Supplementary Information

Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair

Aaron M. Fleming,* Yun Ding, and Cynthia J. Burrows*

*correspondence to: burrows@chem.utah.edu or afleming@chem.utah.edu

Table of Contents

Item	Page
Materials and Methods	S2
Supplementary Figure 1. Schematic depicting how site specific modifications were synthesized into the plasmids.	S6
Supplementary Figure 2. Characterization of the G-quadruplex negative sequence yet SP1 positive.	S7
Supplementary Figure 3. Characterization of the G-quadruplex negative SP1 negative sequence and results of the transfection experiment into glioblastoma cells.	S8
Supplementary Figure 4. Expression of a plasmid containing OG at position 12 in supercoiled vs. relaxed plamids.	S9
Supplementary Figure 5. CD spectra for the 5-track *VEGF* sequence with an abasic site analog F.	S10
Supplementary Figure 6. Transfection results G-quadruplex negative SP1 positive sequence with the 5`-MeO-PS-F analog.	S11
Supplementary Figure 7. Characterization of the *NTHL1* G-quadruplex.	S12
Supplementary Figure 8. Sanger sequencing chromatogram to demonstrate insertion of a site-specific modification into the reporter plasmid.	S13
Supplementary References	S14

Materials and Methods

Plasmid Construction

The plasmids were constructed from the psiCHECK2 plasmid (Promega) that has coding sequences for the Renilla luciferase (Rluc) and firefly luciferase (luc) genes. The luc gene is regulated by the HSV-TK promoter that was not modified and used as the internal standard to conduct the Dual-Glo luciferase assay (Promega). The Rluc gene was originally regulated by the SV40 early enhancer/promoter, which we modified by removing the TATA-box and replacing it with the potential G-quadruplex sequences (PQS) of interest. Additionally, the PQSs were flanked by recognition sequences for the Nt.BspQ1 nicking endonuclease. Insertion of the PQS and nicking endonuclease recognition sequences was achieved using restriction free cloning. The cloning experiment was conducted in a 25-µL reaction with 1x Phusion reaction buffer (NEB), 200 µM dNTPs, 1 µM each of the forward and reverse primers, 1,000 ng of psiCHECK2 plasmid, 3% DMSO, and 0.5 units of Phusion DNA polymerase. The PCR was initiated with a 98 °C denaturation step for 30 sec followed by 30 cycles of PCR consisting of a 98 °C denaturation step for 20 sec, a 55 °C annealing step for 30 sec, a 72 °C extension step for 5 min, and a post step consisting of 98 °C for 8 sec followed by 72 °C for 5 min. After the 30 cycles, a final extension step of 72 °C for 5 min was conducted. Following PCR, the samples were digested with Dpn1 by adding 2.3 µL of Dpn1 reaction buffer and 5 units of Dpn1 to the PCR reaction vessel. The Dpn1 reaction was conducted for 2 hr at 37 °C followed by heat quenching at 80 °C for 20 min.

The PCR and Dpn1 digested reaction was then transformed using NEB 5-alpha competent *E. coli* following the manufacturer's protocol. After transformation, 50-100 µL from the transformation was spread on a pre-warmed agar plate containing ampicillin (100 µg/mL). The plates were incubated overnight at 37 °C. Next, individual colonies were picked and each grown overnight at 37 °C in 5 mL of lysogeny broth media containing ampicillin (100 µg/mL). The plasmid DNA from each grown sample was then purified using a miniprep kit (Qiagen) following the manufacturer's protocols. Last, ~200 ng of plasmid and 1 µM sequencing forward primer in 15 µL of 10 mM Tris (pH 8.0) was submitted for Sanger sequencing at the DNA sequencing core at the University of Utah.

Insertion of site-specific modifications into the plasmids was achieved following literature protocols(1, 2). Specifically, 5 µg of plasmid with Nt.BspQ1 recognition sequences flanking the site in which the DNA base modification will be inserted was placed in 50 µL of 1x Nt.BspQ1 reaction buffer with 5 units of Nt.BspQ1. The reaction was placed at 50 °C for 60 min followed by heat quenching at 80 °C for 20 min. Next, 1 nmole (~1000x) of 5`-phosphorylated oligomer with the site-specific modification was added to the quenched reaction. The modified oligomers were made by solid-phase synthesis in the DNA/Peptide core facility at the University of Utah. They were cleaved, deprotected, and HPLC purified following protocols previously outlined(3). The plasmid oligomer mixture was thermal cycled by heating at 80 °C for 2 min followed by cooling on ice for 2 min, which was repeated four times. The nicks were then sealed using T4 DNA ligase, by adding 6 µL of 10x ligase buffer and 800 units of ligase. The ligation reaction was left at 20 °C for 4 h. To induce supercoiling into the plasmids, they were treated with 5 units of gyrase for 1 h (see Supplementary Fig. 1 for a schematic of this method). The modified, supercoiled plasmids were purified from the protein and excess insert using an Ultra Clean PCR cleanup kit (Mo Bio) following the manufacturer's protocol. Plasmid concentrations were determined by nanodrop UV-vis measurements.

To confirm the DNA modifications were introduced into the plasmid, we applied a protocol established in our laboratory, in which the modification was removed by a DNA glycosylase to yield a ligatable gap (4). By ligating the gap, Sanger sequencing provided a characteristic nucleotide loss at the modification site to confirm the presence of the modification. Complete details of the method can be found in our recent publication(4).

Primer Sequences

Insertion of 5 track *VEGF* PQS
5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACG CTC TTC CGG GGC GGG CCG GGG GCG GGG TCC CGG CGG GGC GCT CTT CTG CAG AGG CCG AGG CCG CCT CGG CCT CTG AGC T

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCA GAA GAG CGC CCC GCC GGG ACC CCG CCC CCG GCC CGC CCC GGA AGA GCG TCA GCC ATG GGG CGG AGA ATG GGC GGA ACT G

Insertion of 4 track *VEGF* PQS
5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACG CTC TTC CGG GGC GGG CCG GGG GCG GGG TGC TCT TCT GCA GAG GCC GAG GCC GCC TCG GCC TCT GAG CT

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCA GAA GAG CAC CCC GCC CCC GGC CCG CCC CGG AAG AGC GTC AGC CAT GGG GCG GAG AAT GGG CGG AAC TG

Insertion of modified PQS negative *VEGF*
5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACG CTC TTC CTG GGC GGA CCT TGG GCG GAA TCC CGG CTG GGC GCT CTT CTG CAG AGG CCG AGG CCG CCT CGG CCT CTG AGC T

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCA GAA GAG CGC CCA GCC GGG ATT CCG CCC AAG GTC CGC CCA GGA AGA GCG TCA GCC ATG GGG CGG AGA ATG GGC GGA ACT G

Insertion of *NTHL1* PQS
5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACG CTC TTC TCG GGT TGC AGT GGG CGC GGG TGA GGG CCC GGG ACG CTC TTC TGC AGA GGC CGA GGC CGC CTC GGC CTC TGA GCT

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCA GAA GAG CGT CCC GGG CCC TCA CCC GCG CCC ACT GCA ACC CGA GAA GAG CGT CAG CCA TGG GGC GGA GAA TGG GCG GAA CTG

Sequencing Primer
5`- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTC ATG CAT CTC AAT TAG TCA GCA ACC ATA GT

Cell Culture Studies

All cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 20 µg/mL gentamicin, 1x glutamax, and 1x non-essential amino acids. The cells were grown at 37 °C with 5% CO_2 at ~80% relative humidity and were split when they reached 70-80% confluence. The OGG1$^{-/-}$ MEF cells were previously developed (5) and deposited in the Ximbio sharing program. The wild type glioblastoma cells (U-87 MG) were purchased from ATCC.

The transfection experiments were conducted in white, 96-well plates by seeding 2 x 10^4 cells per well and then allowing them to grow for 24 h. After 24 h, the cells were transfected with 200-400 ng of plasmid using X-tremeGene HP DNA transfection agent (Roch) following the manufacturer's protocol in Opti-MEM media. All transfection experiments were conducted at least 4 times. Next, 48 h post transfection, the Dual-Glo luciferase (Promega) assay was conducted following the manufacturer's protocol.

The APE1 inhibitor studies were conducted similarly to all studies with the addition of 0.1–1 µM APE1 inhibitor from a DMSO stock solution. Controls in which only DMSO was added along with the modified plasmids were conducted to insure the data obtained resulted from the inhibitor and not the DMSO.

To conduct the siRNA knock down of APE1, FlexiTube siRNAs (Qiagen) were transfected into the glioblastoma cells with x-tremeGene HP DNA transfection agent (Roche) at 1 nM, 5 nM and 50 nM concentration along with the plasmid of interest.

The data were analyzed by converting the luminescence measured into normalized relative response ratios (RRR), which is the luminescence of Rluc divided by the luminescence of luc (i.e., RRR = Rluc/luc). To obtain the normalized expression values reported, each RRR was divided by the RRR for the wild type sequence in that data set, for example, normalized expression = RRR_{OG12}/RRR_{WT}. The error bars represent 95% confidence intervals obtained from the data. A two-tailed Student's t-test was conducted on the data to determine p-values.

Characterization of G-Quadruplexes

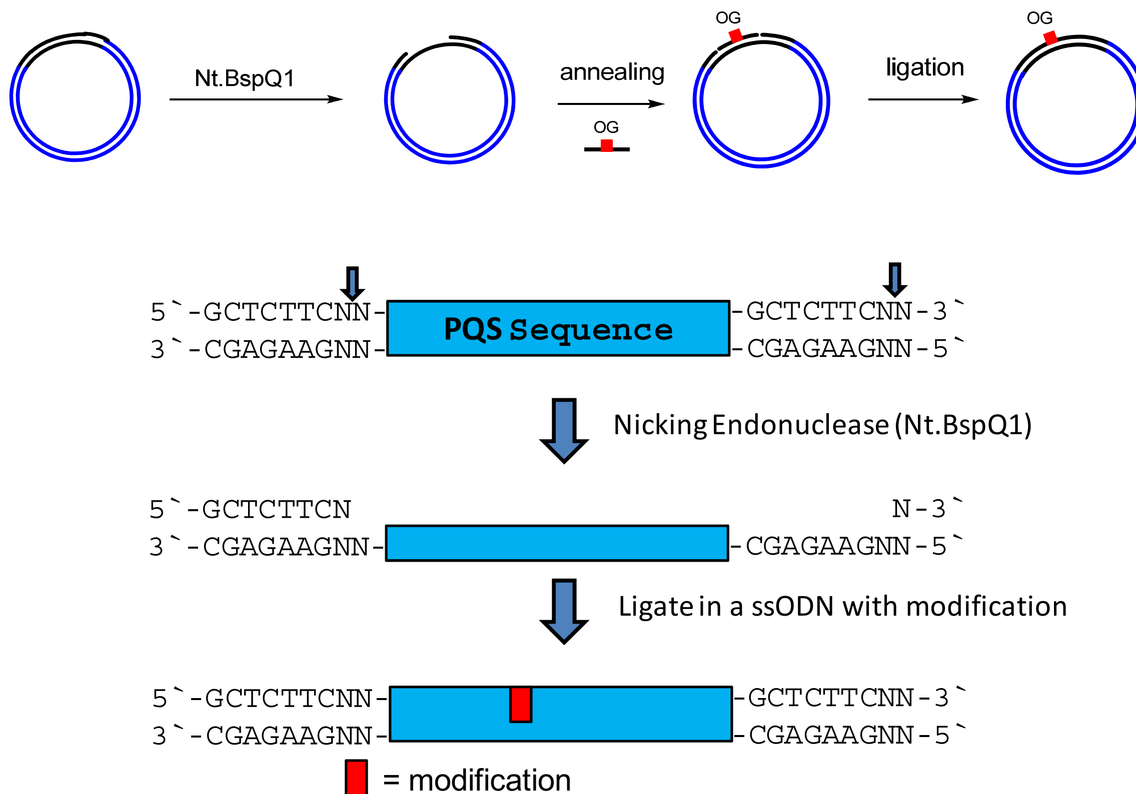
Oxidation Reaction Conditions. 10 µM *NTHL1*-4 at 37 °C 20 mM cacodylic acid (pH 7.4) + 140 mM KCl +12 mM NaCl + 200 mM $KHCO_3$ + 20 mM $K_2S_2O_8$ and 254 nm light for 20 min to produce $CO_3^{\bullet-}$ when NAC was present; 3 mM SIN-1 and 25 mM $KHCO_3$ for reactions without NAC; or, 50 µM riboflavin + 350 nm light for 10 min. Sites of reaction were determined by hot piperidine cleavage for 2 h, which revealed <20% reactivity. These piperidine conditions should uncover ~75% of the Sp reaction sites, ~100% of the Gh sites, and the OG sites were determined by further oxidation with Na_2IrCl_6 before piperidine treatment(6).

CD analysis. The G-quadruplex samples were annealed at 10 µM concentrations in 20 mM lithium cacodylate buffer (pH 7.4) with 140 mM KCl and 12 mM NaCl. The samples were placed in a 0.2 cm quartz cuvette for CD analysis at 20 °C. The recorded data were background subtracted and then normalized on the y-axis to units of molar ellipticity for plotting and comparative purposes.

Thermal melting analysis. The UV melting studies were conducted on a Shimadzu UV-vis spectrometer running their software. The thermal melting values were determined on samples of 1 µM quadruplex in buffered solutions with physiological salt concentrations (20 mM

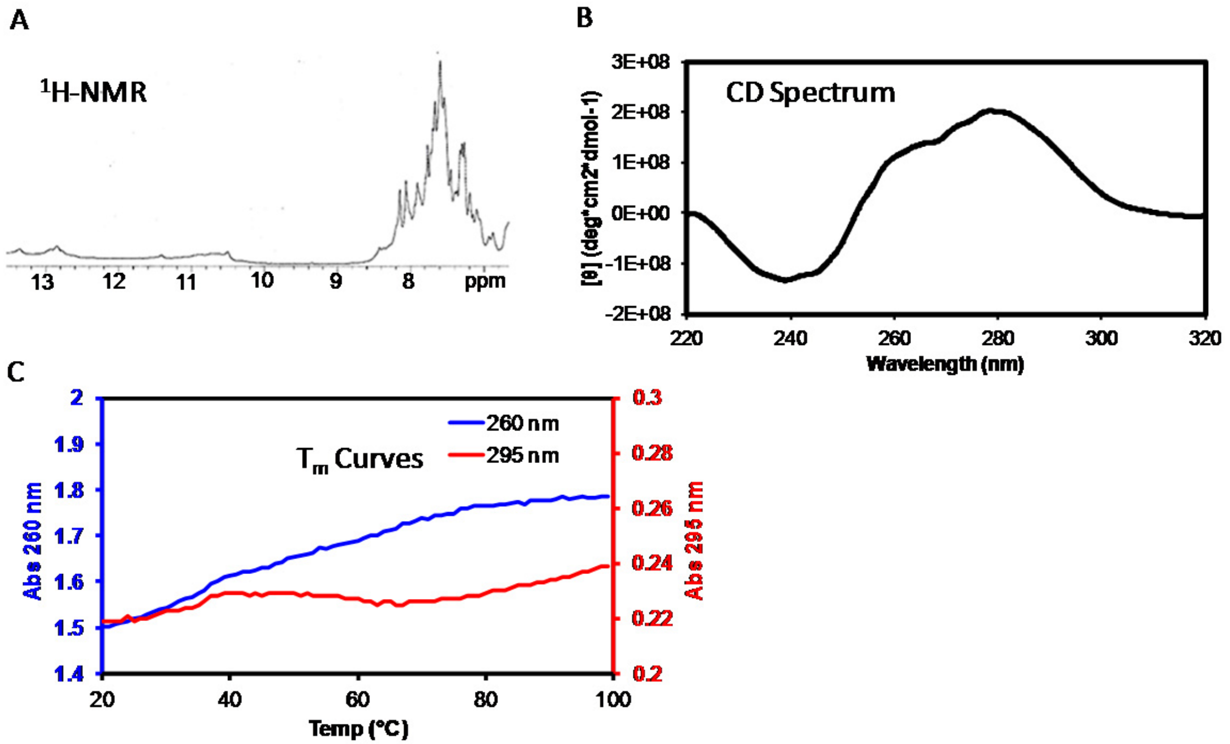
cacodylate pH 7.4, 140 mM KCl, and 12 mM NaCl). The melting experiments were initiated by thermally equilibrating the samples at 20 °C for 10 min followed by heating the samples at 0.5 °C/min and equilibrating at each 1 °C increment for 1 min. A reading at 260 and 295 nm was taken after each 1 °C change in the temperature. The heating phase involved increasing the temperature to 95 °C followed by a 10 min hold and then a return to 20 °C via the same change in temperature and increment time while making absorbance readings at each 1 °C change. Plots of absorbance at 295 nm vs. temperature were constructed, and the thermal melting value was determined from the first derivative of the curve, as well as by a two-point analysis protocol using the instrument's software.

^{1}H-NMR analysis. The G-quadruplex samples were annealed in 300 μL at a 300 μM concentration in 20 mM KP_i (pH 7.0) and 50 mM KCl in 9:1 H_2O:D_2O. The annealed samples were placed in a D_2O-matched Shigemi NMR tube. The samples were analyzed on an 800-MHz NMR spectrometer with the temperature set to 24 °C. Each sample was scanned 1024 times using the Watergate solvent suppression pulse sequence.

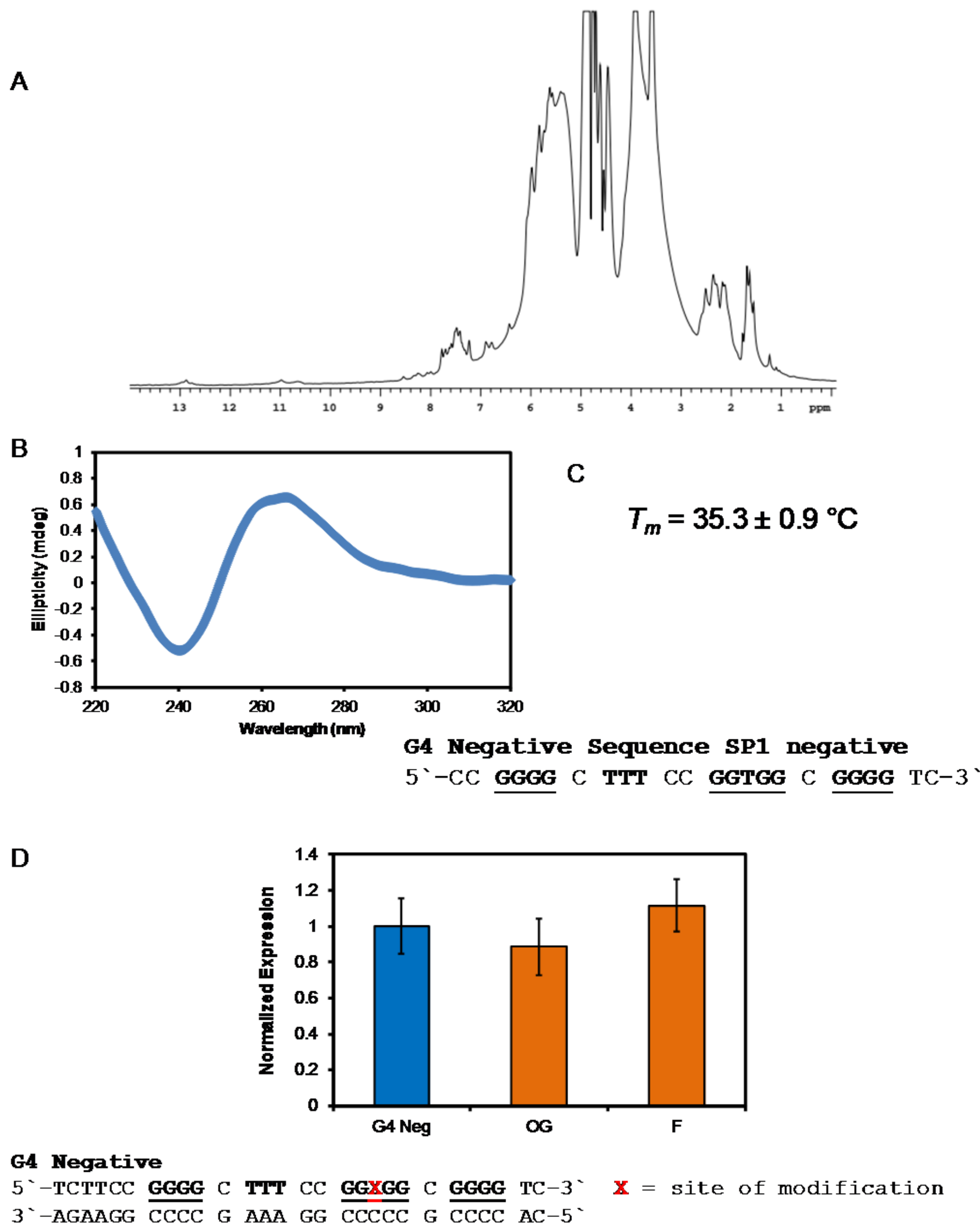


Supplementary Figure 1. Schematic depicting how site specific modifications were synthesized into the plasmids.

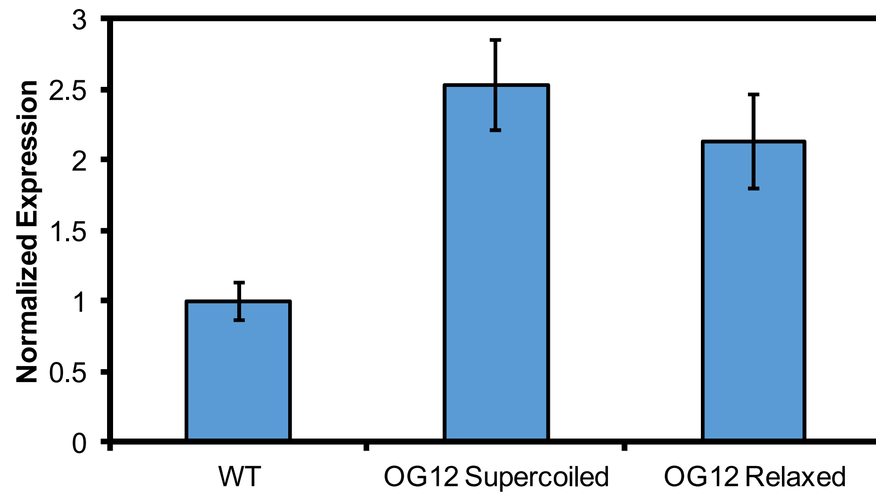
The 5 track *VEGF* sequence inserted = 5`-pGGG GCG GGC CGG GGG CGG GGT CCC GGC GGG GCG CTC TTC T; the 4 track *VEGF* sequence inserted = 5`-pGGG GCG GGC CGG GGG CGG GGT GCT CTT CT; the G-quadruplex negative sequence = 5`-pTG GGC GGA CCG TGG GCG GAA TCC CGG CTG GGC GCT CTT CT; and the 5 track *NTHL1* sequence inserted = 5`-pTG GGC GCG GGT GAG GGC CCG GGA CGC TCT TCT. The underlined Gs are the sites in which modifications were introduced and the p is a 5` phosphate required for ligation.



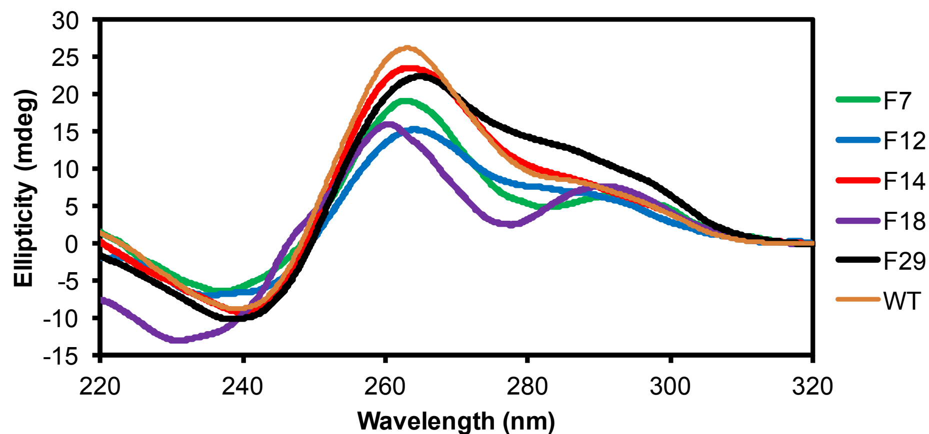
Supplementary Figure 2. Characterization of the G-quadruplex negative sequence. The G-quadruplex negative sequence was 5`-TG GGC GGA CCT TGG GCG GAA TCC CGG CTG GGC GCT CTT CT. (A) Analysis of this sequence by ^{1}H-NMR in 20 mM KP_i (pH 7.0) and 50 mM KCl at 24 °C did not provide characteristic imino proton peaks between 9.5–11.0 ppm indicating G•G Hoogsteen base pairs that should be found in G-tetrads (7). (B) The circular dichroism (CD) spectrum for the G-quadruplex negative strand produced a λ_{max} at ~280 nm indicative of a single-stranded or hairpin-like structure at 20 °C (8). (C) The thermal melting curve for this sequence monitored at 260 (duplex or hairpin melting transitions) and 295 nm (G-quadruplex melting transitions) found a weak transition observed at 260 nm (~35 °C) and no transition at 295 nm(9). The CD and thermal melting experiments were conducted in 20 mM lithium cacodylate buffer (pH 7.4) with 140 mM KCl and 12 mM NaCl. On the basis of these three experiments, the G-quadruplex negative sequence was indeed G-quadruplex negative but may adopt a low stability hairpin.



Supplementary Figure 3. Characterization of the G-quadruplex negative SP1 negative sequence and results of the transfection experiment into glioblastoma cells. (A) ^{1}H-NMR spectrum illustrating that G:G Hoogsteen imino protons (10–11 ppm) indicating G-quadruplex formation were not observed. (B) CD spectrum for the sequence providing λ_{max} at ~268 nm supporting a poorly defined non-G-quadruplex structure. (C) The T_m for this sequence was below the cellular experiments conducted. (D) Results of transfecting this sequence into glioblastoma cells after 48 h to demonstrate without a G-quadruplex OG and F do not impact transcription relative to the WT sequence.

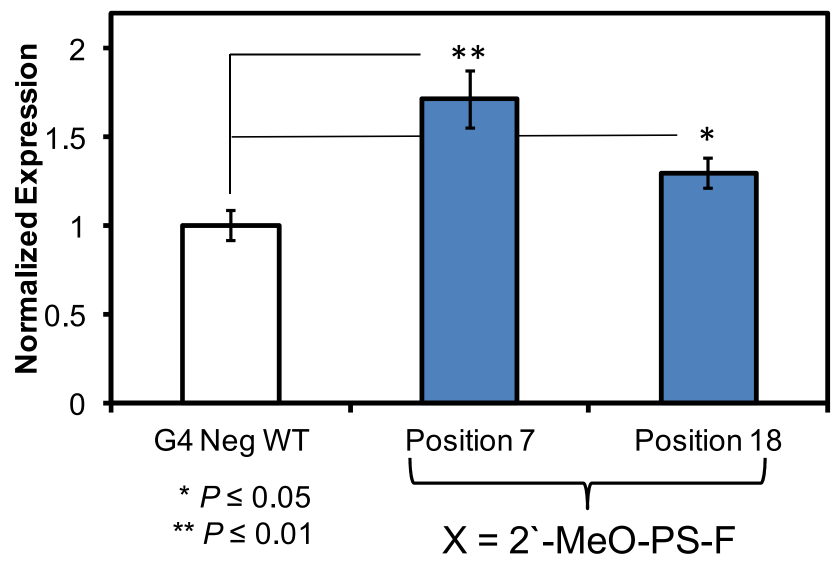


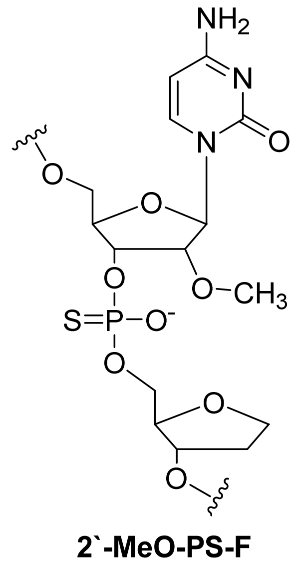
Supplementary Figure 4. Expression of a plasmid containing OG at position 12 in supercoiled vs. relaxed plamids.



Supplementary Figure 5. CD spectra for the 5-track *VEGF* sequence with an abasic site analog F.

The CD spectra were recorded in 20 mM lithium cacodylate buffer (pH 7.4) with 140 mM KCl and 12 mM NaCl at 20 °C at 10 μM DNA. The 5 track studies were conducted in the sequence 5`-CCGGGGC**G**^{7}GGCC**G**^{12}G**G**^{14}GGC**G**^{18}GGGTCCCGGCG**G**^{29}GGCG. These data support G-quadruplex formation via extrusion of the damaged G-track when the 5th G run is present. On the basis of these spectra, the G-quadruplexes are parallel folded when there is a single maximum at 260 nm and a mixture of folds when there are two maxima at 260 and 295 nm(8).

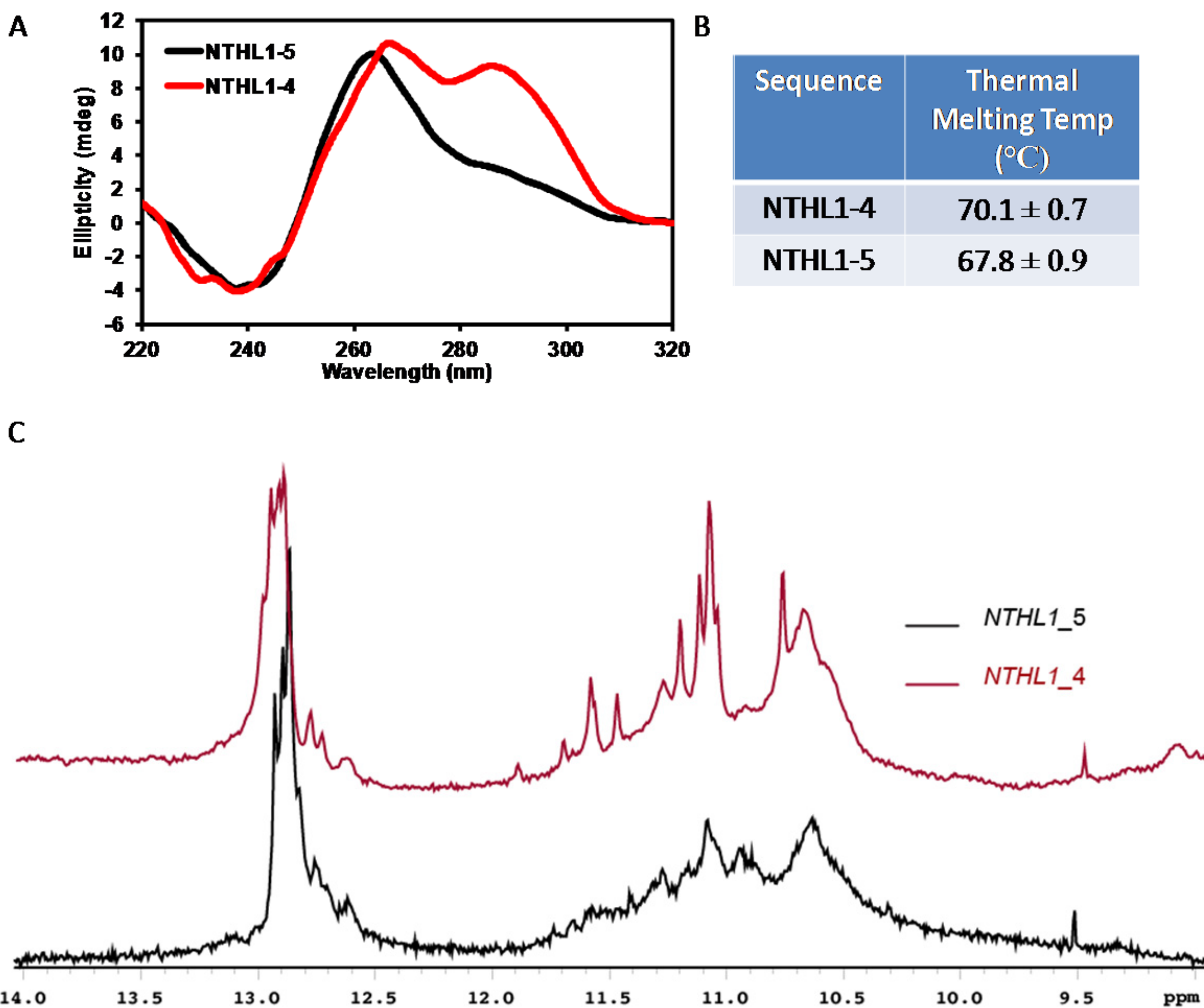




G-quadruplex negative SP1+

5`-CCC **TGGG** C **X7GA** CC **TTGGG** C **X18GAA** TCCGGC **TGGG** CGGAG-3`

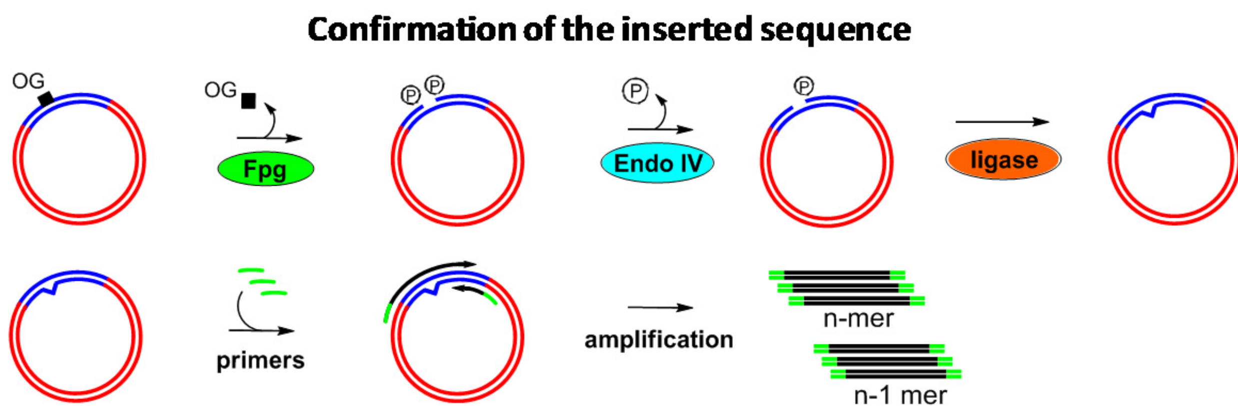
Supplementary Figure 6. Transfection results G-quadruplex negative SP1 positive sequence with the 5`-MeO-PS-F analog.

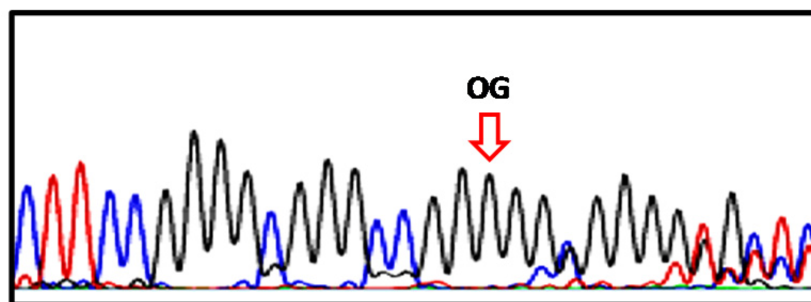


Sequence	Thermal Melting Temp (°C)
NTHL1-4	70.1 ± 0.7
NTHL1-5	67.8 ± 0.9

Supplementary Figure 7. Characterization of the *NTHL1* G-quadruplex.

(A) CD spectra for the 4-track *NTHL1* sequence (5`-GTG GGC GCG GGT GAG GGC CCG GGA C) and the 5-track *NTHL1* sequence (5`-TCG GGT TGC AGT GGG CGC GGG TGA GGG CCC GGG AC) recorded at 20 °C in 20 mM lithium cacodylate buffer (pH 7.4) with 140 mM KCl and 12 mM NaCl. (B) Thermal melting temperatures for both sequences. (C) ^{1}H-NMR spectra recorded for the two sequences in 20 mM KP_i (pH 7.0) and 50 mM KCl at 24 °C. These preliminary data support both 4 and 5 track PQSs to adopt G-quadruplex folds.





Supplementary Figure 8. Sanger sequencing chromatogram to demonstrate insertion of a site-specific modification into the reporter plasmid. To confirm insertion of the site-specifically modified oligomer into the plasmid we utilized a method previously reported from our laboratory (4). The modification was removed with a DNA glycosylase followed by ligation of the gapped product (see top of the figure). Upon Sanger sequencing after the site of modification removal all peaks are doubled in the sequencing chromatogram. Because the modifications were inserted in G runs, the peak doubling starts at the end of the G run as illustrated above (see bottom of the figure).

Supplementary References

1. Allgayer J, Kitsera N, Bartelt S, Epe B, & Khobta A (2016) Widespread transcriptional gene inactivation initiated by a repair intermediate of 8-oxoguanine. *Nucleic Acids Res.*:doi: 10.1093/nar/gkw1473.

2. You C*, et al.* (2012) A quantitative assay for assessing the effects of DNA lesions on transcription. *Nat. Chem. Biol.* 8(10):817-822.

3. Fleming AM, Zhou J, Wallace SS, & Burrows CJ (2015) A role for the fifth G-track in G-quadruplex forming oncogene promoter sequences during oxidative stress: Do these "spare tires" have an evolved function? *ACS Cent. Sci.* 1:226-233.

4. Riedl J, Fleming AM, & Burrows CJ (2015) Sequencing of DNA lesions facilitated by site-specific excision via base excision repair DNA glycosylases yielding ligatable gaps. *J. Am. Chem. Soc.* 138(2):491-494.

5. Klungland A*, et al.* (1999) Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage. *Proc. Natl. Acad. Sci. U. S. A.* 96(23):13300-13305.

6. Fleming AM, Alshykhly O, Zhu J, Muller JG, & Burrows CJ (2015) Rates of chemical cleavage of DNA and RNA oligomers containing guanine oxidation products. *Chem. Res. Toxicol.* 28:1292-1300.

7. Adrian M, Heddi B, & Phan AT (2012) NMR spectroscopy of G-quadruplexes. *Methods* 57(1):11-24.

8. Karsisiotis AI*, et al.* (2011) Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem., Int. Ed.* 50(45):10645-10648.

9. Mergny JL, Phan AT, & Lacroix L (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett* 435(1):74-78.