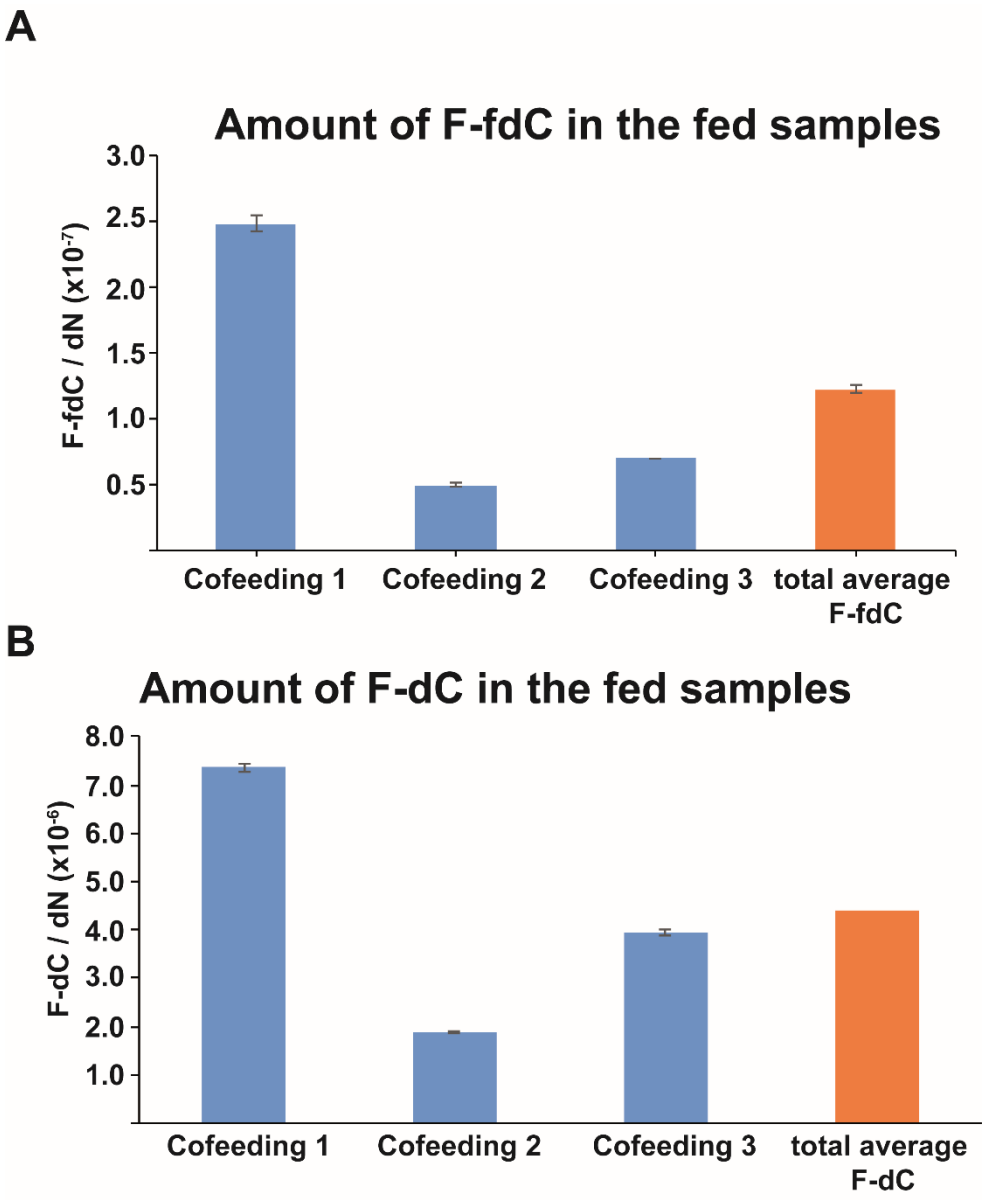


Figure 7 Measured amount of F-fdC (A) and F-dC (B) in the fed samples of N2a cells

Furthermore we were able to calculate the amount of deformylation of Fluoro-fdC by the amounts of Fluoro-fdC and Fluoro-dC. This was able to be done by the unnatural Fluoro compound. In Figure 8 the deformylation rates are shown.

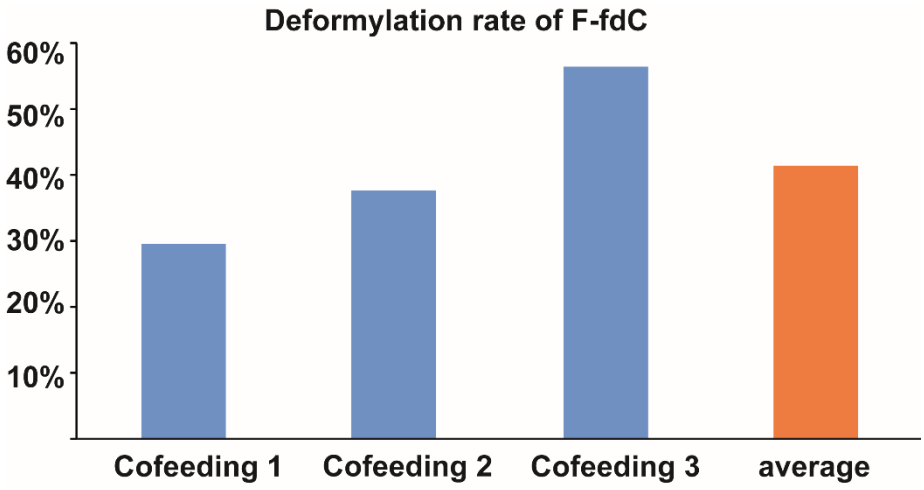
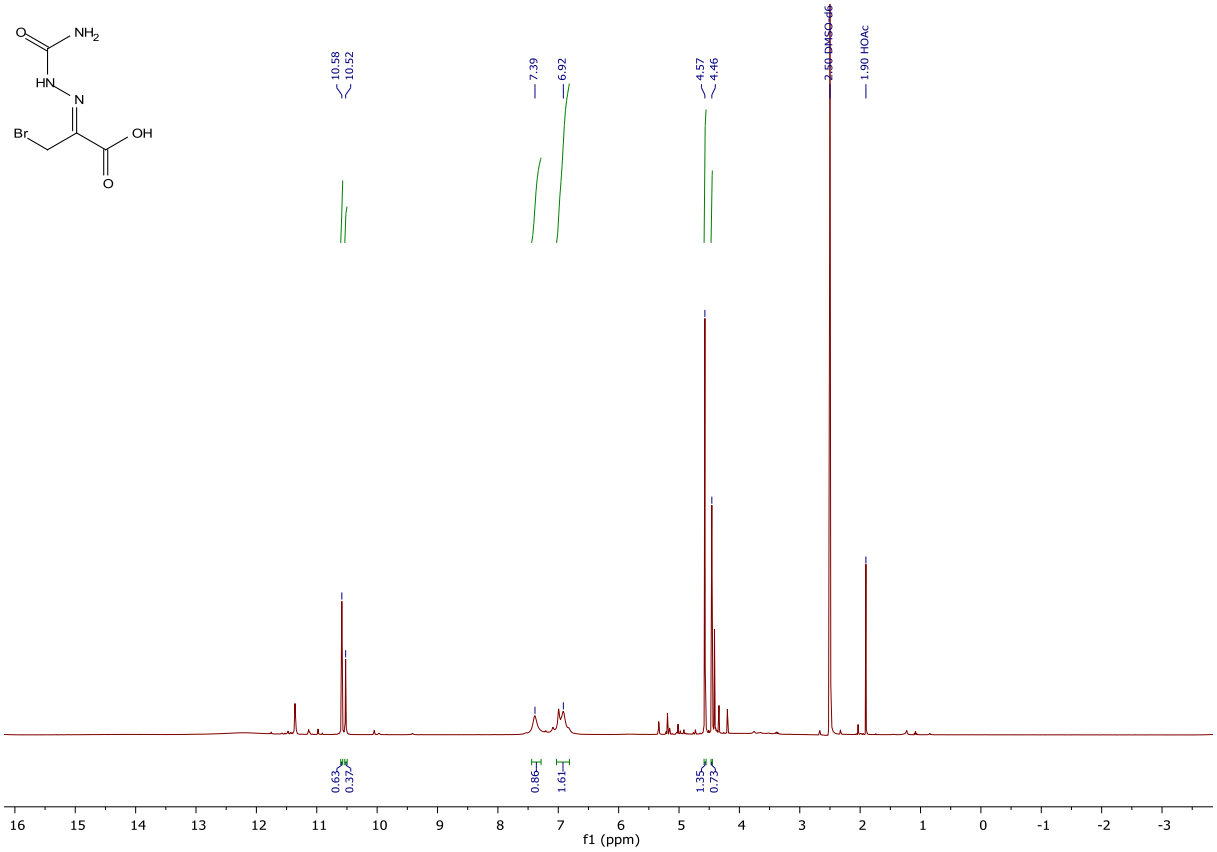
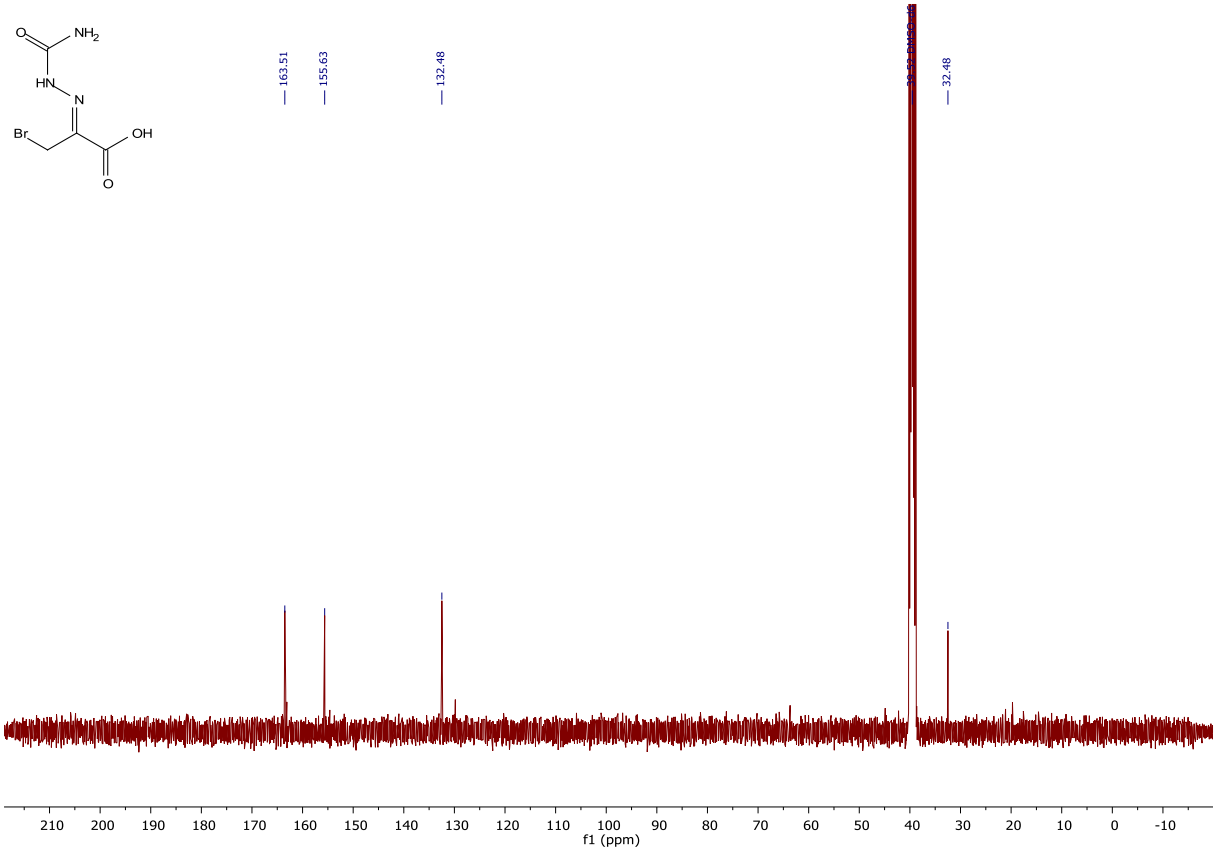
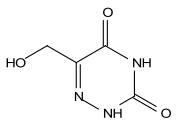


Figure 8 Calculated deformylation rate of F-fdC in N2a cells

NMR spectra of synthesized compounds

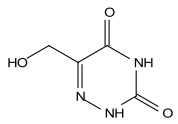





O
HO
NH
N
NH
O

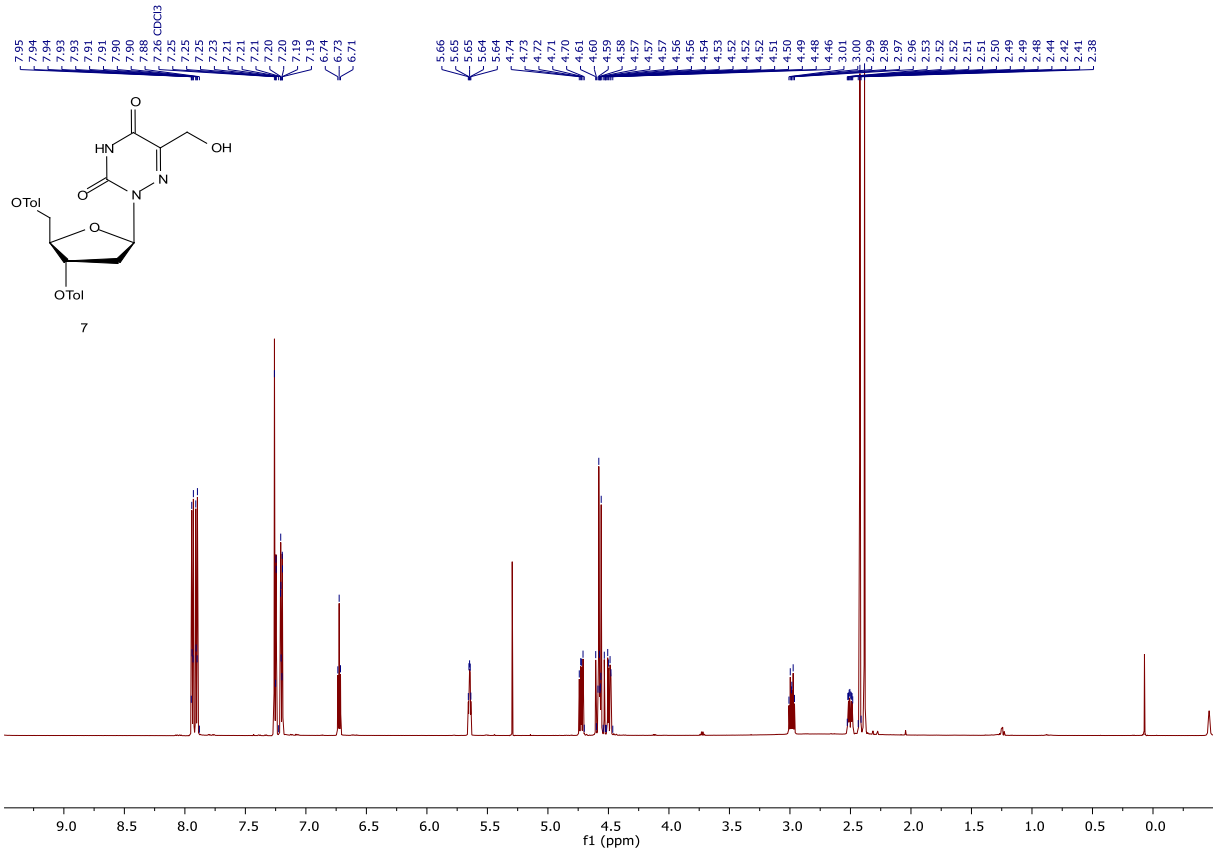
6

12.16
11.93
4.25
2.50 DMSO-d6
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)

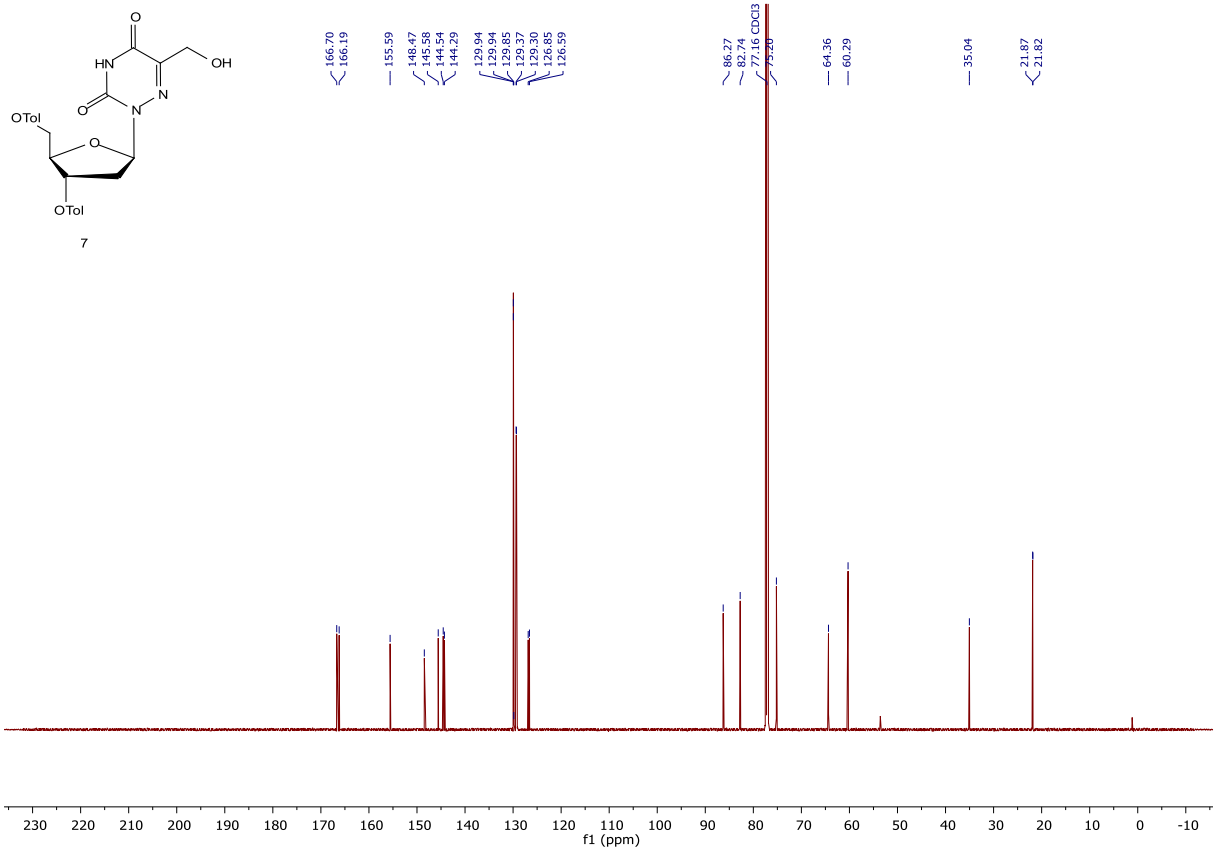

O
HO
NH
N
NH
O

6

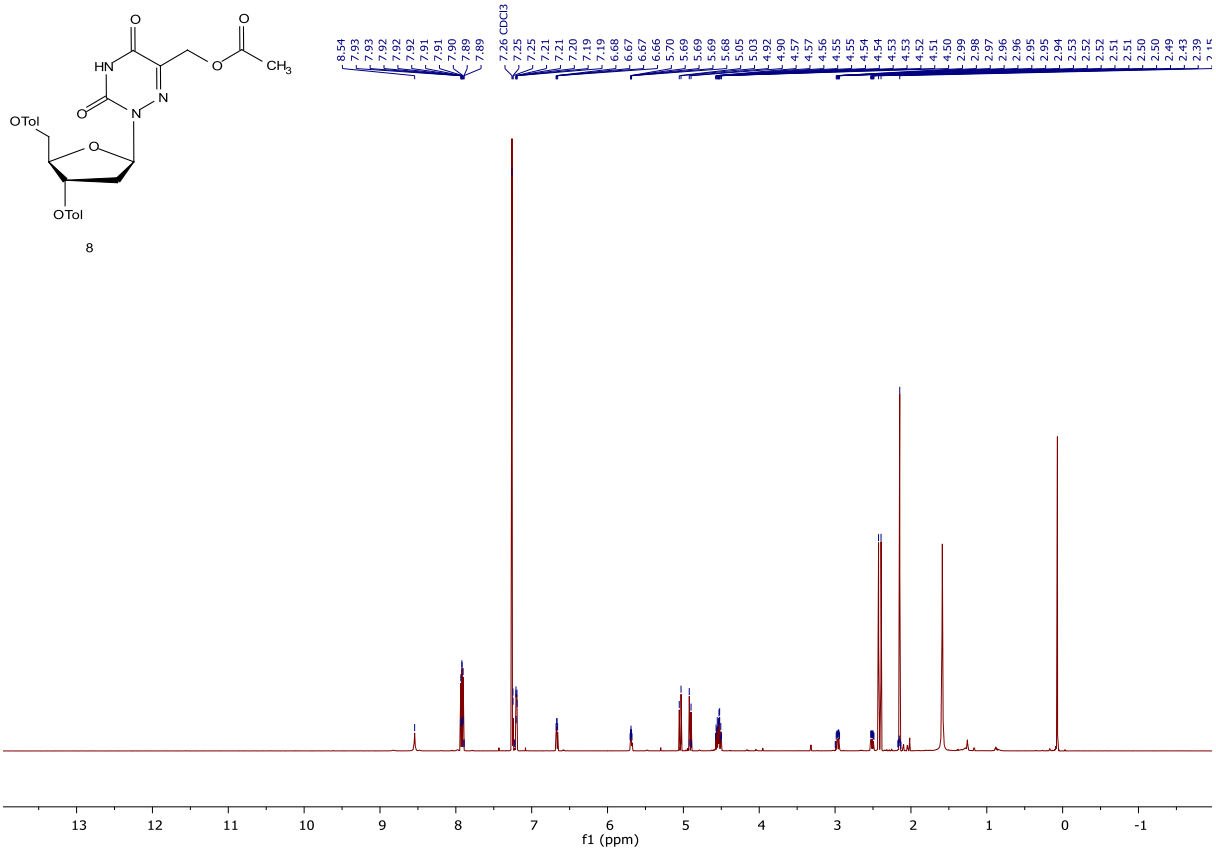
156.75
149.55
144.33
58.02
39.52 DMSO-d6
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-1
f1 (ppm)

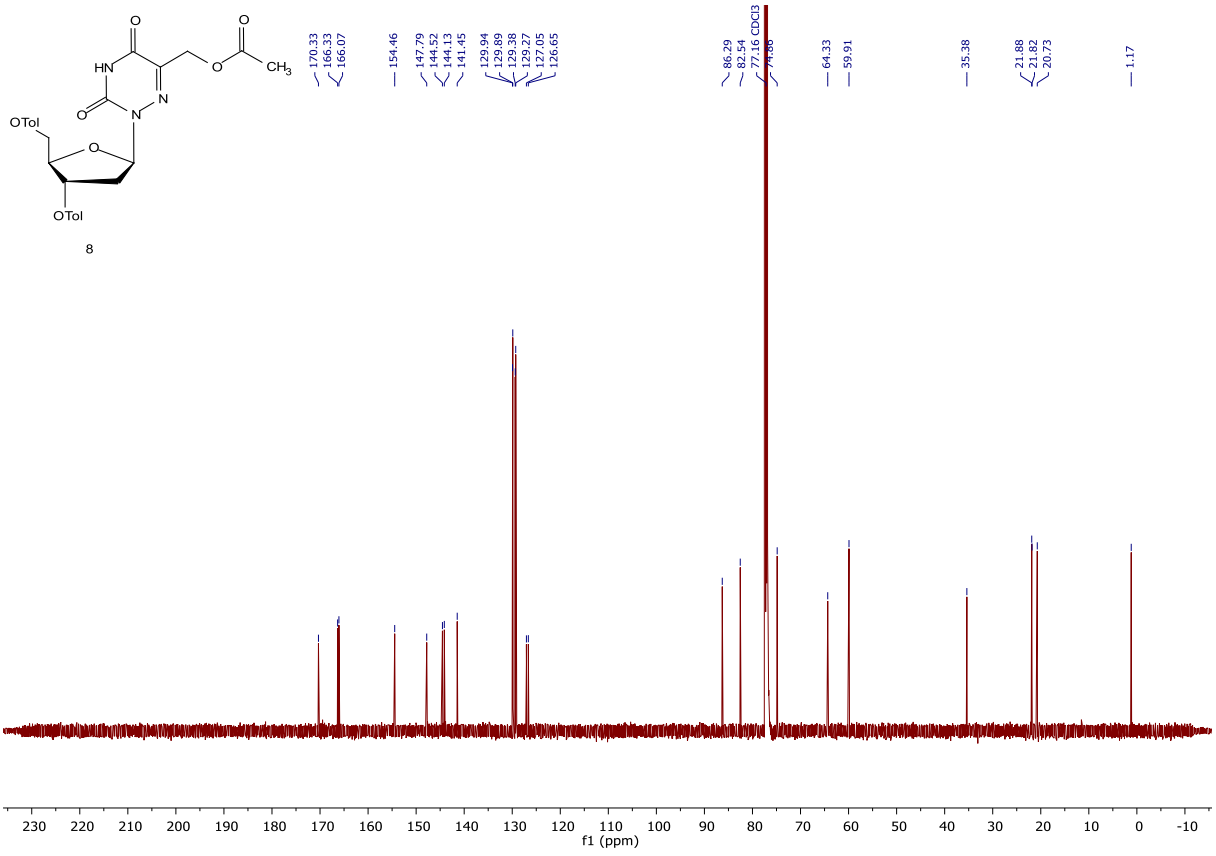

7.95 7.94 7.94 7.93 7.93 7.91 7.91 7.90 7.90 7.88 7.26 CDCl3 7.25 7.25 7.25 7.23 7.21 7.21 7.21 7.20 7.20 7.19 7.19 6.74 6.73 6.71
5.66 5.65 5.65 5.64 5.64 4.74 4.73 4.72 4.71 4.70 4.61 4.60 4.59 4.58 4.57 4.57 4.57 4.56 4.56 4.54 4.53 4.52 4.52 4.52 4.51 4.50 4.49 4.48 4.46 3.01 3.00 2.99 2.98 2.97 2.96 2.53 2.52 2.52 2.51 2.51 2.50 2.49 2.49 2.48 2.44 2.42 2.41 2.38
HN
O
N
OH
OTol
OTol
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

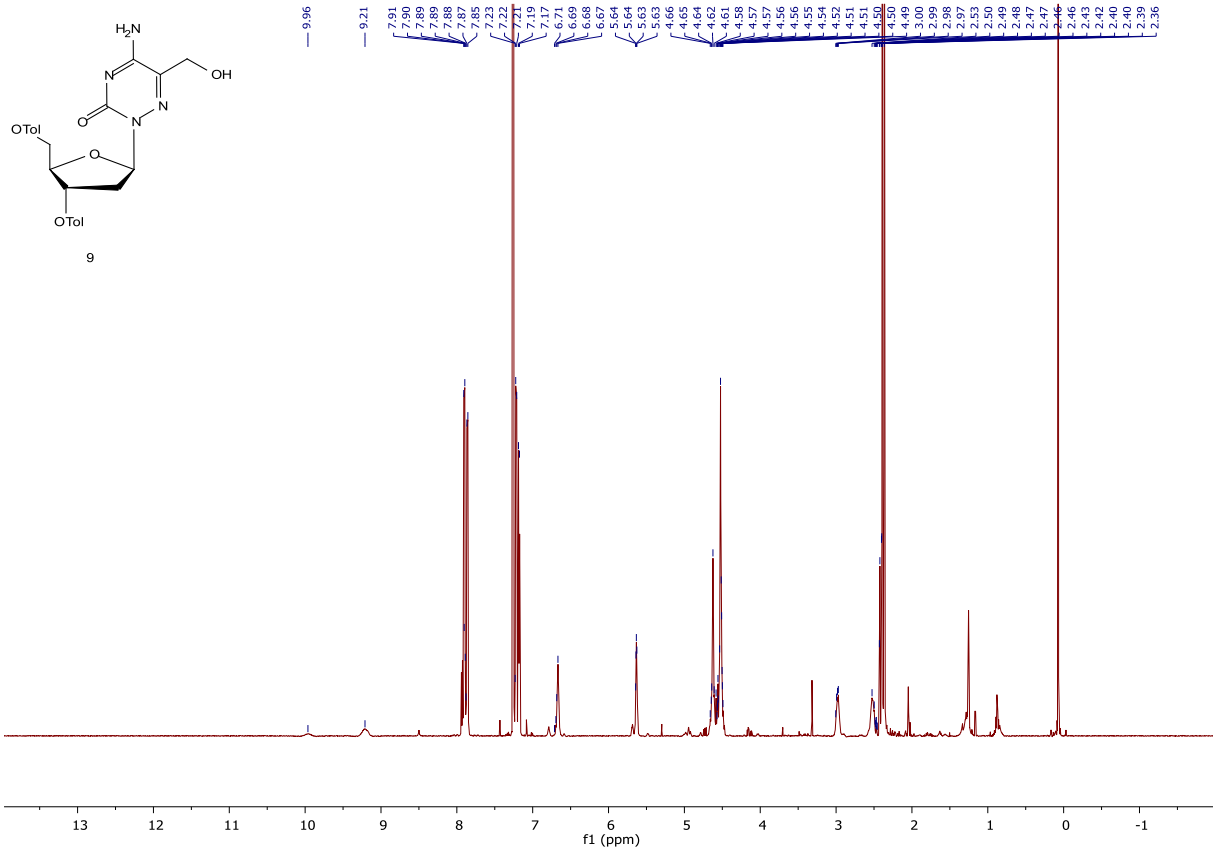
7

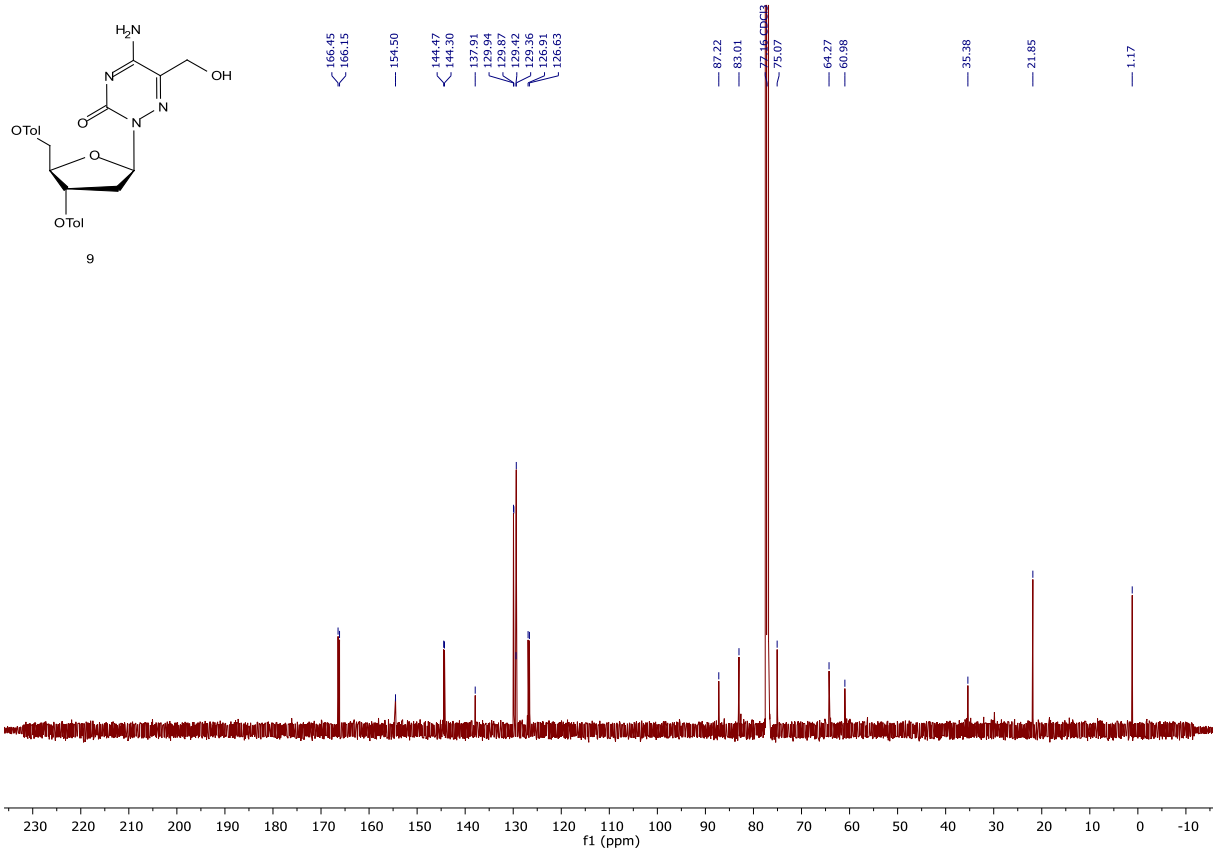

166.70 166.19 155.59 148.47 145.58 144.54 144.29 129.94 129.94 129.85 129.37 129.30 126.85 126.59 86.27 82.74 77.16 CDCl3 75.20 64.36 60.29 35.04 21.87 21.82
HN
O
N
OH
OTol
OTol
230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

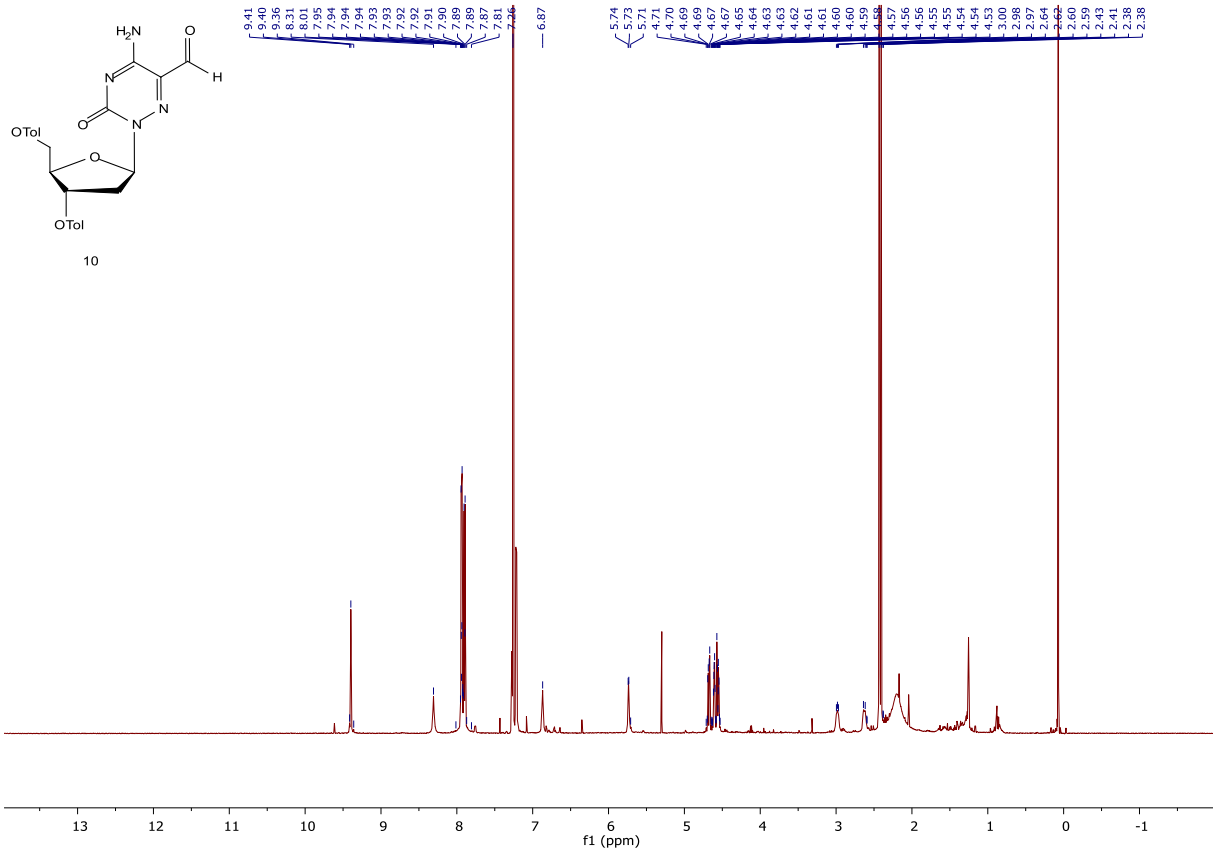
7

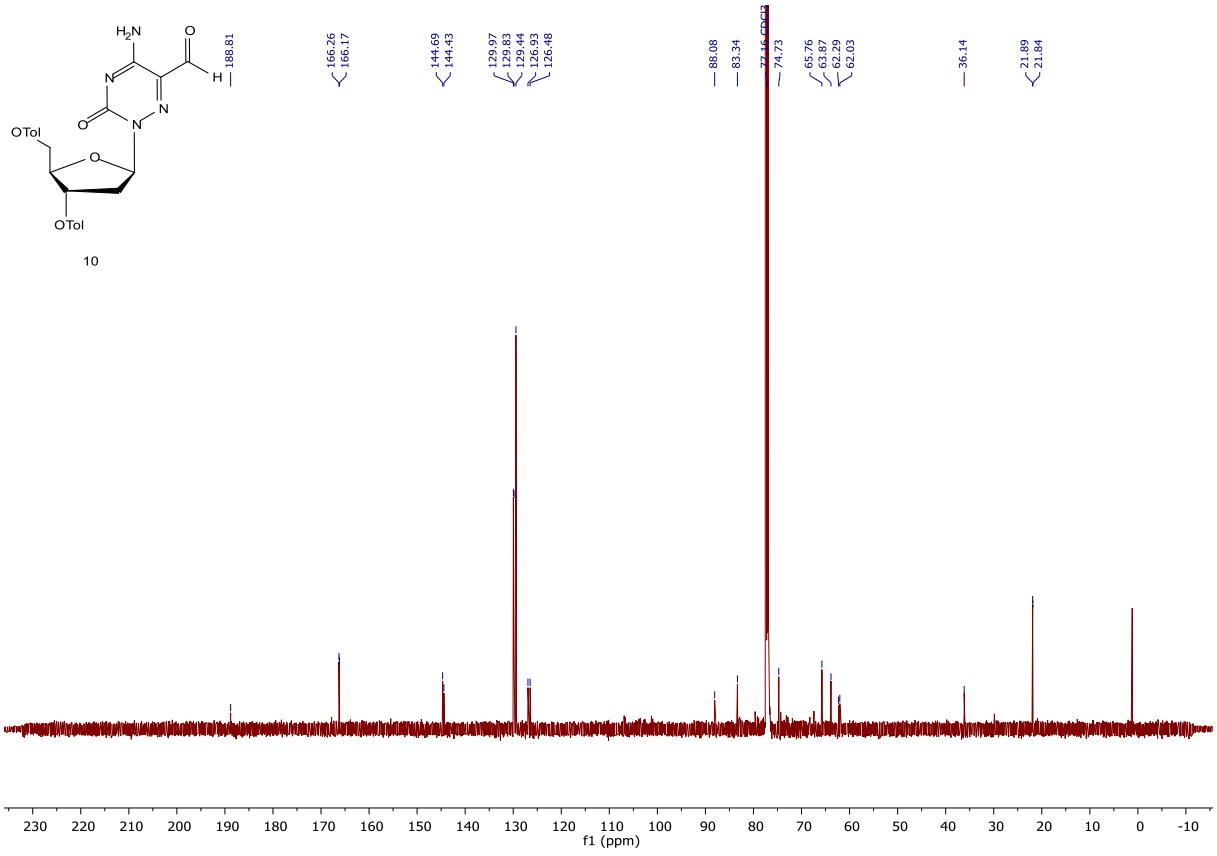

OTol
OTol
8
8.54
7.93
7.93
7.92
7.92
7.92
7.91
7.91
7.90
7.89
7.89
7.26 CDCl3
7.25
7.25
7.21
7.21
7.20
7.19
7.19
6.68
6.67
6.67
6.66
5.70
5.69
5.69
5.69
5.68
5.05
5.03
4.92
4.90
4.57
4.57
4.56
4.55
4.55
4.54
4.54
4.53
4.53
4.52
4.51
4.50
2.99
2.98
2.97
2.96
2.96
2.95
2.95
2.94
2.53
2.52
2.52
2.51
2.51
2.50
2.50
2.49
2.43
2.39
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
f1 (ppm)

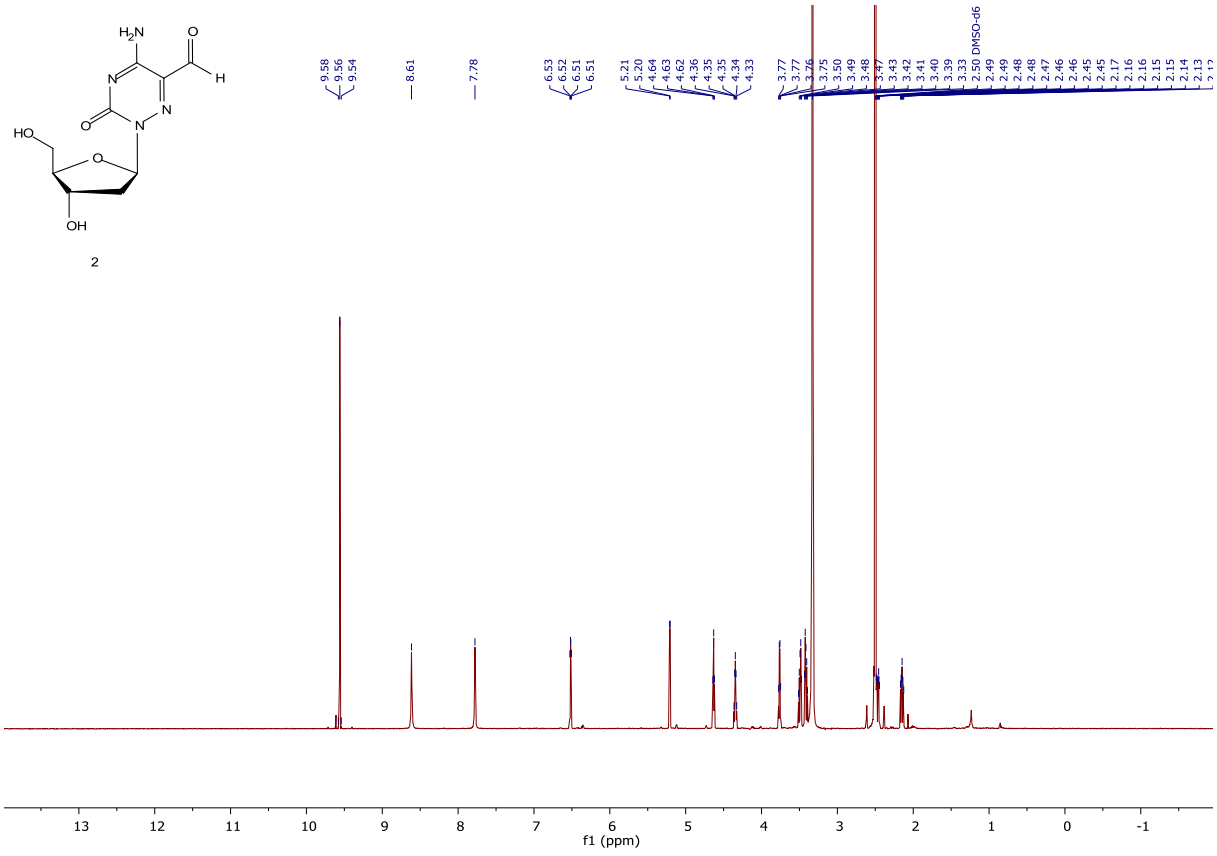

OTol
OTol
8
170.33
166.33
166.07
154.46
147.79
144.52
144.13
141.45
129.94
129.89
129.38
129.27
127.05
126.65
86.29
82.54
77.16 CDCl3
74.88
64.33
59.91
35.38
21.88
21.82
20.73
1.17
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


9.96
9.21
7.91
7.90
7.89
7.89
7.88
7.87
7.85
7.23
7.22
7.21
7.19
7.17
6.71
6.69
6.68
6.67
5.64
5.64
5.63
5.63
4.66
4.65
4.64
4.62
4.61
4.58
4.57
4.57
4.56
4.56
4.55
4.54
4.52
4.51
4.51
4.50
4.50
4.49
3.00
2.99
2.98
2.97
2.53
2.50
2.49
2.48
2.47
2.47
2.46
2.46
2.43
2.42
2.40
2.40
2.39
2.36
H2N
OH
N
N
O
N
OTol
O
OTol
9
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
f1 (ppm)

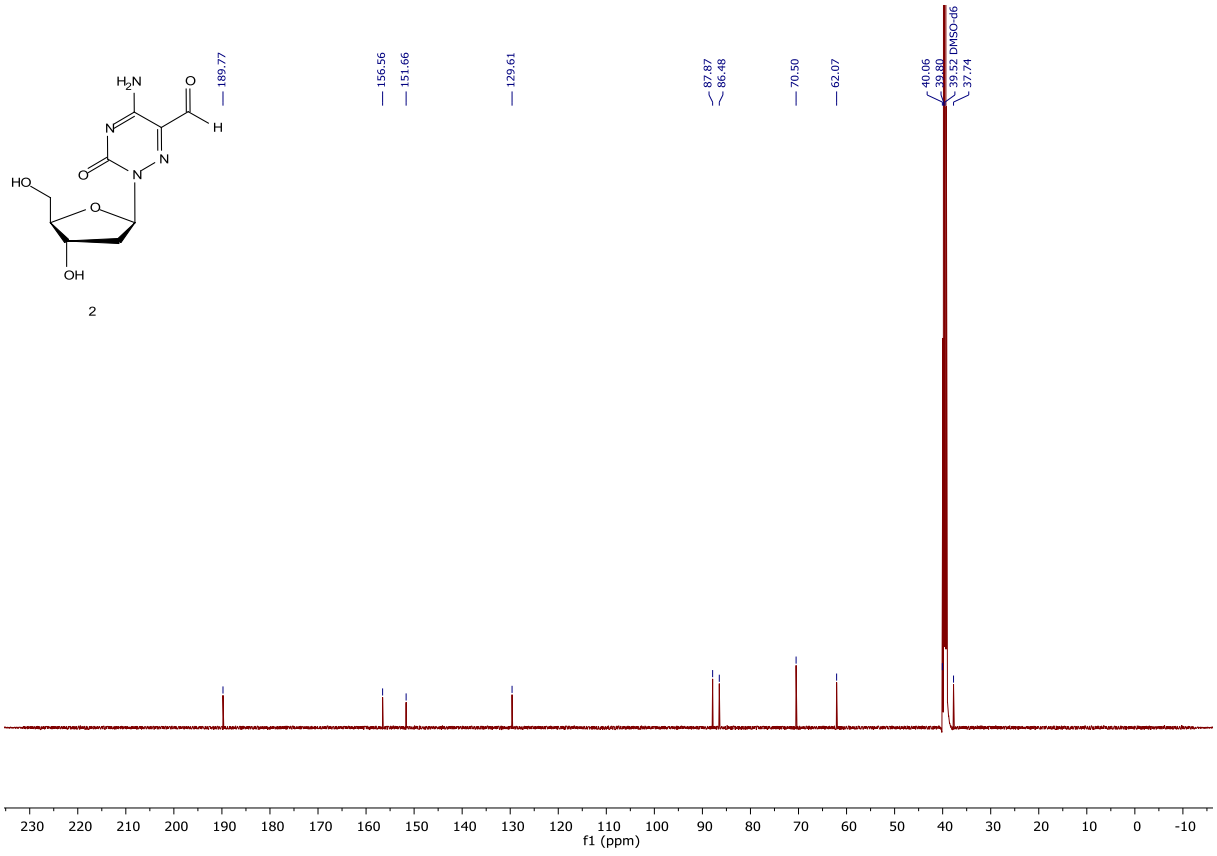

166.45
166.15
154.50
144.47
144.30
137.91
129.94
129.87
129.42
129.36
126.91
126.63
87.22
83.01
77.16 CDCl3
75.07
64.27
60.98
35.38
21.85
1.17
H2N
OH
N
N
O
N
OTol
O
OTol
9
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


9.41
9.40
9.36
8.31
8.01
7.95
7.94
7.94
7.93
7.93
7.92
7.92
7.91
7.90
7.89
7.89
7.87
7.81
7.26
6.87
5.74
5.73
5.71
4.71
4.70
4.69
4.69
4.67
4.67
4.65
4.64
4.63
4.63
4.62
4.61
4.61
4.60
4.60
4.59
4.58
4.57
4.56
4.56
4.55
4.55
4.54
4.54
4.53
3.00
2.98
2.97
2.64
2.62
2.60
2.59
2.43
2.41
2.38
2.38
OTol
OTol
10
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
f1 (ppm)

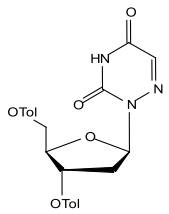

188.81
166.26
166.17
144.69
144.43
129.97
129.83
129.44
126.93
126.48
88.08
83.34
77.16 CDCl3
74.73
65.76
63.87
62.29
62.03
36.14
21.89
21.84
OTol
OTol
10
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


9.58
9.56
9.54
8.61
7.78
6.53
6.52
6.51
6.51
5.21
5.20
4.64
4.63
4.62
4.36
4.35
4.35
4.34
4.33
3.77
3.77
3.76
3.75
3.50
3.49
3.48
3.47
3.43
3.42
3.41
3.40
3.39
3.33
2.50 DMSO-d6
2.49
2.49
2.48
2.48
2.47
2.46
2.46
2.45
2.45
2.17
2.16
2.16
2.15
2.15
2.14
2.13
HO
O
OH
H
O
N
N
N
O
H_2N
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
f1 (ppm)

2

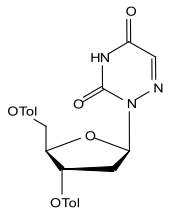

189.77
156.56
151.66
129.61
87.87
86.48
70.50
62.07
40.06
39.80
39.52 DMSO-d6
37.74
HO
O
OH
H
O
N
N
N
O
H_2N
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

2


O
HN
O
N
N
OTol
O
OTol

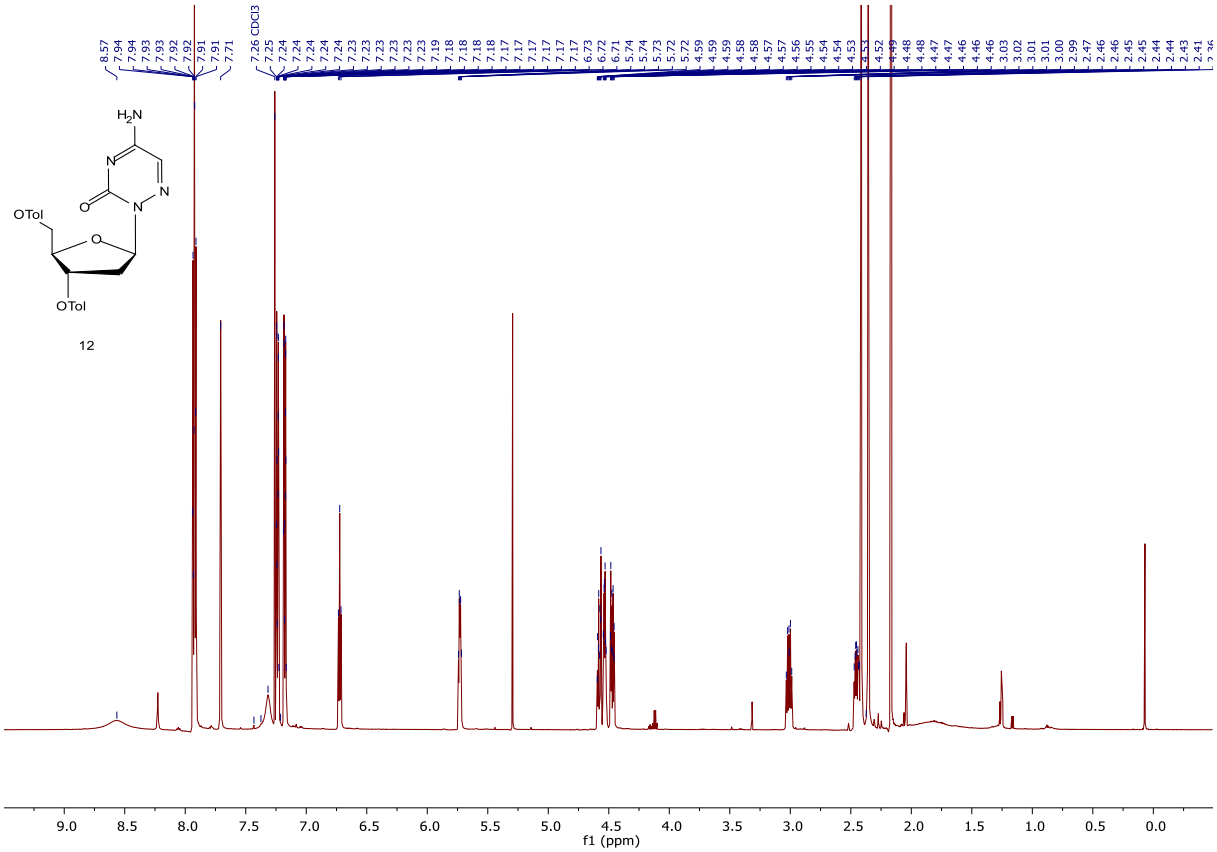
11

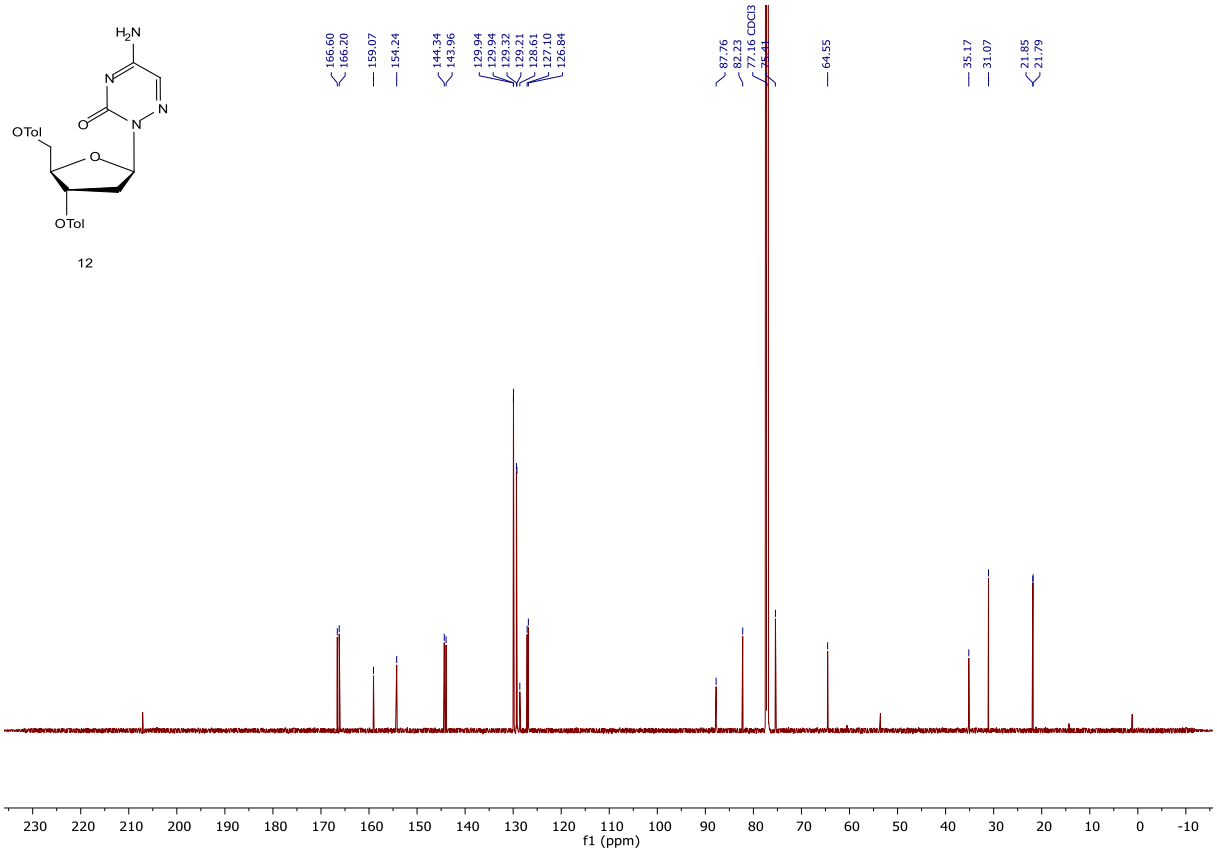
8.86
8.82
8.82
7.96
7.95
7.94
7.93
7.92
7.91
7.90
7.26 CDCl3
6.70
6.68
6.67
5.73
5.73
5.72
5.71
5.70
4.68
4.66
4.65
4.64
4.55
4.53
4.52
4.51
4.48
4.47
4.45
4.44
3.02
3.00
2.98
2.97
2.95
2.51
2.50
2.50
2.49
2.48
2.47
2.46
2.45
2.43
2.41
16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0 -1 -2 -3
f1 (ppm)

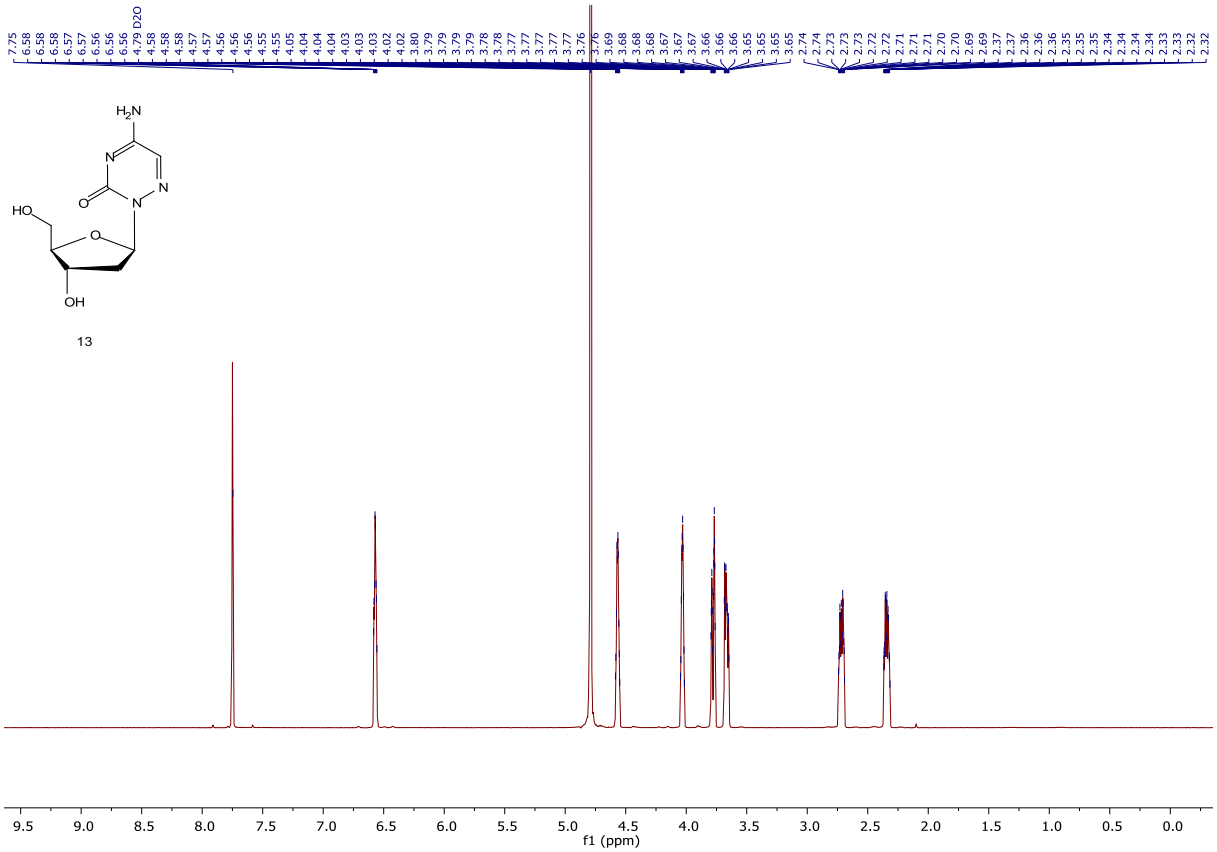

O
HN
O
N
N
OTol
O
OTol

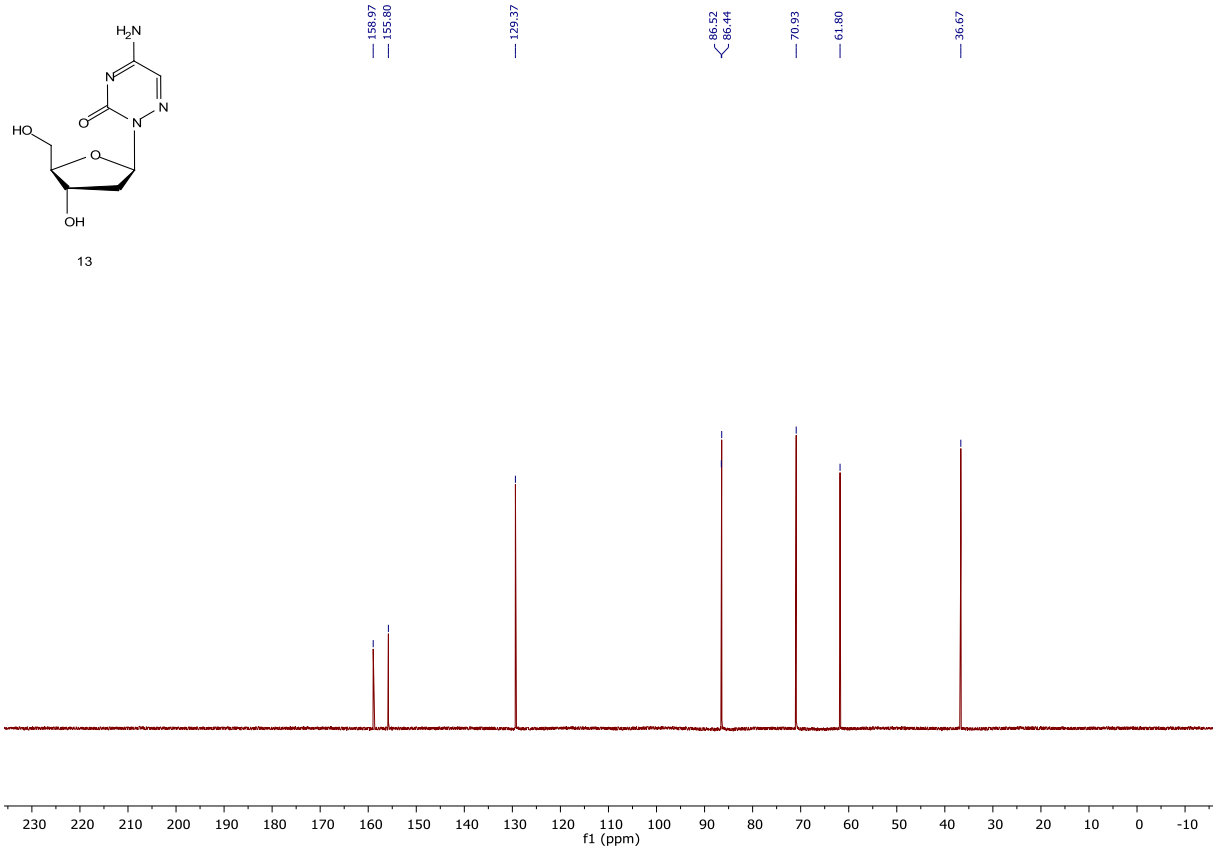
11

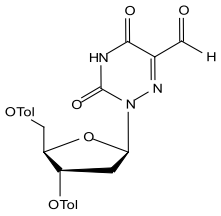
166.36
166.16
155.30
147.77
144.56
144.11
136.10
129.94
129.90
129.39
129.27
127.11
126.61
86.05
82.73
77.16 CDCl3
75.03
64.06
35.17
21.90
21.86
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)


8.57
7.94
7.94
7.93
7.93
7.92
7.92
7.91
7.91
7.71
7.26 CDCl3
7.25
7.24
7.24
7.24
7.24
7.24
7.23
7.23
7.23
7.23
7.23
7.23
7.19
7.18
7.18
7.18
7.18
7.17
7.17
7.17
7.17
7.17
7.17
6.73
6.72
6.71
5.74
5.74
5.74
5.73
5.72
5.72
4.59
4.59
4.59
4.58
4.58
4.57
4.57
4.56
4.55
4.54
4.54
4.53
4.52
4.48
4.48
4.47
4.47
4.46
4.46
4.46
3.03
3.02
3.01
3.01
3.00
2.99
2.47
2.46
2.46
2.45
2.45
2.44
2.44
2.43
2.41
H2N
N
N
O
N
OTol
O
OTol
12
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


H2N
N
N
O
N
OTol
O
OTol
12
166.60
166.20
159.07
154.24
144.34
143.96
129.94
129.94
129.32
129.21
128.61
127.10
126.84
87.76
82.23
77.16 CDCl3
76.41
64.55
35.17
31.07
21.85
21.79
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

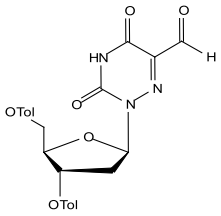

7.75
6.58
6.58
6.58
6.57
6.57
6.56
6.56
6.56
4.79 D2O
4.58
4.58
4.58
4.57
4.57
4.56
4.56
4.56
4.55
4.55
4.05
4.04
4.04
4.04
4.03
4.03
4.03
4.02
4.02
3.80
3.79
3.79
3.79
3.79
3.78
3.78
3.77
3.77
3.77
3.77
3.77
3.76
3.76
3.69
3.68
3.68
3.68
3.67
3.67
3.67
3.66
3.66
3.66
3.65
3.65
3.65
3.65
2.74
2.74
2.73
2.73
2.73
2.72
2.72
2.71
2.71
2.71
2.70
2.70
2.70
2.69
2.69
2.37
2.37
2.36
2.36
2.36
2.35
2.35
2.35
2.34
2.34
2.34
2.34
2.33
2.33
2.32
2.32
H2N
N
N
N
O
HO
O
OH
13
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


158.97
155.80
129.37
86.52
86.44
70.93
61.80
36.67
H2N
N
N
N
O
HO
O
OH
13
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


O
O
HN
H
N
O
N
OTol
O
OTol

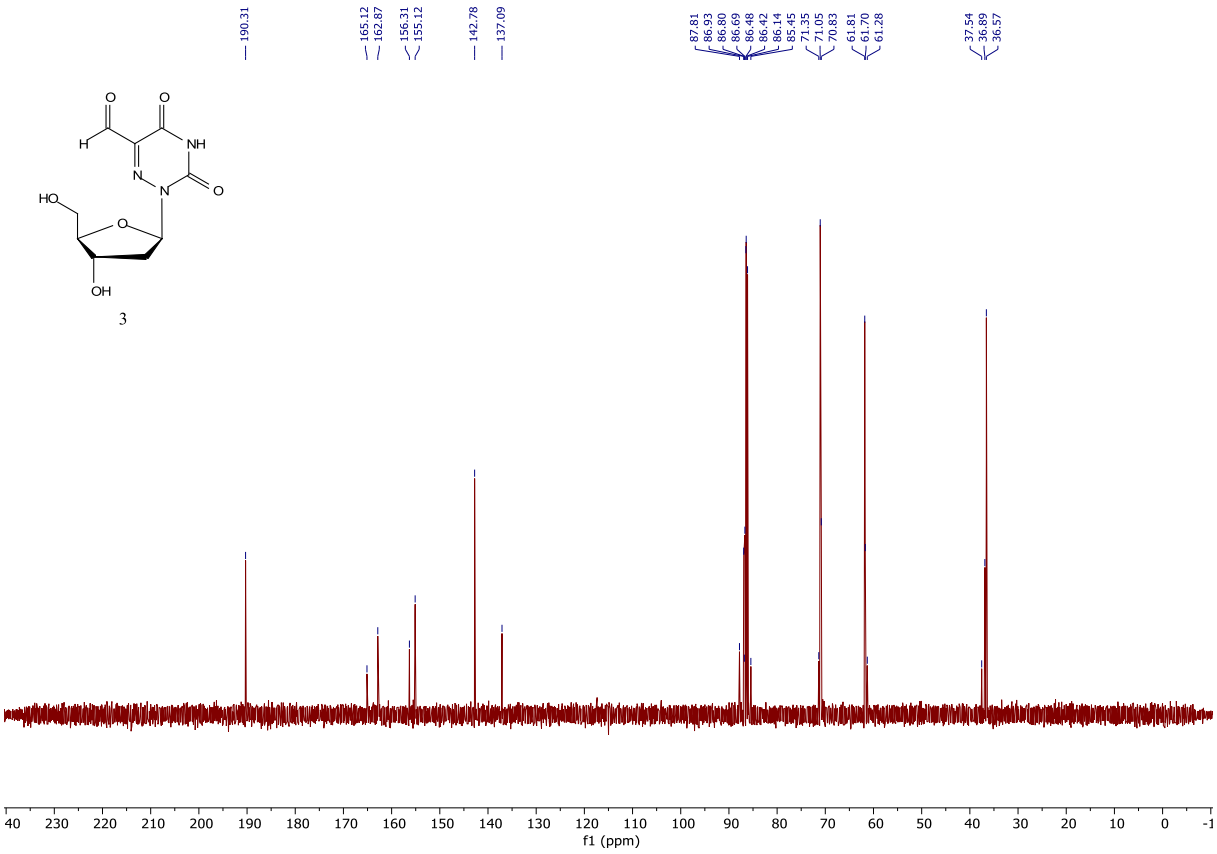
14

9.60
9.58
9.55
9.02
7.95
7.94
7.93
7.93
7.92
7.92
7.91
7.90
7.90
7.89
7.88
7.87
7.85
7.26 CDCl3
6.73
6.71
6.70
6.68
6.67
6.66
5.73
5.72
5.72
5.71
4.67
4.66
4.64
4.64
4.62
4.62
4.60
4.59
4.58
4.57
4.57
4.56
4.54
4.53
3.10
3.08
3.07
3.06
3.05
3.03
2.61
2.60
2.59
2.58
2.57
2.56
2.56
2.55
2.45
2.43
2.41
2.41
2.40
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)

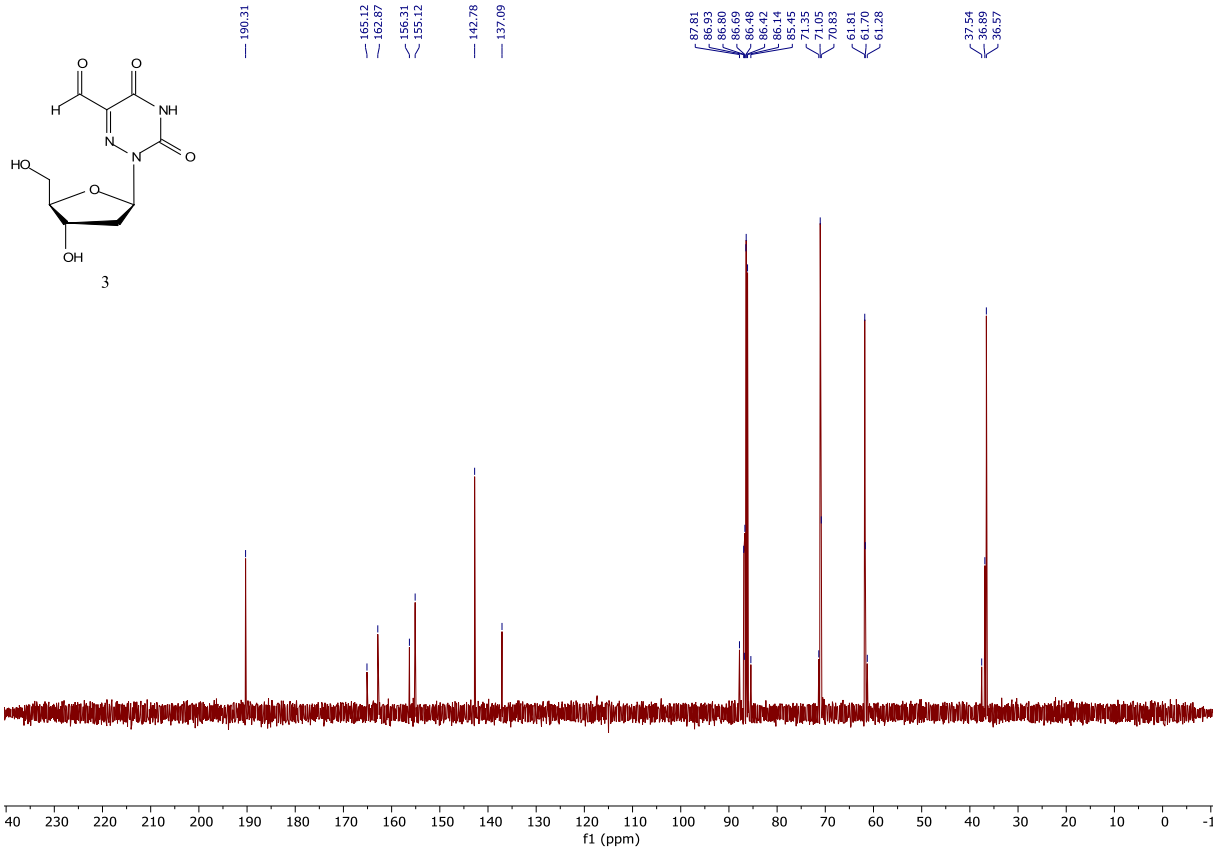

O
O
HN
H
N
O
N
OTol
O
OTol

14

184.16
166.33
166.14
152.79
147.15
144.68
144.40
137.22
129.94
129.84
129.81
129.70
129.54
129.42
129.40
126.81
126.41
87.16
84.08
83.20
77.48 CDCl3
77.36
77.16
76.84 CDCl3
74.59
63.74
35.40
29.85
21.91
21.90
21.85
1.17
40
230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-1
f1 (ppm)


190.31
165.12
162.87
156.31
155.12
142.78
137.09
87.81
86.93
86.80
86.69
86.48
86.42
86.14
85.45
71.35
71.05
70.83
61.81
61.70
61.28
37.54
36.89
36.57
O
O
H
NH
N
N
O
HO
O
OH
40 230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1
f1 (ppm)

3


190.31
165.12
162.87
156.31
155.12
142.78
137.09
87.81
86.93
86.80
86.69
86.48
86.42
86.14
85.45
71.35
71.05
70.83
61.81
61.70
61.28
37.54
36.89
36.57
O
O
H
NH
N
N
O
HO
O
OH
40 230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1
f1 (ppm)

3

Bibliography

[1] I. V. Alekseeva, A. S. Shalamai, V. S. Shalamai, V. P. Chernetskii, *Ukr. Khim. Zh.* **1976**, *42*, 398-401.

[2] W. L. Mitchell, P. Ravenscroft, M. L. Hill, L. J. S. Knutsen, B. D. Judkins, R. F. Newton, D. I. C. Scopes, *J. Med. Chem.* **1986**, *29*, 809-816.

[3] F. R. Traube, S. Schiffers, K. Iwan, S. Kellner, F. Spada, M. Müller, T. Carell, *Nat. Protoc.* **2019**, *14*, 283-312