Supporting information

Dess-Martin Periodinane-mediated oxidation of the primary alcohol of cytidine into a carboxylic acid

Alexandra Serre[1], Vibhu Jha[2,3], Adèle Rivault[1], Leif A. Eriksson[2], Goreti Ribeiro Morais[1], Robert A. Falconer[1]*

[1]*Institute of Cancer Therapeutics, Faculty of Life Sciences, University of Bradford, Bradford BD7 1DP, U.K.*

[2]*Department of Chemistry and Molecular Biology, University of Gothenburg, 405 30 Göteborg, Sweden.*

[3]*Current address: Institute of Cancer Therapeutics, University of Bradford.*

Email: r.a.falconer1@bradford.ac.uk

Table of contents

General information

All chemicals were provided by Sigma Aldrich or Fluorochem. All reactions were monitored by thin-layer chromatography (TLC silica gel 60 F_{254}, aluminium) and/or by LC/MS (Waters Alliance e2695 separation modules, Waters 2998 photodiode array (PDA) detector 280 nm, Waters Acquity QDA mass detector, Hichrom C_{18} 3.5 µM, 2.1 x 150 mm) according to the following methods (**Figure S19**, supporting data). Novel compound **2** was synthesized similarly to compound **1a**.[25] NMR spectra were recorded on a Bruker AMX 400 NMR spectrometer and are reported in parts per million (ppm) on the δ scale relative to residual $CDCl_3$ (δ 7.25 or δ 77.0), CD_3OD (δ 3.31 or δ 49.00) or DMSO-*d*6 (δ 2.50 or δ 39.52). Spectral assignments were accomplished using 2D COSY and HSQC experiments. High-resolution mass spectrometry was recorded using a Thermo scientific, LTQ Orbitrap no. 01289B, Electrospray. Column chromatography was performed using silica gel (230-400 mesh)

General Procedure for the oxidation of 5'-hydroxy cytidine to 5'-aldehyde cytidine: To a stirred solution of alcohol (25 mg, 1 eq.) in dry CH_2Cl_2 (2.5 mL) was added DMP (2 eq.) and the mixture was stirred at room temperature. After 2 hours, the mixture was diluted with CH_2Cl_2 (20 mL) and washed with sodium thiosulfate (2 x 20 mL), water (2 x 20 mL) and brine (2 x 20 mL). The organic phases were combined and dried over anhydrous Mg_2SO_4, filtered, and concentrated under reduced pressure. LCMS analysis of the crude indicated the following yields of **3a** 68%; **3b** 93%; **3c** > 95%; **7** 59%.

General Procedure for the oxidation of 5'-hydroxy cytidine to 5'-carboxylic acid cytidine: To a stirred solution of protected 5'-hydroxy cytidine (25 mg,1 eq.) in dry CH_2Cl_2 (2.5 mL) was

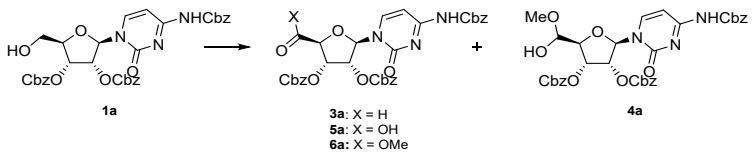
added DMP (4 eq.) and the mixture was stirred at room temperature. After 2 hours, the mixture was further diluted with CH_2Cl_2 (20 mL) and washed with sodium thiosulfate (2 x 20 mL), water (2 x 20 mL) and brine (2 x 20 mL). The organic phases were combined and dried over anhydrous Mg_2SO_4, filtered, and concentrated under reduced pressure. LCMS analysis of the crude indicated the following yields of **5a** 79%; **5b** 96%; **5c** 42%; **8** 95%.

(2S,3S,4R,5R)-5-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)-3,4-bis(((benzyloxy)carbonyl)oxy)tetrahydrofuran-2-carboxylic acid (**5a**) – Column chromatography on silica gel CH_2Cl_2/MeOH/AcOH (97/2/1) afforded **5a** as white solid, η = 67%; ^{1}H NMR (400 MHz, DMSO-*d*6) δ 10.98 (s, 1H), 8.37 (d, 1H, H_6, *J* = 7.6 Hz, 1H), 7.44 – 7.30 (m, 15H, Cbz-Ph), 7.09 (d, 1H, H_5, *J* = 7.6 Hz, 1H), 6.06 (d, 1H, $H_{1'}$, *J* = 4.1 Hz), 5.74 (dd, 1H, $H_{3'}$, *J* = 5.7, 4.7 Hz), 5.54 (dd, 1H, $H_{2'}$, *J* = 5.7, 4.1 Hz), 5.20 (s, 2H, Cbz-(N)-CH_2), 5.17 (s, 2H, Cbz-(O)-$CH_{2(1)}$), 5.10 (s, 2H, Cbz-(O)-$CH_{2(2)}$), 4.74 (d, 1H, $H_{4'}$, *J* = 4.7 Hz); ^{13}C{^{1}H} NMR (101 MHz, DMSO) δ 169.8 (s, $C_{5'}$), 146.7 (s, C_6), 128.5 (m, Cbz-Ph), 95.1 (s, C_5), 90.7 (s, $C_{1'}$), 79.7 (s, $C_{4'}$), 76.4 (s, $C_{2'}$), 75.6 (s, $C_{3'}$), 69.8 (s, Cbz-(O)CH2), 69.7 (s, Cbz-(O)CH_2), 66.7 (s, Cbz-(N)CH_2); Rt = 19.97 min (method 1, UV 254 nm, $[M+H]^+$ = 660); HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{33}H_{29}N_3O_{12}$ 660.1830; found 660.1849.

Synthesis of methyl (2S,3S,4R,5R)-5-(4-(((benzyloxy)carbonyl)amino)-2-oxopyrimidin-1(2H)-yl)-3,4-bis(((benzyloxy)carbonyl)oxy)tetrahydrofuran-2-carboxylate (**6a**) – crude of compound **5a** (25 mg) was dissolved in methanol (2 ml) and stirred with Amberlite IR-120 at room temperature overnight. Then the resin was filtered, and the filtrate evaporated. Column chromatography on silica gel afford **6a** as a white solid, η = 91%; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.43 (d, 1H, H_6, *J* = 7.5 Hz, 1H), 7.42 – 7.29 (m, 16H, Cbz-Ph and H_5), 6.27 (d, 1H, $H_{1'}$, *J* = 3.4 Hz), 5.45 (m, 2H, $H_{2'}$ and $H_{3'}$), 5.23 (s, 2H, Cbz-(N)-CH_2), 5.14 (s, 2H, Cbz-(O)$CH_{2(1)}$), 5.12 (m, 2H, Cbz-(O)$CH_{2(2)}$), 4.75 (d, 1H, $H_{4'}$, *J* = 3.9 Hz), 3.81 (s, 3H, OCH_3); ^{13}C{^{1}H} NMR (101 MHz, $CDCl_3$) δ 169.7 (s, $C_{5'}$), 145.3 (s, C_6), 128.7 (m, Cbz-Ph), 95.5 (s, C_5), 89.9 (s, $C_{1'}$), 79.7 (s, $C_{4'}$), 77.4 (s, $C_{2'}$), 75.6 (s, $C_{3'}$), 70.9 (s, Cbz-(O)CH_2), 70.8 (s, Cbz-(O)CH_2), 68.3 (s, Cbz-(N)CH_2); Rt = 11.69 min (method 2, UV 254 nm, $[M+H]^+$ = 674); HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{34}H_{31}N_3O_{12}$ 674.1986; found 674.1979.

Benzyl (1-((2R,4S,5R)-4-(((benzyloxy)carbonyl)oxy)-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-4-yl)carbamate (**2**) - Column chromatography on silica gel afford **2** as a white solid, overall η = 59% (3 steps); ^{1}H NMR (400 MHz, $CDCl_3$/MeOD) δ 8.31 (d, 1H, H_6, *J* = 7.5 Hz, 1H), 7.40 – 7.25 (m, 11H, Cbz-Ph and H_5), 6.22 (dd, 1H, $H_{1'}$, *J* = 7.9, 5.8 Hz), 5.24 – 5.20 (m, 2H, $H_{2'}$), 4.26 – 4.19 (m, 1H, $H_{3'}$), 5.18 (s, 2H, Cbz-(N)-CH_2), 5.14 (s, 2H, Cbz-(O)CH_2), 3.81 (m, 1H, $H_{4'}$), 2.67 (m, 1H, $H_{5'}$), 2.25 (m, 1H, $H_{5'}$); ^{13}C{^{1}H} NMR (101 MHz, $CDCl_3$) δ 144.9 (s, C_6), 128.8 (m, Cbz-Ph), 96.3 (s, C_5), 87.8 (s, $C_{1'}$), 86.2 (s, $C_{4'}$), 79.0 (s, $C_{2'}$), 78.0 (s, $C_{3'}$), 70.4 (s, Cbz-(O)CH_2), 68.1 (s, Cbz-(N)CH_2), 62.0 (s, $C_{5'}$); Rt = 10.70 min (method 2, UV 254 nm, $[M+H]^+$ = 496); HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{25}H_{25}N_3O_3$ 496.1720; found 496.1708.

Table S1. Oxidation of compound **1a.** Detected masses and AUC of products by LCMS chromatogram under different conditions. **A**) 2 eq. of previously opened DMP (>3 months); **B**) 2 eq. of new DMP or 4 eq. of previously opened DMP; **C**) 4 eq. of new DMP



1a **3a**: X = H **5a**: X = OH **6a:** X = OMe **4a**

Figure	Compound	Retention time (min.)	$[M+H]^+$	% AUC[1]
A	**1a**	20.53	646	20
	4a	20.77, 21.10	676	80
B	**5a**	19.65	660	31
	1a	20.48	646	3
	4a	20.78, 21.09	676	66
C	**5a**	19.84	660	79
	4a	20.78, 20.99	676	15

[1]Conversion was calculated based on the area under curve (AUC) of compounds from the HPLC chromatogram below and method 1 (**Figure S19**).

Figure S1. LCMS chromatogram (diode array detection, method 1, UV 280 nm) of products of oxidation of compound **1a** under different conditions

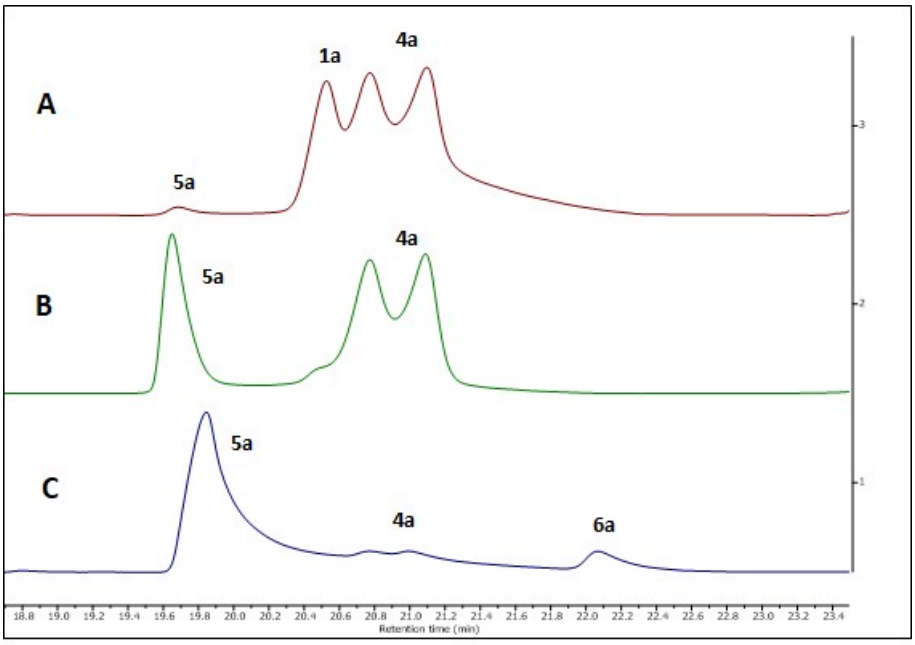


Table S2. Oxidation of compound **1a** with 2 eq. of DMP. Detected masses and AUC of products by LCMS chromatogram before extraction. LCMS sample prepared in **A**) MeOH; **B**) MeCN

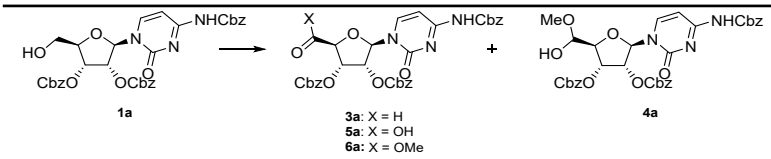


Figure	Compound	Retention time (min.)	$[M+H]^+$	% AUC[1]
A	**5a**	20.12	660	18
	4a	21.11, 21.37	676	67
	6a	22.44	674	15
B	**5a**	20.12	660	36
	4a	21.11, 21.36	676	64

[1]Conversion was calculated based on the area under curve (AUC) of compounds from the HPLC chromatogram below and method 1 (**Figure S19**).

Figure S2. LCMS chromatogram (diode array detection, method 1, UV 280 nm) of products of oxidation of compound **1a** before extraction. LCMS sample prepared in **A**) MeOH; **B**) MeCN

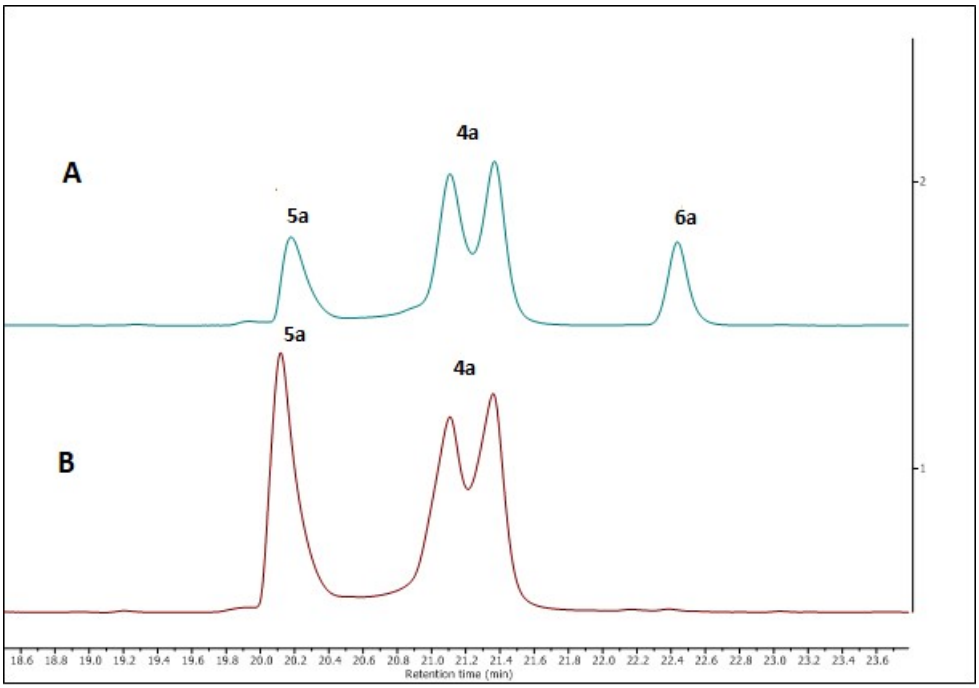
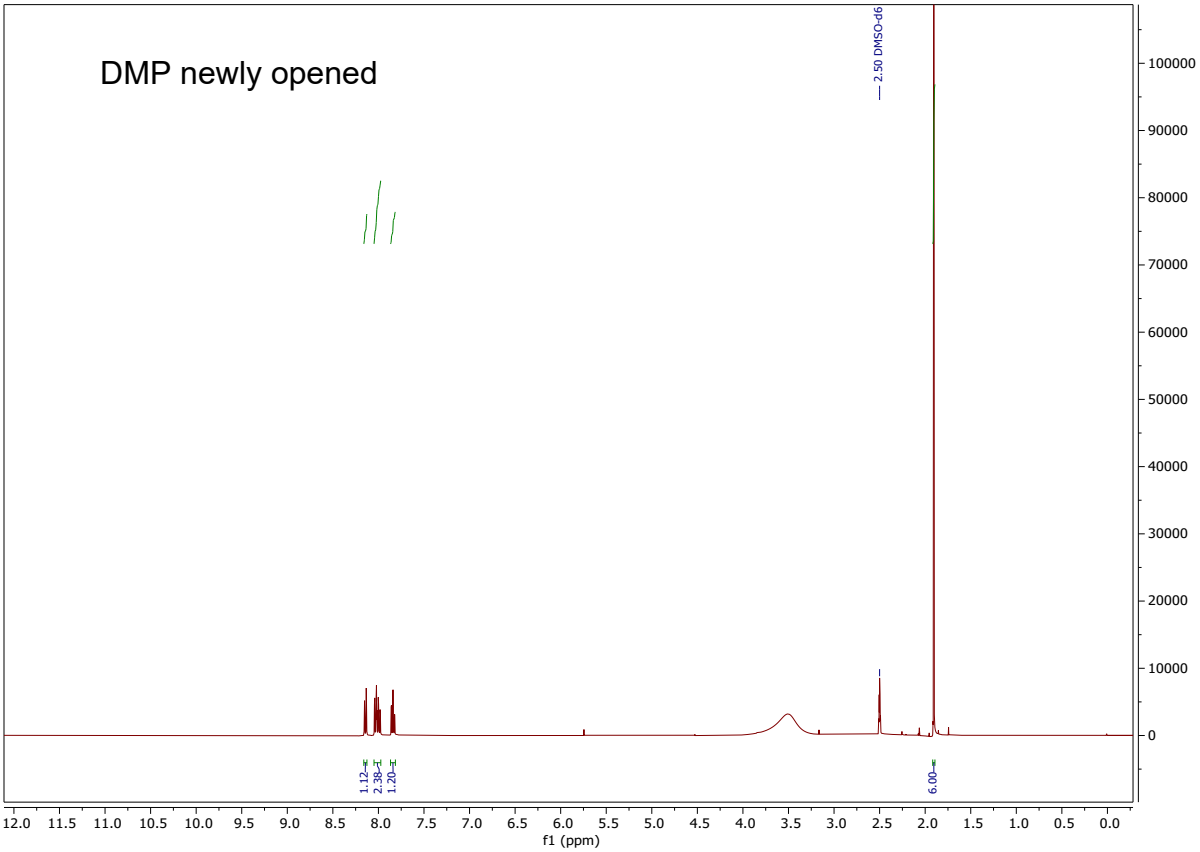


Figure S3. ^{1}H NMR (400 MHz, DMSO-*d*6) spectra of DMP from different batches



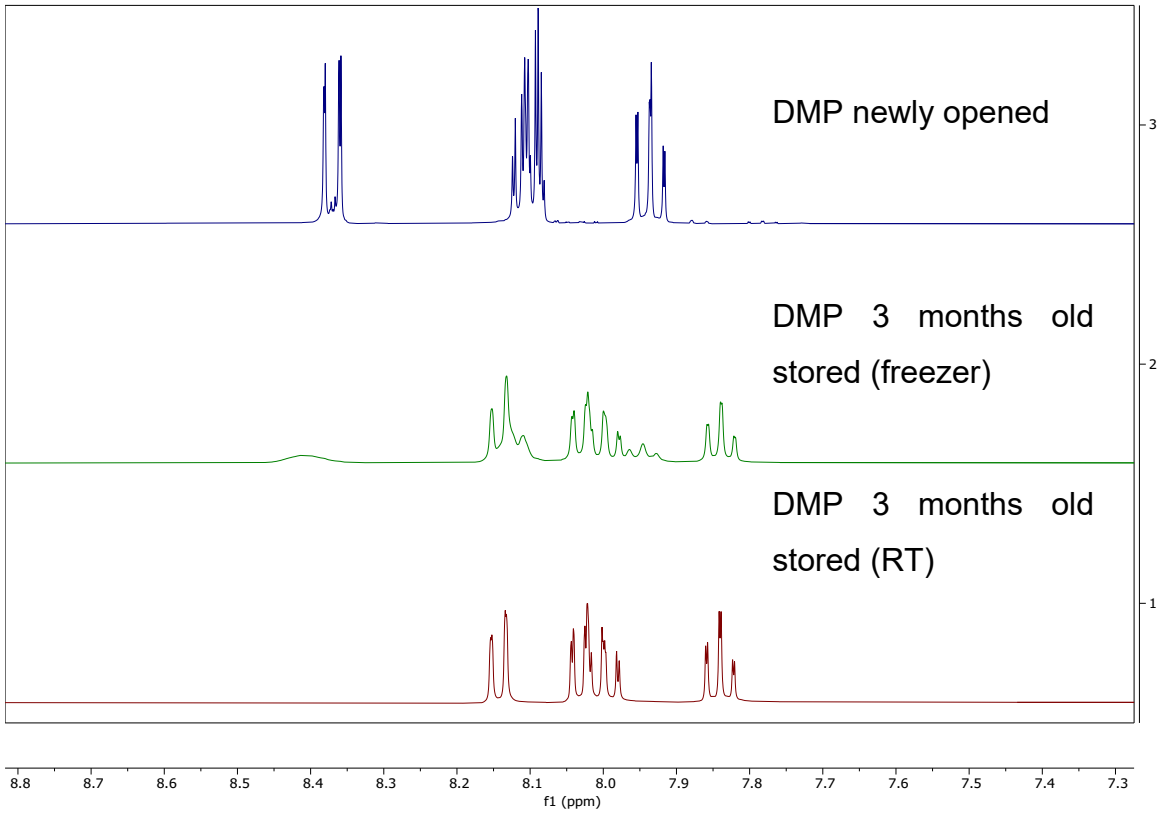


Figure S4. Oxidation of compound **1a** with DMP under different conditions.

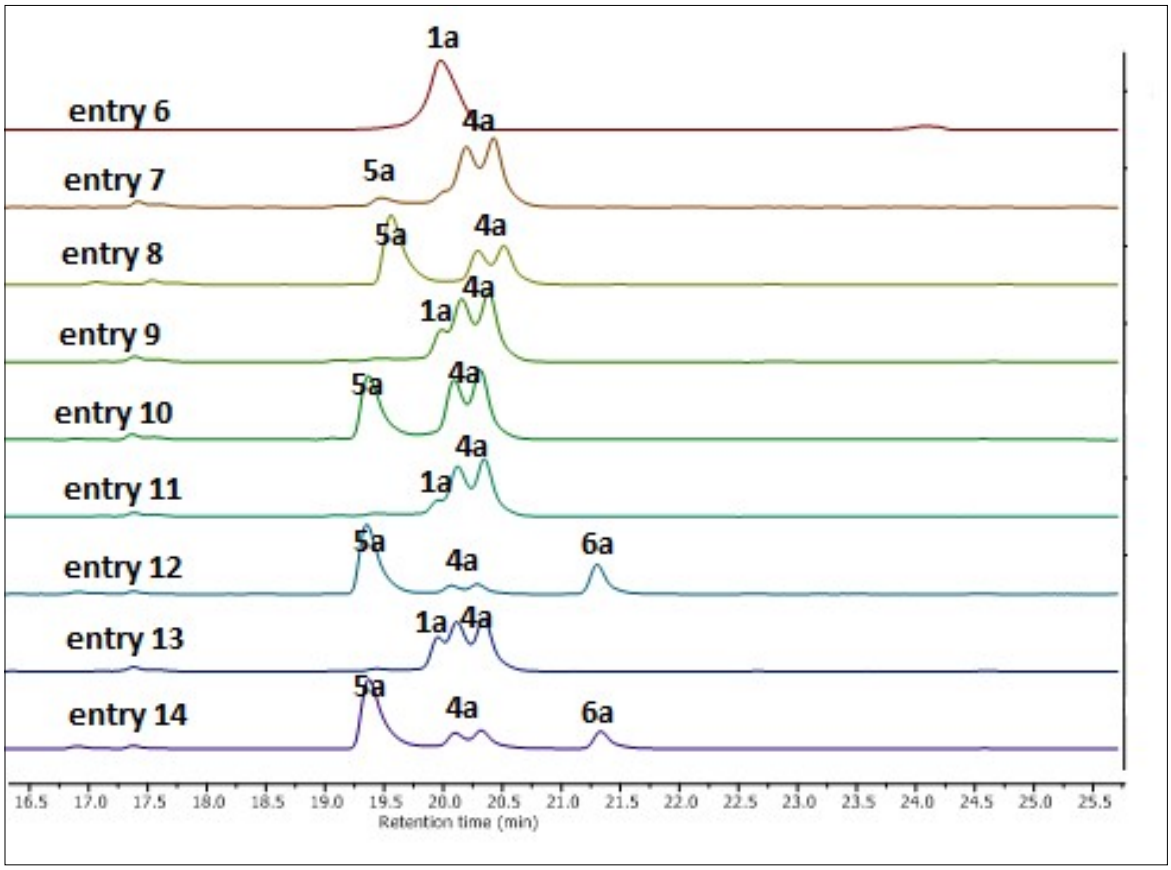
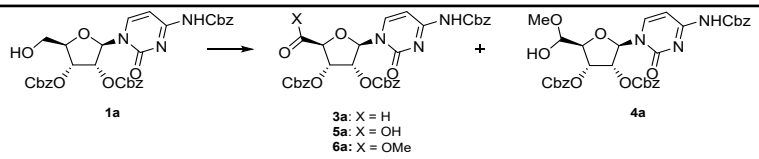


Table S3. Oxidation of compound **1a** with 2 eq. DMP. Detected masses and AUC of products by LCMS chromatogram after extraction. LCMS sample prepared in **A**) MeOH; **B**) MeCN



1a → **3a**: X = H; **5a**: X = OH; **6a**: X = OMe + **4a**

Figure	Compound	Retention time (min)	$[M+H]^+$	% AUC[1]
A	**5a**	20.01	660	37
	4a	20.98, 21.24	676	62
B	**5a**	20.14	660	39
	4a	21.00, 21.24	676	61

[1]Conversion was calculated based on the area under curve (AUC) of compounds from the HPLC chromatogram below and method 1 (**Figure S19**).

Figure S5. LCMS chromatogram (diode array detection, method 1, UV 280 nm) of products of oxidation of compound **1a** after extraction. LCMS sample prepared in **A**) MeOH; **B**) MeCN

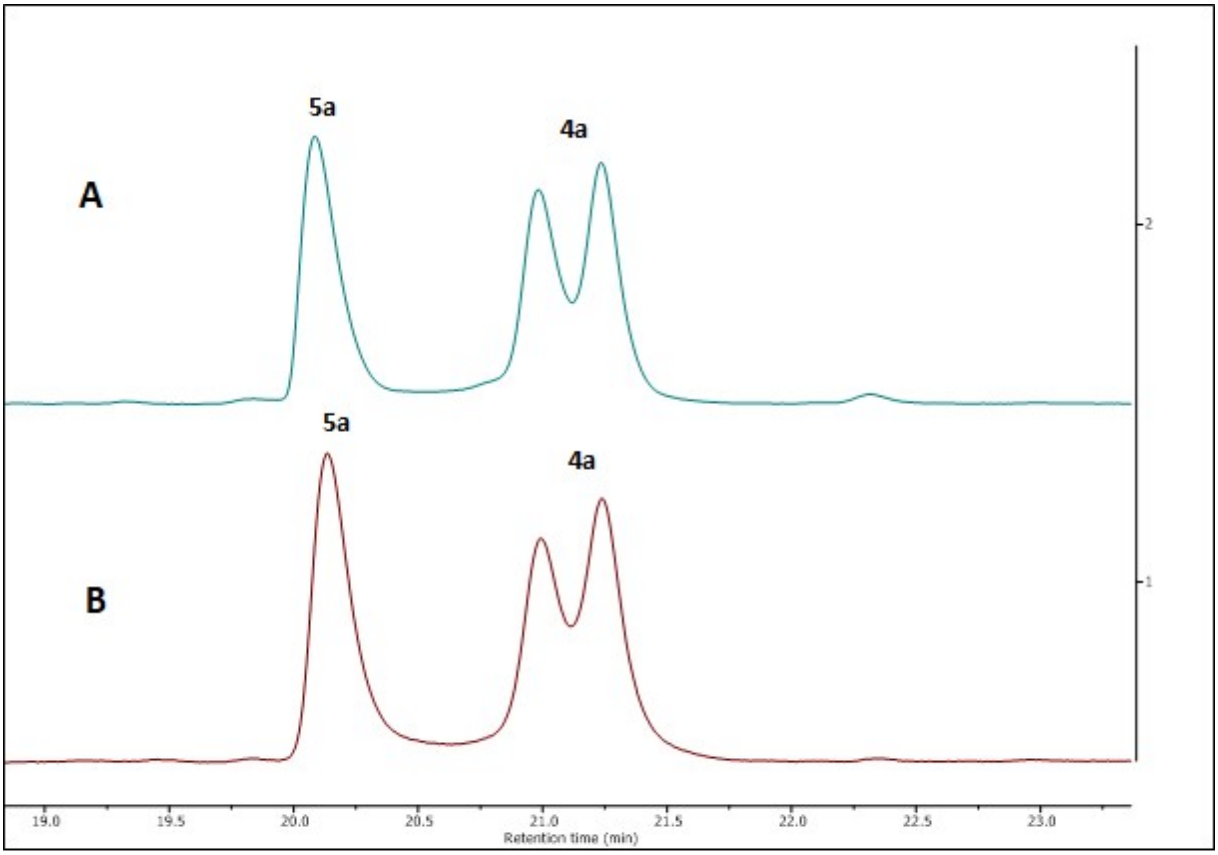


Figure S6. LCMS chromatogram of oxidation of compound **1a** with 2 eq. DMP using mobile phase A: 90%water, 10% acetonitrile, 0.1% formic acid; mobile phase B: 10% water, 90% acetonitrile, 0.1% formic acid (diode array detection, method 2, Figure S19, UV 280 nm).

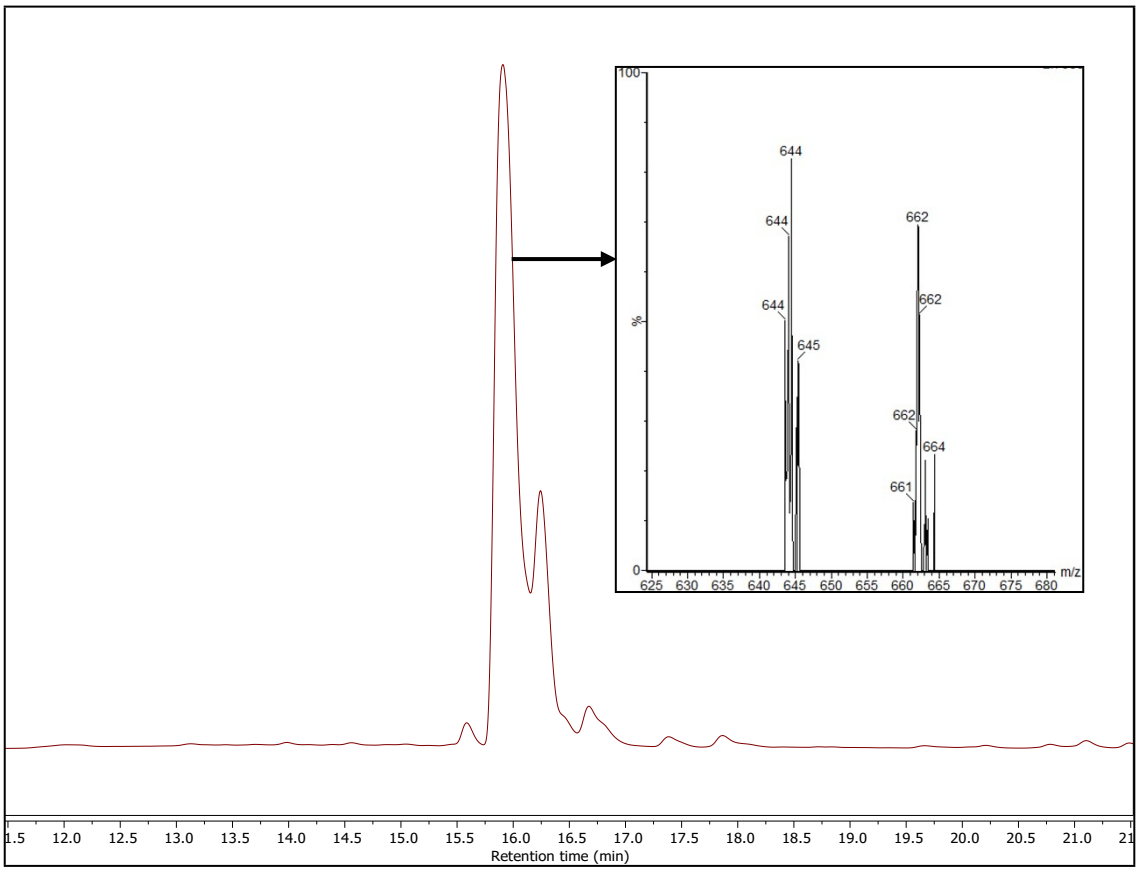


Figure S7. Comparison between initial coordinates (low-energy 3D structures) and the coordinates after DFT calculation of **1a** and hydrate intermediate of **3a** showing conformational changes. Initial coordinates of **A**) **1a** (green); **B**) hydrate intermediate of **3a** (cyan); Coordinates after DFT calculation of **C**) **1a** (magenta); **D**) hydrate intermediate of **3a** (purple). Superposed initial coordinates and the coordinates after DFT calculation of **E**) **1a**; **F**) hydrate intermediate of **3a.**

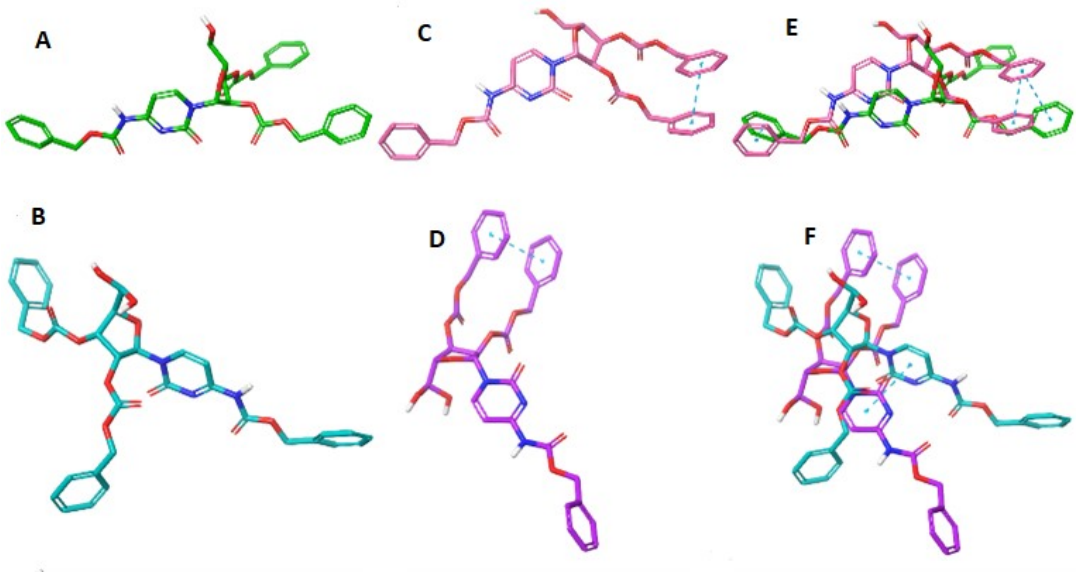


Figure S8. RMSD analysis from the 200 ns MD simulations of **1a** (magenta) and the hydrate intermediate of **3a** (purple).

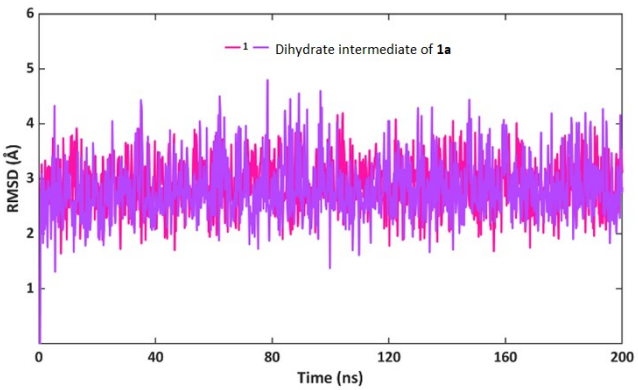


Figure S9. MD simulation analysis of **1a** (magenta). **A)** MD pose at the 85th ns showing π-π stacking between the cytidine ring and one of the Cbz-phenyl rings; **B)** MD pose at the 145th ns showing π-π stacking between the two phenyl rings of Cbz.

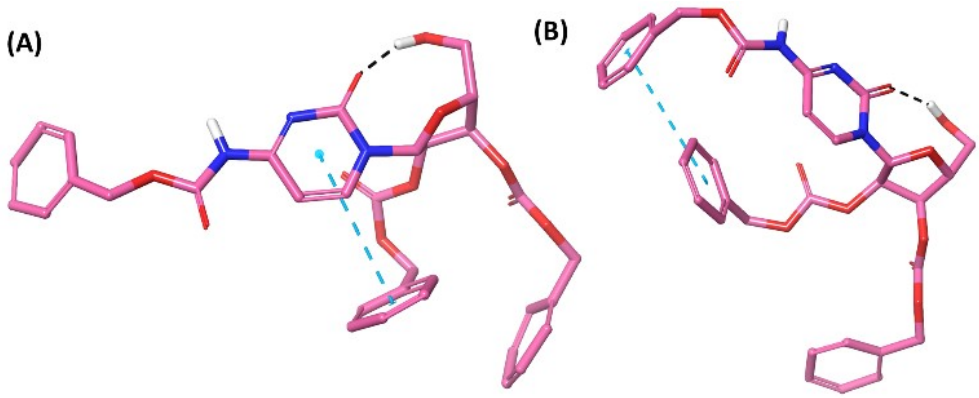


Figure S10. MD simulation analysis of the hydrate intermediate of **3a** (purple). MD poses at the **A**) 96th ns; **B**) 116th ns, showing π-π stacking between the two phenyl rings of Cbz.

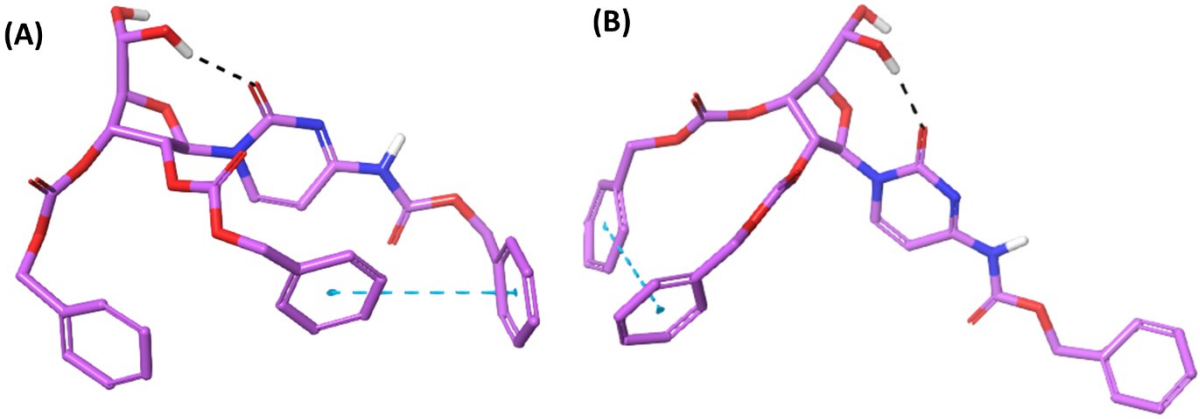
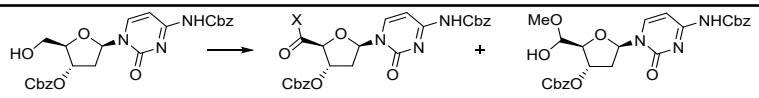


Table S4. Oxidation of compound **2.** Detected masses and AUC of products by LCMS chromatogram under different conditions. **A**) 2 eq. of new DMP, **B**) 4 eq. of new DMP



2 **7**: X = H **8**: X = OH **9**: X = OMe **10**

Figure	Compound	Retention time (min)	$[M+H]^+$	% AUC[1]
A	**8**	10.76	510	41
	10	10.78, 10.84	526	59
B	**8**	10.74	510	> 95

[1]Conversion was calculated based on the area under curve (AUC) of compounds from the HPLC chromatogram below and method 2 (**Figure S19**).

Figure S11. LCMS chromatogram (diode array detection, method 2, UV 280 nm) of products of oxidation of compound **2** under different conditions

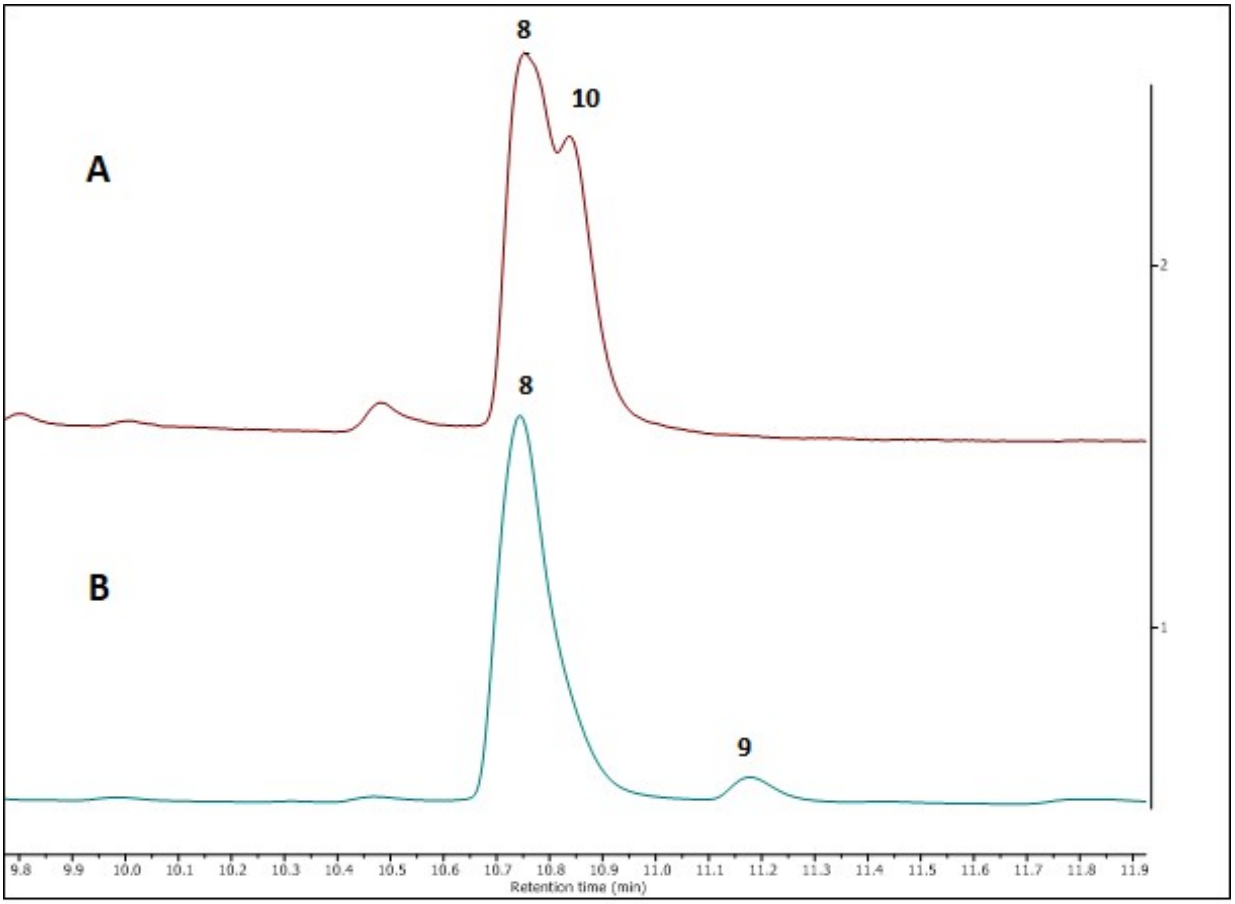
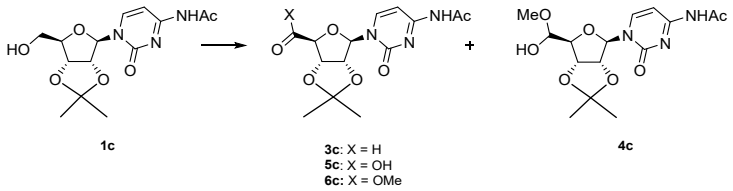


Table S5. Oxidation of compound **1c**. Detected masses and AUC of products by LCMS chromatogram under different conditions. **A**) 2 eq. of new DMP, **B**) 4 eq. of new DMP



1c **4c**

3c: X = H
5c: X = OH
6c: X = OMe

Figure	Compound	Retention time (min)	$[M+H]^+$	% AUC[1]
A	**4c**	9.18, 9.28	356	100
B	**5c**	8.96	340	42
	4c	9.16, 9.38	356	58

[1]Conversion was calculated based on the area under curve (AUC) of compounds from the HPLC chromatogram below and method 1 (**Figure S19**).

Figure S12. LCMS chromatogram (diode array detection, method 1, UV 280 nm) of products of oxidation of compounds **1c** under different conditions.

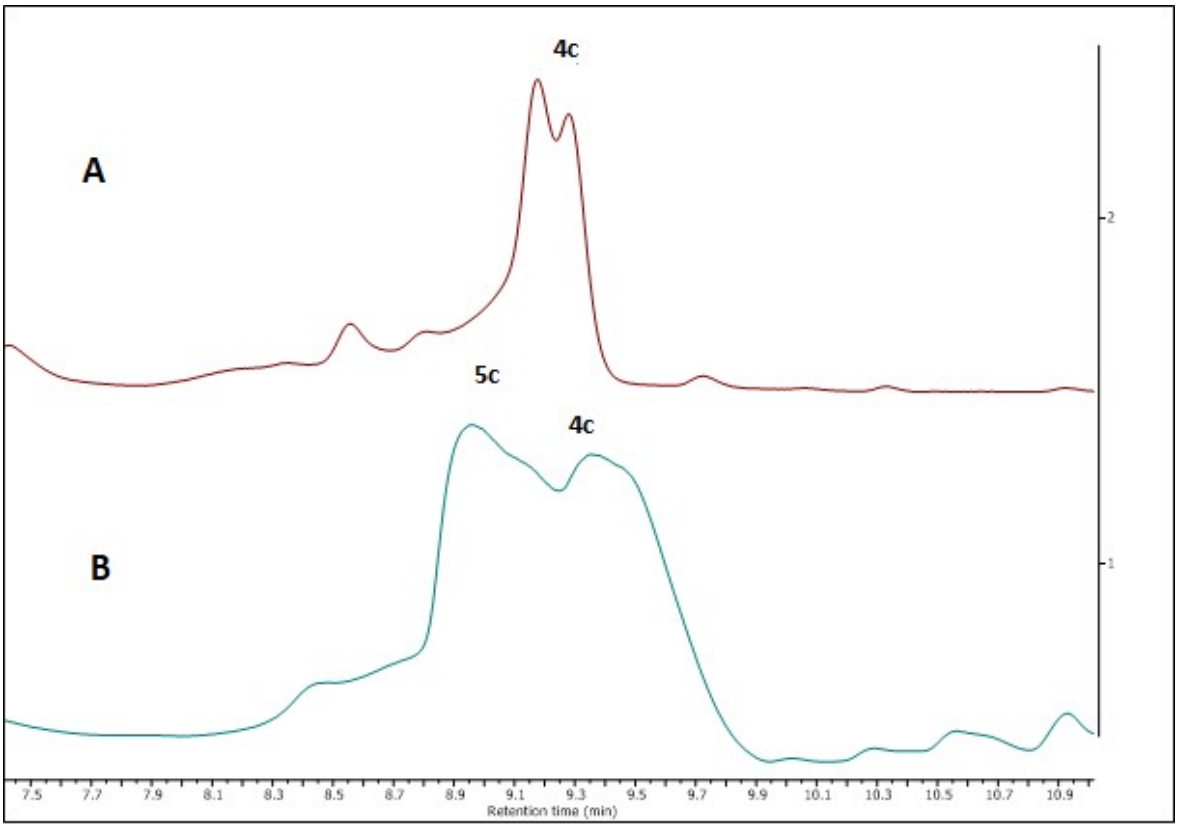


Figure S13A. ^{1}H NMR spectra (400 MHz, $CDCl_3$) of aldehyde **3c** (purified by column chromatogram; residue of org. phase added directly to the column chromatography on silica gel)

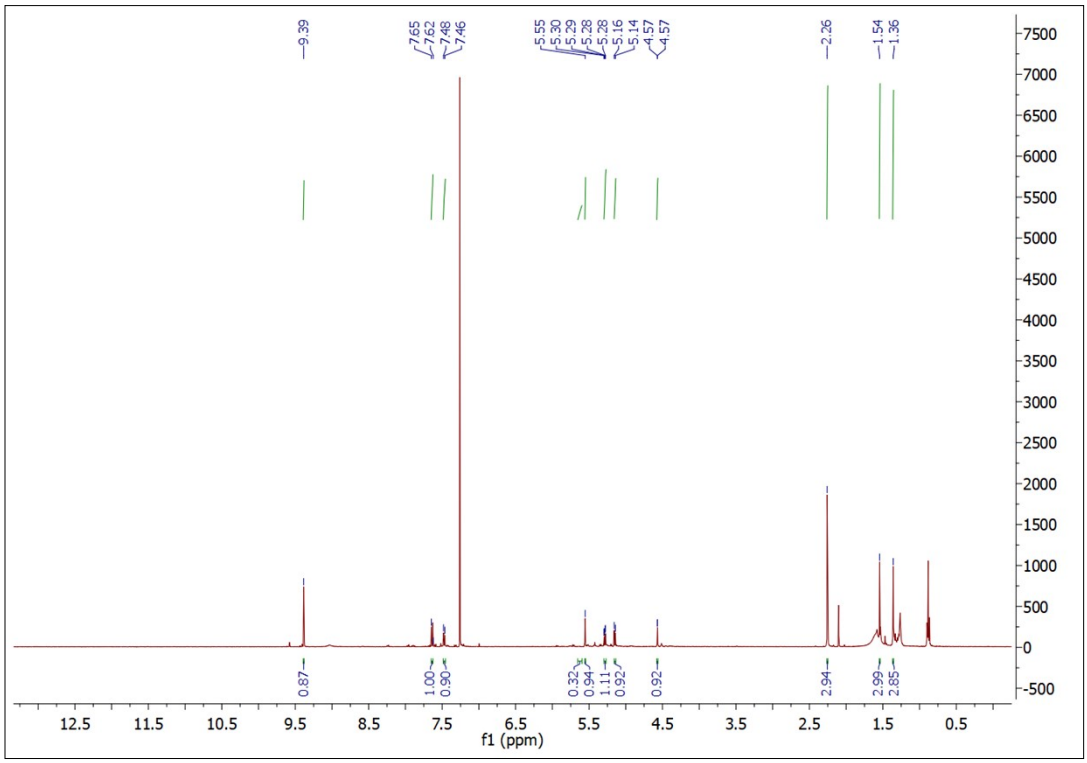


Figure S13B. ^{1}H NMR spectra (400 MHz, $CDCl_3$) of mixture aldehyde **3c** and methyl hemiacetal **4c** (purified by column chromatogram; residue of org. phase redissolved in dichloromethane and methanol with silica gel before addition to the column chromatographic on silica gel)

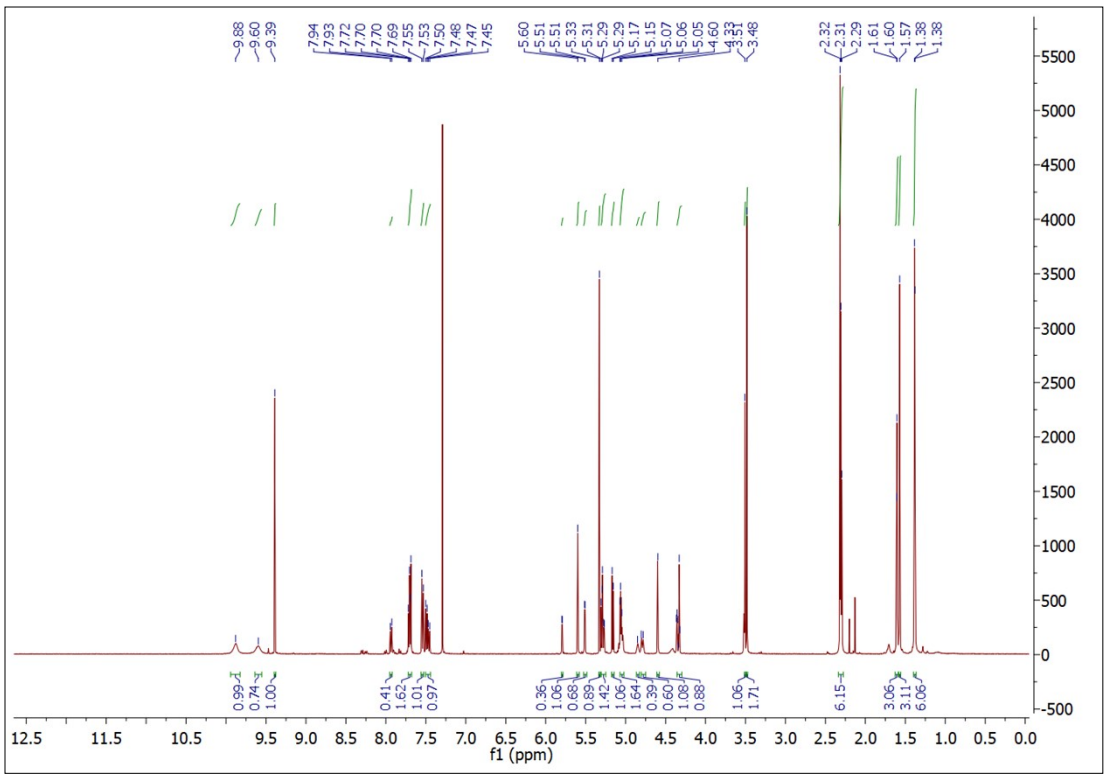


Figure S14. Comparison between **A**) initial coordinates (low-energy 3D structures) of cytidine analogue **1c** and B) its coordinate after DFT calculation; C) superposed initial coordinates and the coordinates after DFT calculation of **1c.**

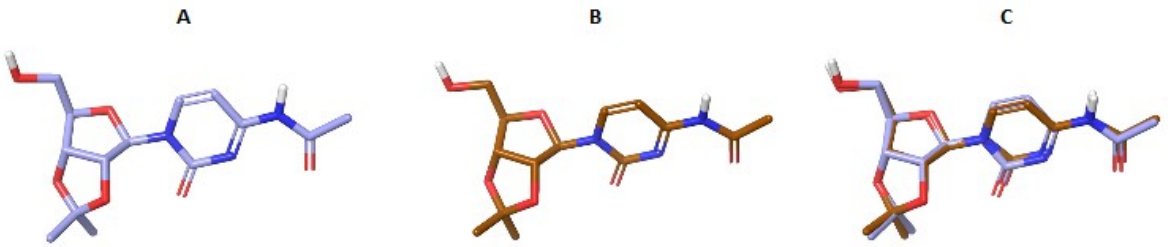


Table S6. Oxidation of compound **1b**. Detected masses and AUC of products by LCMS chromatogram under different conditions **A**) 2 eq. of new DMP, **B**) 4 eq. of new DMP

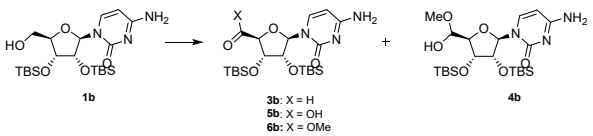


Figure	Compound	Retention time (min)	$[M+H]^+$	% AUC[1]
A	**3b**	11.96	470	50
	5b	12.14	486	7
	4b	12.23	502	43
B	**5b**	12.05	486	96
	4b	12.30	502	4

[1]Conversion was calculated based on the area under curve (AUC) of compounds from the HPLC chromatogram below and method 2 (**Figure S19**).

Figure S15. LCMS chromatogram (diode array detection, method 2, UV 280 nm) of products of oxidation of compounds **1b** under different conditions.

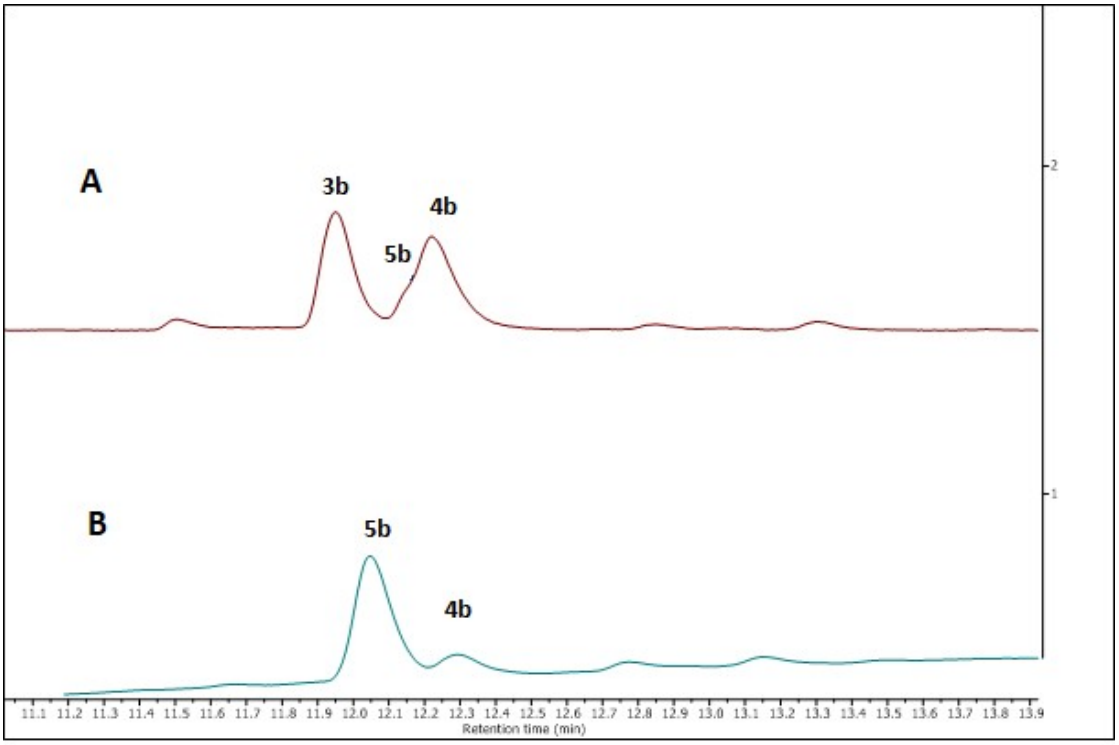


Figure S16. Representative structures from the MD simulations of **A**) cytidine analogue **1a**; **B**) hydrate intermediate of **3a**; **C**) cytidine analogue **1c**

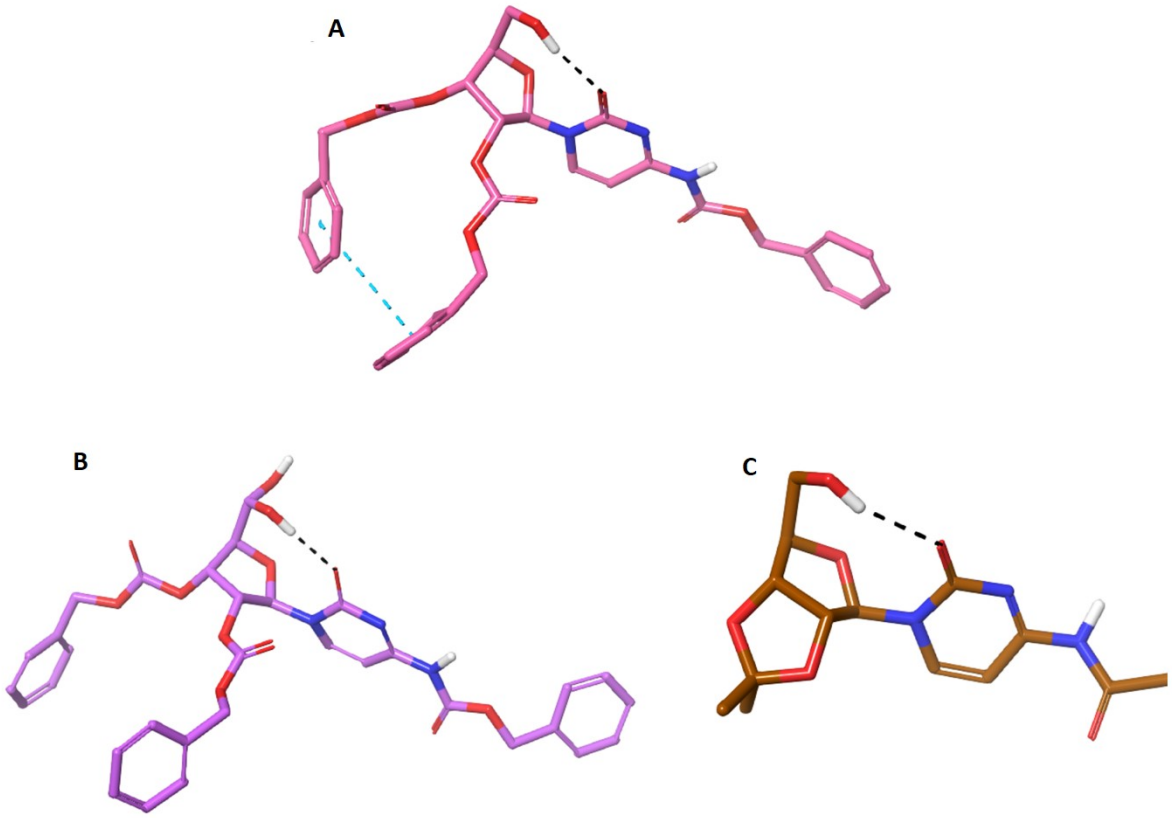


Figure S17. ^{1}H NMR spectra (400 MHz, $CDCl_3$) of **A**) methyl riboside **11**; **B**) methyl riboside **11** with 2 eq. DMP; **C**) methyl riboside **11** with 4 eq. DMP.

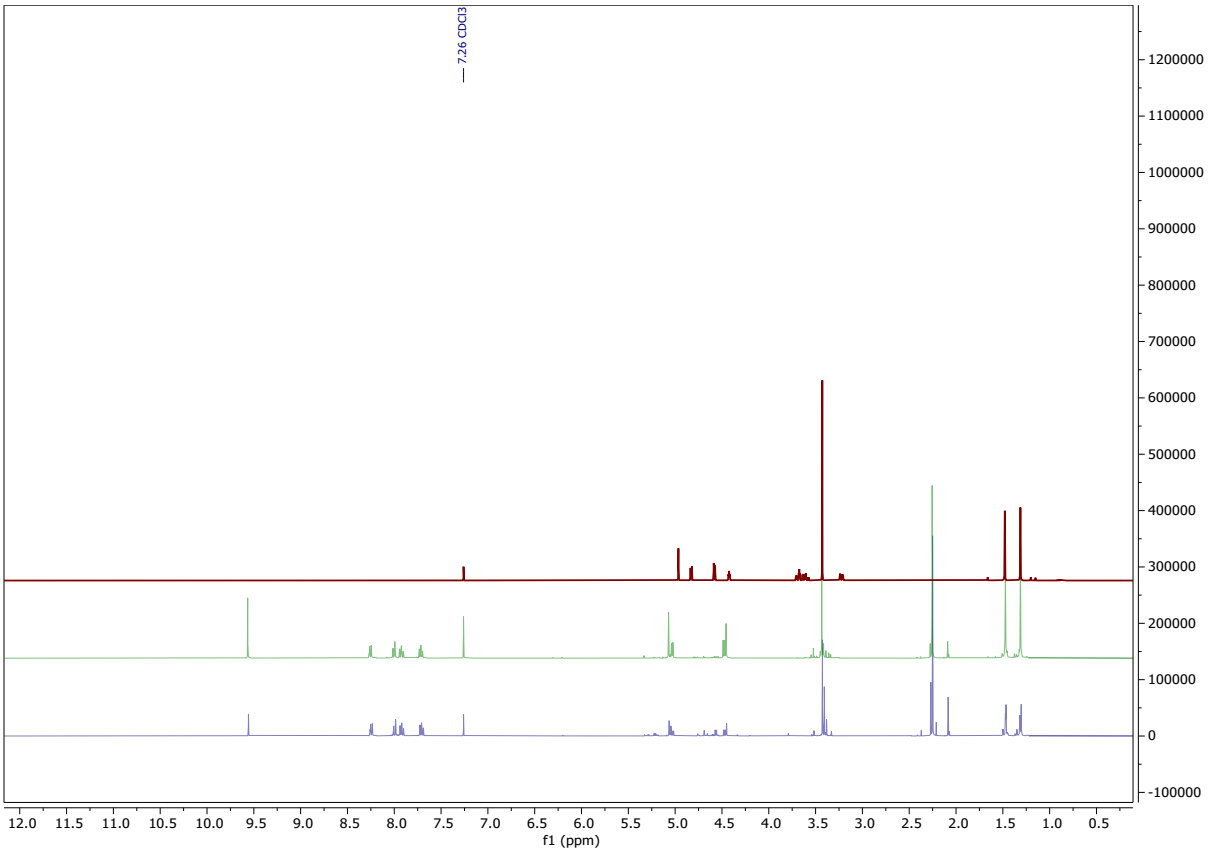


Table S7. Oxidation of methyl riboside **12**. Detected masses and AUC of products by LCMS chromatogram under different conditions **A**) 2 eq. of new DMP, **B**) 4 eq. of new DMP

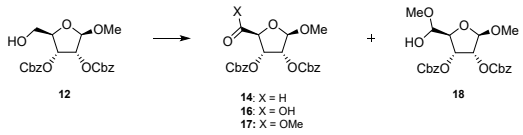


Figure	Compound	Retention time (min)	$[M+H]^+$	% AUC[1]
A	**18**	18.86, 19.28	463	43
	16	19.13	447	57
B	**18**	18.83	463	9
	16	19.04	447	81

[1]Conversion was calculated based on the area under curve (AUC) of compounds from the HPLC chromatogram below and method 1 (**Figure S19**)

Figure S18. LCMS chromatogram (diode array detection, method 1, UV 280 nm) of products of oxidation of methyl riboside **12** under different conditions.

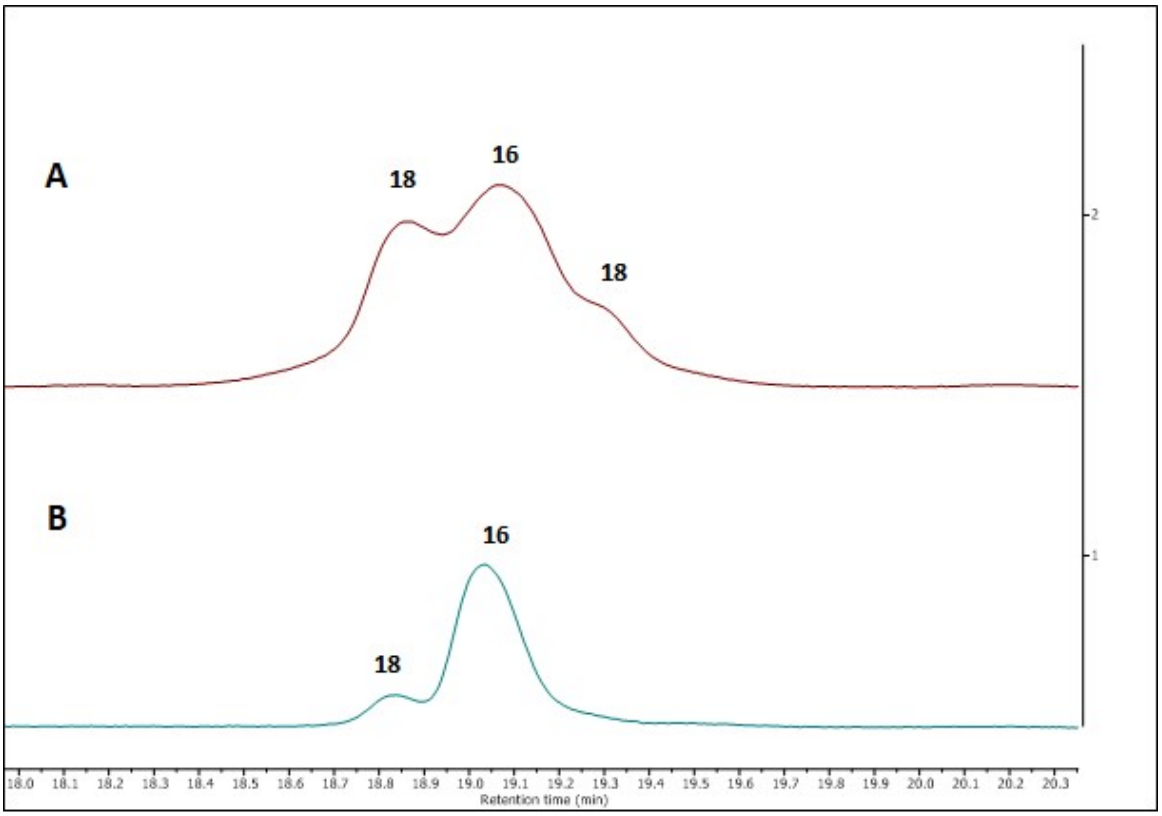


Figure S19. LCMS gradient

mobile phase A: 90% water, 10% MeOH, 0.1% formic acid;

mobile phase B: 90% MeOH, 10% water, 0.1% Formic acid;

flow: 0.25 ml/min;

Method 1:

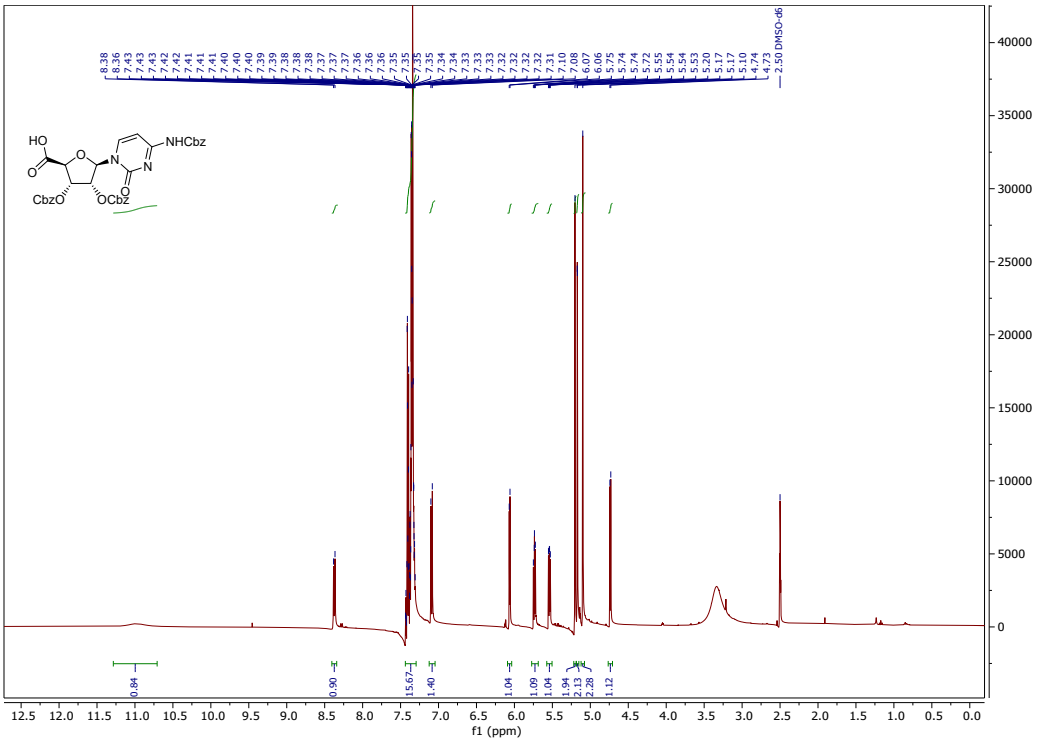
Time (min)	Mobile Phase A (%)	Mobile Phase B (%)
0	95	5
5	95	5
10	30	70
20	10	90
25	0	100
30	0	100
32	95	5
35	95	5

Method 2:

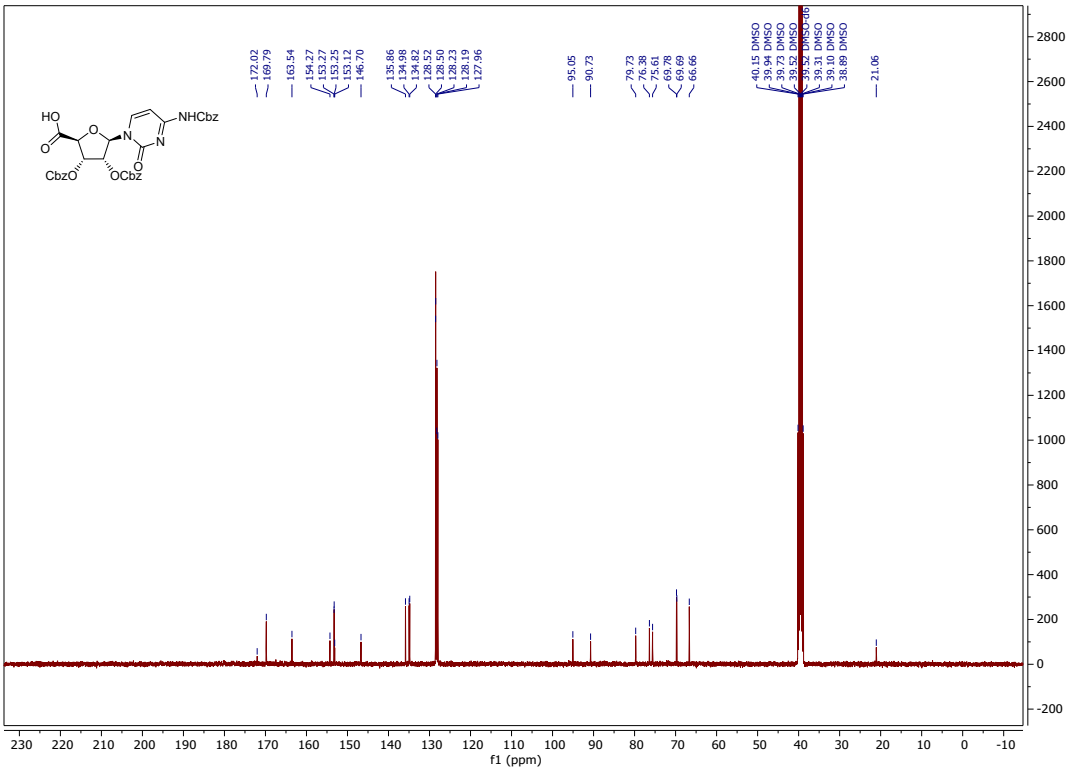
Time (min)	Mobile Phase A (%)	Mobile Phase B (%)
0	95	5
3	95	5
5	0	100
30	0	100
32	95	5
35	95	5

Figure S20 Characterisation of compound **5a**

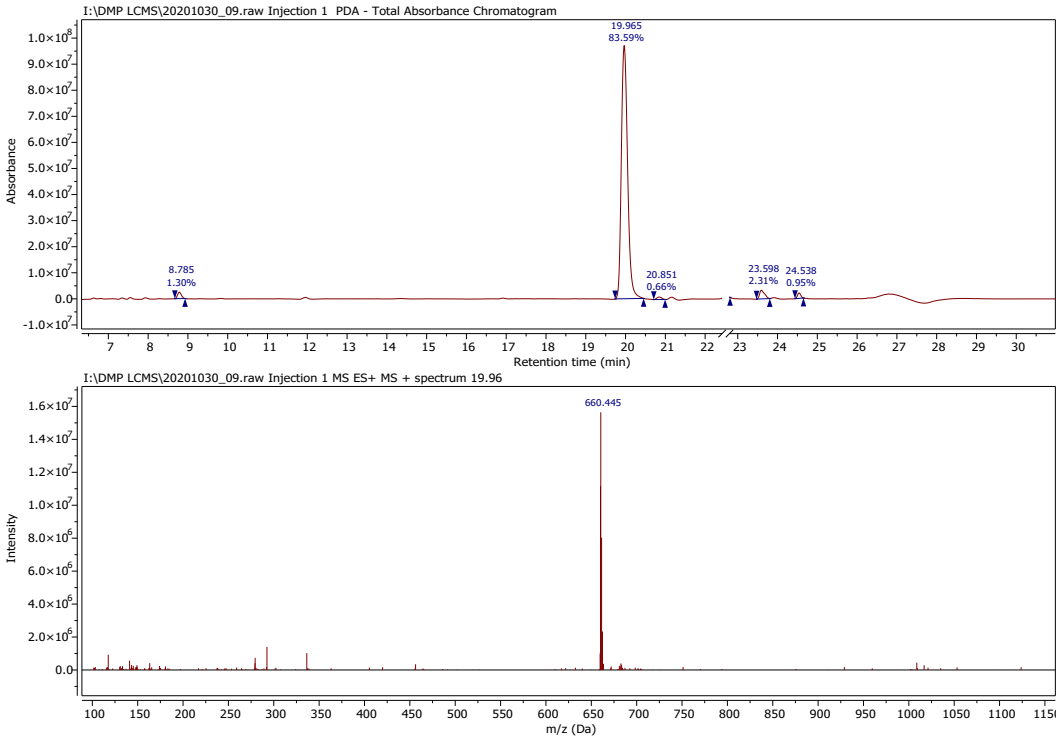
^{1}H NMR spectra of compound **5a** (400 MHz, DMSO-*d*6)



^{13}C{^{1}H} NMR of compound **5a** (100 MHz, DMSO-*d*6)



LC chromatogram and MS spectrum of compound **5a**



HRMS spectrum of compound **5a**

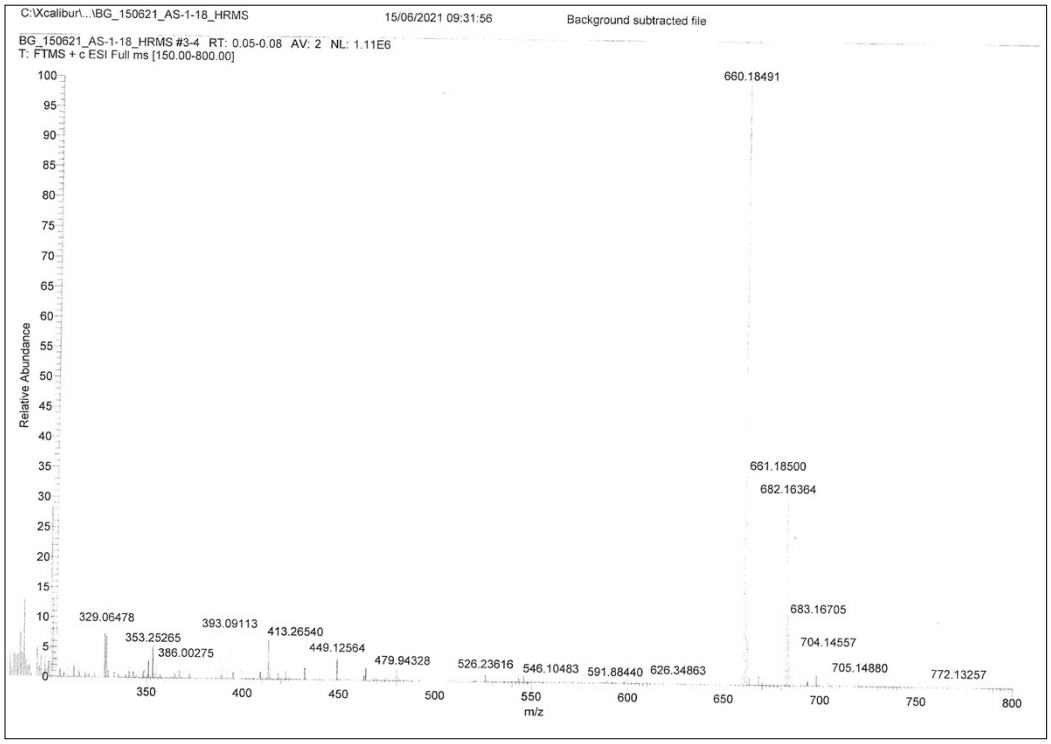
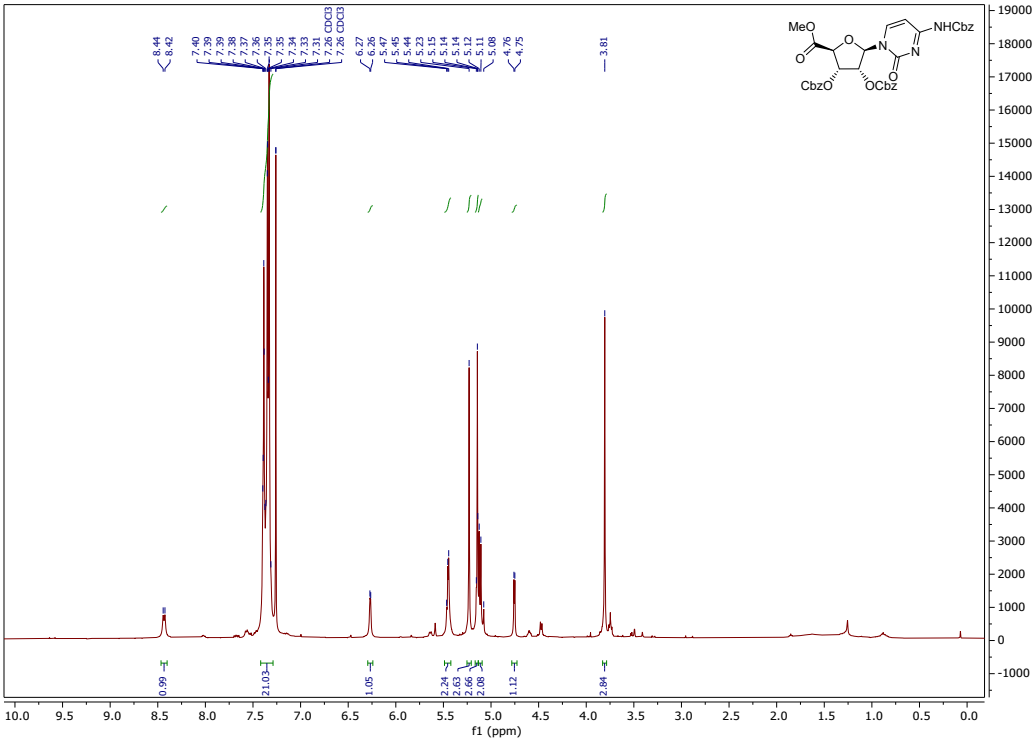
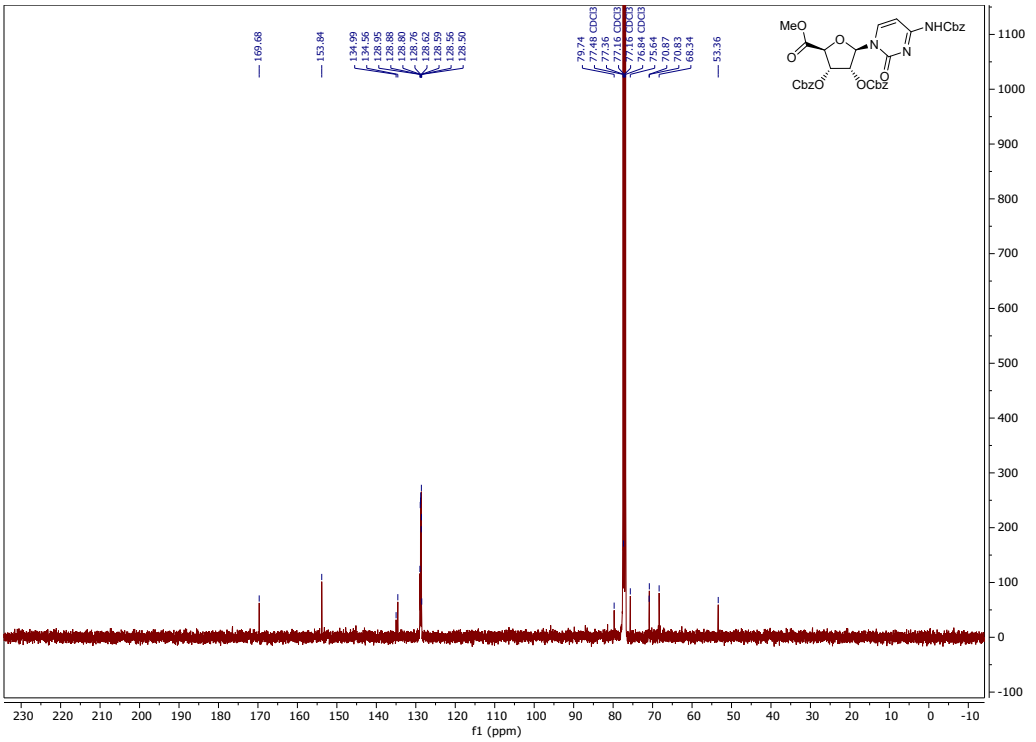


Figure S21 Characterisation of compound **6a**

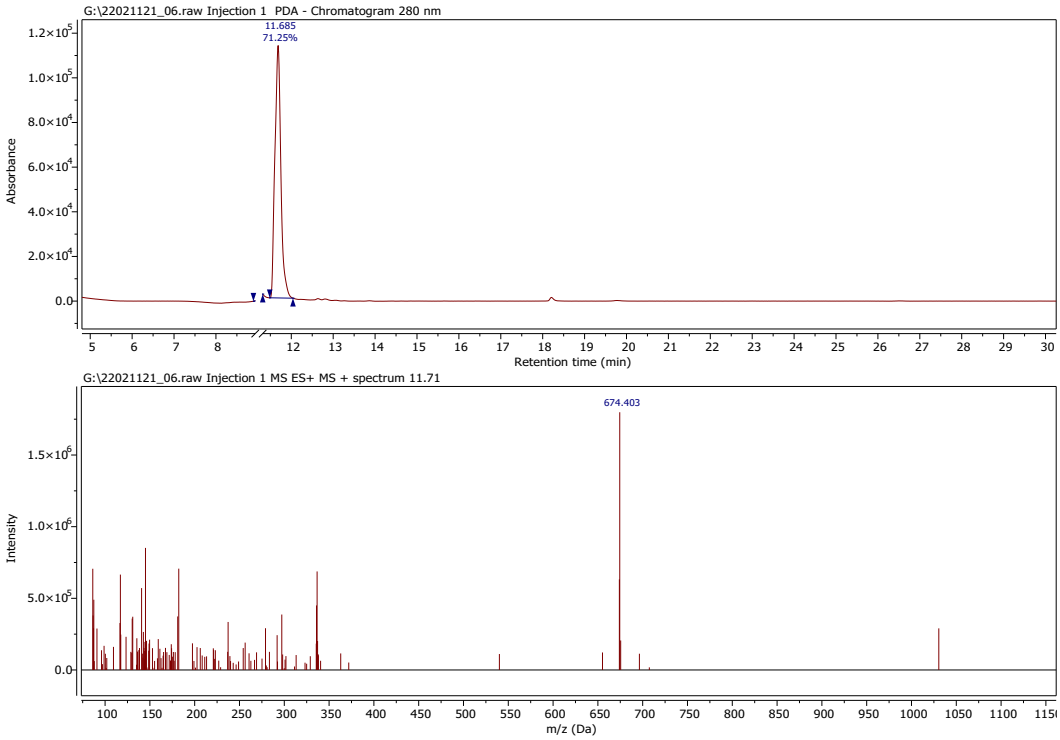
^{1}H NMR spectra of compound **6a** (400 MHz, $CDCl_3$)



^{13}C{^{1}H} NMR of compound **6a** (100 MHz, $CDCl_3$)



LC chromatogram and MS spectrum of compound **6a**



HRMS spectrum of compound **6a**

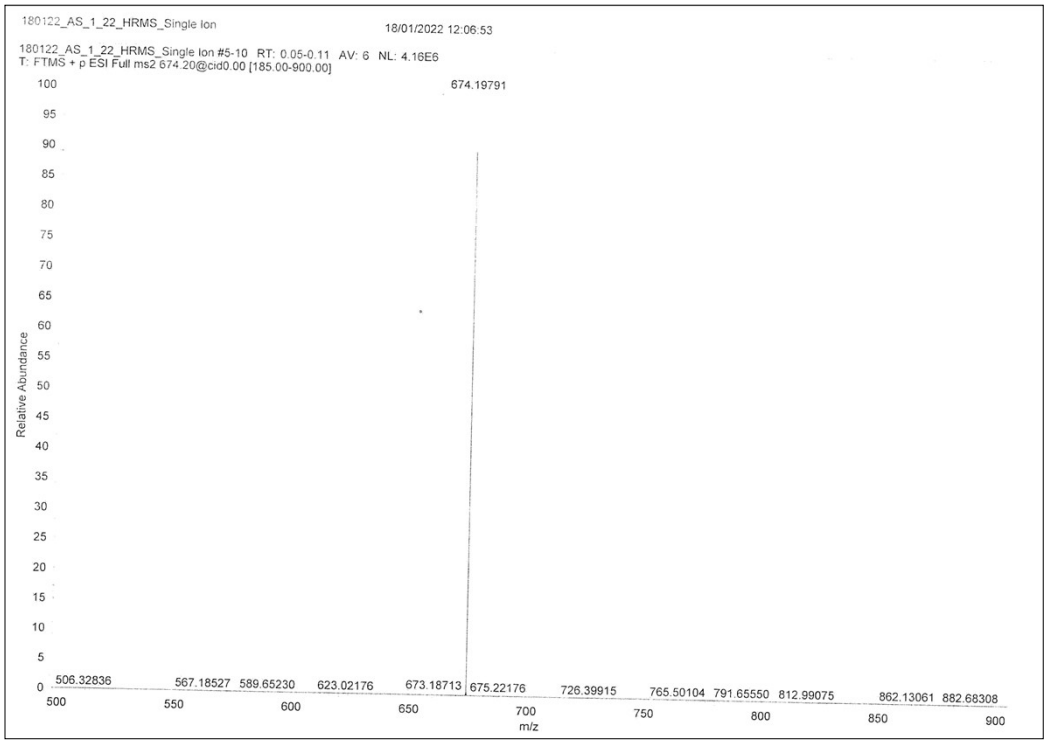
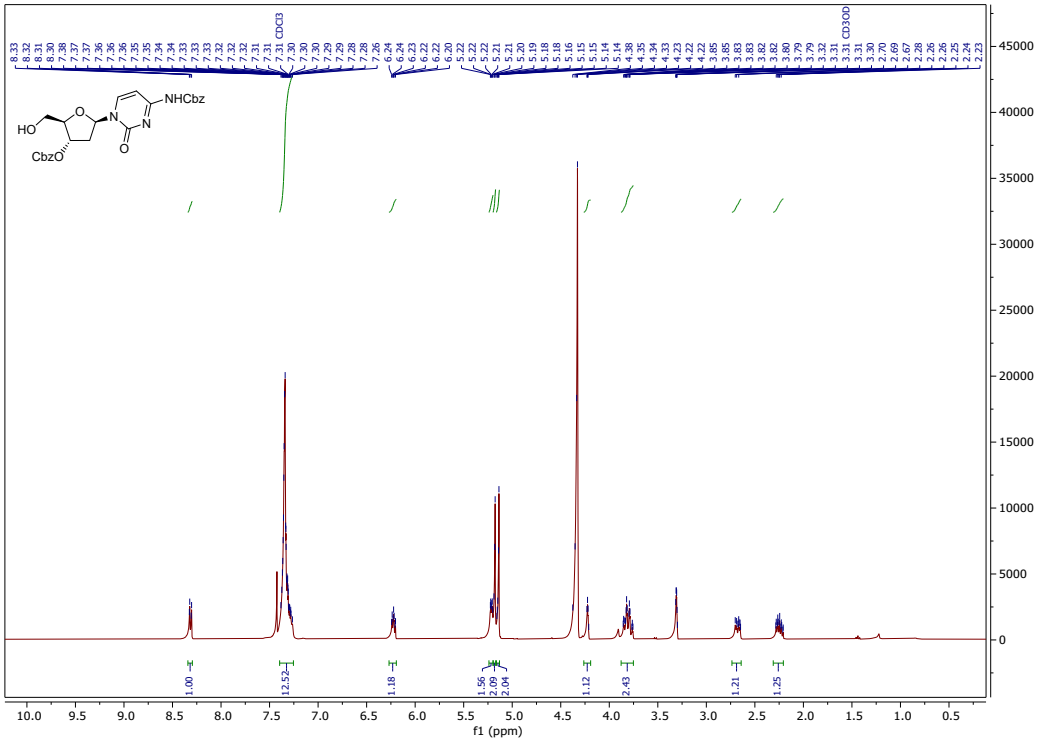
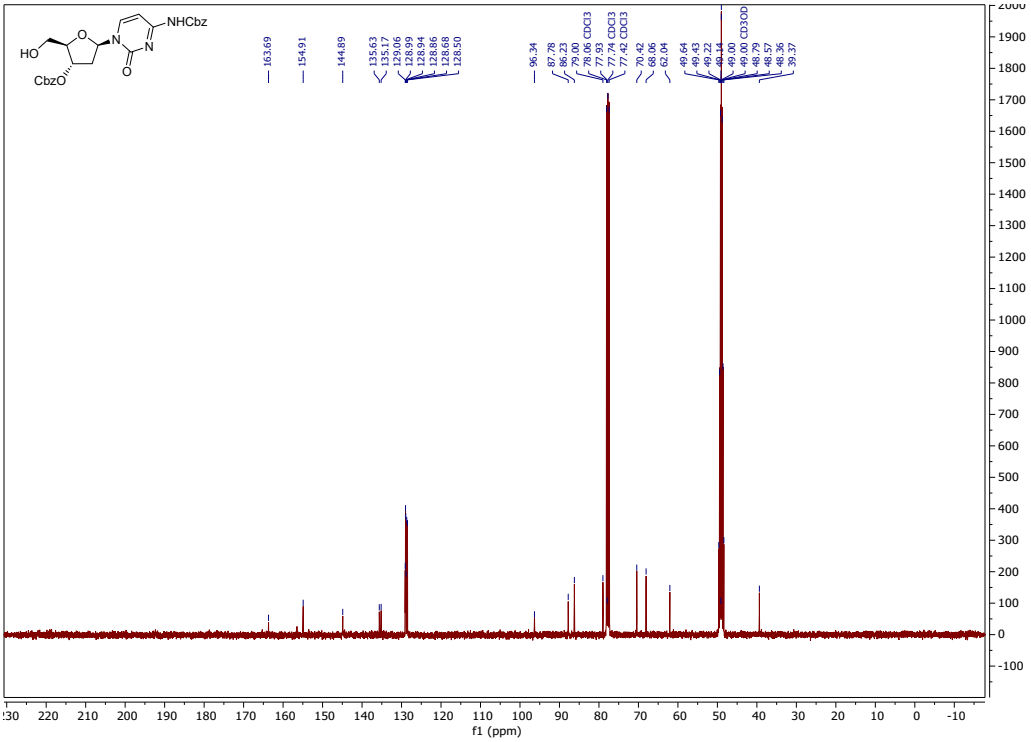


Figure S22 Characterisation of compound **2**

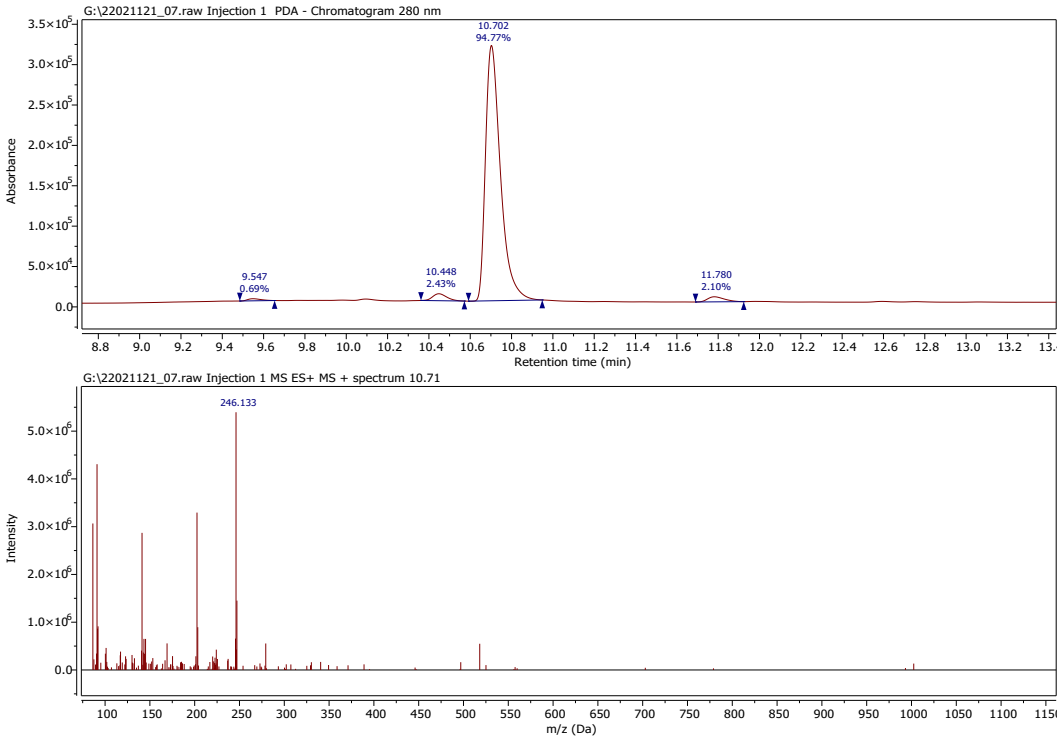
^{1}H NMR spectra of compound **2** (400 MHz, $CDCl_3$)



^{13}C{^{1}H} NMR of compound **2** (100 MHz, $CDCl_3$)



LC chromatogram and MS spectrum of compound **2**



HRMS spectrum of compound **2**