Supporting Information

**Biomimetic Iron Complex Achieves TET Enzyme Reactivity**

David Schmidl*, Niko S. W. Jonasson*, Eva Korytiaková†, Thomas Carell, and Lena J. Daumann*

anie_202107277_sm_miscellaneous_information.pdf
Supporting Information
©Wiley-VCH 2019
69451 Weinheim, Germany
Table of Contents

1. Materials and Methods ............................................................................................................. 4
   1.1. Solvents and chemicals ..................................................................................................... 4
   1.2. Methods and manipulations .............................................................................................. 4
   1.3. Gas chromatography-mass spectroscopy (GC-MS) .......................................................... 4
   1.4. NMR spectroscopy .......................................................................................................... 4
   1.5. ESI and EI Mass spectrometry .......................................................................................... 4
   1.6. Nucleoside Quantification in Oligonucleotide Experiments ........................................... 4
   1.7. High performance liquid chromatography (HPLC) in Nucleoside Experiments ............. 5
   1.8. UV/Vis Oligonucleotide Quantification ............................................................................ 6
   1.9. Matrix assisted laser desorption ionization mass spectrometry (MALDI MS) ............... 6
   1.10. Bond Dissociation Energies and Measured Kinetic Constants ...................................... 6

2. Experimental Section .................................................................................................................. 7
   2.1. General Procedure for HPLC Experiments .................................................................... 7
   2.2. HPLC Traces .................................................................................................................... 7
      2.2.1. Overview .................................................................................................................... 7
      2.2.2. HPLC trace of 5mdC, 5hmdC, 5fdC, and 5cadC ....................................................... 8
      2.2.3. HPLC trace of 5mC, 5hmC, 5fC, and 5caC ............................................................... 8
      2.2.4. HPLC trace of 1 after filtration ................................................................................ 9
      2.2.5. Spike Experiments – No Additions ....................................................................... 9
      2.2.6. Spike Experiments – 5mdC or 5mC ...................................................................... 10
      2.2.7. Spike Experiments – 5hmdC or 5hmC .................................................................. 10
      2.2.8. Spike Experiments – 5fdC or 5fC ........................................................................ 11
      2.2.9. Spike Experiments – 5cadC or 5caC ..................................................................... 11
   2.3. GC-MS Reference Samples .............................................................................................. 12
      2.3.1. 2′-Deoxyribomo-γ-lactone ...................................................................................... 12
      2.3.2. 5-methyl-2′deoxyribo-cytidine ............................................................................... 13
      2.3.3. Overlay of References ............................................................................................. 14
   2.4. Reactions of 1 with substrates ......................................................................................... 15
      2.4.1. Reaction of 1 with 5mdC ......................................................................................... 15
      2.4.2. Reaction of 1 with 5hmdC ..................................................................................... 18
      2.4.3. Reaction of 1 with d3-5mdC .................................................................................. 20
      2.4.4. Filtration and control experiments ......................................................................... 21
   3. Oligonucleotide Experiments ................................................................................................. 22
      3.1. General Procedure for Oligonucleotide Experiments .................................................. 22
      3.2. Oligonucleotide Concentration ................................................................................... 22
SUPPORTING INFORMATION

1. Materials and Methods

1.1. Solvents and chemicals

Chemicals were purchased from commercial sources (Sigma Aldrich, ABCR, Acros Organics, Alfa Aesar, TCI Chemicals, Oakwood Chemicals) or the LMU Munich chemical supply and used without further purification except for dichloromethane, diethyl ether and hexanes. The solvents were purchased from the LMU Munich chemical supply and distilled once under reduced pressure prior to use.

1.2. Methods and manipulations

All manipulations were carried out in flame-dried glassware under a positive pressure of nitrogen gas and magnetically stirred if not stated otherwise. Air- and moisture-sensitive chemicals and absolute solvents were transferred via stainless-steel cannula or syringe. Organic solutions were concentrated by rotary evaporation at 40 °C. Analytical thin layer chromatography (TLC) was performed on pre-coated (silica gel, 0.25 mm, 60 Å pore-size, 230–400 mesh, Merck KGA) aluminum plates or which were impregnated with a fluorescent indicator (254 nm). TLC plates were visualized by exposure to ultraviolet light.

1.3. Gas chromatography-mass spectroscopy (GC-MS)

Reference samples were used as received, reaction samples were dried under high vacuum ($< 1 \times 10^{-2}$ mbar) prior to derivatization. 0.5 mg of each sample were suspended in 400 µl acetonitrile, 100 µl BSTFA were added and the mixture heated to 70 °C for 30 min. The samples were filtered and injected (1 µl, split or splitless mode depending on substrate concentration) onto an Agilent® 7920 GC equipped with a 30 m HP5-MS column (Agilent® 19091S-439UI) coupled to an Agilent® 5970 EI mass spectrometer. The injector temperature was set to 280 °C and the temperature of the ion source 230 °C. The initial oven temperature was 80 °C, held there for 2 min, ramped to 240 °C at 5 K/min and then held there for 20 min. Mass spectra were recorded in scan mode between 70-400 m/z.

1.4. NMR spectroscopy

$^1$H NMR, $^2$H NMR and $^{13}$C NMR spectra were recorded at room temperature on Bruker Avance III (400MHz) operating at 400MHz for proton nuclei, 60 MHz for deuterium nuclei, and 100MHz for carbon nuclei. $^1$H-chemical shifts are recorded in ppm units relative to CDCl$_3$ (δ$_H$ = 7.26), CD$_2$CN (δ$_H$ = 1.94), or D$_2$O (δ$_H$ = 4.79). $^{13}$C chemical shifts are given in ppm units relative to CDCl$_3$ (δ$_C$ = 77.16), CD$_2$CN (δ$_C$ = 1.32) or DMSO-d$_6$ (δ$_C$ = 39.52). The following abbreviations were used: s = singlet, d = doublet, t = triplet, dd = doublet of doublets, dt = doublet of triplets, m = multiplet, br = broad. Coupling constants (J) are given in Hertz. The software used for data processing was MNova Version 12.0.1-20560.

1.5. ESI and EI Mass spectrometry

ESI spectra were recorded with a Thermo Finnigan LTQ FT Ultra Fourier Transform Ion Cyclotron Resonance mass spectrometer with acetonitrile/water as the carrier solvent. Quantitative UHPLC-MS/MS analysis of digested DNA samples was performed using an Agilent 1290 UHPLC system equipped with a UV detector and an Agilent 6490 triple quadrupole mass spectrometer. The source-dependent parameters were as follows: gas temperature 80 °C, gas flow 15 L/min (N$_2$), nebulizer 30 psi, sheath gas heater 275 °C, sheath gas flow 11 L/min (N$_2$), capillary voltage 2.500 V in the positive ion mode, capillary voltage −2.250 V in the negative ion mode and nozzle voltage 500 V. The fragmentor voltage was 380 V / 250 V. Delta EMV was set to 500 (positive mode) and 800 (negative mode). Compound-dependent parameters were summarized in Supplementary Table S1. Chromatography was performed by a Poroshell 120 SB-C8 column (Agilent, 2.7 µm, 2.1 mm × 150 mm) at 35 °C using a gradient of water and MeCN, each containing 0.0085% (v/v) formic acid, at a flow rate of 0.35 mL/min: 0 → 4 min; 0 → 3.5% (v/v) MeCN; 4 → 7.9 min; 3.5 → 5% MeCN; 7.9 → 8.2 min; 5 → 80% MeCN; 8.2 → 11.5 min; 80% MeCN; 11.5 → 12 min; 80 → 0% MeCN; 12 → 14 min; 0% MeCN. The autosampler was cooled to 4 °C. The injection volume was amounted to 39 µL.

1.6. Nucleoside Quantification in Oligonucleotide Experiments

In the case of the oligonucleotide experiments, nucleosides of interest were quantified using the stable isotope dilution technique.[2] Internal standards (isotopically labelled 5mdC*, 5hmC*, 5fcC*, 5cadC*, and 8-oxo-dG*) were added to a sample and the samples were then digested using Degradase Plus from Zymo Research, and subsequently measured on a tandem UHPLC-MS/MS system: Agilent Technologies, model no. 1290 Infinity II LC; MS system: Triple quadrupole LC/MS system with iFunnel technology (Agilent Technologies, model no. 6490). The exact method was described here.[2] The sample data were analyzed by Quantitative MassHunter Software from Agilent. Nucleosides dC and dG were quantified using the signal of the UV-trace with known calibration curves.[2] Used internal calibration curves, as well as UV-calibration curves are shown in Table S1.
### Table S1: Calibration curves for UHPLC-MS/MS quantification of nucleosides after enzymatic digestion of the oligonucleotide product mixture.

| Nucleoside    | Linear Regression          |
|---------------|----------------------------|
| UV-dC         | $y = 0.28558x - 0.02812$   |
| UV-dG         | $y = 0.54214x + 0.1554$    |
| mdC           | $y = 0.75157x - 3.70649 \times 10^{-4}$ |
| hmdC          | $y = 0.89129 - 3.96 \times 10^{-3}$ |
| fdC           | $y = 1.18055x - 0.00206$   |
| cadC          | $y = 0.94585x - 0.01218$   |
| 8-oxo-dG      | $y = 0.9355 - 0.00138$     |

### Table S2: Calibration curves for UV absorption quantification in HPLC traces.

#### HPLC 1

| Substrate | Linear Regression          |
|-----------|----------------------------|
| 5mdC      | $y = 445.99x - 15.244$     |
| 5hmdC     | $y = 451.86x - 3.3986$     |
| 5fdC      | $y = 550.06x - 14.941$     |
| 5cadC     | $y = 289.74x - 5.9377$     |

#### HPLC 2

| Substrate | Linear Regression          |
|-----------|----------------------------|
| 5mdC      | $y = 860.96x + 78.134$     |
| 5hmdC     | $y = 905.75x + 85.083$     |
| 5fdC      | $y = 1016.5x + 89.150$     |
| 5cadC     | $y = 611.05x + 55.824$     |
| 5mC       | $y = 609.33x + 41.662$     |
| 5hmC      | $y = 601.37x + 41.525$     |

#### Figure S1: HPLC (UV, $\lambda = 280$ nm) Quantification curves for the nucleosides 5mdC, 5hmdC, 5fdC and 5cadC in the nucleoside experiments.

1.7. High performance liquid chromatography (HPLC) in Nucleoside Experiments

HPLC traces were recorded on two Agilent 1260 Infinity II HPLC systems, each equipped with a flexible quaternary pump (G7104C), a vialsampler with heatable column compartment (G7129C) and a multi-wavelength detector (G7165A). An ACE C-18/PFP column (150×4.6mm; 5 μm; 100 Å) was used with a binary solvent system grading from 100% A/0% B to 99% A/1% B over 18 min (A: 0.1% TFA in H₂O; B: 0.1% TFA in MeCN). Quantifications were performed on both instruments, the respective distinctive calibration curves and regression calculations are given below (Table S2, Figure S1).
Table S3: Proportionality factors for the quantification of nucleobases and D$_2$-5mdC in the nucleobase experiments based on the nucleoside standard calibrations given in Table S2 and Figure S1.

| Nucleobase/Nucleoside | Proportionality Factor |
|-----------------------|------------------------|
| 5mC/5mdC              | 0.692915425            |
| 5hmC/5hmdC            | 0.652580968            |
| 5fC/5fdC              | 0.606517414            |
| 5caC/5cadC            | 0.409907652            |
| D$_2$-5mdC/5mdC       | 0.755318881            |

1.8. UV/Vis Oligonucleotide Quantification

The concentration of oligonucleotide reaction samples were measured after filtration, lyophilisation and re-suspension. The absorption at $\lambda = 260$ nm was measured on a NanoDrop (NanoDrop Technologies, cat. no. ND-1000) and the concentration determined by the following equation (Equation 1).

$$c = \frac{A_{260}}{\varepsilon_{260}}$$

Equation 1

c [$\text{mM}$] is the concentration, $A_{260}$ [cm] is the absorption measured at $\lambda = 260$ nm corrected for $d = 1$ cm path length, $\varepsilon_{260}$ [$\text{mM}^{-1}\text{cm}^{-1}$] is the absorption coefficient of the relevant oligonucleotide (as supplied by the manufacturer Sigma Aldrich).

Table S4: Abbreviations for the used oligonucleotide strands and corresponding absorption coefficients at $\lambda = 260$ nm.

| Strand sequence (5'->3')             | Abbreviation | $\varepsilon_{260}$ [$\text{mM}^{-1}\text{cm}^{-1}$] |
|--------------------------------------|--------------|-----------------------------------------------------|
| HO-CGUAAAC5mdCG-OH                   | O            | 90                                                  |
| HO-CGUAAACCG-OH                      | No5mC        | 87                                                  |
| HO$_2$PO-CGUAAACCG-OH                | No5mC-P      | 87                                                  |

1.9. Matrix assisted laser desorption ionization mass spectrometry (MALDI MS)

MALDI mass spectra were recorded on a Bruker Daltonics Autoflex ll time-of-flight mass spectrometer equipped with a nitrogen laser ($\lambda = 337$ nm). 3-hydroxypicolinic acid was used as matrix substance. The acceptable deviation of the instrument is $\pm 3$ m/z and peaks are observed with peak widths of 5-15 m/z. Therefore, the acceptable deviation in m/z differences was deemed to be 10-20 m/z. It should be noted that this means that cytosine/cytidine cannot be differentiated from uracil/uridine. Differentiation between guanine/guanosine and adenine/adenosine should be possible.

1.10. Bond Dissociation Energies and Measured Kinetic Constants

Table S5: Calculated bond dissociation energies of various relevant carbon atoms in nucleoside substrates and rate constants previously measured for the reaction of 5mC to 5hmC to 5fC to 5caC as nucleobases (using 1) or in DNA context (using TET2).

| Hydrogen Atom[a] | BDE [kcal mol$^{-1}$][b] | $k_0$ (1) [L mol$^{-1}$ s$^{-1}$][c] | $k_{cat}$ (TET2) [10$^{-3}$ s$^{-1}$][d] |
|------------------|--------------------------|-------------------------------------|-----------------------------------------|
| 5mC (CH$_3$)     | 90.4                     | 7.37                                | 2.12                                    |
| 5hmC (CH$_3$:OH) | 86.2                     | 63.56                               | 0.63                                    |
| 5fC (CHO)        | 92.9                     | 12.81                               | 0.46                                    |
| Deoxyribose 1'   | 87.7                     | -                                   | -                                       |
| Deoxyribose 5'   | 92.4                     | -                                   | -                                       |

[a] The indicated hydrogen atom refers to the free nucleobase in the case of 5mC, 5hmC, and 5fC. For deoxyribose the relevant hydrogens atoms were calculated for deoxy-guanosine. [b] Nucleobase BDEs calculated by Hu et al.[20], deoxyribose BDEs calculated by Steenken et al.[21] [c] Measured by Jonasson and Daumann [d] Measured by Hu et al.[20]
Experimental Section

2. Nucleoside Experiments

2.1. General Procedure for HPLC Experiments

Reactions of the respective nucleoside substrate with 1 were conducted in water at room temperature. The iron(IV)-oxo complex \([\text{Fe}^{IV}(O)(\text{Py}3\text{Me}_2\text{H})][\text{Ce(NO}_3)_6]\) was treated as described by Jonasson and Daumann to remove the cerium counter ion to obtain 1 and then diluted to reach a final concentration \(c = 10 \text{ mM}\). This solution was added to a solution of the nucleoside substrate in volumes reflecting the envisioned stoichiometric ratio (final concentration of 1: 5 mM), whereupon the time count was started. After appropriate time points samples were taken from the reaction mixture, filtered through silica (2.0–2.5 mL) to stop the reaction, and the nucleoside and nucleobase products were eluted with 7.0 mL water. The eluted solution was freeze-dried and the resulting solid resuspended in water. The suspension was then centrifuged (4000 rpm, 2 min) and the supernatant subjected to HPLC analysis.

2.2. HPLC Traces

2.2.1. Overview

![HPLC traces of a mixture of nucleosides](image)

**Figure S2.** HPLC references traces of a mixture of 5mC, 5hmC, 5fC, 5cC, 5mdC, 5hmdC, 5fdC, and 5cadC, used for calibration purposes and shown here as reference mixture. The small signal at 4.1 min is an impurity found in 5hmdC.
2.2.2. HPLC trace of 5mdC, 5hmdC, 5fdC, and 5cadC

Figure S3. Superimposed HPLC references traces of 5mdC, 5hmdC, 5fdC, and 5cadC. The small signal at 4.1 min is an impurity found in 5hmdC, the small signal at 2.5 min is assigned to trace amounts of 5caC in the 5cadC sample.

2.2.3. HPLC trace of 5mC, 5hmC, 5fC, and 5caC

Figure S4. Superimposed HPLC references traces of 5mC, 5hmC, 5fC, 5caC.
2.2.4. HPLC trace of 1 after filtration

Figure S5. HPLC reference trace of a sample of 1 that was treated like a regular reaction (filtration through silica, lyophilisation) as described under Section 2.1.

2.2.5. Spike Experiments – No Additions

In these experiments, the expected products were added individually to a product mixture obtained from the reaction of 1 with 5mdC ([1] = 5 mM, [5mdC] = 5 mM, H2O, T = 22 °C) and the resulting mixture analysed using HPLC. In the following, the original HPLC without any additions (Figure S6) as well as the spiked traces are shown. These spiked traces are combined in corresponding nucleoside/nucleobase pairs (Figure S7 - Figure S10).

Figure S6. HPLC trace of a representative reaction sample of the reaction of 1 with 5mdC.
2.2.6. Spike Experiments – 5mdC or 5mC

Figure S7. Superimposed HPLC traces of a representative reaction sample of the reaction of 1 with 5mdC with added 5mdC or 5mC.

2.2.7. Spike Experiments – 5hmdC or 5hmC

Figure S8. Superimposed HPLC traces of a representative reaction sample of the reaction of 1 with 5mdC with added 5hmdC or 5hmC.
2.2.8. Spike Experiments – 5fdC or 5fC

Figure S9. Superimposed HPLC traces of a representative reaction sample of the reaction of 1 with 5mdC with added 5fdC or 5fC.

2.2.9. Spike Experiments – 5cadC or 5caC

Figure S10. Superimposed HPLC traces of a representative reaction sample of the reaction of 1 with 5mdC with added 5cadC or 5caC.
2.3. GC-MS Reference Samples

2.3.1. 2'-Deoxyribo-γ-lactone

**Figure S11.** GC-MS trace of 2'-deoxyribo-γ-lactone as reference sample.

**Figure S12.** EI MS Spectrum at a retention time of 18.0 min, assigned to 2'-deoxyribo-γ-lactone, confirmed by 2020 NIST database.
2.3.2. 5-methyl-2'-deoxyribo-cytidine

Figure S13. GC-MS trace of 5mdC as reference sample.

Figure S14. EI MS Spectrum at a retention time of 12.2 min, assigned to 2’-deoxyribose.
Figure S15. El MS Spectrum at a retention time of 19.1 min, assigned to 5mC.[7]

2.3.3. Overlay of References

Figure S16. Superimposed GC-MS traces (compare Figures S13 and S14) of 2'-deoxy-ribo-γ-lactone and 5mdC as reference samples.
2.4. Reactions of 1 with substrates

2.4.1. Reaction of 1 with 5mdC

Scheme S1: Reaction of 5mdC with 1 in water and observed products. Conditions: [5mdC] = 1 mM; [1] = 5 mM; T = 22 °C; t = 15 min or 30 min (reagent ration experiments) or t = up to 70 min (time-resolved reaction).

Reagent ratio experiments
To a solution of 5mdC in water (2 mM, 2.5 mM, 3.3 mM, 5 mM, 10 mM; 500 µL each) was added the appropriately prepared solution of 1 in water (10 mM, 500 µL). Each of the experiments was performed twice. One series was then stirred for 15 min, the other one for 30 min during which a color change from green to slightly brown was observable. After the respective time, the reactions were filtered through silica as described above. The freeze-dried residues were resuspended in water (either 1000 mL for 1:1 and 1:2, or 500 µL for 1:3, 1:4 and 1:5) and subjected to HPLC analysis (5.0 µL injection) as described above.

Time-resolved reaction
To a solution of 5mdC in water (2.0 mM, 1.60 mL, 1.0 equiv.) was added the appropriately prepared solution of 1 in water (10 mM, 1.60 mL, 5.0 equiv.) and the reaction stirred for 70 min during which a color change from green to orange-brown was observable. Every 5 min, samples (200 µL) were taken as described above. The freeze-dried residue was resuspended in water (200 µL) and subjected to HPLC analysis (5.0 µL injection) as described above.

Representative HPLC trace (1:5 equivalents, 30 min)

Figure S17. HPLC trace of a representative reaction sample of the reaction of 1 (5 mM) with 5mdC (1 mM) after 30 min.
Mass spectrometric analysis of the reaction
A reaction between 5mdC and 1 (1.00 mL, final concentrations: [5mdC] = 1.0 mM, [1] = 5.0 mM) was carried out for 70 min and worked up as described above. The freeze-dried residue was utilized for LC-MS and GC-MS measurements.
For LC-MS, the reaction sample was resuspended in water to the original 5mdC starting concentration (max. 1.0 mM). The experiment was carried out with the ACE C-18/PFP column used for HPLC experiments and with 100% aqueous NH₄HCO₃ buffer pH 5.5.

For GC-MS the reaction sample was suspended in acetonitrile (200 uL) and BSTFA (200 uL) added. The mixtures were heated to 70 °C for 30 min and then subjected to GC-MS analysis.

**GC-MS trace (1:5 equivalents, 30 min)**

![GC-MS trace](image)

**Figure S18.** GC-MS trace of a representative reaction sample of the reaction of 1 (5 mM) with 5mdC (1 mM) after 30 min.

**Overlaid with reference samples**

![GC-MS trace overlaid](image)

**Figure S19.** GC-MS trace of a representative reaction sample of the reaction of 1 (5 mM) with 5mdC (1 mM) after 30 min overlaid with references.
Figure S20. EI MS Spectrum at a retention time of 19.1 min, assigned to 2'-deoxyribose.

Figure S21. EI MS Spectrum at a retention time of 19.1 min, assigned to 2'-deoxyribono-γ-lactone, confirmed by 2020 NIST database.

Figure S22. EI MS Spectrum at a retention time of 19.1 min, assigned to 5mC[5], confirmed by 2020 NIST database.
2.4.2. Reaction of 1 with 5hmC

Scheme S2: Reaction of 5hmC with 1 in water and observed products. Conditions: [5hmC] = 1 mM; [1] = 5 mM; T = 22 °C; t = up to 70 min (time-resolved reaction).

This reaction was performed identical to the reaction with 5mdC as substrate, as described in Section 2.1. The observed color change was significantly faster than in the case of 5mdC.

Representative HPLC trace (1:5 equivalents, 30 min)

Figure S23. HPLC trace of a representative reaction sample of the reaction of 1 (5 mM) with 5hmC (1 mM) after 30 min.
Nucleoside quantification

Figure S24. Detected amounts of 5hmC, 5fC, and 5cadC in the reaction between 1 and 5hmC after certain time intervals (duplicate measurement). Conditions: [5hmC] = 1 mM, [1] = 5mM, H₂O, T = 22 °C. Data points are averaged from two replicates.
2.4.3. Reaction of 1 with d3-5mdC

Scheme S3: Reaction of D3-5mdC with 1 in water and observed products. Conditions: [5hmC] = 1 mM; [1] = 5 mM; T = 22 °C; t = up to 70 min (time-resolved reaction).

This reaction was performed identical to the reaction with 5mdC as substrate, as described in Section 2.1. The observed color change was significantly slower than in the case of 5mdC.

Representative HPLC trace (1:5 equivalents, 30 min)

Figure S25. HPLC trace of a representative reaction sample of the reaction of 1 (5 mM) with D3-5mdC (1 mM) after 30 min.
2.4.4. Filtration and control experiments

To test whether all products were completely eluted (to ensure accuracy of quantification) under the filtration conditions described above, one series with three samples (200 µL) of mixtures of the eight monitored reaction products (either 0.02 mM, 0.20 mM or 1.0 mM) was filtered through silica (2.4 mL). A second identical series was directly freeze-dried. The freeze-dried residues were resuspended in water (200 µL) and the supernatant subjected to HPLC analysis (5.0 µL injection) as described above.

Similarly, it was tested whether the nucleoside products would hydrolyze during the work-up procedure. A solution of all four nucleoside products (200 µL, 0.50 mM each) was filtered through silica (2.4 mL) and worked up as described above. The freeze-dried residue was resuspended in water (200 µL) and the supernatant subjected to HPLC analysis (5.0 µL injection).
3. Oligonucleotide Experiments

3.1. General Procedure for Oligonucleotide Experiments

The oligonucleotides were dissolved in MilliQ water so that a concentration of \(c = 0.5 \text{ mM}\) or \(c = 1 \text{ mM}\) was reached. The required amount of oligo was then transferred into an Eppendorf reaction tube and diluted with water to reach the desired concentration (see individual experiments for details). The iron(IV)-oxo complex \([\text{Fe}^6\text{O}(\text{Py}_3\text{Me}_2\text{H})][\text{Ce}(\text{NO}_3)_3]\) was treated as described by Jonasson and Daumann to remove the cerium counter ion to obtain 1 with a mixture of \(\text{F}^-\), \(\text{OH}^-\), and \(\text{NO}_3^-\) counter ions and diluted to reach a final concentration \(c = 5 \text{ mM}\). Then, the required amount of 1 was added to the Eppendorf reaction tubes containing the oligo sample so that a final volume of \(v = 50 \mu\text{l}\) was reached. The samples were then incubated at room temperature (22-25 °C) for \(t = 45\) min. Subsequently, the samples were diluted with another 50 \(\mu\text{l}\) of water and filtered through a pad of silica as described previously by Jonasson and Daumann (0.3 ml silica in a 1 ml syringe, washed with 2 x 400 \(\mu\text{l}\) water). The samples were lyophilized to dryness and then re-suspended in 250 \(\mu\text{l}\) of water. After syringe filtration (PTFE, 45 µm) the oligonucleotide concentration of the samples was determined via UV/Vis absorption at \(\lambda = 260\) nm (see chapter Materials and Methods). For time resolved reaction monitoring, a sample \(\text{O} ([\text{O}] = 0.25 \text{ mM})\) in 250 \(\mu\text{l}\) water was prepared. An aqueous solution of 1 (250 \(\mu\text{l}\), [1] = 1 mM) was added, so that a final concentration of \([\text{O}] = 0.125 \text{ mM}\) and [1] = 0.5 mM was reached. The mixture was incubated at room temperate (22 °C) and samples (50 \(\mu\text{l}\)) were taken after reaction times of 0.5 min, 5 min, 10 min, 17 min, 26 min, 37 min and treated as described above. The samples were then treated as described above. The obtained solid product mixtures were then resuspended in 500 \(\mu\text{l}\) of water.

3.2. Oligonucleotide Concentration

The concentration of the oligonucleotide was determined as described above (see Materials and Methods). Two measurements of \(A_{260}\) were taken for each sample and the result averaged.

From this, the concentration was calculated \((c_{\text{calc.}})\) using Equation 1. Then, the background generated during the experimental procedure (see sample \(\text{H}_2\text{O}\)) was subtracted to obtain \(c_{\text{corr.}}\) which was then used for calculation of the necessary amounts for the digestion procedure.

Table S6: Overview of the conducted oligonucleotide experiments, the respective measured absorption values at \(\lambda = 260\) nm and the thereby calculated concentrations. For concentration calculations the absorption coefficient \(\varepsilon_{260}\) supplied in Table S4 was used.

| Oligo 2 | Experimental Conditions | Concentration | A_{260} | c_{corr} [µM] | c_{calc} [µM] | Meas. 2 | Meas. 1 | Cc [µM] | Cc [µM] |
|---------|-------------------------|---------------|---------|--------------|--------------|---------|---------|----------|----------|
| Untreated|                         |               |         |              |              |         |         |          |          |
| O-2 + Fe1 |                         |               |         |              |              |         |         |          |          |
| O-2 + Fe2 |                         |               |         |              |              |         |         |          |          |
| TD1-0.5 min |                         |               |         |              |              |         |         |          |          |
| TD1-1 min |                         |               |         |              |              |         |         |          |          |
| TD1-10 min |                         |               |         |              |              |         |         |          |          |
| TD1-17 min |                         |               |         |              |              |         |         |          |          |
| TD2-26 min |                         |               |         |              |              |         |         |          |          |
| TD1-37 min |                         |               |         |              |              |         |         |          |          |
| TD2-0.5 min |                         |               |         |              |              |         |         |          |          |
| TD2-2.5 min |                         |               |         |              |              |         |         |          |          |
| TD2-5 min |                         |               |         |              |              |         |         |          |          |
| TD2-10 min |                         |               |         |              |              |         |         |          |          |
| TD2-17 min |                         |               |         |              |              |         |         |          |          |
| TD2-26 min |                         |               |         |              |              |         |         |          |          |
| TD2-37 min |                         |               |         |              |              |         |         |          |          |
| ds-O* |                         |               |         |              |              |         |         |          |          |
| ss-O* |                         |               |         |              |              |         |         |          |          |
| ss-C* |                         |               |         |              |              |         |         |          |          |
| No5mC-C |                         |               |         |              |              |         |         |          |          |
| No5mC-CP |                         |               |         |              |              |         |         |          |          |
| O-2 control |                         |               |         |              |              |         |         |          |          |
| Fe control |                         |               |         |              |              |         |         |          |          |
| H2O** |                         |               |         |              |              |         |         |          |          |

*) These reactions contained additionally 2.5 mM NaCl. **) This sample was treated as all reaction samples (incubated, filtered, lyophilized, re-suspended, syringe-filtered) and used for determination of the background absorbance at \(\lambda = 260\) nm.

3.3. Digestion and Quantification Procedure

0.1 µg of oligonucleotide in 35 µL \(\text{H}_2\text{O}\) were digested as follows: 1 µL of Degradase plus (Zymo research), 2.5 µL of Degradase plus reaction buffer (Zymo research), 0.05 µL of Benzonase nuclease (Merck, Serratia marcescens) and a specific amount of isotopically labelled internal standards were added. The mixture was incubated for 4 h at 37 °C and then stored at -20 °C. Prior to LC-MS/MS analysis, samples were filtered by using an AcroPrep Advance 9 filter plate 0.2 µm Supor (Pall Life Science).
3.4. Quantification Results

Here we present some additional data on the detection of 8-oxo-dG in the reaction samples of 1 with O (Figure S27, lower graph). Also, we show an additional set of graphs for the individual nucleoside levels detected in the time-dependent reaction of 1 with O (Figure S28). The data is the same as in Figure 4 in the manuscript, it is simply another mode of presentation meant to provide more clarity.

**Figure S27:** Detected amounts of 5mdC, 5hmdC, 5fdC, 5cadC (upper graph) and 8-oxo-dG (lower graph) in reaction of 1 with O. Conditions: [O] = 0.25 mM, [1] = 0.25 mM / 1.0 mM, H2O, T = 22 °C, reaction time = 30 min. UL0D refers to under level of detection.
3.5. Product Ratios

In the reaction of 1 with the oligonucleotide O, we detected different product ratios depending on the amount of 1 that was used. Figure S29 shows this for the reaction of O with 1 and 4 equiv. of 1, respectively.

As expected, a higher amount of 1 leads to more oxidized species. It remains to be tested whether the continuous addition of 1 leads to accumulation of 5hmC or 5fdC over 5cadC. Whereas the ratios 5cadC/5hmC and 5cad/5fdC quickly approach a constant value (0.08-0.1) the ratio between 5fdC/5hmC shows a maximum after around 15 min of reaction time. This could prove useful to stir the reaction towards producing mostly 5fdC formation.
3.6. Side-Product Analysis

In the following, characteristic MALDI mass spectra are discussed in detail and possible fragmentation pathways are proposed.

**O control**

A signal at ~2620 m/z was observed in all measurements, including the O control (Figure S30, treated as reaction mixture without exposure to 1) oligonucleotide sample. This signal probably stems from production and was attributed to a 9-mer oligonucleotide. Additionally, a signal at ~3205 m/z was also observed in all measurements, including the O control oligonucleotide sample. This signal probably also stems from production and was attributed to an 11-mer oligonucleotide. Very small amounts of signals between 2930-3100 m/z were also observed but could not be assigned to any species.

**O + 1 equiv. 1**

In the reaction of O with 1 equiv. of 1 several additional signals compared to the control sample were observed. In Figure S31 a possible fragmentation pathway that would result in additional amounts of the signal at 2626 m/z is shown. During this pathway, a cytidine-3’-phosphate residue is generated. It is unclear whether this reaction is actually occurring or whether only residual amounts of the 9-mer (vide supra) are detected. In Figure S32, two oxidative pathways are shown resulting in the loss of guanine or cytosine and the emergence of the signals at 2810 m/z and 2834 m/z, respectively. These proposed pathways are further discussed in Scheme S4.

---

**Figure S30**: Excerpt of the MALDI MS spectrum obtained for the control reaction of O without 1. Small signals are observed at 2626 m/z and 3205 m/z, these were assigned to 9-mer and 11-mer contaminants stemming from production.

**Figure S31**: Excerpt of the MALDI MS spectrum obtained for the reaction of O with 1 equiv. of 1. A possible fragmentation pathway resulting in the observed signal at 2626 m/z is indicated. The signal at 3205 m/z was observed in all samples and attributed to an 11-mer oligonucleotide contamination from production of the oligonucleotide. The signals between 2800-3000 m/z are shown enlarged in Figure S32.
Figure S32: Zoomed in excerpt of the MALDI MS spectrum obtained for the reaction of O with 1 equiv. of 1. Possible fragmentation pathway resulting in the observed signals at 2810 m/z and 2834 m/z are indicated (loss of guanine or cytosine, blue and red, respectively). The signal at 2946, 2949 and 2965 m/z were attributed to 5hmC, 5fcC, and 5cadC-containing 10-mer oligonucleotides stemming from the reaction of O with 1. The signal at 2822 m/z could not be assigned clearly to any side reaction, however, loss of adenine might present a possible explanation.

O + 4 equiv. 1
In this sample, the same fragmentation pattern is detected in MALDI MS as in the one that was only exposed to 1 equiv. of 1 (Figure S31 and Figure S32), albeit higher amounts of reacted/decomposed oligonucleotides are observed. Additional small signals were detected that could not be assigned.

Figure S33: Excerpt of the MALDI MS spectrum obtained for the reaction of O with 4 equiv. of 1. A possible fragmentation pathway resulting in the observed signal at 2626 m/z is indicated (green). The signals between 2800-3000 m/z are shown enlarged in Figure S34, those between 2860-2940 m/z are shown enlarged in Figure S35.
**Supporting Information**

**Figure S34:** Zoomed excerpt of the MALDI MS spectrum obtained for the reaction of O with 4 equiv. of 1. Possible fragmentation pathway resulting in loss of guanine or cytosine (blue and red, respectively) are indicated.

**Figure S35:** Zoomed excerpt of the MALDI MS spectrum obtained for the reaction of O with 4 equiv. of 1. Possible fragmentation pathway resulting in loss of an additional guanine or cytosine (blue: 2834 m/z to 2703 m/z and red: 2834 m/z to 2741 m/z or 2810 m/z to 2717 m/z) are indicated. Fragments with m/z at 2834 and 2810 resulted from a previous loss of a cytosine or guanidine fragment, respectively (Figure S34).

**Control Reactions: O-2 and O-2P**

In order to screen for additional side reactions, we exposed the “control strand” O-2 HO-CCUUAACCCG-OH to 4 equiv. of 1. Using MALDI MS (Figure S36 and Figure S37) we detected a very significant amount of decomposition reactions – much more than if the regular oligonucleotide O is used. Loss of guanine and cytosine is detected from the starting material and from the decomposition product that resulted from loss of cytosine or cytidine-3’-phosphate in the first place.

We also wanted to see if the presence of a phosphate cap on the 5' end of the oligonucleotide had any effect on the decomposition reactions, we therefore exposed O-2P HO3POCCUUAACCCG-OH to 4 equiv. of 1. We did the same fragmentation patterns as already discussed, however, it seems that less loss of cytosine is observed compared to HO-CCUUAACCCG-OH (Figure S37).
Figure S36: Excerpt of the MALDI MS spectrum obtained for the reaction of the “control strand” O-2 HO-CCUAACCG-OH with 4 equiv. of 1. A possible fragmentation pathway resulting in the observed signal at 2620 m/z is indicated (green). The signals between 2650-2950 m/z are shown enlarged in Figure S37.

Figure S37: Zoomed in excerpt of the MALDI MS spectrum obtained for the reaction of the “control strand” O-2 HO-CCUAACCG-OH with 4 equiv. of 1. Possible fragmentation pathway resulting in loss of guanine or cytosine (blue and red, respectively) are indicated.
Figure S38: Excerpt of the MALDI MS spectrum obtained for the reaction of the “control strand” O-2P HO₂PO-CCUUAACCCG-OH with 4 equiv. of 1. A possible fragmentation pathway resulting in the observed signal at 2708 m/z is indicated (green). The signals between 2750-3050 m/z are shown enlarged in Figure S37.

Figure S39: Zoomed in excerpt of the MALDI MS spectrum obtained for the reaction of the “control strand” O-2P HO₂PO-CCUUAACCCG-OH with 4 equiv. of 1. Possible fragmentation pathway resulting in loss of guanine or cytosine (blue and red, respectively) are indicated. The signal at 2782 m/z is also probably formed by loss of guanine from the signal at 2895-2907 m/z (not indicated by an arrow).

Mechanistic Proposal: Hydroxylation of Position 5’ or 1’.
Scheme S4: A) Hydroxylation/oxidation of the 5' position of the second-to-last nucleotide on the 5' end of the oligonucleotide strand resulting in loss of a cytidine 3'-phosphate fragment and formation of an aldehyde function on the nucleotide (green). B) Hydroxylation/oxidation of the 1' position of the last nucleotide on the 5' end of the strand resulting in loss of a cytosine nucleobase fragment and formation of a lactone (red). C) Hydroxylation/oxidation of the 1' position of the last nucleotide on the 3' end of the strand resulting in loss of a guanine nucleobase fragment and formation of a lactone (blue). D) Oxidation of the 5 position of the 5mdC residue on the second-to last position on the 3' end of the strand resulting the formation of 5hmC, 5fdC, and 5cadC residues.
4. Synthetic Procedures

4.1. Synthesis of 1

$[\text{Ce(NO}_3\text{)}_6]$ was synthesized according to published literature procedures. Exchange of the cerium counterion to a mixture of $\text{F}^-$, $\text{OH}^-$ and $\text{NO}_3^-$ was performed according to the literature procedure.

Scheme S5: Four step synthesis of $[\text{Fe(O)(Py}_{5}\text{-Me}_{2}\text{-H})][\text{Ce(NO}_3\text{)}_6]$ as performed in this work, adapted from Cong et al. and Chang et al.

A detailed discussion of the synthesis of $[\text{Ce(NO}_3\text{)}_6]$ as well as the exchange of the counterion can be found in our previous publication involving 1.

4.2. Synthesis of 5-formyl-2’deoxy-cytidine (5fdC)

Scheme S6: Four step synthesis of 5fdC as performed in this work, adapted from the literature.
4.2.1. Synthesis of silyl protected 2'-deoxy-cytidine (2)

According to a modified standard literature procedure: [8]
Both 2'-deoxycytidine (dC; 1.00 g, 4.40 mmol, 1.00 equiv.) and imidazole (1.50 g, 22.0 mmol, 5.00 equiv.) were dissolved in anhydrous N,N-dimethylformamide (8.8 mL) and the solution was cooled to 0 °C. tert-Butyldimethylsilyl chloride (1.99 g, 13.2 mmol, 3.00 equiv.) was added in three portions and the reaction was allowed to slowly warm to room temperature. After 16 h, the mixture was diluted with ethyl acetate (30 mL) and saturated aqueous sodium bicarbonate solution (80 mL) was added slowly. The layers were separated and the aqueous layer was extracted with ethyl acetate (3×50 mL). The combined organic layers were washed with saturated aqueous sodium chloride solution (50 mL) and dried over sodium sulfate. The dried solution was filtered and the filtrate was concentrated in vacuo. The resulting yellow viscous oil was purified by flash column chromatography (2% to 5% methanol in dichloromethane) to afford silyl-protected 2'-deoxy-cytidine 2 (1.93 g, 4.23 mmol, 96%) as a foaming colorless solid.

\(^1\)H NMR (400 MHz, CDCl\(_3\), 295 K): δ = 8.03 (dd, J = 7.2, 3.2 Hz, 1H), 7.78 (s, 0H), 7.12 (s, 1H), 6.26 (dd, J = 6.4, 5.0 Hz, 1H), 5.66 (d, J = 7.3 Hz, 1H), 4.36 (q, J = 5.8 Hz, 1H), 3.98 – 3.84 (m, 2H), 3.83 – 3.70 (m, 1H), 2.43 (dt, J = 12.5, 6.0 Hz, 1H), 2.17 – 2.02 (m, 1H), 0.92 (s, 9H), 0.87 (s, 9H), 0.10 (d, J = 2.0 Hz, 5H), 0.05 (s, 6H).

HR-MS (ESI): calculated for C\(_{21}\)H\(_{42}\)N\(_3\)O\(_4\)Si\(_2\) [M+H]\(^+\): 456.2708 found: 456.2705.

4.2.2. Synthesis of silyl protected 5-iodo-2'-deoxy-cytidine (3)

According to a modified standard literature procedure: [9]
Silyl-protected 2'-deoxycytidine 2 (1.79 g, 3.92 mmol, 1.00 equiv.) and iodine (2.19 g, 8.62 mmol, 2.20 equiv.) were dissolved in anhydrous acetonitrile (175 mL). Ceric ammonium nitrate (4.73 g, 8.62 mmol, 2.20 equiv.) was added, upon which the solution turned dark brown, and the reaction was heated to 60 °C under exclusion of light for 2.5 h. Then, the reaction was cooled to room temperature and saturated aqueous sodium bicarbonate solution (50 mL), followed by aqueous sodium thiosulfate solution (1.0 M, 30 mL) were added, giving a slightly yellow suspension. The suspension was diluted with water (100mL) and acetonitrile was evaporated under reduced pressure. The residual suspension was filtered through a plug of celite and the filter cake was extensively washed with ethyl acetate (300 mL). The layers were separated and the aqueous layer was extracted with ethyl acetate (4×80 mL). The combined organic layers were washed with saturated aqueous sodium chloride solution (100 mL) and dried over sodium sulfate. The dried solution was filtered and the filtrate was concentrated in vacuo. The resulting yellow solid was purified by flash column chromatography (0% to 2% to 4% methanol in dichloromethane) to afford silyl-protected 5-iodo-2'-deoxy-cytidine 3 (1.20 g, 2.06 mmol, 53%) as a foaming colorless solid.

\(^1\)H NMR (400 MHz, CDCl\(_3\), 295 K): δ = 8.02 (s, 1H), 6.21 – 6.13 (m, 1H), 5.37 (s, 1H), 4.28 (td, J = 6.0, 3.0 Hz, 1H), 3.91 (q, J = 2.7 Hz, 1H), 3.82 (dd, J = 11.4, 2.6 Hz, 1H), 3.69 (dd, J = 11.4, 2.6 Hz, 1H), 2.40 (dd, J = 13.3, 5.9, 3.1 Hz, 1H), 1.89 (ddd, J = 13.3, 7.1, 6.1 Hz, 1H), 0.84 (d, J = 19.8 Hz, 17H), 0.07 (d, J = 4.5 Hz, 6H).

HR-MS (ESI): calculated for C\(_{21}\)H\(_{42}\)I\(_2\)N\(_3\)O\(_4\)Si\(_2\) [M+H]\(^+\): 582.1675 found: 582.1680.
4.2.3. Synthesis of silyl protected 5-formyl-2'-deoxy-ctydine (4)

According to a modified standard literature procedure,[10] silyl-protected 5-iodo-2'-deoxycytidine 3 (600 mg, 1.03 mmol, 1.00 equiv.), [Pd₂(dba)₃] chloroform adduct (107 mg, 0.103 mmol, 0.100 equiv.) and triphenylphosphine (162 mg, 0.618 mmol, 0.600 equiv.) were dissolved in anhydrous toluene (15 mL) in a high-pressure reactor. The reactor was flushed with carbon monoxide gas (3.5 bar) while stirring the reaction solution and the pressure was released after 10 min. This flushing procedure was repeated two more times to achieve carbon monoxide saturation of the reaction solution. Finally, the reactor was filled with carbon monoxide gas (3.5 bar) and the reaction mixture was heated to 65 °C. A solution of tributyltin hydride (0.33 mL, 1.2 mmol, 1.2 equiv.) in anhydrous toluene (0.33 mL) was added dropwise via syringe over 18 h. After 13 h of the addition, the pressure had dropped to normal pressure which is why the reactor was refilled with carbon monoxide gas (3.5 bar). After another 7 h, the reaction mixture was cooled to room temperature and the pressure was released. The greenish-black suspension was diluted with ethyl acetate (20 mL) and filtered through a plug of silica. The filter cake was extensively washed with ethyl acetate (200 mL) and the filtrate was concentrated in vacuo. The resulting orange oil was purified by flash column chromatography on a mixture of silica gel and potassium carbonate (9:1; 20% to 33% to 50% ethyl acetate in hexanes) to afford silyl-protected 5-formyl-2'-deoxycytidine 4 (232 mg, 0.480 mmol, 47%) as a yellow solid.

**1H NMR (400 MHz, CDCl₃, 295 K):**
\[ \delta = 9.49 \text{ (s, } 1\text{H}), 8.56 \text{ (s, } 1\text{H}), 8.19 \text{ (s, } 1\text{H}), 6.21 \text{ (t, } J = 6.1 \text{ Hz, } 1\text{H}), 6.10 \text{ (s, } 1\text{H}), 4.35 \text{ (dt, } J = 6.4, 3.7 \text{ Hz, } 1\text{H}), 4.05 \text{ (q, } J = 2.7 \text{ Hz, } 1\text{H}), 3.96 \text{ (dd, } J = 11.6, 2.6 \text{ Hz, } 1\text{H}), 3.78 \text{ (dd, } J = 11.6, 2.5 \text{ Hz, } 1\text{H}), 2.62 \text{ (ddd, } J = 13.6, 6.2, 3.9 \text{ Hz, } 1\text{H}), 2.14 – 2.02 \text{ (m, } 1\text{H}), 0.90 \text{ (s, } 9\text{H}), 0.89 \text{ (s, } 8\text{H}), 0.10 \text{ (s, } 3\text{H}), 0.09 \text{ (s, } 3\text{H}), 0.08 \text{ (s, } 3\text{H}), 0.07 \text{ (s, } 4\text{H)} \]

**HR-MS (ESI):**
calculated for C₂₂H₄₂N₃O₅Si₂ [M+H]⁺: 484.2658
found: 484.2650.

4.2.4. Synthesis of 5-formyl-2'deoxy-cytidine (5fdC)

According to a modified standard literature procedure,[10] silyl-protected 5-formyl-2'-deoxycytidine 4 (229 mg, 0.473 mmol, 1.00 equiv.) was dissolved in anhydrous ethyl acetate (14 mL) in a plastic tube open to air and hydrogen fluoride pyridine complex (70wt% HF, 0.37 mL, 30 equiv.) was added via syringe at room temperature. After 22 h, the yellow suspension was cooled to 0 °C and methoxytrimethylsilane (4.2 mL, 30 mmol, 64 equiv.) was added slowly via syringe. The reaction was allowed to slowly warm to room temperature and after additional two days, the reaction mixture was diluted with methanol (15 mL) and all volatiles were removed in vacuo. The yellow, solid crude product was purified by HPLC (Column: Agilent Prep-C18 column 50×10.0mm; 5 µm; 2% MeCN in H₂O for 8 min) to afford 5-formyl-2'-deoxycytidine (5fdC) (99.8 mg, 0.391 mmol, 83%) as a colorless solid.

**1H NMR (400 MHz, D₂O, 295 K):**
\[ \delta = 9.53 \text{ (s, } 1\text{H}), 8.82 \text{ (s, } 1\text{H}), 6.21 \text{ (t, } J = 6.1 \text{ Hz, } 1\text{H}), 4.44 \text{ (dt, } J = 6.6, 4.7 \text{ Hz, } 1\text{H}), 4.15 \text{ (td, } J = 4.7, 3.3 \text{ Hz, } 1\text{H}), 3.92 \text{ (dd, } J = 12.6, 3.3 \text{ Hz, } 1\text{H}), 3.80 \text{ (dd, } J = 12.6, 4.9 \text{ Hz, } 1\text{H}), 2.59 \text{ (ddd, } J = 14.2, 6.5, 4.9 \text{ Hz, } 1\text{H}), 2.38 \text{ (ddd, } J = 14.2, 6.5, 5.6 \text{ Hz, } 1\text{H)} \]

**HR-MS (ESI):**
calculated for C₁₀H₁₄N₃O₅ [M+H]⁺: 256.0928
found: 256.0926.
4.3. Synthesis of 5cadC

According to a modified standard literature procedure,[11,12]

In a high-pressure reactor, silyl protected 5-iodo-2'-deoxycytidine 3 (400 mg, 0.688 mmol, 1.00 equiv.) and bis(acetonitrile)palladium dichloride (8.9 mg, 34 μmol, 0.050 equiv.) were dissolved in anhydrous methanol (8.5 mL). The reactor was flushed with carbon monoxide gas (3.5 bar) while stirring the reaction solution and the pressure was released after 10 min. This flushing procedure was repeated two more times to achieve carbon monoxide saturation of the reaction solution. A solution of N,N-di-iso-propylethylamine (0.24 mL, 1.4 mmol, 2.0 equiv.) in anhydrous methanol (0.25 mL) was added via syringe in one portion, the reactor was filled with carbon monoxide gas (3.5 bar) and the reaction mixture was heated to 65 °C. After 16 h, the reaction was cooled to 21 °C, the pressure was released and the orange suspension was filtered through a plug of silica. The filter cake was extensively washed with ethyl acetate (100 mL) and the filtrate was concentrated in vacuo. The resulting oily orange crude product was combined with the crude product of a second identical experiment and they were purified by flash column chromatography (33% to 50% to 67% ethyl acetate in hexanes) to afford silyl-protected 5-methoxycarbonyldeoxycytidine 5 (525 mg, 1.02 mg, 74% combined yield) as a slightly yellow powder.

1H NMR (400 MHz, CDCl3):
\[ \delta = 8.58 \ (s, \ 1H), 7.88 \ (s, \ 1H), 6.12 \ (dd, \ J = 7.3, 5.8 \ Hz, \ 1H), 5.81 \ (s, \ 1H), 4.29 \ (dt, \ J = 6.3, 2.4 \ Hz, \ 1H), 4.00 \ (q, \ J = 2.9 \ Hz, \ 1H), 3.79 \ (dd, \ J = 11.3, 3.1 \ Hz, \ 1H), 3.76 \ (s, \ 3H), 3.70 \ (dd, \ J = 11.3, 3.1 \ Hz, \ 1H), 2.56 \ (ddd, \ J = 13.5, 5.9, 2.5 \ Hz, \ 1H), 1.91 \ (ddd, \ J = 13.4, 7.4, 6.1 \ Hz, \ 1H), 0.82 \ (d, \ J = 6.7 \ Hz, \ 18H), 0.01 \ (s, \ 3H), 0.01 \ (s, \ 3H), 0.00 \ (s, \ 2H), -0.00 \ (s, \ 3H). \]

HR-MS (ESI):
calculated for C_{23}H_{42}N_{2}O_{4}Si_{2} [M+H]^+: 514.2763
found: 514.2767.
4.3.2. Synthesis of 5-carboxy-2'-deoxycytidine (5cadC)

According to a modified standard literature procedure,[12] silyl-protected 5-methoxycarbonyldeoxycytidine 5 (520 mg, 1.01 mmol, 1.00 equiv.) was suspended in anhydrous ethyl acetate (30 mL) in a plastic tube open to air and hydrogen fluoride pyridine complex (70 wt% HF, 0.40 mL, 15 mmol, 15 equiv.) was added whereupon a clear solution resulted. After 14 h, the reaction suspension was cooled to 0 °C and methoxymethylsilane (4.5 mL, 32 mmol, 32 equiv.) was added slowly via syringe. After 3 h, the reaction was allowed to warm to room temperature and was stirred for additional 3 h. The resulting yellow suspension was diluted with methanol and all volatiles were removed in vacuo. The yellow crude product was dissolved in a mixture of acetonitrile and water (1:1, 100 mL) and lithium hydroxide (290 mg, 12.1 mmol, 12.0 equiv.) was added. After 18 h, the suspension was acidified to pH 3 with 3.0 M hydrochloric acid and all volatiles were removed in vacuo to give 856 mg of a dark brown solid. A portion of 227 mg was purified by HPLC (Column: Agilent Prep-C18 column 50×10.0 mm; 5 µm; 2% MeCN in H2O for 8 min, MeCN and H2O both contained 0.1% trifluoracetic acid) to afford 5-carboxy-2'-deoxycytidine (5cadC) (30.6 mg, 0.113 mmol, 42% extrapolated yield over two steps) as a colorless solid.

\[ ^1H \text{ NMR (400 MHz, D}_2\text{O)}: \delta = 8.94 \text{ (s, 1H), 8.49 (s, 1H), 6.25 (t, J = 6.1 Hz, 1H), 4.48 (dt, J = 6.5, 4.6 Hz, 1H), 4.13 (td, J = 4.5, 3.3 Hz, 1H), 3.91 (dd, J = 12.6, 3.3 Hz, 1H), 3.80 (dd, J = 12.6, 4.6 Hz, 1H), 2.55 (ddd, J = 14.2, 6.6, 4.9 Hz, 1H), 2.42 (dt, J = 14.2, 6.1 Hz, 1H) ppm.} \]

\[ \text{HR-MS (ESI): calculated for C}_{10}\text{H}_{14}\text{N}_3\text{O}_6 \text{[M+H]}^+: 272.0877 \text{ found: 272.0875.} \]

4.4. Synthesis of nucleobases 5-formylcytosine (5fC) and 5-carboxycytosine (5caC)

4.4.1. Synthesis of 5-formylcytosine (5fC)

5-Formyl-2'-deoxycytidine (5fdC; 119 mg, 0.446 mmol, 1.00 equiv.) was suspended in water (7.0 mL) and concentrated hydrochloric acid (37%, 1.0 mL) was added. The suspension was stirred at room temperature open to air for two days and then neutralized with 25% aqueous ammonia solution. The mixture was diluted with methanol (10 mL) and concentrated in vacuo. The crude product was purified by HPLC (Column: Agilent Prep-C18 column 50×10.0 mm; 5 µm; 100% H2O for 4 min, MeCN and H2O both contained 0.1% trifluoracetic acid) to afford 5-formylcytosine (5fC; 48.6 mg, 0.349 mmol, 78%) as a slightly tan solid.

\[ ^1H \text{ NMR (400 MHz, DMSO, 295 K):} \delta = 9.49 \text{ (s, 1H), 8.50 (s, 1H), 8.49 (s, 1H).} \]

\[ \text{HR-MS (ESI): calculated for C}_{5}\text{H}_{4}\text{N}_3\text{O}_2 \text{[M-H]}^-: 138.0309. found: 138.0308.} \]
4.4.2. Synthesis of 5caC

5-Methoxycarbonyl-2'-deoxycytidine (5cadC; 49.2 mg, 0.172 mmol, 1.00 equiv.) was suspended in water (2.6 mL) and concentrated hydrochloric acid (37%, 0.5 mL) was added. The resulting clear solution was stirred at room temperature open to air for 2 days, then the mixture was heated to 45 °C. After 15 h, the reaction was allowed to cool to room temperature, diluted with methanol (5 mL) and concentrated in vacuo to afford a yellowish solid. This crude product was combined with 5-ethoxycarbonyl-2'-deoxycytidine (37.3 mg, 0.375 mmol combined, 1.00 equiv.) and suspended in water (19 mL). Lithium hydroxide (108 mg, 4.50 mmol, 12.0 equiv.) was added upon which a clear solution resulted. After stirring at room temperature for 24 h, the solvent was removed in vacuo and the crude product purified by HPLC (Column: Agilent Prep-C18 column 50×10.0mm; 5 µm; 1% MeCN in H2O for 4 min, H2O contained 0.1% trifluoracetic acid) to afford 5-carboxycytosine (5caC; 36.2 mg, 0.233 mmol, 62%).

^1H NMR (400 MHz, DMSO, 295 K):
δ = 8.24 (s, 1H).

HR-MS (ESI):
Calculated for C_{10}H_{11}N_{3}O_{2} [M–H]–: 154.0257
Found: 154.0257

4.5. Synthesis of D3-5mdC

According to a modified literature procedure,[13] 5-ido-2'-deoxycytidine 3 (302 mg, 0.518 mmol, 1.00 equiv.) was dissolved in anhydrous tetrahydrofuran (11 mL) and the solution was cooled to −78 °C. A 1.0 M solution of D3-methylmagnesium iodide in diethyl ether (0.52 mL, 0.52 mmol, 1.0 equiv.) was added dropwise via syringe and the clouded solution stirred for 5 min. Tetrais(triphenylphosphine)palladium (59.9 mg, 0.10 mol, 0.10 equiv.) and copper(I) chloride (256 mg, 2.59 mmol, 5.00 equiv.) were added, followed by more D3-methylmagnesium iodide solution (2.6 mL, 2.6 mmol, 5.0 equiv.). The flask was covered with aluminum foil and the reaction mixture stirred at −78 °C for another 50 min. Then the reaction was allowed to warm to 0 °C and stirred overnight, during which it was allowed to warm to room temperature. After 18 h, the yellowish-green suspension was diluted with dichloromethane (30 mL), saturated aqueous ammonium chloride solution (15 mL) and water (15 mL). The layers were separated and the aqueous layer extracted with dichloromethane (3×40 mL). The combined organic layers were washed with saturated aqueous sodium chloride solution (50 mL) and dried over sodium sulfate. The dried solution was filtered and the filtrate was concentrated in vacuo. The crude product was purified by flash column chromatography (0% to 9% to 19% methanol in dichloromethane + 1% triethylamine) to afford an inseparable mixture of TBS-protected 2'-deoxycytidine and TBS-protected D3-5-methyl-2'-deoxycytidine as a yellowish solid. This mixture was then diluted in anhydrous ethyl acetate (15 mL) in a plastic tube in open air and hydrogen fluoride pyridine complex (70% HF, 0.37 mL, 7.3 mmol, 15 equiv.) was added dropwise via syringe. The yellowish solution was stirred at room temperature overnight. After 22 h, the resulting suspension was cooled to 0 °C and methoxytrimethylsilane (2.2 mL, 16 mmol, 32 equiv.) was added. After 30 min, the reaction mixture was allowed to warm to room temperature and was stirred for another 2.5 h. Then, the suspension was diluted with methanol (10 mL) and concentrated in vacuo. The crude product was purifed by HPLC (Column: Agilent Prep-C18 column 50×10.0mm; 5 µm; 1% MeCN in H2O for 7 min, MeCN and H2O both contained 0.1% trifluoracetic acid) to afford D3-5-methyl-2'-deoxycytidine (D3-5mdC; 21.5 mg, 88.0 µmol, 17%) as a colorless solid.

^1H NMR (400 MHz, D2O):
δ = 7.95 (s, 1H), 6.28 (t, J = 6.5 Hz, 1H), 4.48 (dt, J = 6.5, 4.2 Hz, 1H), 4.10 (dt, J = 4.9, 3.8 Hz, 1H), 3.89 (dd, J = 12.6, 3.4 Hz, 1H), 3.79 (dd, J = 12.6, 4.9 Hz, 1H), 2.54 – 2.32 (m, 2H).

^2H NMR (61 MHz, H2O):
δ = 1.98 (s, 3H).

^13C NMR (101 MHz, D2O):
δ = 159.6, 148.8, 141.5, 103.7, 87.0, 86.2, 70.1, 60.9, 39.1.

HR-MS (ESI):
Calculated for C_{10}H_{12}D_{3}N_{3}NaO_{4} [M+Na]^+: 267.1143
Found: 267.1139
5. Appendix
5.1. NMR spectra
5.1.1. TBS-dC 2

\[
\delta [\text{ppm}] \\
6.33 \\
6.24 \\
9.52 \\
9.08 \\
0.99 \\
0.95 \\
1.00 \\
1.02 \\
0.97 \\
1.02 \\
1.00 \\
1.11 \\
7.26 \text{CDCl}_3
\]

5.1.2. 5-iodo-TBS-dC 3

\[
\delta [\text{ppm}] \\
5.70 \\
5.51 \\
8.52 \\
8.31 \\
1.12 \\
1.08 \\
1.10 \\
1.98 \\
1.03 \\
1.29 \\
1.10 \\
1.01 \\
7.26 \text{CDCl}_3
\]
5.1.3. 5-Formyl-TBS-dC 4
5.1.4. 5-Formyl-2'-deoxycytidine (5fdC)
5.1.5. 5-(Methoxycarbonyl)-TBS-dC 5
5.1.6. 5-carboxyl-2'-deoxycytidine (5cadC)
5.1.8. 5-carboxylcytosine (5caC)
5.1.9. \( \text{D}_3\text{-5}-\text{methyl-2'-deoxycytidine (D}_3\text{-5mdC)} \)

\( ^1\text{H} \)
References

[1] G. R. Fulmer, A. J. M. Müller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, Organometallics 2010, 29, 2176–2179.
[2] F. R. Traube, S. Schiffrers, K. Iwan, S. Kellner, P. Spada, M. Müller, T. Carell, Nat. Prot. 2019, 14, 283–312.
[3] L. Hu, J. Lu, J. Cheng, Q. Rao, Z. Li, H. Hou, Z. Lou, L. Zhang, W. Li, W. Gong, M. Liu, C. Sun, X. Yin, J. Li, X. Tan, P. Wang, Y. Wang, D. Fang, Q. Cui, P. Yang, C. He, H. Jiang, C. Luo, Y. Xu, Nature 2015, 527, 118–122.
[4] S. Steenken, S. V. Jovanovic, L. P. Candeias, J. Reynisson, Chem. Eur. J. 2001, 7, 2829–2833.
[5] N. S. W. Jonasson, L. J. Daumann, Chem. Eur. J. 2019, 25, 12091–12097.
[6] T. Chantarojsiri, Y. Sun, J. R. Long, C. J. Chang, Inorg. Chem. 2015, 54, 5879–5887.
[7] J. Cong, D. Kinschel, Q. Daniel, M. Safdari, E. Gabrielsson, H. Chen, P. H. Svensson, L. Sun, L. Kloos, J. Mater. Chem. A 2016, 4, 14550–14554.
[8] D. B. Gophane, S. T. Sigurdsson, Beilstein J. Org. Chem. 2015, 11, 219–227.
[9] M. Münzel, D. Globisch, C. Trindler, T. Carell, Org. Lett. 2010, 12, 5671–5673.
[10] D. Globisch, M. Münzel, M. Müller, S. Michalakis, M. Wagner, S. Koch, T. Brückl, M. Biel, T. Carell, PLoS One 2010, 5, e15367.
[11] M. Münzel, U. Lischke, D. Stathis, T. Pfaffeneder, F. A. Gnerlich, C. A. Deiml, S. C. Koch, K. Karaghiosoff, T. Carell, Chem. Eur. J. 2011, 17, 13782–13788.
[12] B. Steigenberger, S. Schieser, B. Hackner, C. Brandmayr, S. K. Laube, J. Steinbacher, T. Pfaffeneder, T. Carell, Org. Lett. 2013, 15, 366–369.
[13] M. Münzel, D. Globisch, T. Brückl, M. Wagner, V. Welzmiiller, S. Michalakis, M. Müller, M. Biel, T. Carell, Angew. Chem. Int. Ed. 2010, 49, 5375–5377.