Supporting Information

5-Hydroxymethylcytosine-selective oxidation with peroxotungstate

Akimitsu Okamoto,[1,2,*] Kaori Sugizaki,[1] Akiko Nakamura,[1] Hiroyuki Yanagisawa,[1] Shuji Ikeda[1]

[1]Advanced Science Institute, RIKEN, Wako, Saitama 351-0198, Japan, [2]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

Experimental Section

Preparation of ^{hm}C-containing DNA. Artificial DNA was synthesized using the conventional phosphoramidite method employing an NTS H-6 DNA/RNA synthesizer. The phosphoramidite of ^{hm}C was prepared according to a previously reported protocol (K. Sugizaki, S. Ikeda, H. Yanagisawa and A. Okamoto, *Org. Biomol. Chem.*, 2011, **9**, 4176). The synthesized DNA was purified using reversed-phase HPLC on a 5-ODS-H column (10 × 150 mm, elution with a solvent mixture of 0.1 M triethylammonium acetate (pH = 7.0), linear gradient over 20 min from 5% to 20% acetonitrile at a flow rate of 3.0 mL/min).

Metal oxidation and hot piperidine treatment. The dinuclear peroxotungstate (**1**) and dinuclear peroxomolybdate were prepared according to the reported synthetic procedures (K. Kamata, K. Yamaguchi, N. Mizuno, *Chem.–Eur. J.*, 2004, **10**, 4728; N. J. Campbell, A. C. Dengel, C. J. Edwards and W. P. Griffith, *J. Chem. Soc. Dalton Trans.*, 1989, 1203). The fluorescein-labeled DNA (5 μM) to be examined was incubated in a solution of 5 mM of the metal complex (plus 50 mM hydrogen peroxide as an option) and 100 mM sodium chloride in 50 mM sodium phosphate (pH = 7.0) at 50 °C for 5 h. The reaction solution was filtered to deionize it using Micro Bio-Spin Columns with Bio-Gel P-6 (Bio-Rad). After drying *in vacuo*, the precipitated DNA was redissolved in 50 μL of 10% piperidine (*v/v*), heated at 90 °C for 2 h, and then evaporated to dryness using a vacuum rotary evaporator.

Tungsten oxidation of $d^{hm}C$. The oxidant **1** (128 mg, 195 μmol) and 30% hydrogen peroxide (45 μL) were added to a solution of $d^{hm}C$ (100 mg, 389 μmol) in water (1 mL) at 50 °C for 24 h. After the mixture had cooled to room temperature, the precipitated inorganic salt was removed by centrifugation and decantation. After purification using reversed-phase HPLC (3% to 10% acetonitrile in 0.1 M triethylammonium acetate, pH = 7.0), the collected fractions were evaporated to dryness under reduced pressure. The residue was dissolved in methanol and diethyl ether was added to the solution to give a white precipitate. The precipitate was washed with diethyl ether and dried to obtain the product as a white powder (28 mg, 22%), which was a mixture of diastereomeric isomers with the same molecular weight. The product was analyzed using ESI mass spectrometry (Fig. S2 and S3), ^{1}H NMR spectroscopy (Fig. S4), and ^{13}C NMR spectroscopy (Fig. S5). HRMS (ESI) $[M + Na]^+$, $C_{10}H_{16}N_2O_8Na$, calcd. 315.0804, found 315.0814; ^{1}H NMR (500 MHz, DMSO-d_6) of the major isomer, δ 10.39 (m, 1H, NH), 6.30 (d, 1H, 6-OH, J_{OH-6H} = 5.0 Hz), 6.06 (dd, 1H, 1′-H, $J_{1'H-2'H}$ = 6.0 Hz, $J_{1'H-2''H}$ = 8.5 Hz), 5.42 (s, 1H, 5′-OH), 5.11 (d, 1H, 3′-OH, $J_{OH-3'H}$ = 4.0 Hz), 4.82 (d, 1H, 6-H, J_{6H-OH} = 5.0 Hz), 4.78 (t, 1H, CH_2–<u>OH</u>, J_{CH2-OH} = 5.5 Hz), 4.12 (m, 1H, 4′-H), 3.60 (m, 1H, 3′-H), 3.41 (m, 4H, 5′-CH_2 and <u>CH_2</u>–OH), 2.18 (m, 2H, 2′-CH_2 (β)), and 1.82 (m, 2H, 2′-CH_2 (α)); the minor isomer, δ 10.37 (m, 1H, NH), 6.02 (t, 1H, 1′-H, $J_{1'H-2'H}$ = 6.5 Hz), 5.92 (d, 1H, 6-OH, J_{OH-6H} = 4.0

Hz), 5.35 (s, 1H, 5′-OH), 5.17 (d, 1H, 3′-OH, $J_{OH-3'H}$ = 4.5 Hz), 4.91 (d, 1H, 6-H, J_{6H-OH} = 4.0 Hz), 4.90 (t, 1H, CH_2–OH, J_{OH-CH2} = 6.5 Hz), 4.16 (m, 1H, 4′-H), 3.57 (m, 1H, 3′-H), 3.49 (m, 2H, 5′-CH_2), 3.44 (m, 2H, CH_2–OH), 2.12 (m, 2H, 2′-CH_2 (β)), and 1.80 (m, 2H, 2′-CH_2 (α)); ^{13}C NMR (125 MHz, DMSO-d_6) of the major isomer, δ 172.3 (4-C=O), 151.8 (2-C=O), 86.0 (4′-CH), 83.3 (1′-CH), 76.3 (5-C), 76.2 (6-CH), 70.9 (3′-CH), 65.1 (5′-CH_2), 62.2 (CH_2-OH), 36.9 (2′-CH); the minor isomer, δ 172.5 (4-C=O), 152.4 (2-C=O), 85.9 (4′-CH), 83.9 (1′-CH), 76.6 (5-C), 75.5 (6-CH), 69.7 (3′-CH), 64.9 (5′-CH_2), 60.7 (CH_2–OH), 38.4 (2′-CH).

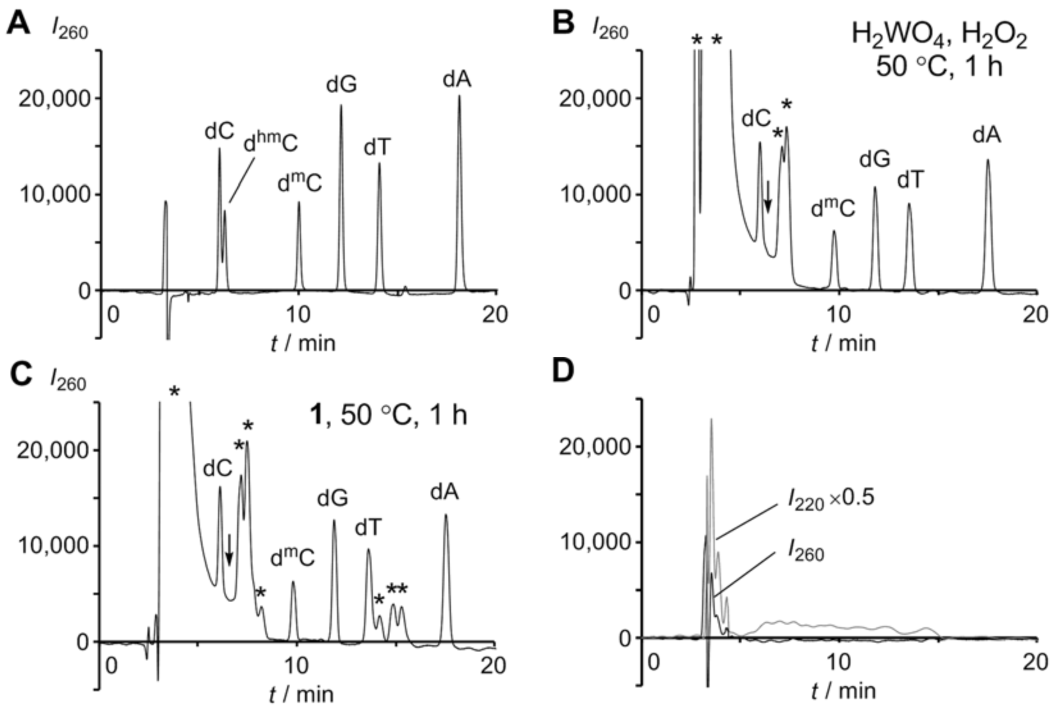

Fig. S1 HPLC profiles of nucleosides before and after tungsten oxidation. The nucleotides were analysed using reverse phase HPLC on a 5-ODS-H column (10 × 150 mm, elution carried out using a solvent mixture of 0.1 M triethylammonium acetate (pH = 7.0), and a linear gradient over 20 min from 3% to 10% acetonitrile). Profile A: Before oxidation. Profile B: Treatment with tungstic acid and hydrogen peroxide. Profile C: Treatment with **1**. The signal from $d^{hm}C$ disappeared (denoted by the arrow). The asterisks denote the signals originating from degradation of oxidants. Profile D shows the reaction products from $d^{hm}C$ (black = 260 nm monitoring, and gray = 220 nm monitoring).

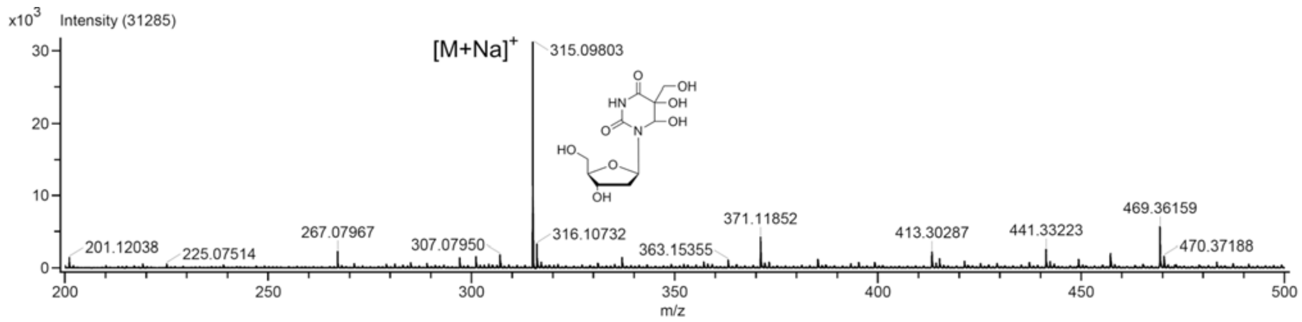

Fig. S2 ESI mass spectrum of the peroxotungstate **1** oxidation products of d^{hm}C. The sodium salt of trihydroxylated dT is shown in the figure inset.

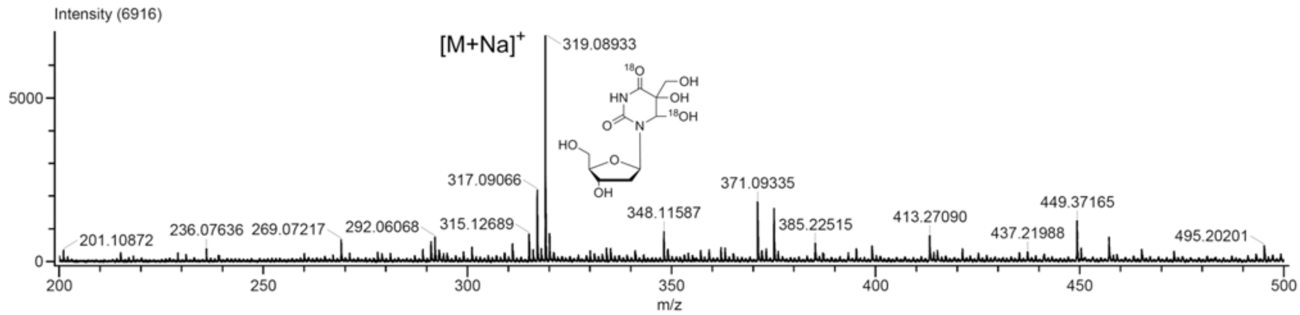

Fig. S3 ESI mass spectrum of the peroxotungstate **1** oxidation products of d^{hm}C (Oxidation in 90% $H_2{}^{18}O$). The sodium salt of trihydroxylated dT (2^{18}O) is shown in the figure inset.

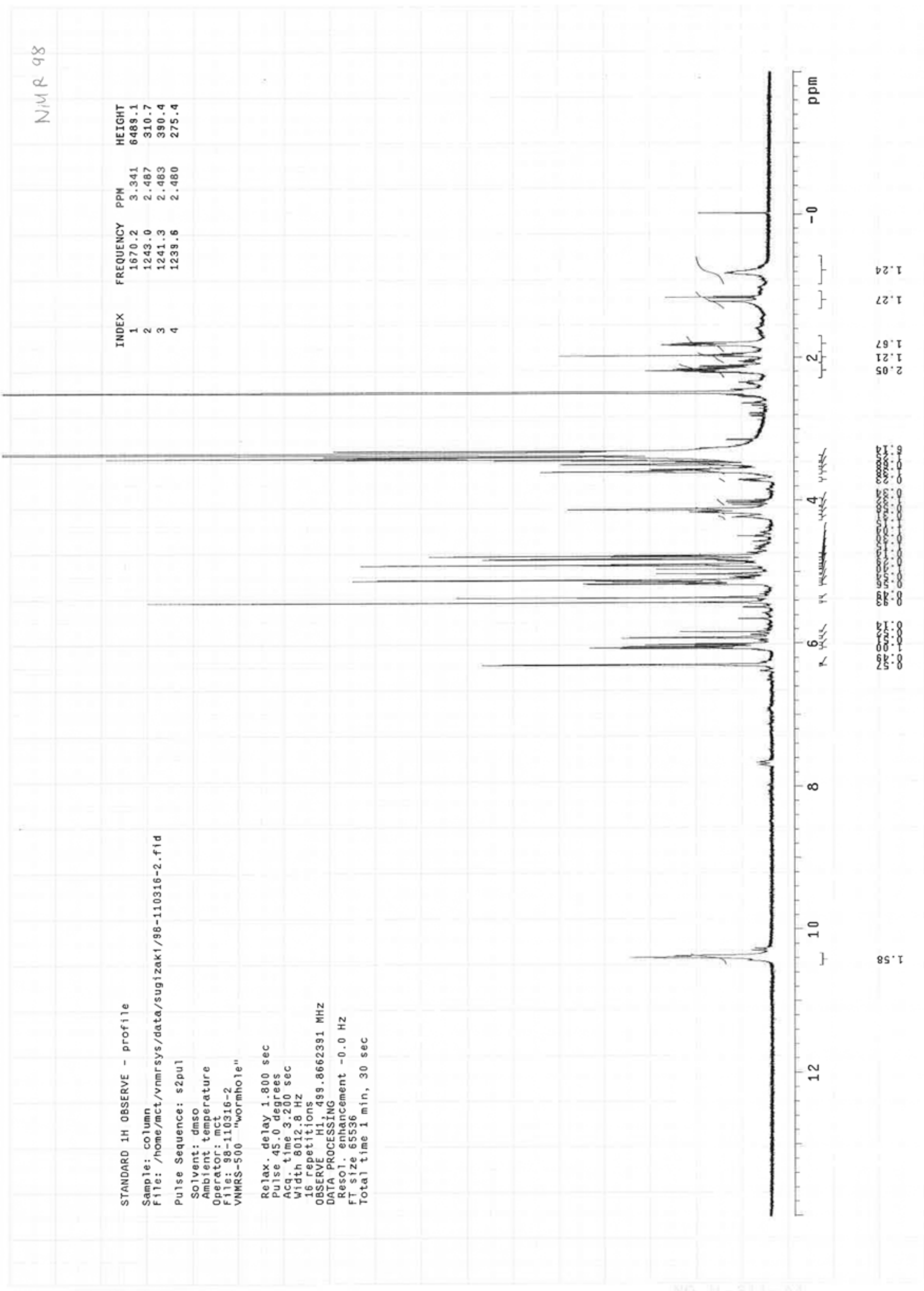

Fig. S4 ¹H NMR spectra of the peroxotungstate **1** oxidation products of d^{hm}C.

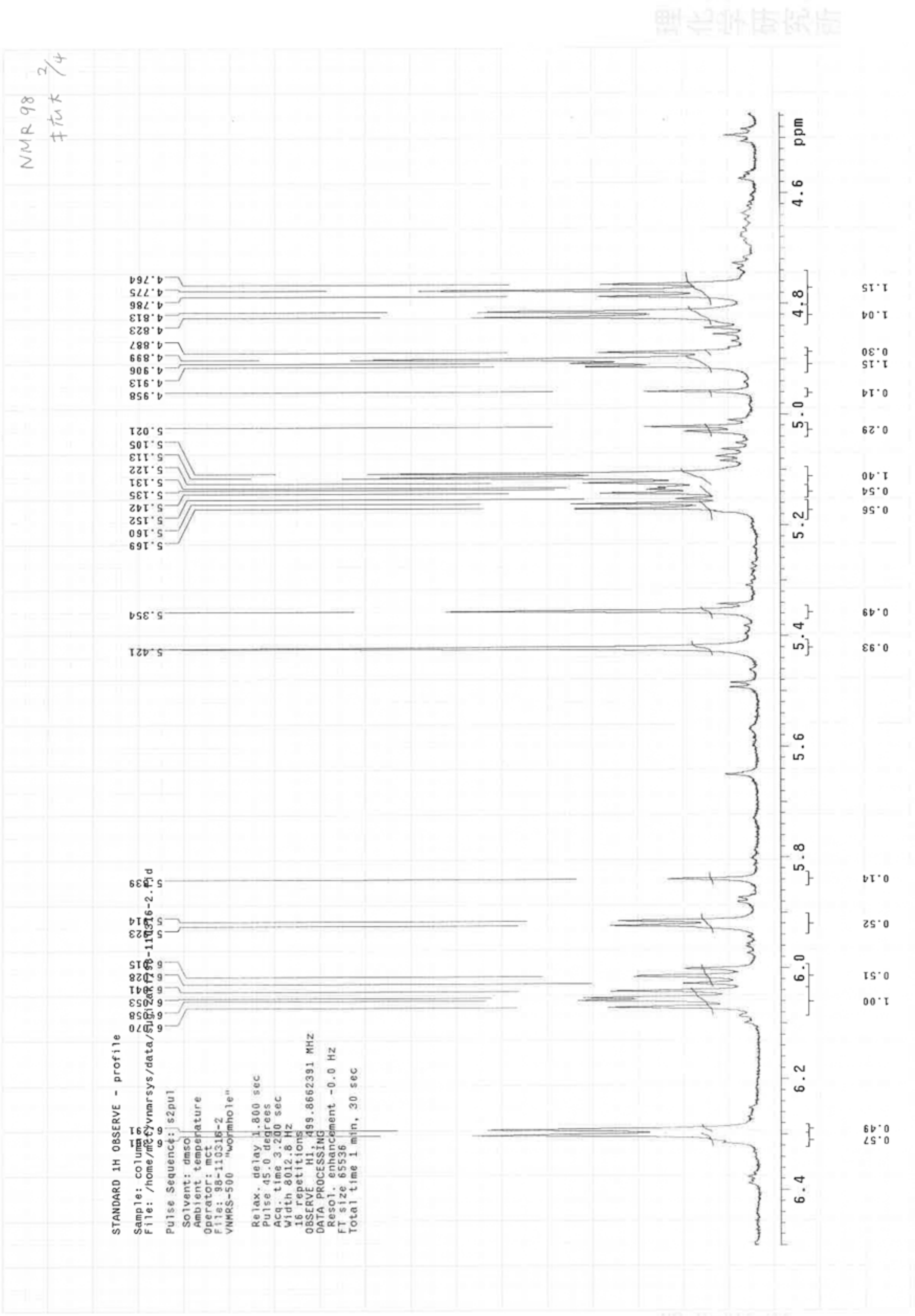

Fig. S4 (continued).

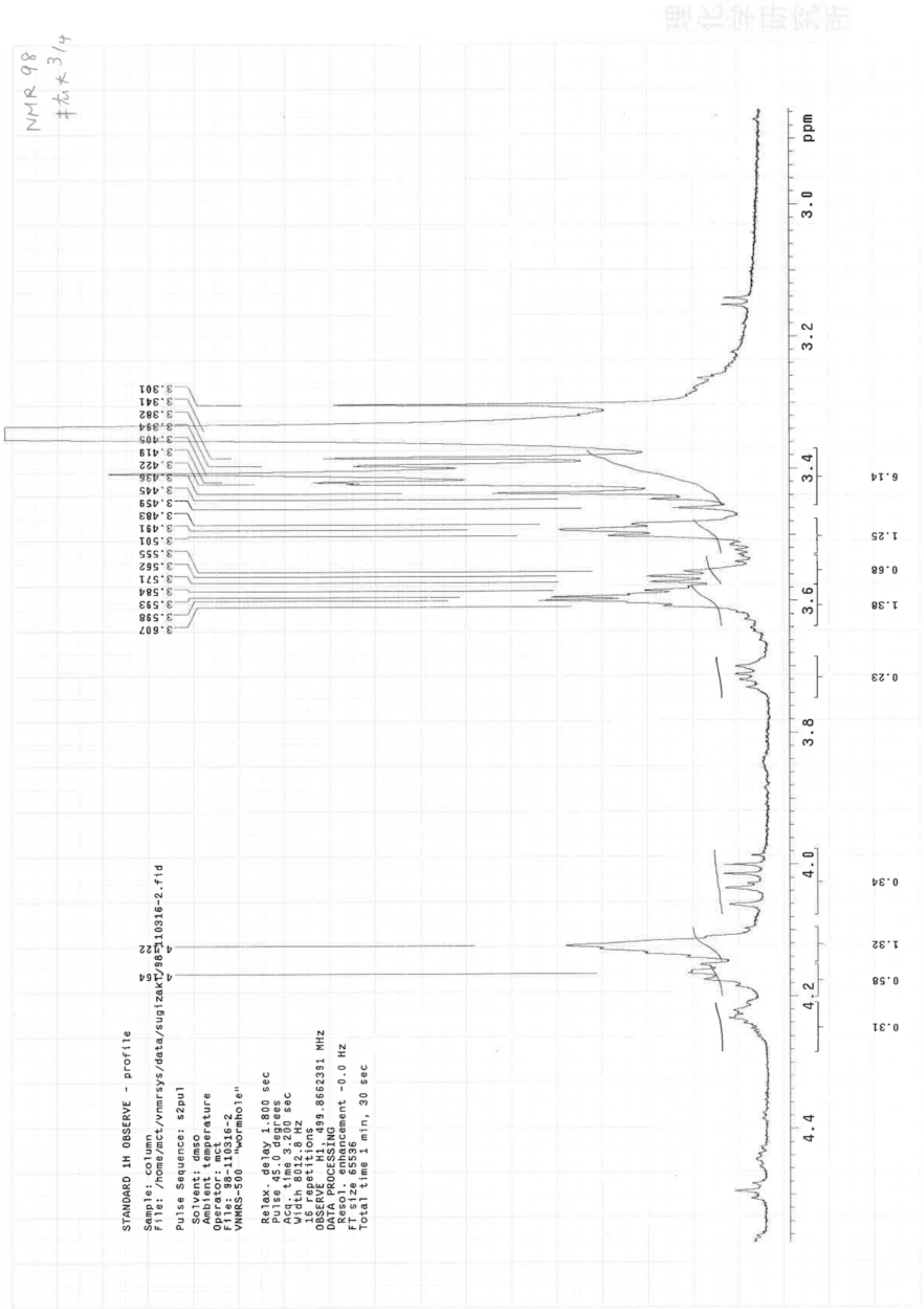

Fig. S4 (continued).

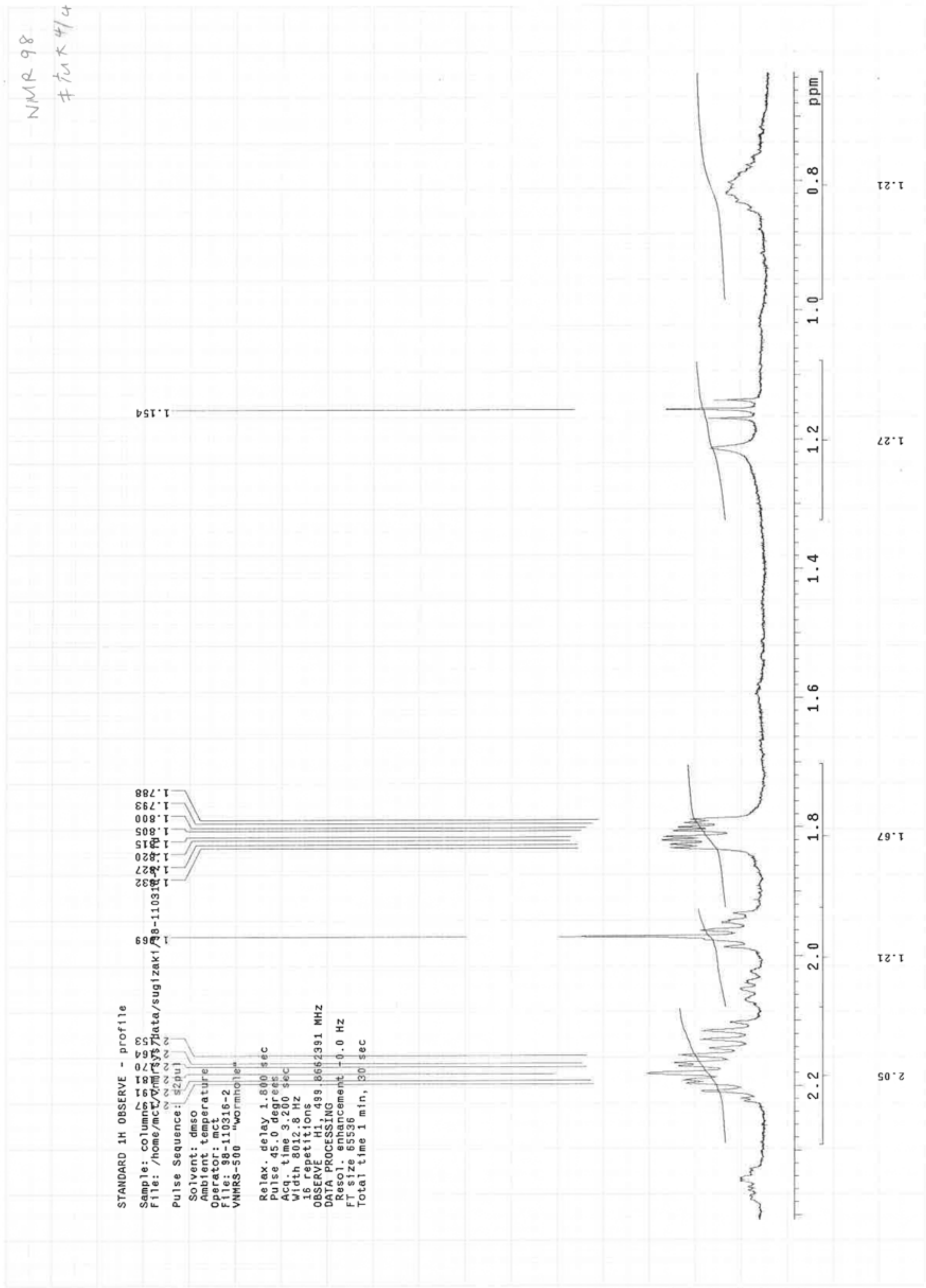

Fig. S4 (continued).

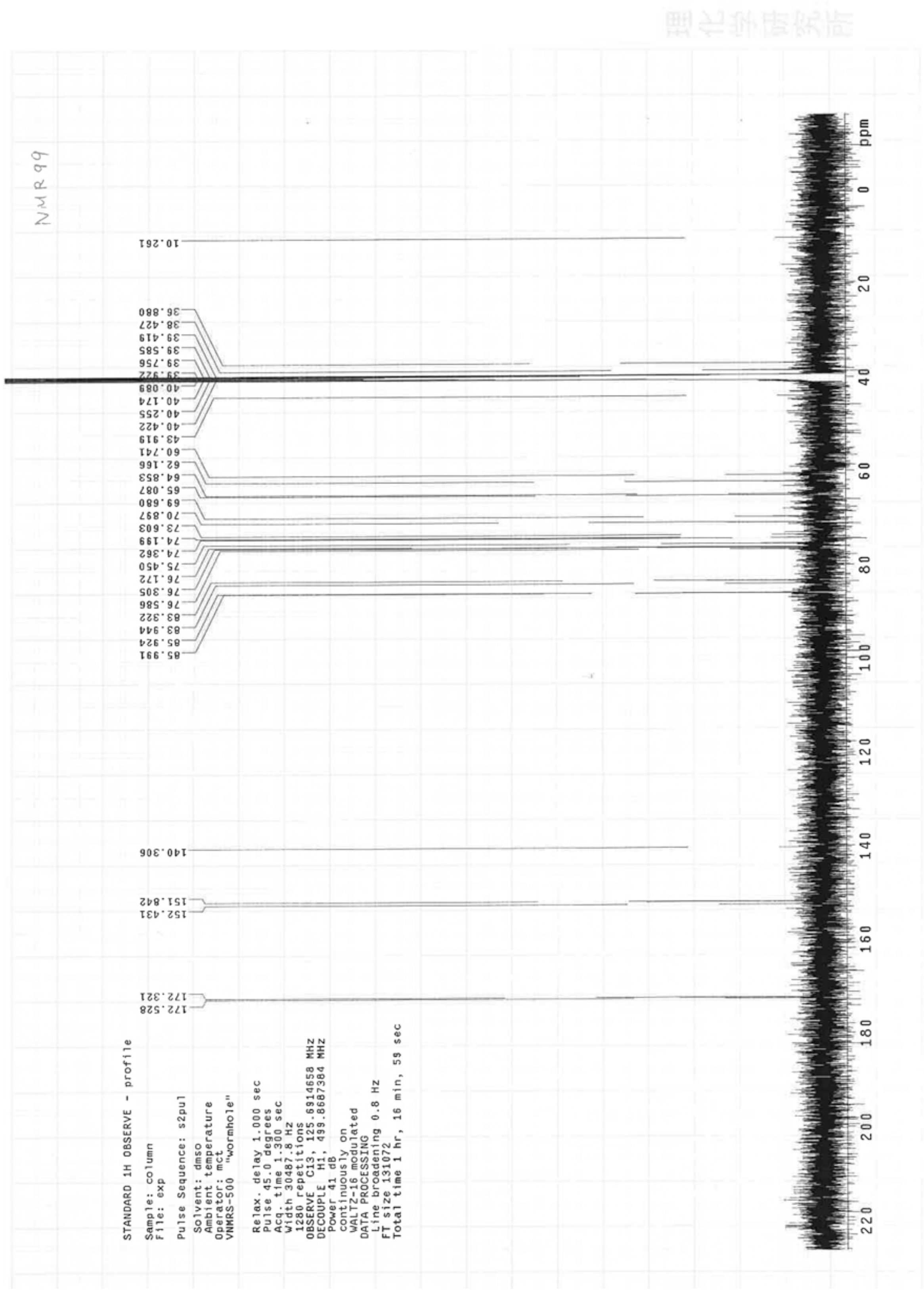

Fig. S5 ^{13}C NMR spectra of the peroxotungstate **1** oxidation products of d^{hm}C.

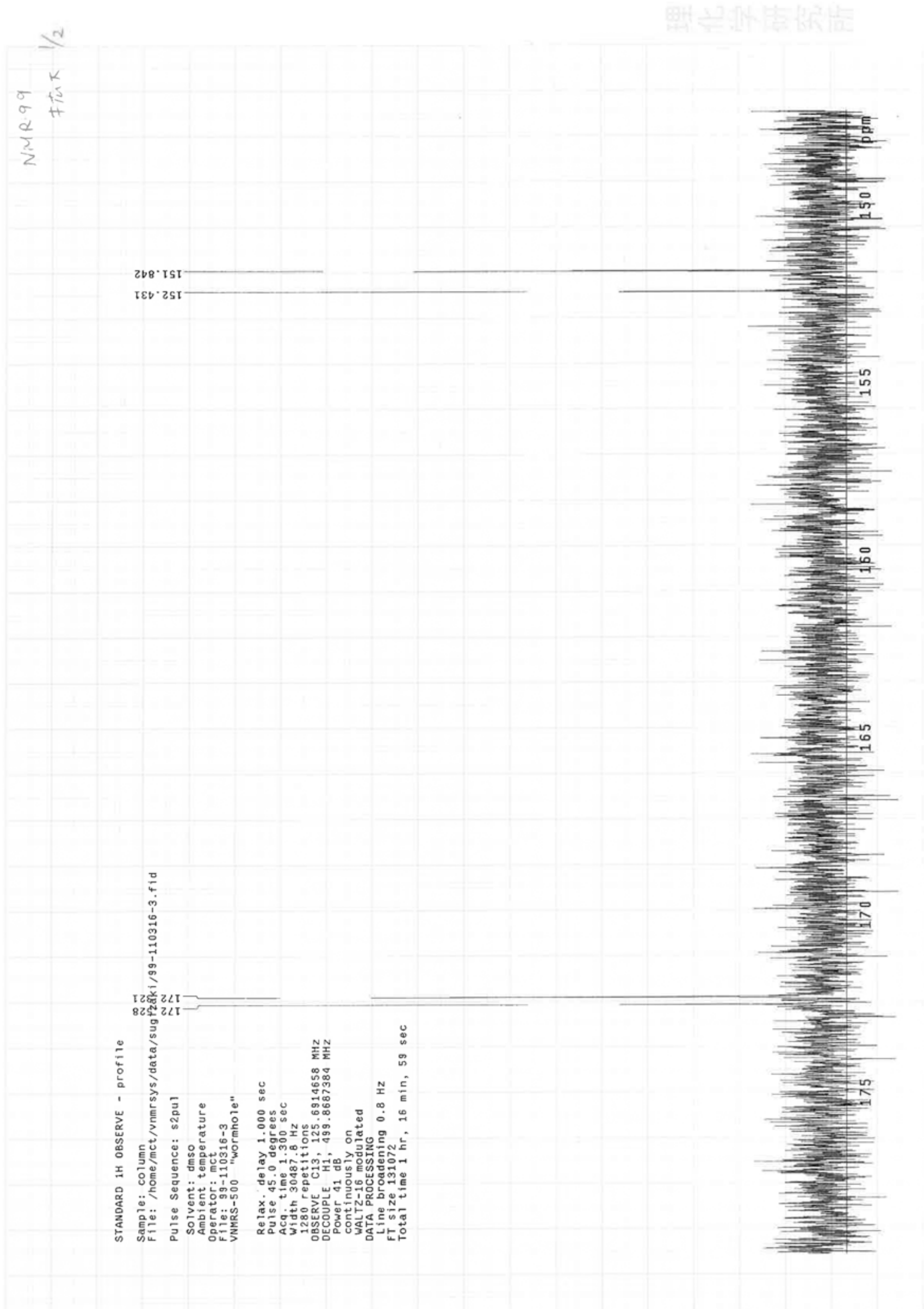

Fig. S5 (continued).

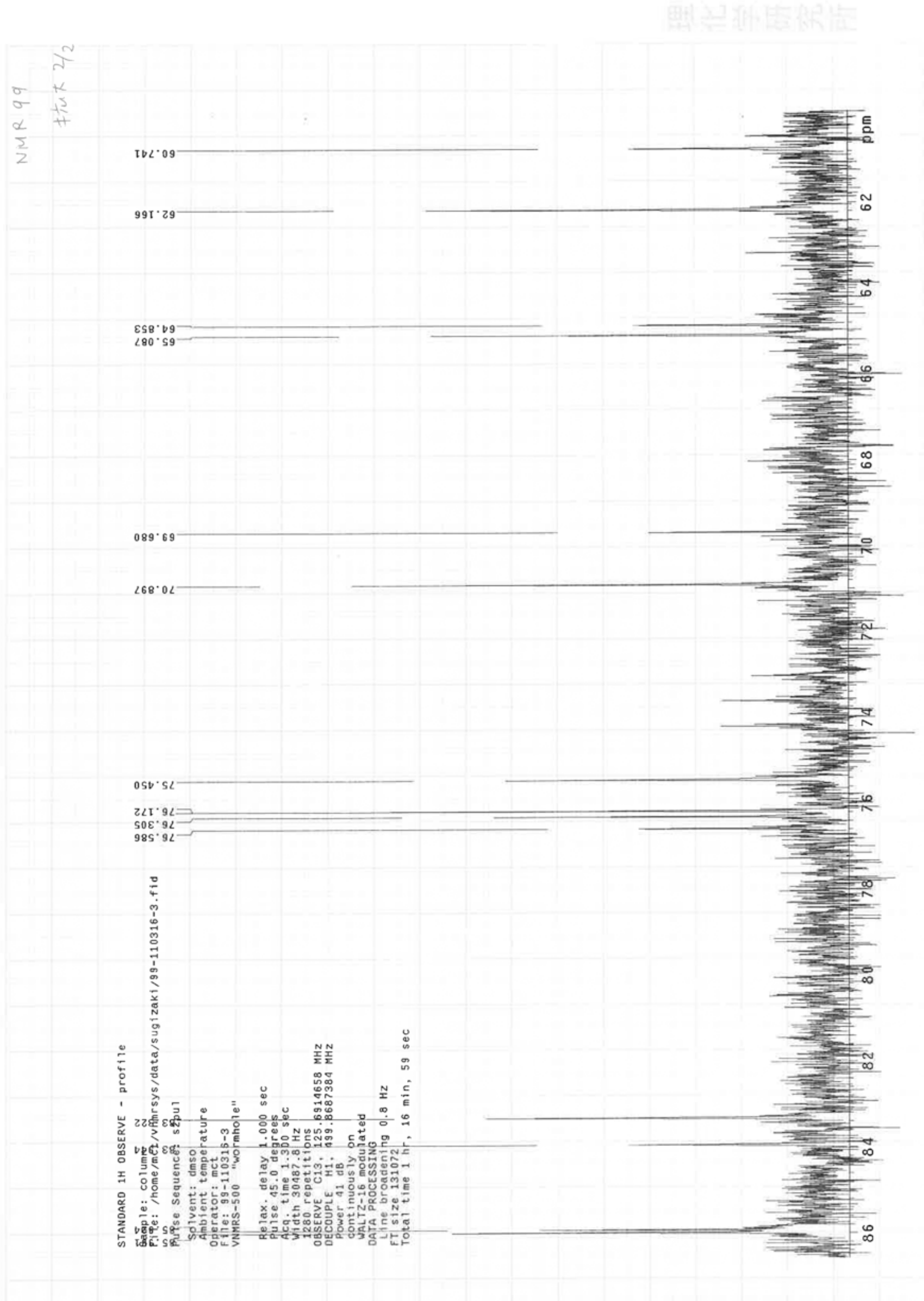

Fig. S5 (continued).

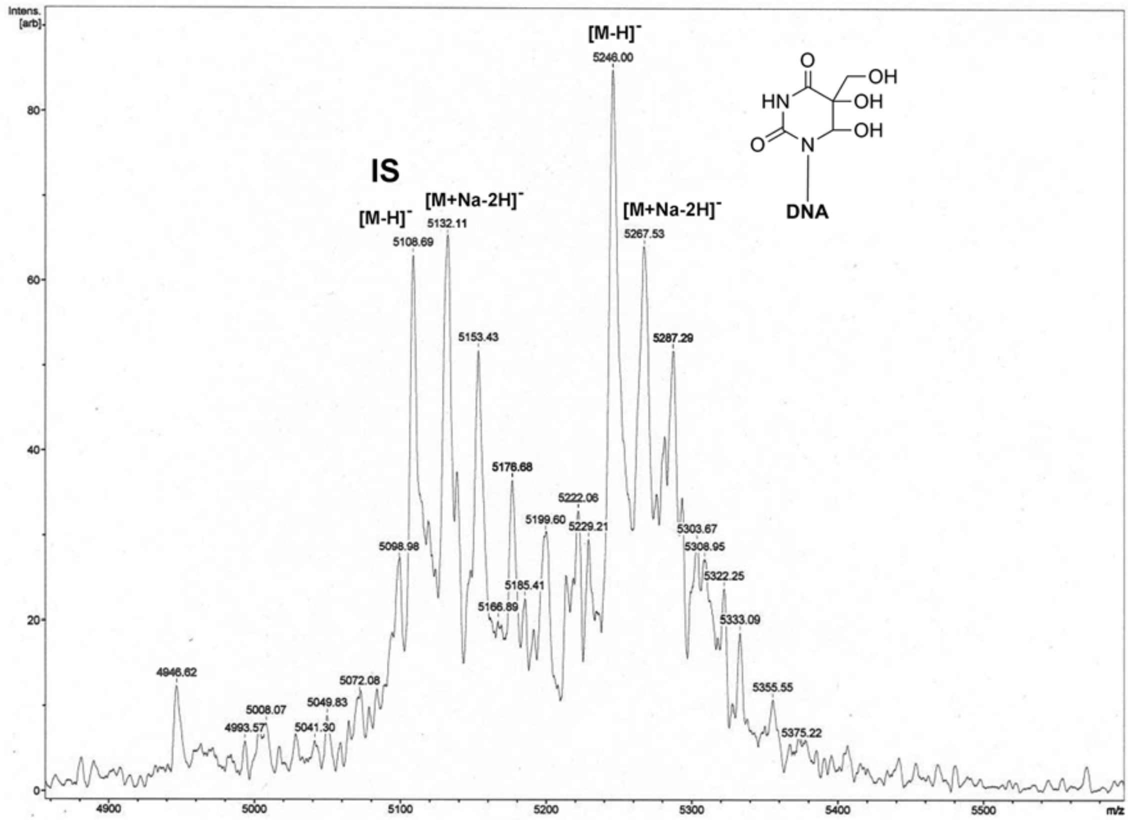

Fig. S6 MALDI-TOF mass spectrum of the peroxotungstate **1** oxidation products of **DNA1(^{hm}C)**. Matrix = 2,4,6-trihydroxyacetophenone and diammonium hydrogen citrate. Negative mode. Dihydroxylated product: [M – H]$^-$, $C_{177}H_{207}N_{73}O_{89}P_{15}$, calc. = 5245.59, found = 5246.00. IS denotes internal standard, $(dT)_{17}$.

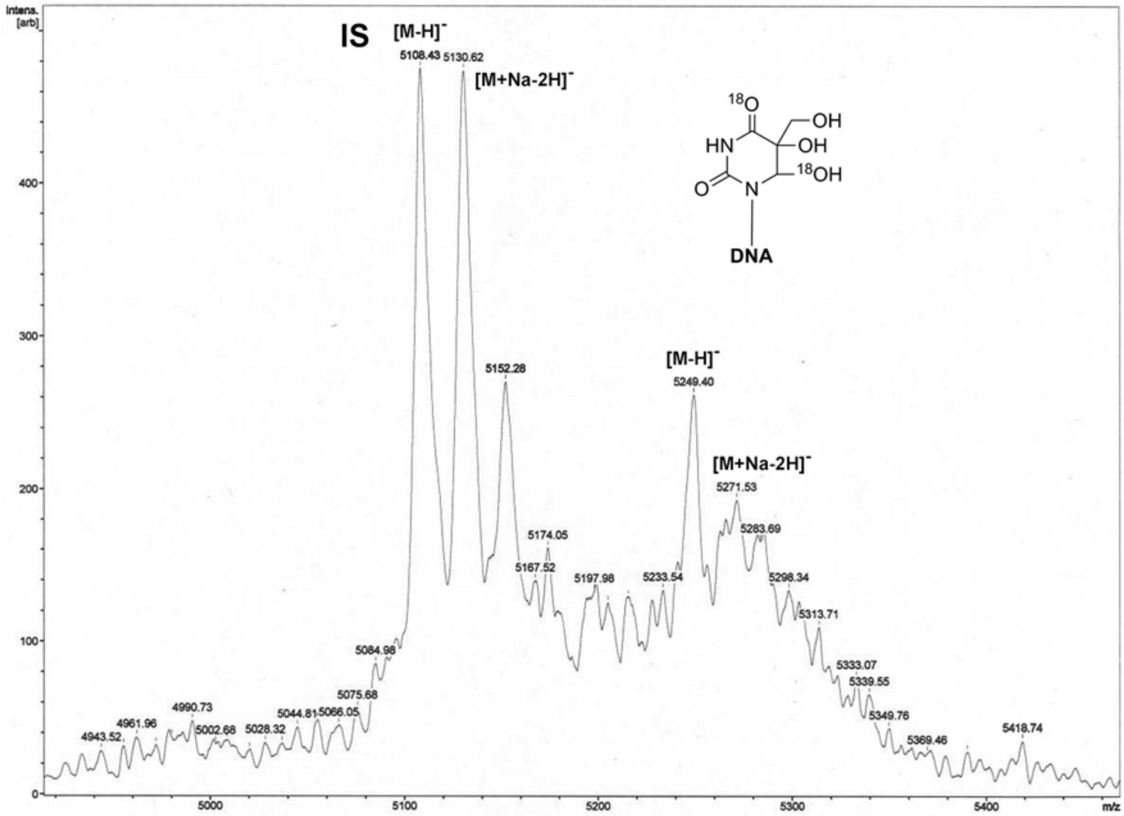

Fig. S7 MALDI-TOF mass spectrum of the products of **DNA1(^{hm}C)** oxidized with peroxotungstate **1** in $H_2{}^{18}O$. Matrix = 2,4,6-trihydroxyacetophenone and diammonium hydrogen citrate. Negative mode. Dihydroxylated product: $[M-H]^-$, $C_{177}H_{207}N_{73}{}^{16}O_{87}{}^{18}O_2P_{15}$, calc. = 5249.59, found = 5249.40. IS denotes internal standard, $(dT)_{17}$.

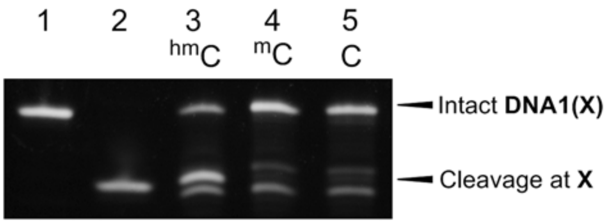

Fig. S8 A gel image after PAGE analysis of the products after oxidation by tungstic acid in the presence of hydrogen peroxide. Lane 1 = intact DNA1(^{hm}C), Lane 2 = 5′-Fluo-AAAAAAGp-3′, and Lanes 3–5 = oxidation products of DNA1(X).

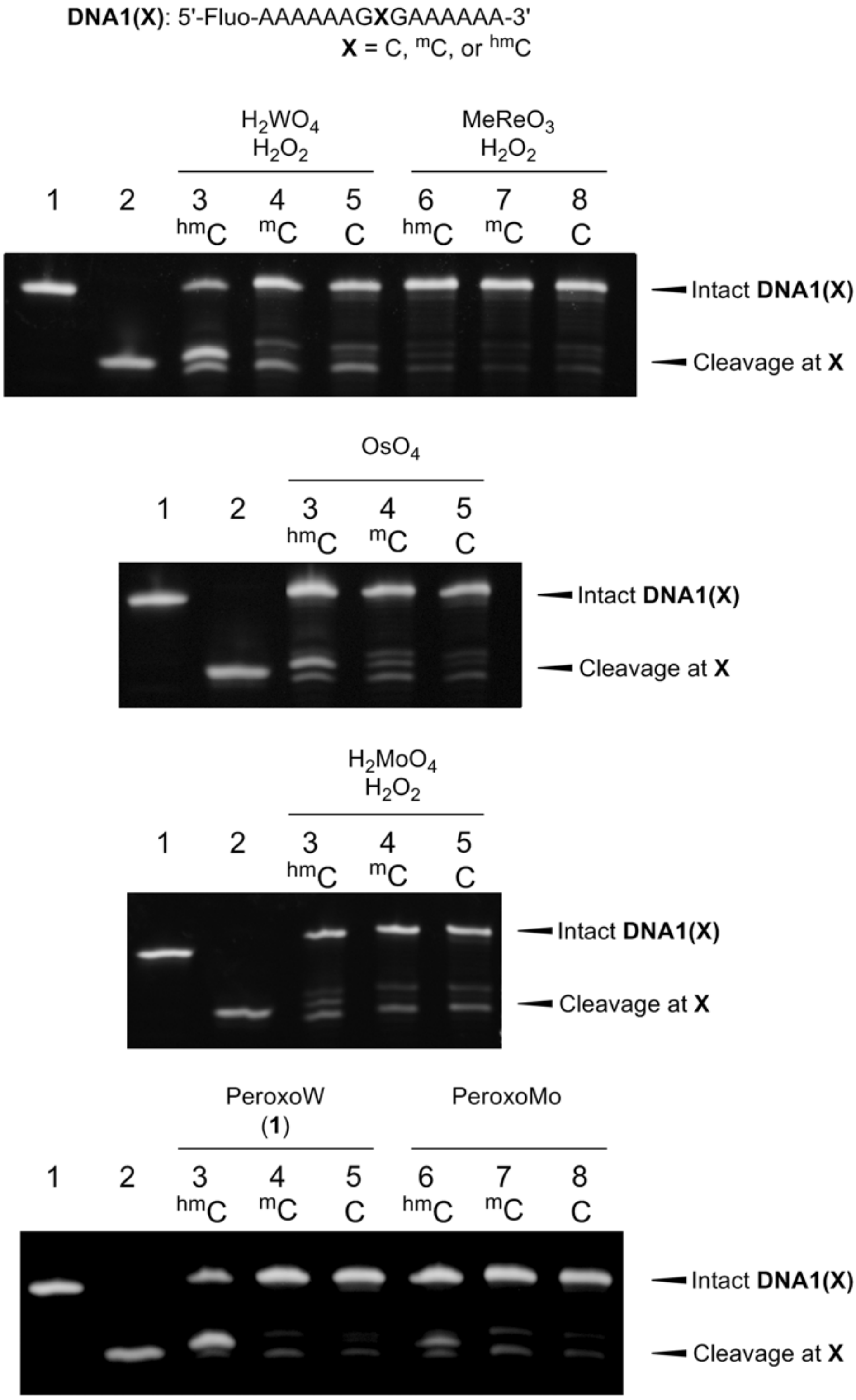

Fig. S9 PAGE results of the oxidation products of **DNA1(X)** (X = C, ^{m}C, or ^{hm}C) using W, Os, Mo, and Re. The band intensities are summarized in Fig. 1b. Lane 1 = intact **DNA1(^{hm}C)** and Lane 2 = 5′-fluo-AAAAAAGp-3′.

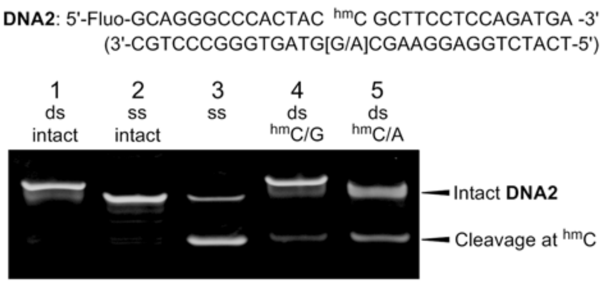

Fig. S10 A gel image after PAGE analysis of the products after oxidation by tungstic acid in the presence of hydrogen peroxide. Lane 1, the hybrid of DNA2 and fully matched complementary DNA (T_m = 75 °C); lane 2, intact DNA2; lane 3, oxidation products of DNA2; lane 4, oxidation products of the hybrid of DNA2 and fully matched complementary DNA; lane 5, oxidation products of the hybrid of DNA2 and ^{hm}C/A-mismatched DNA (T_m = 69 °C).

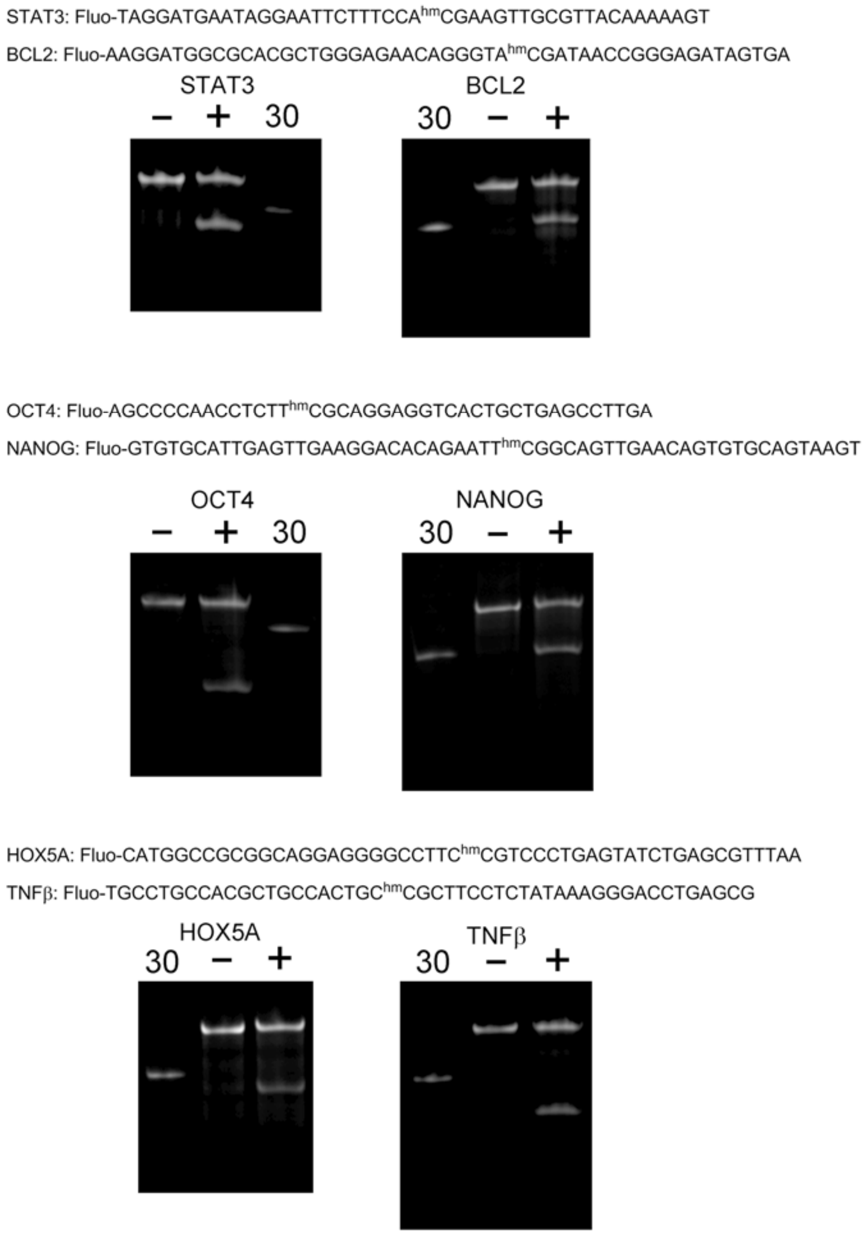

Fig. S11 PAGE results of the oxidation products of the DNA fragments containing ^{hm}CG dinucleotides in the promoter regions of human STAT3, BCL2, OCT4, NANOG, HOX5A and TNF-β genes. Lane '−', before oxidation; lane '+', after oxidation and piperidine treatment; lane '30', a 30-mer DNA marker (5′-Fluorescein-GCAGGGC-CCACTAC^{hm}CGCTTCCTCCAGATGA-3′).