Supporting Information

Trapping and Structural Elucidation of a Very Advanced Intermediate in the Lesion-Extrusion Pathway of hOGG1

Seongmin Lee, Christopher T. Radom, and Gregory L. Verdine*
Department of Chemistry
Harvard University
Cambridge, Massachusetts 02138

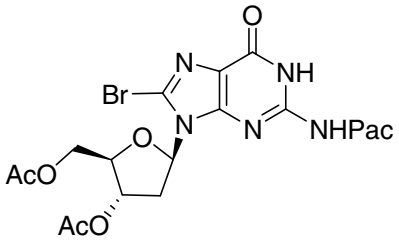
General Methods

All reagents, including NBS, DEAD, DMTrCl, 2′-deoxyguanine, 4,5-dicyanoimidazole, PacCl, *o*-nitrophenethyl alcohol, and 2-cyanoehtyltetraisopropyl phosphordiamidite, purchased were used as received. *o*-Nitrophenisopropyl alcohol was prepared by known procedures.[1] NBS was recrystallized from boiling water. Sodium sulfate (Na_2SO_4) was anhydrous. All solvents used in reactions, recrystallization, chromatography, and workup were distilled.

Unless otherwise indicated, all reactions were carried out under in a positive pressure of argon in anhydrous solvents and the reaction flasks were fitted with rubber septa for the introduction of substrates and reagents via syringe. Progress of reactions was monitored by thin layer chromatography (TLC) in comparison with the starting materials. All TLC analyses were carried out on Merck Silica Gel 60 F254 TLC plates (0.25 mm thickness). The plates were visualized by ultraviolet illumination at 254 nm and immersion in visualizating solution. The two commonly employed TLC visualizing solutions were: (i) *p*-anisaldehyde solution (1350 mL absolute ethanol, 50 mL concentrated H_2SO_4, 37 mL *p*-anisaldehyde), and (ii) permanganate solution (weight percents of 1 % $KMnO_4$ and 2 % Na_2CO_3 in water).

Analytical samples were obtained from flash silica gel chromatography, using silica gel of 230-400 mesh ASTM. ^{1}H NMR, ^{13}C NMR and ^{31}P NMR spectra were recorded on Bruker AM 500 (500 MHz). NMR spectra were determined in chloroform-d_1 ($CDCl_3$), DMSO-d_6 or methanol-d_4 (CD_3OD) solution and are reported in parts per million (ppm) from the residual chloroform (7.24 ppm and 77.0 ppm) standard. Peak multiplicates in ^{1}H-NMR spectra, when reported, are abbreviated as s (singlet), d (doublet), t (triplet), m (multiplet), and/or ap (apparent) and/or br (broad). Mass spectra were all obtained on either a JEOL AX-505 or a JEOL SX-102. HPLC analyses were carried out on an HP 1090 liquid chromatograph equipped with a Beckman C18 reversed-phase column.

Preparation of 8-bromo-2-*N*-phenoxyacetyl-3′, 5′-diacetyl-2′-deoxyguanosine (**2**)

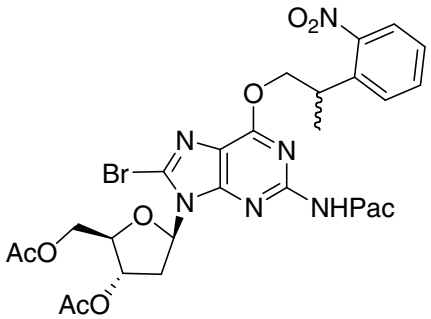


To a suspension of 2'-deoxyguanosine (2.85 g, 10 mmol) in acetonitrile were added acetic anhydride (3 equiv.), triethyl amine (5 equiv.), and DMAP (0.2 equiv.) and the mixture was stirred for 6 hours at ambient temperature. The solid was collected by filtration, washed with cold acetonitrile (50 mL), and dried to provide 2'-deoxy-3',5'-diacetyl guanosine. A solution of 2'-deoxy-3',5'-*O*-diacetyl guanosine (6.56 g) in freshly dried pyridine (100 mL) was cooled to 0 °C and phenoxyacetyl chloride (7.7 mL, 3 equiv.) was added dropwise for 30 minutes. After stirring for 18 hours at 0 °C, the reaction mixture was quenched by adding saturated $NaHCO_3$, extracted with ethyl acetate, washed with brine, dried over anhydrous sodium sulfate, evaporated under reduced pressure, and purified by silica gel chromatography to provide *N*2-phenoxyacetyl-2′-deoxy-3′5′-diacetyl guanosine (5.80 g). ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.8 (1H, br), 9.71 (1H, br), 7.88 (1H, s, C_8-H), 7.30-6.80 (5H, m), 6.22 (1H, s, C1'-H), 5.40 (1H, s, C3'-H), 4.72 (2H, s, C5'-H), 4.36 (1H, C4'-H),

4.27 (2H, s, CH_2OPh), 2.87 (1H, m, C2'-H), 2.56 (1H, m, C2'-H), 2.10 (3H, s, OAc), 1.98 (3H, s, OAc); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.9, 170.5, 170.4, 156.8, 155.5, 147.8, 146.7, 138.0, 130.0, 129.7, 122.9, 122.0, 115.0, 84.8, 82.7, 74.5, 67.1, 63.8, 37.3, 21.1, 20.9.

To an acetonitrile solution of *N*2-phenoxyacetyl-2′-deoxy-3′5′-diacetyl guanosine (4.30 g) were added pyridine (1.4 mL, 2 equiv.) and *N*-bromosuccinimide (4.68 g, 3 equiv.) and the mixture was vigorously stirred for 15 hours at ambient temperature. The reaction mixture was concentrated under reduced pressure and subjected to silica gel chromatography to afford 8-bromo-*N*2-phenoxyacetyl-2′-deoxy-3′5′-diacetyl guanosine (**2**, 2.59g). ^1H NMR (500 MHz, CD_3OD) δ 11.86 (1H, s, br), 9.60 (1H, s, br), 7.36-7.04 (5H, m), 6.32 (1H, t, J = 6.8 Hz, C1'-H), 5.63 (1H, m, C3'-H), 4.71 (2H, s, CH_2OPh), 4.61 (1H, dd, J=5.9, 11.3 Hz), 4.30 (1H, m, C4'-H), 4.26 (1H, dd, J = 6.8, 11.3 Hz), 3.43 (1H, ddd, J = 13.6, 7.3, 6.2 Hz, C2'-H), 2.42 (1H, m, C2'-H), 2.15 (3H, s, OAc), 1.98 (3H, s, OAc); ^{13}C NMR (125 MHz, DMSO-d_6) δ 172.8, 171.6, 172.3, 170.9, 158.2, 157.6, 147.6, 132.8, 130.2, 122.0, 117.5, 115.2, 66.8, 21.6.; LRMS (M+H) 564; HRMS calculated for $C_{22}H_{22}BrN_5O_8$ (M+H) 564.0730, found 564.0750.

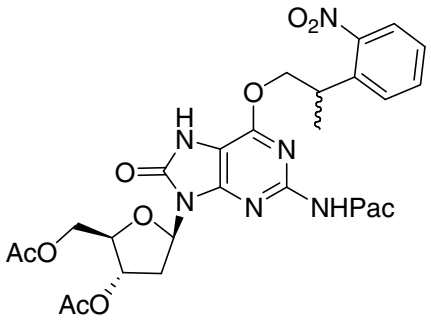
Preparation of 8-bromo-6-*O*-(*o*-nitrophenisopropyl)-2-*N*-phenoxyacetyl-3′, 5′-diacetyl-2′-deoxyguanosine (**3**)



To a THF (10 mL) solution of 8-bromo-*N*2-phenoxyacetyl-2′-deoxy-3′5′-diacetyl guanosine (563 mg), triphenyl phosphine (393 mg, 1.5 equiv.), and *o*-nitrophenisopropyl alcohol (270 mg, 1.5 equiv.) was added diethylazodicarboxylate (234 μL, 1.5 equiv.) dropwise and the mixture was stirred for 12 hours at ambient temperature. The solvent was removed under reduced pressure and the mixture was subjected to silica gel chromatography to give phenisopropyl ether **3** (297 mg, 41%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.97 (1H, s, br, NH), 7.90-7.02 (9H, m, aromatic C-H), 6.38 (1H, t, J = 7.0 Hz, C1'-H), 5.81 (1H, dt, J = 6.6, 3.3 Hz, C3'-H), 4.82 (2H, m), 4.65 (2H, d, J = 4.4 Hz, CH_2OPh), 4.56 (1H, dd, J = 5.8, 11.8 Hz, C5'-H), 4.42 (1H, dd, J = 7.5, 11.4 Hz, C5'-H), 4.33 (1H, dt, J = 3.5, 8.4 Hz, C4'-H), 3.82 (1H, m, C2'-H), 3.47 (2H, t, J = 7.0 Hz), 2.45 (1H, ddd, J = 3.0, 7.0, 14.0 Hz, C2'-H), 2.12 (3H, s, OAc), 2.00 (3H, s, OAc); ^{13}C NMR (125 MHz, $CDCl_3$) δ 171.1, 170.5, 166.4, 159.6, 157.4, 150.9, 149.6, 133.4, 132.9, 130.0, 125.1, 122.5, 117.2, 115.6, 86.5, 83.1, 74.8, 68.3, 66.7, 64.3, 62.5, 34.8, 33.0, 21.5, 21.4; LRMS (M+H) 713; HRMS calculated for $C_{30}H_{29}BrN_6O_{10}$ (M+H) 713.1207, found 713.1195.

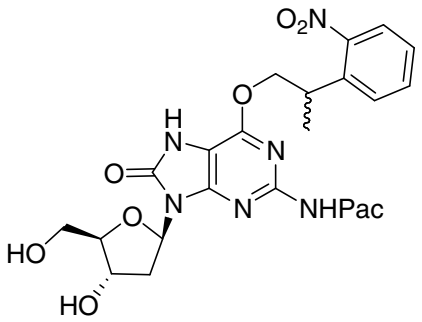
Preparation of 8-oxo-6-*O*-(*o*-nitrophenisopropyl)-2-*N*-phenoxyacetyl-3′, 5′-diacetyl-2′-deoxyguanosine (**4**)



To an acetic acid solution of phenisopropyl ether **3** (287 mg) was added sodium acetate (650 mg, 20 equiv.) in one portion and the mixture was stirred for 12 hours at 80 °C. After evaporation of the solvent under reduced pressure, the reaction mixture was partitioned between ethyl acetate and water. The organic layer was separated, washed with sat.

aqueous $NaHCO_3$, dried over sodium sulfate, evaporated, and subjected to silica gel chromatography to provide **4** (120 mg, 46%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.47 (1H, br, N7-H), 8.83 (1H, br, N1-H), 7.83-7.04 (9H, m), 6.38 (1H, t, J = 6.9 Hz, C1'-H), 5.63 (1H, dt, J = 3.3, 6.2 Hz, C3'-H), 4.73-4.64 (4H, m), 4.56 (1H, dd, J = 5.5, 11.4 Hz, C5'-H), 4.39 (1H, dd, J = 7.0, 11.4 Hz, C5'-H), 4.25 (1H, m, C4'-H), 3.52 (1H, dt, J = 6.9, 11.9 Hz, C2'-H), 3.41 (2H, t, J = 6.3 Hz), 2.33 (1H, ddd, J = 3.3, 7.0, 14.0 Hz, C2'-H), 2.09 (3H, s, OAc), 2.01 (3H, s, OAc); ^{13}C MMR (125 MHz, $CDCl_3$) δ 171.00, 170.6, 157.3, 153.0, 152.3, 150.4, 150.0, 149.7, 133.3, 133.1, 132.8, 132.4, 132.3, 132.2, 130.0, 128.8, 124.8, 122.5, 115.1, 103.0, 82.3, 82.2, 75.2, 67.9, 67.0, 64.3, 60.6, 33.5, 32.5, 21.2, 21.0, 14.4; LRMS (M+H) 651; HRMS calculated for $C_{30}H_{30}N_6O_{11}$ (M+H) 651.2051, found 651.2039.

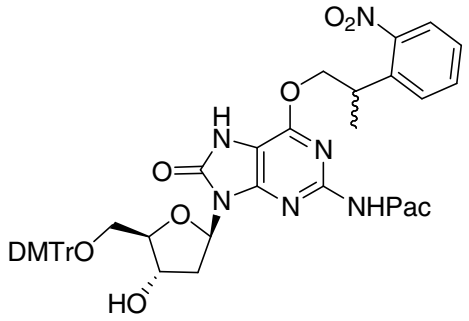
Preparation of 8-oxo-6-*O*-(*o*-nitrophenisopropyl)-2-*N*-phenoxyacetyl-2′-deoxyguanosine



To an ethyl alcohol (10 mL) solution of diacetate **4** (119 mg) was added 1N aqueous sodium hydroxide (5 mL) in one portion at 0 °C and the mixture was vigorously stirred for 10 minutes. The reaction was quenched with 1N HCl, concentrated, and purified by silica gel chromatography to afford the diol (94 mg, 89%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.80 (1H, br, N7-H), 8.96 (1H, br, N1-H), 7.83-6.95 (9H, m), 6.43 (1H, t, J = 6.6 Hz, C1'-H), 4.83 (1H, m, C4'-H), 4.71-4.38 (5H, m), 3.89 (1H, d, J = 10.9 Hz, C5'-H), 3.78 (1H, d, J = 11.1 Hz, C5'-H), 3.41 (2H, t, J = 6.2 Hz), 2.90 (1H, dt, J = 7.3, 13.2 Hz), 2.24 (1H, ddd, J =

3.3, 7.0, 14.0 Hz, C2'-H), 2.09 (3H, s, OAc), 2.01 (3H, s, OAc); LRMS (M+H) 567; HRMS calculated for $C_{26}H_{26}N_6O_9$ (M+H) 581.1996, found 581.1998.

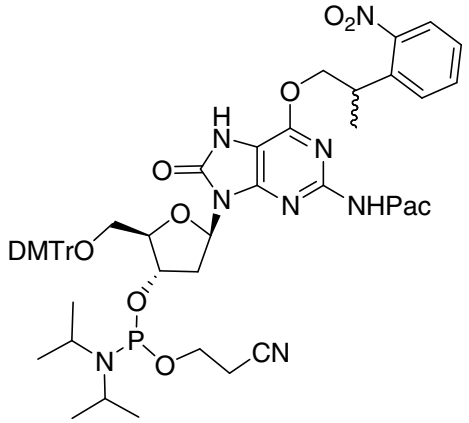
Preparation of 8-oxo-6-*O*-(*o*-nitrophenisopropyl)-2-*N*-phenoxyacetyl-5′-*O*-(4,4′-dimethoxytrityl)-2′-deoxyguanosine



8-oxo-6-*O*-(*o*-nitrophenisopropyl)-2-*N*-phenoxyacetyl-2′-deoxy guanosine (93 mg) was dissolved in pyridine (3 mL), treated with dimethoxytrityl chloride (162 mg, 3 equiv.), and stirred two hours at room temperature. The reaction was stopped by adding saturated aqueous sodium bicarbonate. The aqueous phase was extracted with ethyl acetate and the combined organic layer was washed with brine, dried over sodium sulfate, filtered, and concentrated. Purification by chromatography on silica gel yielded the desired product (72 mg, 51%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.70 (1H, br, N1-H), 7.83-6.61 (22H, m), 6.38 (1H, dd, J = 4.8, 8.1 Hz, C1'-H), 5.09 (1H, dd, J = 6.6, 12.4 Hz, C3'-H), 4.70 (2H, m), 4.56 (1H, dd, J = 5.5, 11.4 Hz, C5'-H), 4.60 (2H, br), 4.00 (1H, dd, J = 5.5, 11.0 Hz, C4'-H), 3.67 (3H, s, OMe), 3.66 (3H, s, OMe), 3.47 (1H, dd, J = 6.3, 9.9 Hz, C5'-H), 3.41 (2H, t, J = 6.2 Hz), 3.35 (1H, dd, J = 4.7, 9.2 Hz, C5'-H), 3.16 (1H, dt, J = 6.9, 12.5 Hz, C2'-H), 2.33 (1H, dt, J = 7.7, 13.4 Hz, C2'-H), 2.09 (3H, s, OAc), 2.01 (3H, s, OAc); ^{13}C MMR (125 MHz, $CDCl_3$) δ 158.5, 157.3, 153.4, 152.2, 150.1, 149.5, 145.2, 136.4, 136.3, 133.5,

133.4, 132.8, 130.3, 130.2, 130.0, 128.4, 128.3, 127.8, 126.8, 124.9, 122.6, 115.1, 113.1, 113.0, 103.0, 86.3, 85.7, 81.4, 72.5, 67.9, 67.0, 64.8, 60.6, 55.3, 37.0, 32.6, 21.3, 14.4.

Preparation of 8-oxo-6-*O*-(*o*-nitrophenisopropyl)-2-*N*-phenoxyacetyl-5′-*O*-(4,4′-dimethoxytrityl)-3′-(2-cyanoethyl *N*, *N*-diisopropylphosphoramidite)-2′-deoxyguanosine (5)



To a CH_2Cl_2 solution of 8-oxo-6-*O*-(*o*-nitrophenisopropyl)-2-*N*-phenoxyacetyl-5′-*O*-(4,4′-dimethoxytrityl)-2′-deoxy guanosine (107 mg) and 4,5-dicyanoimidazole (7 mg, 0.7 equiv.) was added 2-cyanoehtyl tetraisopropyl phosphordiamidite (52 μL, 2 equiv.) in one portion and the mixture was stirred for 30 minutes at 25 °C. The solvent was removed in vacuo and the residue was purified by silica gel chromatography to give the desired phosphoramidite **5** (127 mg, 100%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.87 (1H, br), 8.50 (1H, br), 7.86-6.63 (22H, m), 6.39 (1H, C1'-H), 4.90 (1H, C3'-H), 4.72 (2H), 4.60 (2H), 4.18 (1H, C4'-H), 3.68 (6H, 2-OMe), 3.63-3.24 (5H), 2.72 (1H, m), 2.38 (3H, m), 1.24-1.08 (12H, m).

DNA synthesis and purification

The DNA oligomer 5'-AGCGTCCA**X**GTCTACC-3', where X denotes position of the photocaged oxoG, was synthesized on a MerMade-12 DNA synthesizer (BioAutomation) using standard reagents. Special care was taken to limit exposure of light to the photocaged oxoG phosphoramidite and oligonucleotide from this point forward. DNA oligomer 5'-

TGGTAGACCTGGACGC-3', where the underlined position denotes the site of modification with the thiol-bearing tether, was synthesized using the phosphoramidite derivative of O_4-triazoyl-dU (Glen Research), and functionalized with a two-methylene unit linker. Following DNA synthesis, the ologonucleotide was dried under vacuum and then incubated with 200 uL of a 50% solution of cystamine in water at room temperature for 12 hours. The aqueous solution was pulled off, the resin washed twice with 200 uL water. The washes were combined and neutralized to pH 6.0-pH 6.5 with glacial acetic acid and desalted using a NAP-10 column. The aqueous solution of the oligonucleotide was then dried in a speed vacuum concentrator and finally purifed using urea PAGE. DNA was purified by 20% urea-PAGE, and dissolved in 10mM Tris 8.0, 1mM EDTA. Masses were verified by MALDI-TOF spectrometric analysis.

Cloning, overexpression, and purification of N149C hOGG1

The point mutation at position 149 was introduced into the cDNA encoding the human OGG1 core domain (residues 12-327) by mutagenesis using the megaprimer method. The newly produced fragment was cloned into the pET30a vector (Novagen) using the restriction sites *EcoRI* and *HindIII* and sequenced throughout the hhOGG1-coding sequence. Protein was overexpressed in BL21 (DE3) cells (Novagen) and purified essentially as previously described,[2] differing at two steps. First, the log-phase culture grown at 37 °C was induced by adding IPTG to a final concentration of 25 μM and then incubated for 22-24 hours at 16° C. Second, the protein eluted from the Ni-NTA column in 500 mM NaCl was diluted 10x with 20 mM Tris pH 7.4, 1 mM EDTA, and 10 mM β-ME, loaded onto a Hi-Trap SP column (GE Healthcare) via peristaltic pump, and eluted with a 5 - 80% NaCl gradient over 25 mL in the same buffer.

Preparative crosslinking and crystallization

DNA duplexes were annealed by mixing 100 μM photocaged oxoG containg strand (5′-AGCGTCCA**X**GTCTACC-3′; **X** = O^6-photocaged oxoG) with thiol-tether containing

strand (5′-TGGTAGAC**C**TGGACGG; thiol-tether indicated by **C**) in degassed buffer containing 10 mM Tris pH 8.0 and 1 mM EDTA, incubating at 75 °C for 10 min., then cooling to 4° C by a step gradient of -1 °C/min. Protein for crosslinking was either obtained directly from Sephadex size-exclusion chromatography (GE Healthcare) or by buffer exchange of a protein stock solution containing 10 mM β-ME. Both methods employed degassed buffer containing 20 mM Tris pH 7.4, 1 mM EDTA, and 100 mM NaCl. In all cases degassing was accomplished by argon-purging for 60 – 90 minutes. Crosslinked complex was prepared by mixing 20 μM protein with 10 μM duplex DNA at 4 °C in batches of 1.2 mL. Reactions were incubated under argon for 48 hours. The complex was purified and crystallized by a procedure analogous to that previously described.[3] Briefly, the crosslinked complex was loaded onto a MonoQ column (GE Healthcare) and eluted using a 180 - 600 mM salt gradient in a buffer solution with 20 mM Tris pH 7.4, 1 mM EDTA. Fractions containing crosslinked complex were pooled and concentrated in an Amicon Ultra-4 5,000 MWCO filter unit (Millipore). The complex was buffer exchanged into an argon-degassed 1:1 mixture of 100 mM NaCl, 20 mM Tris pH 7.4, 1 mM EDTA, and 10 mM Tris pH 8.0, 1 mM EDTA then concentrated in excess of 280 μM. Complex was filtered through a 0.22 μm Millex-GV (Millipore) filter, quantified at UV_{260} using a 250-μM duplex of oxoG-containing DNA duplex as a standard, then diluted with the appropriate volume of 1:1 buffer to give 280 μM. Crystallization was achieved via the hanging drop vapor diffusion method by mixing 1 μL of complex with an equal volume of well solution containing 100 mM sodium cacodylate pH 6.0, 200 mM $CaCl_2$, and 12-13% PEG 8000 and equilibrating against 1 mL of well solution. Crystals appeared within a few days and were allowed to grow for 2-3 weeks. Non-photodeprotected crystals were transferred to a cryoprotectant containing 100 mM sodium cacodylate pH 6.0, 200 mM $CaCl_2$, 13% PEG 8000, and 25% glycerol, then frozen in liquid nitrogen for data collection.

Construction of Laser Apparatus

A 6 mW continuous wave diode-pumped solid state laser (CrystaLaser) at 373.5 nm with a $1/e^2$ beam diameter of 0.9 mm was affixed to a breadboard. An anti-reflection coated fused

silica plano-convex singlet (Thorlabs #LA4725-UV, f = 75 mm) was placed in the beampath to focus the light on the sample. The intensity in the 0.060 mm diameter ($1/e^2$) focus was 212 W/cm^2. A three-axis translation stage bearing a magnet from a magnetic CrystalWand (Hampton Research) was placed at the focus of the beam for convenient positioning of the cryoloop. A beamdump was placed at the end of the board to terminate the laser beam. Prior to experimental use, the laser apparatus was thermally equilibrated at 4 °C.

Analytical Photodeprotection Procedure

Crystals were harvested from the hanging drop and washed successively in three solutions containing 100 mM sodium cacodylate pH 6.0, 200 mM $CaCl_2$, 13% PEG 8000, and 25% glycerol. Crystals were mounted on the translator, exposed to laser radiation for various lengths of time, and then transferred to 20 μL of solution containing 20 mM Tris pH 7.4, 1 mM EDTA, 100 mM NaCl, and 10 mM β-ME. Crystals were then incubated at 37 °C for 30 min. DNA was then radioactively labeled with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]-ATP (PerkinElmer). Samples were loaded on a prerun 20% urea polyacrylamide gel in 1X TBE buffer and electrophoresed at 300V for several hours at room temperature. Bands were quantified using ImageQuant TL (GE Healthcare).

Photodeprotection of caged oxoG crystals

Screening for photodeprotection conditions and recovery of hOGG1 activity was performed as follows. The overall procedure for this assay is detailed in Figure 1. A continuous wave diode-pumped solid state laser (CrystaLaser) with an output at 373.5 nm and a focused beam diameter of 0.060 mm was used. Individual hexagonal rod-shaped photocaged complex (PCC) crystals were exposed to the laser beam for varying lengths of time, dissolved in buffer and incubated at 37° C, then radiolabeled and analyzed by gel electrophoresis. Crystals in the path of the beam illuminated with a brilliant blue color, a characteristic that assisted in the accurate positioning of the barely perceptible crystal in the narrow laser beam. It was found that hOGG1 did not process the caged oxoG in unflashed crystals (Figure 1B). On the other hand, laser exposure rescued hOGG1 activity in a semi-

time-dependent manner (Figure 1B-C). The variation in cleavage activity can be attributed to a number of factors, including crystal orientation and centering in the beam as well as the size of the crystal; all of these variables may affect photodeprotection efficiency.

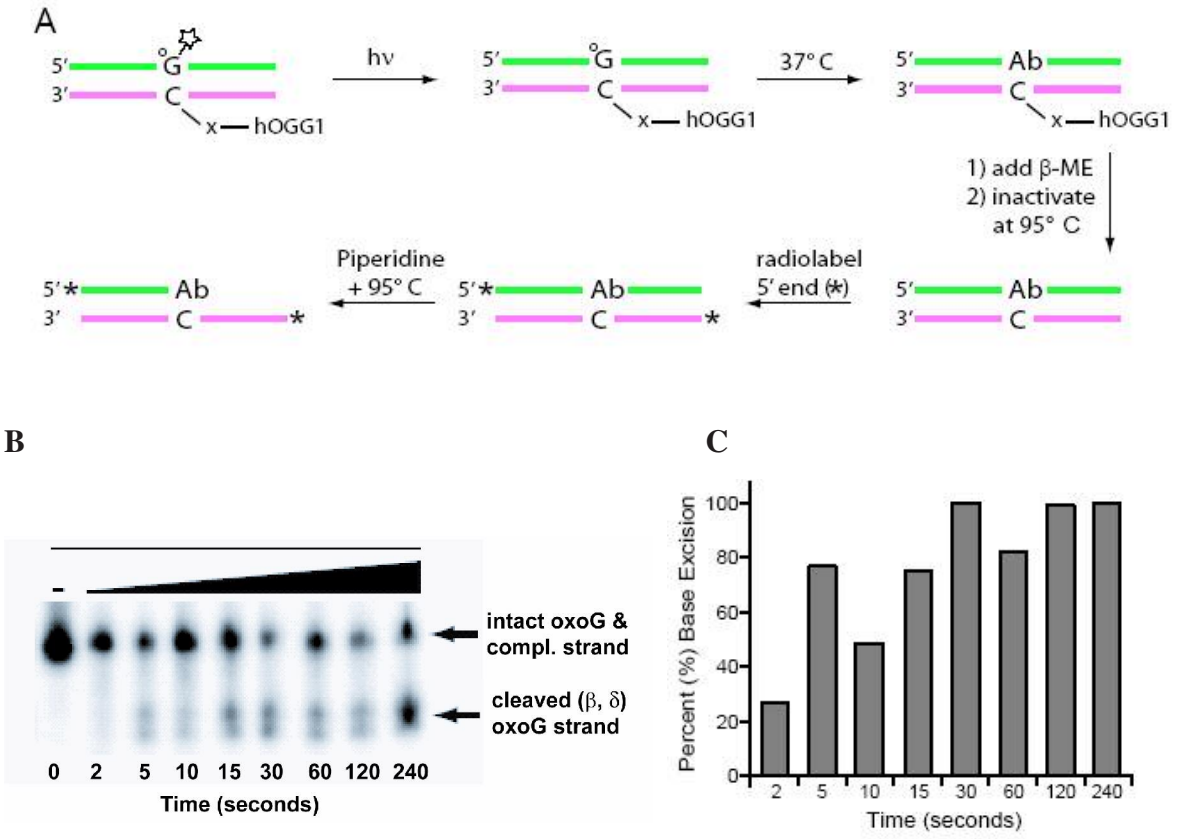


Figure 1. Photodeprotection of photocaged crystals. (A) Details of the photodeprotection assay. Crystals of hOGG1 crosslinked to photocaged oxoG DNA were irradiated with laser light at 373.5 nm to yield oxoG (°G). Crystals were then dissolved in buffer and incubated at 37°C to promote base excision which yields an abasic (Ab) site. Reducing agent (β-ME) was added to reverse the crosslink and solutions were incubated at 95°C to kill enzyme activity. DNA was radiolabeled using T4 polynucleotide kinase (PNK), then piperidine was added and the solutions were heated to 95° C to promote strand cleavage 3' of the abasic site. The 3' elimination product is not radiolabeled because it bears a phosphate at its 5' end. (B) Urea-PAGE analysis of the experiment outlined in *A*. The upper band (lower mobility) consists of both DNA strands of the duplex, the one containing oxoG (or caged oxoG in the case of the – lane) and its complement containing the thiol tether; it is important to note that only the former is cleaved by hOGG1, so at 100% cleavage the intensity of the upper band is reduced by only 50%. Note that the unflashed (-) crystals show negligible excision activity. The doublet of lower (higher mobility) bands correspond to β- (upper) and δ–cleavage (lower) products. (C) Percent (%) base excision of the PCC is represented in bar graph format.

Cryoprotection Procedure

Crystals used for photodeprotection were prepared as described above. After 2-3 weeks of growth, crystals were transferred to a cryoprotectant containing 100 mM sodium cacodylate pH 6.0, 200 mM $CaCl_2$, 13% PEG 8000, 50 mM β-ME, and 25% glycerol; the reducing agent was included in the cryoprotectant buffer to thwart damage to the protein/DNA by the liberated *o*-nitrostyrene. Crystals were mounted on the translator, irradiated with 373.5 nm laser light for 30 seconds, and then quickly plunged into liquid nitrogen in preparation for data collection.

Data Collection and Refinement

X-ray data for the unflashed species (PCC) were collected at 100K at the X25 beamline at the National Synchrotron Light Source. Data for the photodeprotected (FC) species were collected at 100K at the 19-ID beamline at the Structural Biology Center of Advanced Photon Source of Argonne National Laboratory. All data were processed using HKL2000 program suite. Data collection and refinement statistics are summarized in Table 1. For each structure, the coordinates of the crosslinked lesion-recognition complex were used as the initial search model for refinement using CNS. Residues in the active site and those mediating DNA contacts were omitted from the initial search model. A partial model was generated by rigid body fit, energy minimization and simulated annealing. A σA-weighted F_o - F_c map clearly demonstrated electron density for the omitted residues. Iterative rounds of energy minimization, simulated annealing, and grouped B-factor refinement in CNS were performed while R_{free} was monitored. Subsequent rounds of energy minimization, simulated annealing, and individual B-factor refinement were performed. Simulated-annealing omits maps were used to reduce model bias. Manual adjustments to the model were made using Quanta2000 (Accelrys). Water molecules were added to the model using automated (CNS) and manual inspection of the difference maps. Density for some amino acid side-chains was occasionally incomplete; in these cases, only the ordered portion of the side chain was built in to the model. Renderings of the structures were generated using Pymol v. 0.99.

Table 1 - Data collection and Refinement Statistics

Data collection statistics	**PCC**	**FC**
Beamline	X25 of NSLS	19-ID of APS
Space group	$P6_522$	$P6_522$
Unit cell dimensions		
a, *b*, *c* (Å)	91.8, 91.8, 210.8	92.1, 92.1, 211.5
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)[a]	50-2.33 (2.41-2.33)	50-2.78 (2.88-2.78)
Unique reflections [a]	23,327 (2,057)	14,055 (1,361)
Redundancy[a]	10.9 (7.6)	11.3 (11.2)
Completeness (%)[a]	98.5 (90.6)	99.9 (100.0)
R_{merge} [a,b]	0.068 (0.464)	0.076 (0.492)
$< I / \sigma >$[a]	32.2 (2.6)	28.8 (4.8)
Refinement and model statistics		
Limiting resolution (Å)	2.33	2.78
R_{cryst} (%)[a,c]	23.6 (37.1)	22.0 (34.4)
R_{free} (%)[a,d]	27.8 (37.6)	26.7 (39.3)
Mean *B*-value, all atoms (Å^2)[e]	58.6	69.0
R.m.s. deviations from ideality[e]		
Bond lengths (Å)	0.007	0.007
Bond angles (°)	1.22	1.24
Dihedral angles (°)	21.5	21.7
Ramachandran plot (%)[e]		
Most favored	87.8	84.1
Additionally allowed	11.4	15.5
Generously allowed	0.7	0.4
Non-water atoms	2,428	2,787
Water molecules	71	33

[a] Values in parentheses refer to the highest resolution bin.

[b] $R_{merge} = \Sigma_{hkl} |I(hkl) - <I(hkl)>| / \Sigma_{hkl} <I(hkl)>$.

[c] $R_{cryst} = \Sigma_{hkl} |F_o(hkl) - F_c(hkl)| / \Sigma_{hkl} | F_o(hkl)$.

[d] R_{free} was calculated based on 7.5% of the total data randomly selected and omitted during structure refinement.

[e] Values calculated using PROCHECK

Structure of the PCC complex

The structure of photocaged complex (PCC), solved to 2.33 Å (Table 1), shows no density for the caged oxoG residue in the electron density map. Superposition of the PCC with the G-complex (Figure 2A) reveals that the active site residues demonstrate high correspondence (Cα r.m.s.d. of 0.150 Å, Figure 2C). In both structures, the Lys249 side chain angles toward the DNA backbone. The DNA on the 3' side of the caged oxoG residue occupies a similar, albeit slightly altered, path compared to DNA 3' of G. The modest disparity in the position of the left flank is localized to the stretch immediately 3' of caged oxoG (Figure 2C). Likewise, the DNA on the 5' side also sits in a slightly different space compared to the G-complex. Nevertheless, the extensive correspondence between the two complexes implies that the caged residue likely resides in the same exo-site as does G.

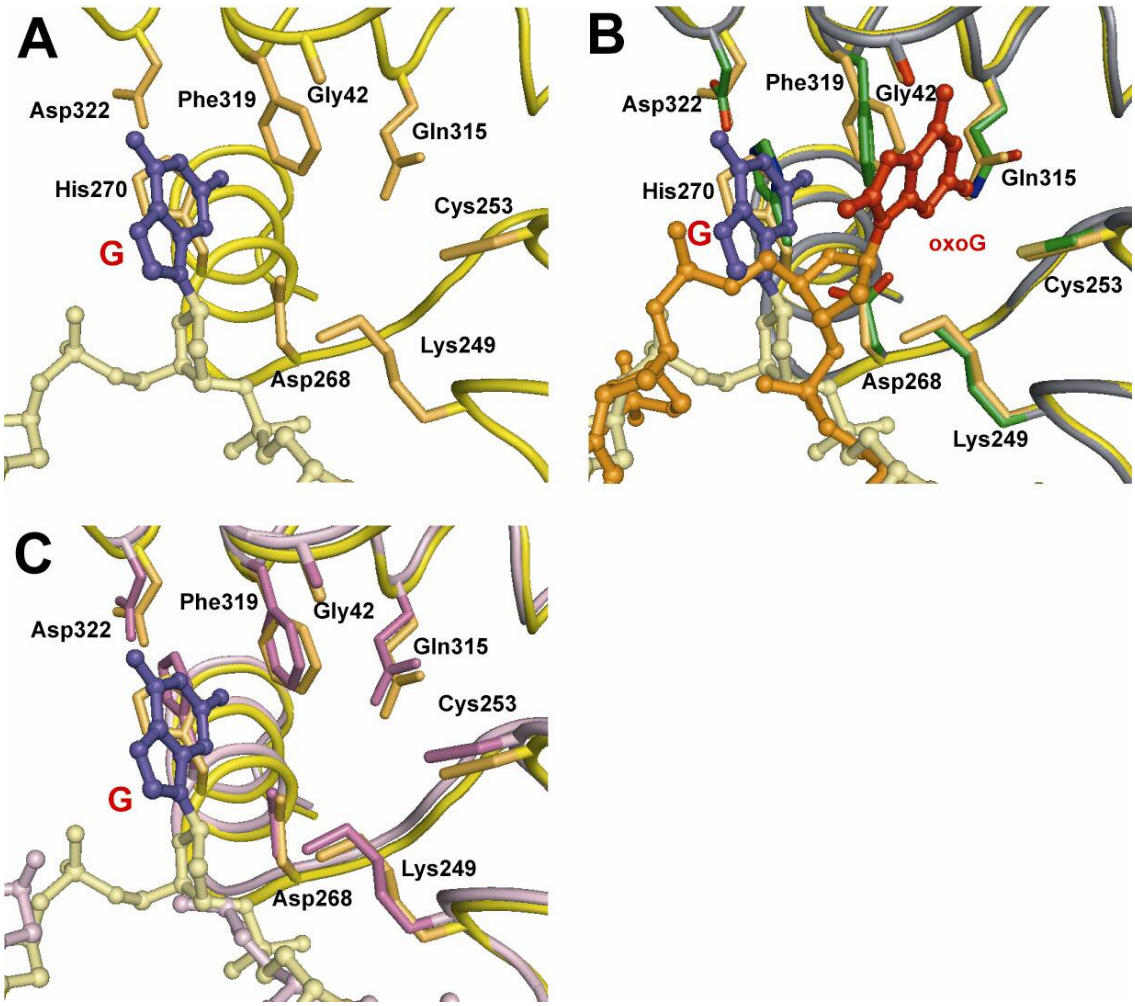


Figure 2. (A) Active site view of hOgg1/G interrogation complex (PDB code 1YQK), with the protein backbone in yellow, side-chains in light orange, G in blue, and DNA in pale yellow. (B) Heavy atom superposition of the G-interrogation complex (color scheme as in A) with the FC (protein backbone in dark gray; side-chains, dark green; oxoG, red; DNA, gold) (C) Heavy atom superposition of the G-interrogation complex (color scheme as in A) with the PCC (protein backbone and DNA in light pink; side-chains, pink).

Fit of the model to the electron density map of the FC complex

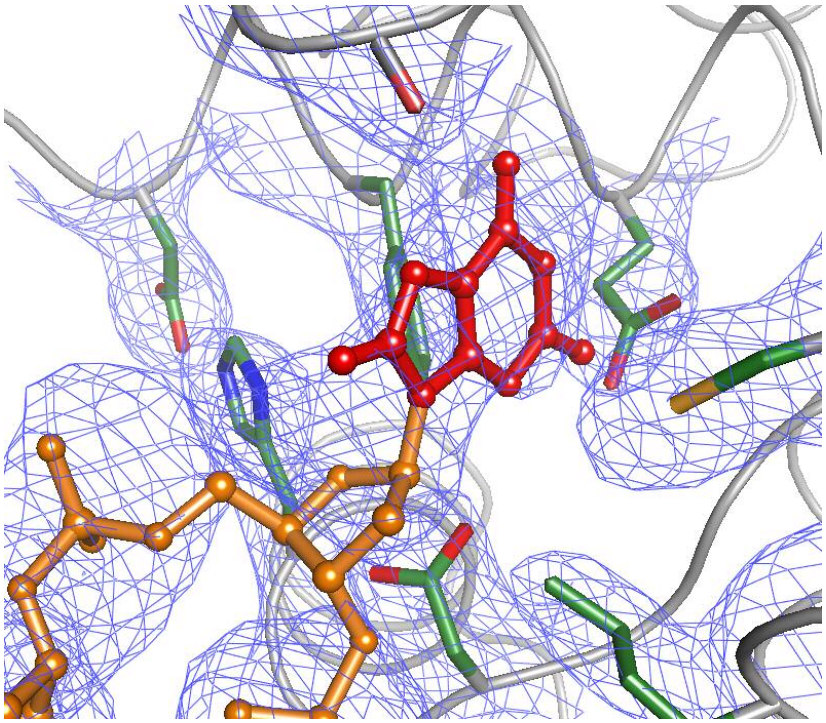
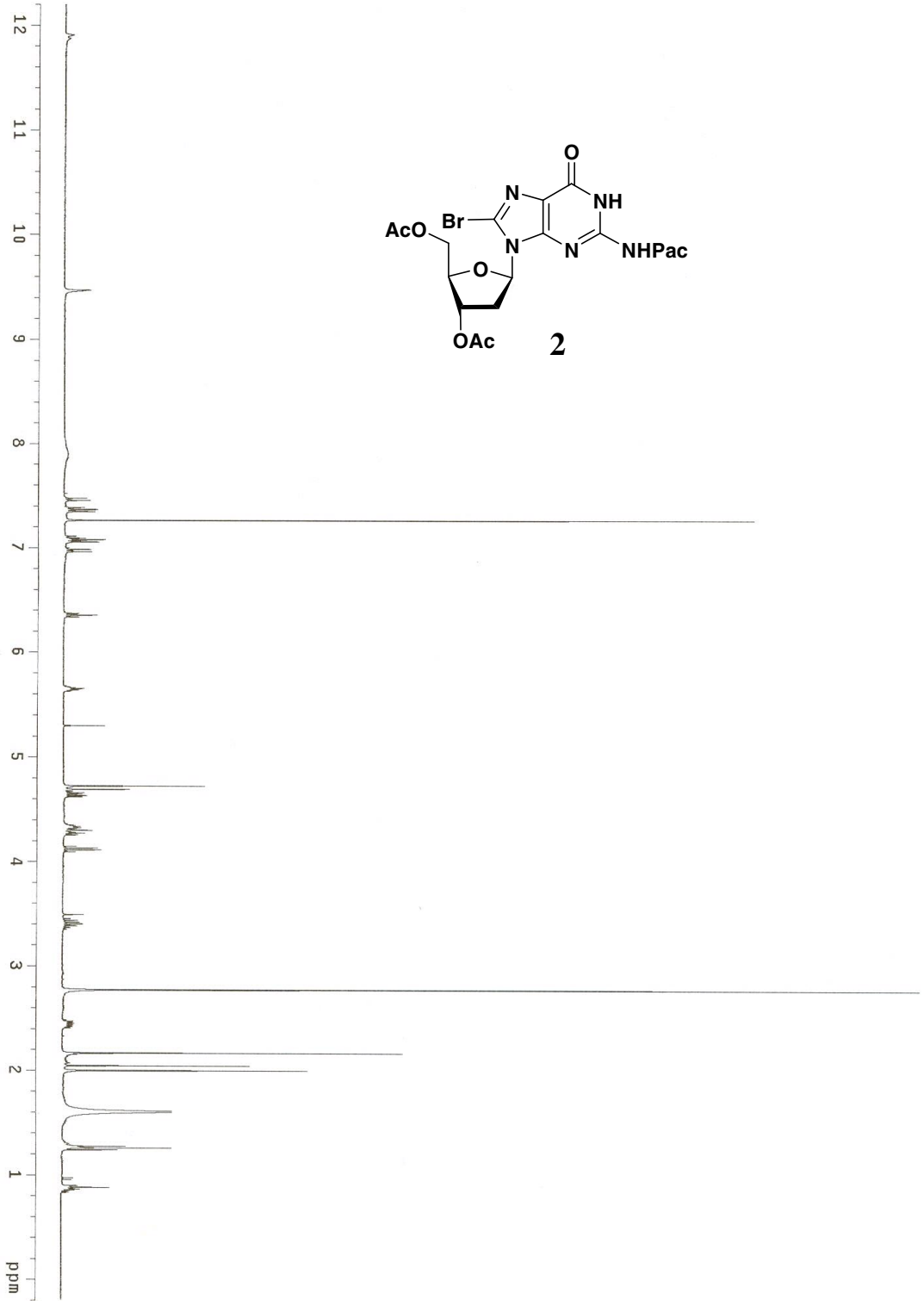


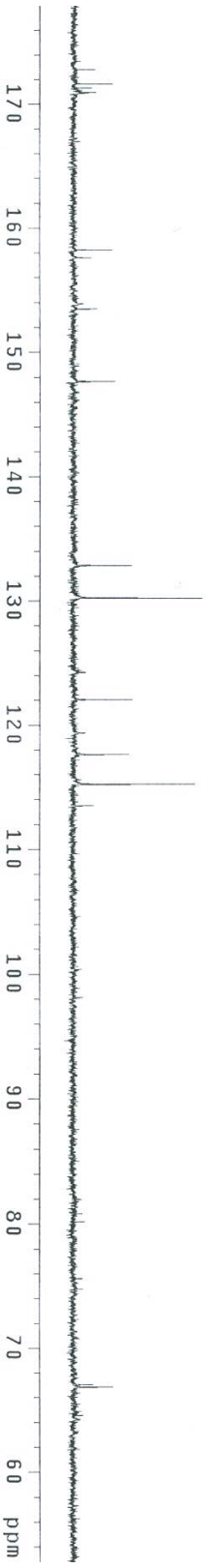
Figure 3. Active site view showing a σA-weighted 2*F*o - *F*c electron density map contoured at 0.75 σ. with protein backbone in gray, side chains in green, oxoG in red, and DNA in gold.

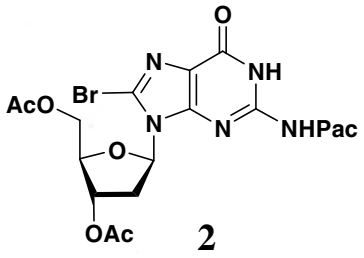
References

1. Bhusan, K. R.; DeLisi, C.; Laursen, R. A. *Tetrahedron Lett*. **2003,** *44*, 8585.
2. Norman, D. P.; Chung, S. J.; Verdine, G. L. *Biochemistry* **2003,** *42*, 1564.
3. Bruner, S. D.; Norman, D. P.; Verdine, G. L. *Nature* **2000,** *403*, 859.

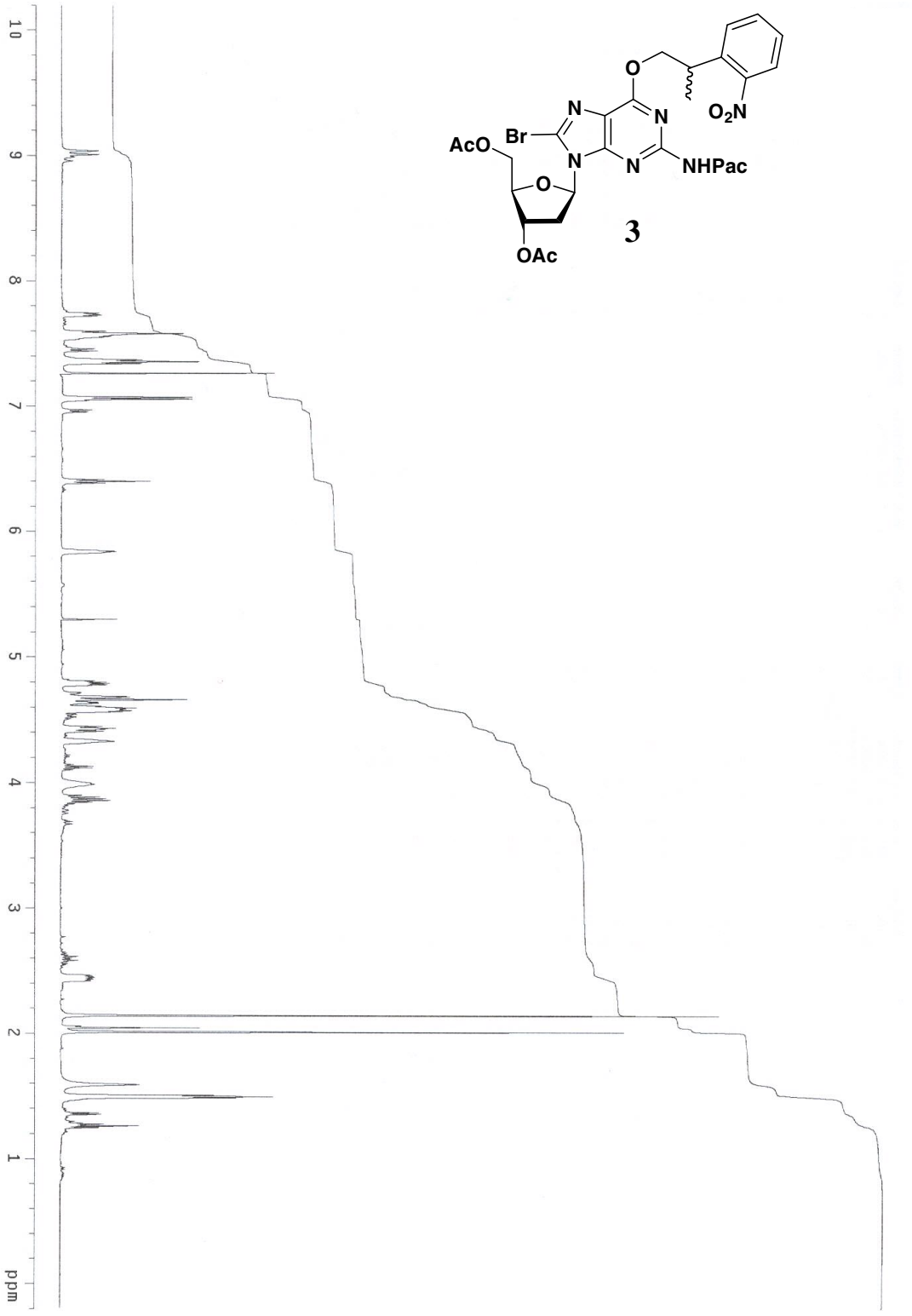


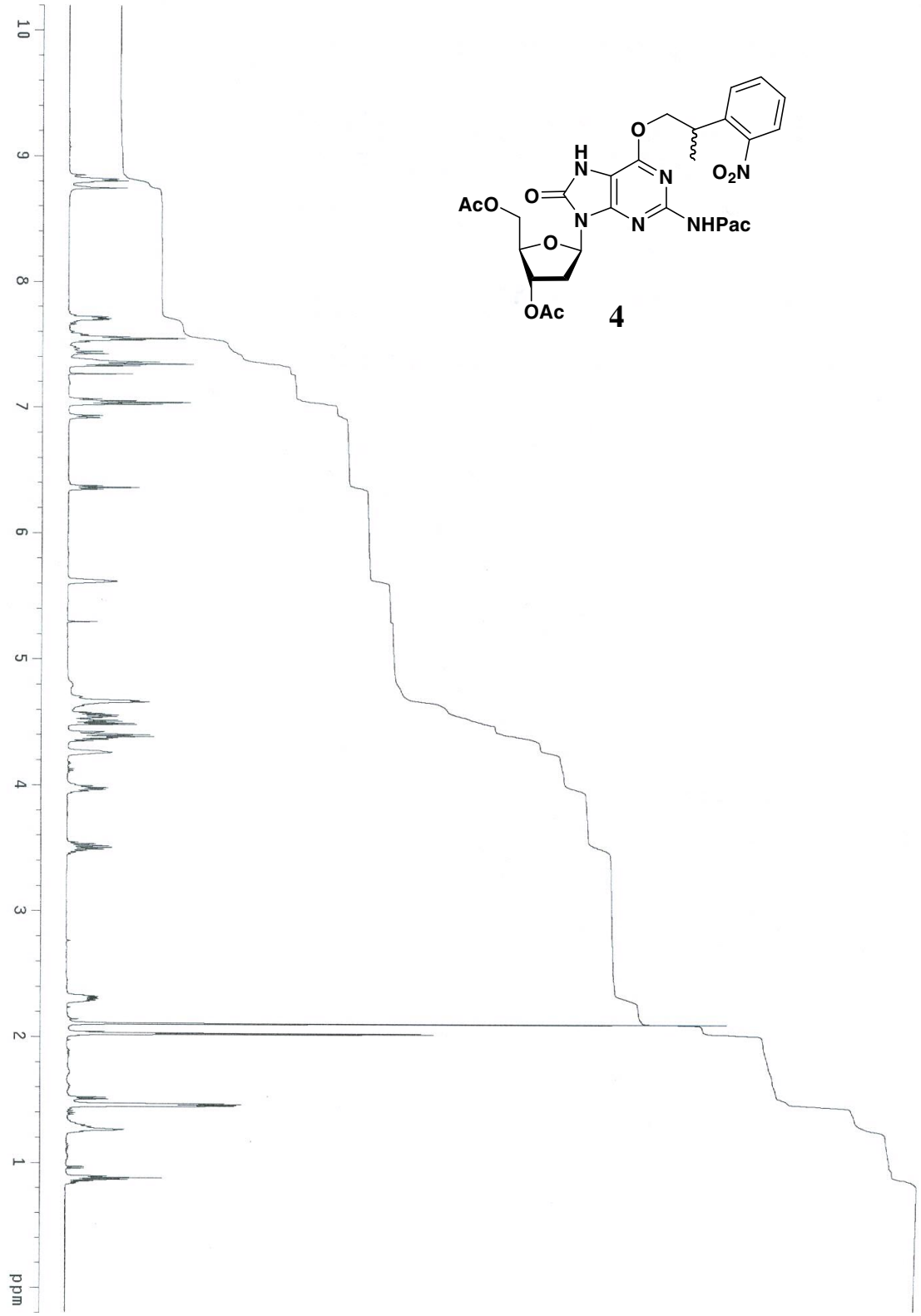
2


170
160
150
140
130
120
110
100
90
80
70
60
ppm

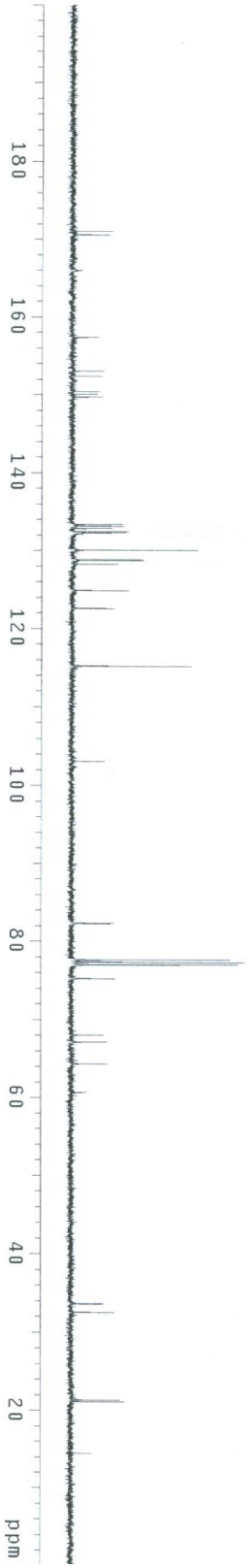

O
N
NH
Br
AcO
N
N
NHPac
O
OAc

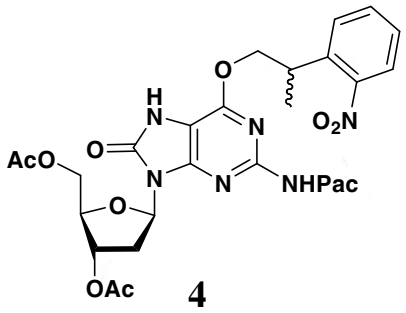
2


10
9
8
7
6
5
4
3
2
1
ppm
O
N
N
O2N
AcO
Br
N
N
NHPac
O
OAc
3

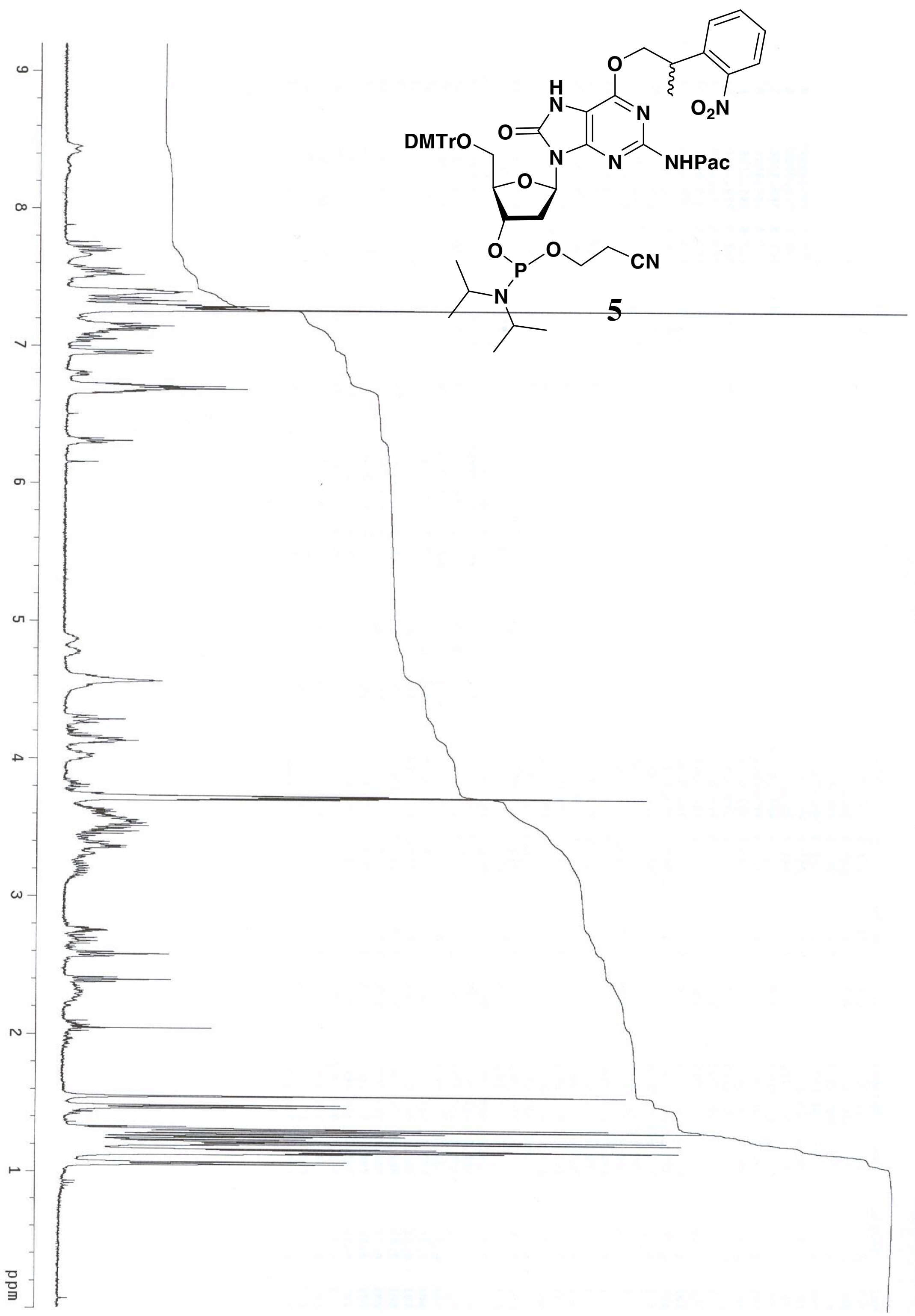


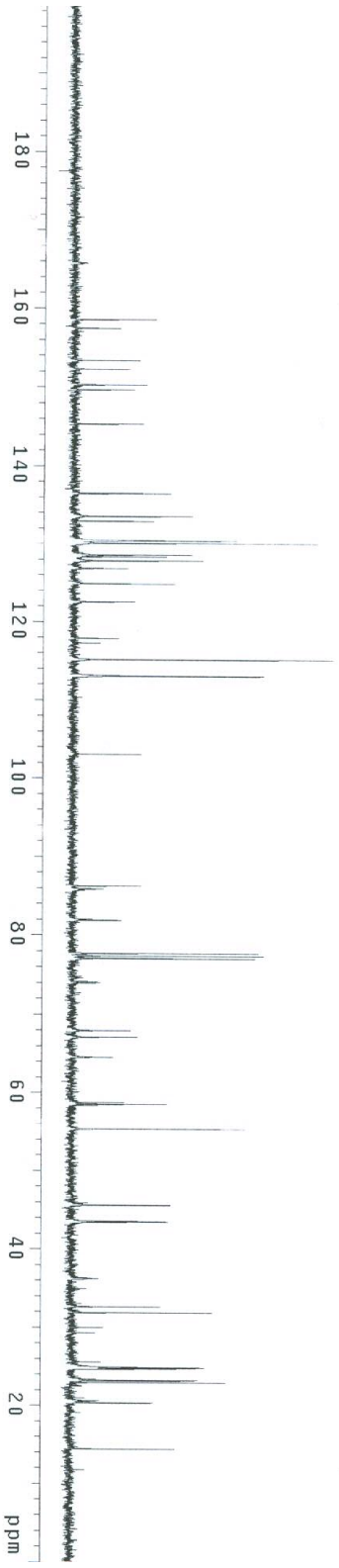
4

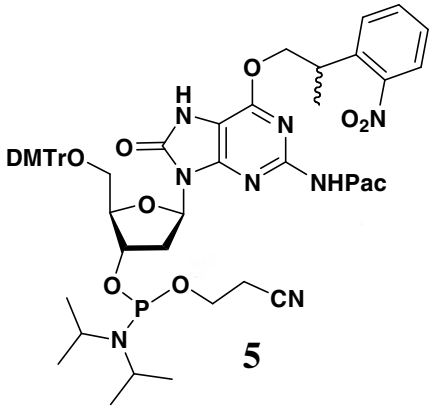




4


DMTrO
NHPac
O_2N
CN
5
9
8
7
6
5
4
3
2
1
ppm


180
160
140
120
100
80
60
40
20
ppm


O
H
N
N
O2N
O
DMTrO
N
N
NHPac
O
O
O
P
O
CN
N

5