Published in final edited form as:

Chem Res Toxicol. 2019 May 20; 32(5): 899–909. doi:10.1021/acs.chemrestox.9b00041.

Oxidative modification of guanine in a potential Z-DNA-forming sequence of a gene promoter impacts gene expression

Aaron M. Fleming*, Judy Zhu, Yun Ding, Selma Esders, and Cynthia J. Burrows*

315 S 1400 East, Dept. of Chemistry, University of Utah, Salt Lake City, UT 84112-0850

Abstract

One response to oxidation of guanine (G) to 8-oxo-7,8-dihydroguanine (OG) in a gene promoter is regulation of mRNA expression suggesting an epigenetic-like role for OG. A proposed mechanism involves G oxidation within a potential G-quadruplex-forming sequence (PQS) in the promoter enabling a structural shift from B-DNA to a G-quadruplex fold (G4). When OG was located in the coding vs. template strand, base excision repair led to an on/off transcriptional switch. Herein, a G-rich, potential Z-DNA-forming sequence (PZS) comprised of a d(GC)_n repeat was explored to determine whether oxidation in this motif was also a transcriptional switch. Bioinformatic analysis found 1650 PZSs of length >10 nts in the human genome that were overrepresented in promoters and 5'-UTRs. Studies in human cells transfected with a luciferase reporter plasmid in which OG was synthesized in a PZS context in the promoter found that a coding strand OG increased expression, and a template strand OG decreased expression. The initial base excision repair product of OG, an abasic site (AP), was also found to yield similar expression changes as OG. Biophysical studies on model Z-DNA strands found OG favored a shift in the equilibrium to Z-DNA from B-DNA, while an AP disrupted Z-DNA to favor a hairpin placing AP in the loop where it is a poor substrate for the endonuclease APE1. Overall, the impact of OG and AP in a PZS on gene expression was similar to that in a PQS but reduced in magnitude.

Graphical Abstract

*To whom correspondence should be addressed burrows@chem.utah.edu or afleming@chem.utah.edu.

Notes

The authors declare no competing financial interests in this work.

Introduction

Oxidative stress represents a shift in the balance between the cellular levels of reactive oxygen species (ROS) and a cell's ability to counteract the threats imposed by these species via antioxidant and repair pathways.¹ Oxidation of cellular components by ROS track with the initiation and progression of cancer, inflammation, age-related diseases, cardiovascular disease, and traumatic brain injury.²⁻⁴ On the other hand, support is growing for the concept that evolution harnessed ROS as signaling agents for cellular regulation.⁵⁻⁸ Pathways responsible for enzymatic handling of ROS,⁹ inflammatory response,^{8,10} DNA repair,¹¹ and some oncogenes¹²⁻¹⁵ may be regulated in part by ROS and their processing. Many knowledge gaps and critical questions still remain in our understanding of the chemical and biological details regarding how cells engage ROS for regulating the response to oxidative stress.

Pivotal for biology to respond to oxidative stress or inflammation is the capability to change mRNA levels and ultimately protein concentrations for key response genes. Initiation of these changes can occur on the genome via alterations in the epigenetic chemical modification status.¹⁶ Further, when these chemical marks are modulated within the vicinity of gene promoters, mRNA synthesis can be directly impacted. An additional component to gene regulation is restriction of protein regulators to promoters in heterochromatin that can be opened by the actions of chromatin remodelers allowing the transcriptional machinery to gain entry for gene regulation.¹⁷ Perillo and coworkers found that the flavin-dependent chromatin remodeler LSD1 demethylates H2K4me2 or H3K9me3 and delivers H₂O₂ at the site of reaction in a gene promoter for oxidation of a guanine (G) heterocycle to yield 8-oxo-7,8-dihydroguanine (OG; Scheme 1A).¹⁴ The introduction of OG in a gene promoter focuses the base excision repair (BER) pathway glycosylase OG-glycosylase 1 (OGG1) for removal of the oxidized heterocycle to set off a cascade of DNA repair-mediated events resulting in induction of transcription of the *BCL-2* gene. This seminal work identified chromatin remodeling-mediated oxidation of a genome focuses BER to regulatory regions for upregulation of transcription suggesting that DNA repair and transcriptional regulation are correlated.^{14,18} This report and studies conducted since then have identified many instances of G oxidation to OG in a gene promoter regulating transcription,^{12,13,15,19,20} leading to the realization that OG may be an epigenetic-like chemical modification to DNA.²¹⁻²³

Linking G oxidation in G-rich promoter elements with gene activation via BER or nucleotide excision repair (NER) has been documented in the *SIRT1*,²⁰ *VEGF*,^{13,15} *TNF- α* ,¹⁹ *KRAS*,¹² *CASP9*,²⁴ *CYP26A1*,²⁴ and many proinflammatory genes.⁸ Research in our laboratory used chemical tools to identify unequivocally that the oxidatively modified heterocycle OG in the G-rich, potential G-quadruplex forming sequence (PQS) naturally located in the coding strand of the *VEGF* promoter can induce transcription.¹³ We demonstrated BER removal of OG to form an abasic site (AP) in duplex DNA unmasks a G-quadruplex (G4) fold (Scheme 1B). As a result of the natural *VEGF*PQS possessing additional G runs beyond those needed for G4 formation, the non-canonical fold is stabilized by looping the AP out and replacing the modified G run with the backup G run (or "spare tire").²⁵ Placement of the AP in a loop stalls the actions of apurinic/apyrimidinic

endonuclease 1 (APE1) on its substrate for gene induction via the *trans*-acting function of this protein (Scheme 1B).²⁶ The proposal was supported by control studies conducted in OGG1-deficient cells, APE1 knockdown by siRNAs, and experiments in oxidatively modified plasmids incapable of G4 formation.^{13,27} A key finding in our work was the observation that DNA structure in a gene promoter can be malleable as a result of G oxidation in a PQS context, and the non-canonical structure formed (i.e., a G4) provides a platform for APE1 to bind but poorly cleave leading to gene activation.

Additional studies moved the *VEGFPQS* to the template strand of the promoter and found that OG in this context diminished gene expression (Scheme 1B).²⁷ These findings suggest OG in promoter PQS contexts can be an up or down regulator of transcription. This was further documented in the *NTHL1*,¹³ *RAD17*,²⁸ and *PCNA*²⁹ promoter PQSs. Finally, to further probe the plasticity of DNA when oxidatively modified with OG to impact mRNA synthesis, OG was synthetically incorporated into the *VEGFi*-motif (iM) context to determine the impact on gene expression.²⁷ This additional study in the iM context found OG impacts gene regulation in non-canonical structures beyond G4s. We note G oxidation is highly favored in G runs found in PQSs^{30–32} relative to any Gs in an iM-forming sequence, and the iM example is likely not relevant to biological regulation, although the findings were instructive in exploring the proposed mechanism for gene activation. Lastly, the Tell laboratory demonstrated APE1 stalls on an AP in the loop of a hairpin leading to induction of the *SIRT1* gene in oxidatively stressed human cells.²⁰ Collectively, these studies suggest that non-canonical structures in gene promoters, when formed during oxidative modification of DNA, can drive phenotypic changes.

Mounting evidence supports formation of G4s,^{33,34} iMs,^{35,36} cruciform DNA,³⁷ triplexes,³⁸ and Z-DNA^{38,39} in human genomic DNA, and examples exist supporting the idea that these non-canonical structures can regulate transcription in human cells,^{34,38,40,41} and prokaryotic cells.^{42,43} An open question remaining is whether non-canonical structures beyond G4s,^{12,13} iMs,²⁷ and hairpins,²⁰ once oxidatively modified at G, can regulate mRNA expression during oxidative stress conditions. Regions with the potential to adopt Z-DNA are G rich, and should be susceptible to oxidation. Potential Z-DNA forming sequences (PZSs) are found in alternating pyrimidine-purine repeats of 5'-(CpG)_n-3'.⁴⁴ In solution, Z-DNA is observed at high ionic strength or in highly supercoiled plasmids.⁴⁵ Some variations in the sequence can be accommodated in Z-DNA.⁴⁵ Further, PZSs can be found in some CpG islands that are important for introduction of the epigenetic marker 5-methylcytosine; additionally, 5-methylcytosine favors the transition from B- to Z-form DNA at lower ionic strength.^{45,46} Herein, a bioinformatic screen of the human genome was conducted to locate PZSs and determine their distributions and locations. Next, studies were conducted to address the impact of G oxidation to OG within a PZS context in a gene promoter on transcription of a reporter gene in a human glioblastoma cell line (U87). The in cellulo observations led to a series of in vitro experiments to begin to understand how Z-DNA duplexes respond to the oxidatively modified G nucleotide OG, and how an AP product generated upon release of OG via OGG1 impacts the duplex. The final study inspected how APE1-mediated cleavage of an AP model (tetrahydrofuran, THF, or F) is impacted in various structural contexts. The findings are compared to prior studies regarding non-canonical structures in the human genome and gene regulation upon G oxidation to OG.

Materials and Methods

Bioinformatic Analysis for Z-DNA in the Human Genome.

Inspection for the potential Z-DNA-forming sequences (PZSs) was conducted on the hg38 version of the human genome. A script modified from Quadparser⁴⁷ was developed in-house to search for PZSs with d(CG)_n repeats of $n \geq 5$ (i.e., lengths = 10, 12, 14...; see Supporting Information for a copy of the code). Only one strand of the genome was inspected for the sequences to avoid double counting of the PZSs identified. The location for each PZS was determined via the PAVIS tool,⁴⁸ and the script to determine the distribution around the TSS and visualize the data was previously reported.⁴⁹ The promoter region for the genes was defined as 2000-nt upstream of the transcription start site (TSS).

DNA Strand Preparation.

All DNA strands were synthesized and deprotected by the DNA/Peptide core facility at the University of Utah following standard protocols. The site-specific introduction of OG or the tetrahydrofuran abasic site model (tetrahydrofuran or F) was achieved using commercially available phosphoramidites. After synthesis and deprotection via standard protocols, the crude oligomers were purified using an anion-exchange HPLC column running a mobile phase system consisting of A (1 M NaCl, 20 mM NaP_i at pH 7 in 1:9 MeCN:ddH₂O) and B (1:9 MeCN:ddH₂O). The method was initiated at 20% B and increased via a linear gradient to 100% B over 30 min with a flow rate of 1 mL/min while monitoring the absorbance at 260 nm. The purified samples were dialyzed against ddH₂O for 48 h, lyophilized to dryness, and resuspended in ddH₂O to make stock solutions. The concentrations of the samples were determined by measuring the absorbance at 260 nm, in which the nearest-neighbor approximation model was used to estimate the extinction coefficient. The extinction coefficients for the modified DNA strands were estimated by replacing OG for G and omitting a nucleotide for F. The oligomers were studied at the specified concentrations and buffers indicated for each experiment, as described for each experiment below.

Plasmid Construction.

Modification of the plasmid to contain the PZS in the promoter of a luciferase gene was achieved using a method previously outlined.¹³ Synthesis of plasmids containing site-specifically incorporated OG or the AP model F was achieved following a previously established protocol.¹³ Confirmation of the successful incorporation of the modification into the plasmids was performed by a gap ligation and Sanger sequencing protocol that we have reported.⁵⁰ The complete details of the synthesis and PCR primers used can be found in the Supporting Information.

Cell Studies.

The human U87-MG glioblastoma cells (U87) were obtained from ATCC. All cells were grown in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum, 20 µg/mL gentamicin, 1x glutamax, and 1x non-essential amino acids. The cells were grown at 37 °C with 5% CO₂ at ~80% relative humidity and were split when they reached ~75% confluence. The transfection experiments were conducted in white, 96-well plates by

seeding 3×10^4 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 per well using X-tremeGene HP DNA transfection agent (Roche) following the manufacturer's protocol in Opti-MEM media. The dual-glo luciferase assay (Promega) was conducted following the manufacturer's protocol 48 h post transfection. The transfection experiments were conducted at least four times, and the errors reported represent 95% confidence intervals.

APE1 Activity Assays.

The DNA samples were annealed in 20 mM Tris (pH 7.9), 50 mM KOAc, 10 mM Mg(OAc)₂, and 1 mM DTT by heating the solutions at 90 °C for 5 min followed by slow cooling to room temperature. The APE1 reactions were conducted on a 10-μL aliquot of the annealed DNA solution containing 0.04 pmol ³²P-labeled DNA and 0.06 pmol unlabelled DNA at 37 °C. The reactions were initiated by addition of 0.5 U of APE1 (New England Biolabs) and allowed to progress for 0.5 min, 1 min, 2 min, 5 min, 30 min, or 60 min. The reactions were terminated by adding 10 μL of stop buffer (95 % formamide, 10 mM NaOH, 0.1 % xylene cyanol, 0.1 % bromphenol blue, and 10 mM EDTA) followed by heat denaturation at 65 °C for 20 min. The reactions were performed in triplicate. Next, 12 μL of the stop buffer and heat denatured reaction was analysed on a 20% denaturing polyacrylamide gel electrophoresis (PAGE; 75 W and ~3.5 h run time). The electrophoresed gels were applied to a phosphor screen for 18 h followed by storage-phosphor autoradiography visualization. The bands were analysed using ImageQuant™ image analysis software. Reaction yields were calculated from the product and reactant band intensities measured using the image analysis software.

Circular Dichroism Analysis.

The Z-DNA samples were annealed at a 10 μM concentration in 20 mM NaP_i (pH 7.4) with 0.15–4.5 M NaCl. The samples were placed in a 0.2-cm quartz cuvette for circular dichroism (CD) analysis at 20 °C. The recorded data were solvent background subtracted and then normalized on the y-axis to units of molar ellipticity ([Θ]) for plotting and comparative purposes.

Thermal Melting Analysis.

The thermal melting (T_m) values were determined on samples of 3 μM oligomer in 20 mM NaP_i (pH 7.4) with 20 mM NaCl. The melting experiments were initiated by thermally equilibrating the samples at 20 °C for 10 min followed by heating at 0.5 °C/min and equilibrating at each 1 °C increment for 1 min. Readings at 260 nm were taken after each 1 °C change in the temperature from 20 °C to 100 °C. Plots of absorbance at 260 nm vs. temperature were constructed, and the T_m values were determined by a two-point analysis protocol using the instrument's software.

Results and Discussion

Bioinformatic inspection of the human genome for PZSs.

The bioinformatic analysis focused on determination of where PZSs reside in the human genome, and whether or not they are enriched in key regulatory regions of the genome.

These questions were addressed by development of an algorithm to search computationally the human genome to locate all PZSs of the repeat sequence d(CG)_n, where n ≥ 5. The analysis was initiated at n = 5 because this was the minimum repeat length found to be capable of adopting Z-DNA within the context of a supercoiled plasmid in *E. coli* cells.⁵¹ The inspection identified 1650 PZSs in the human genome. The number of PZSs found in our search is much lower than previously reported because our search only inspected for classical d(CG)_n PZSs while omitting other motifs capable of Z-DNA formation.^{52,53} In the population of sequences, 886 had a length of 10 nt (n = 5) that decreased to 2 sequences with a length of 26 nt (n = 13; Figure 1A and Tables S1 and S2). A comparison was made to the number of PZSs found at each length to the expected number based on a random distribution weighing all four nucleotides distributed equally. The expected number of PZSs in the human genome (3×10⁹ bp) with a length of 10 nt (n = 5) is 2861; in contrast, 886 sequences of this length were found that identified an underrepresentation of 10 nt long PZSs (Figure 1A). On the other hand, inspection of longer repeat lengths found an enrichment of PZSs relative to the expected count. For lengths 12 nt (n = 6), 14 nt (n = 7), and 16 nt (n = 8), the expected counts were 178, 11, and 0.7 sequences, while 361, 192, and 118 sequences were found, respectively (Figure 1A). An enrichment of the actual number found compared to the expected number holds for all PZSs of length greater than or equal to 12 nt.

For all of the PZSs identified their distribution in known human genomic elements were determined (Figure 1B). First, 38.2% of the identified PZSs were in introns that is similar to the number expected on the basis of introns comprising 38.6% of the human genome. Regions found underrepresented in PZSs include the intergenic and 3'-UTRs with 0.7-fold and 0.6-fold, respectively, of the number expected based on the size of these regions in the human genome. Regions enriched in PZSs include exons with 3.4-fold more, promoters with 7.9-fold more, and 5'-UTRs with 13.3-fold more PZSs (Figure 1C). The observation that PZSs were overly enriched in promoters and 5'-UTRs was quite surprising as a consequence of these regions being key regulatory locations for transcriptional and translational regulation. The PZS pool was then restricted to those found at a greater frequency than expected (n ≥ 6) to identify their genomic locations. Comparison of the complete pool of sequences (n ≥ 5) to the pool restricted to longer sequences found similar distributions throughout the human genome (Figure 1B blue vs. orange). There existed small differences in the promoter and 5'-UTR distributions, in which there were more in the promoter and fewer in the 5'-UTR with the pool of longer sequences. In a final analysis, a histogram was made of the number of PZSs and their location relative to the TSS within a window of 2000 bp upstream and downstream of this point (Figure 1D). Enrichment of PZSs was observed within ~300 nt of the TSS on the promoter side and within ~750 nt of the TSS on the 5'-UTR side of the TSS. The window in which PZSs were enriched positions them for possible regulatory elements of transcription and translation. The present findings for enrichment of classical PZSs is consistent with prior reports inspecting for a broader range of PZS motifs.⁵²

Oxidative modification of G in a PZS context impacts transcription.

A prior study in the Burrows and Rokita laboratories mapped G nucleotides sensitive to oxidation in the context of model Z-DNA strands using Co²⁺ and KHSO₅ to generate ROS.

⁵⁴ In Z-DNA, the G nucleotides throughout the model sequence studied provided similar levels of reactivity toward oxidation suggesting reactivity is determined by exposure to oxidant; this contrasts G oxidation in B-DNA that shows strong dependency on the sequence as a result of π stacking in the duplex.^{30–32} Consequently, in the experiments described next, OG was first synthesized at a central G of a PZS in the promoter of a luciferase gene in a reporter plasmid. The reporter plasmid analyzed possessed the Renilla luciferase (Rluc) gene bearing the PZS-modified promoter, as well as the firefly luciferase (luc) gene that was left in the native state to be used as an internal standard for transcription level quantification. The promoter for the Rluc gene is the SV40 early/enhancer sequence that has a TATA box positioned from –16 to –23 relative to the TSS,⁵⁵ which was replaced with the PZS d(CG)₁₅ (Table 1). The methods to synthesize the PZS and OG in the promoter were previously reported.¹³ The longer PZS was selected to ensure the favorability of Z-DNA formation within the plasmid on the basis of prior cell studies.^{51,56} Synthetic incorporation of OG in either the coding (i.e., non-template) or template strand of the PZS addressed the question of how the strand of OG occupancy impacts transcription. The plasmids were transfected into human glioblastoma (U87) cells. Comparisons of the results in the PZS context were made to the potential G-quadruplex-forming sequence *VEGFG4*, potential i-motif forming sequence *VEGF*iM, and a sequence that adopts B-DNA (i.e., neg. control) with OG studied in the same cell line (Table 1). The *VEGFG4*, iM, and B-DNA studies were previously reported.^{13,27}

The transcription levels from the reporter plasmid were measured via a dual-glo luciferase assay to independently measure Rluc and luc protein levels in the cells after a 48-h incubation. The luminescence values were used to compute relative response ratio (RRR = Rluc/luc), after which the RRRs were normalized to plasmids bearing the structure-switching sequence without OG (normalized RRR = RRR_{OG} / RRR_G; Figure 2). For the plasmids with the PZS in the coding strand, the presence of OG led to small but significant 1.6-fold increase in expression (Figure 2A). Values for the potential structural switches with OG in the *VEGFPQS* or potential iM sequence contexts in the coding strand previously showed that^{13,27} the PQS context led to a 3-fold increase and the potential iM context led to a 2.1-fold increase in expression (Figure 2A). Lastly, OG in a B-form sequence incapable of adopting a non-canonical structure was silent in the coding strand of the promoter with respect to the impact on mRNA synthesis (Figure 2A). These comparisons demonstrate that OG in the coding strand of a promoter close to the TSS in a possible Z-DNA, G4, or iM structure switching context can increase mRNA expression. The ranking of expression increase observed for OG in the structure switching contexts was G4 > iM > Z-DNA >> B-DNA. In the PZS, OG was synthesized at another position leading to a similar change in Rluc expression that suggests expression is impacted similarly when other G nucleotides are oxidatively modified (Figure S1).

The d(GC)_n repeat nature of Z-DNA places an equal number of reactive sites on both strands. Therefore, OG was synthesized in the PZS context in the template strand of the promoter for Rluc in the reporter plasmid. Transfection of the OG-modified plasmid into U87 cells found a ~2-fold reduction in Rluc expression when OG was in the PZS context in the template strand (Figure 2B). This observation was compared to a previous study²⁷ in which OG in the *VEGFPQS*, *VEGF*potential iM, and a B-DNA-sequence context all of

which led to a similar 2–3-fold reduction in mRNA synthesis (Figure 2B). A similar change in transcription for OG at another position in the PZS context yielded a similar decrease in Rluc expression (Figure S1). Thus, regardless of the structure-switching capacity of the sequence, when G is oxidatively modified in the template strand of the promoter near the TSS, mRNA expression is downregulated.

Our prior studies with the *VEGFPQS* found oxidative modification of G to OG in the coding strand was acted on by OGG1 to yield an AP as part of the activation process.¹³ For OG in the template strand of the promoter near the TSS, gene activation occurred via an APE1-independent mechanism, and transcription-coupled NER was implicated for repair in this strand.²⁷ In vivo, OGG1 appears to be mainly a monofunctional glycosylase yielding an AP that is passed on to the next BER member APE1.⁵⁷ Once OG is released from the coding strand, the resulting AP destabilizes B-DNA, and provides the thermodynamic drive to yield a G4 fold,¹³ which is facilitated by the *VEGF* sequencing possessing five G tracks allowing extrusion of the AP into a large loop.²⁵ The AP in a loop of a G4 fold is bound by APE1 and poorly cleaved, on the basis of literature reports and our unpublished results.⁵⁸ Therefore, to probe the importance of an AP in impacting transcription in the PZS context, an analog of an AP (i.e., THF or F) was synthesized in the plasmid to be transfected to U87 cells in order to measure the Rluc expression. The F analog of an AP was selected because this model retains the ability to be acted upon by APE1 while being stable to β -elimination reactions resulting in strand breaks during synthesis, to which natural APs are prone.⁵⁹

When F was synthetically incorporated at the same sites as OG and the corresponding plasmids were transfected into U87 cells, a similar direction in the change of mRNA synthesis was observed for the coding or template strands of the promoter (Figure 3). That is, F in the coding strand led to an increase in gene expression of ~1.6 fold while F in the template strand was found to decrease expression by >2 fold. These observations indicate that when OG is processed in the coding strand, the AP formed can increase gene expression, while an AP in the template strand is disruptive to RNA pol II for mRNA synthesis. In the coding strand, the observation with the F-containing plasmid supports the conclusion that modifications induce mRNA synthesis occur after OG is released from the DNA by OGG1. This finding is consistent with our prior work regarding AP in the *VEGF* PQS context in the coding strand.¹³ On the template strand of the promoter near the TSS, modification of G to OG or F consistently decreases gene activity.

Because the magnitude and direction of changes in mRNA expression were similar for OG in the PZS context as compared to the *VEGFPQS*, and *VEGF* potential iM sequence context in both strands (Figures 2A and 2B), we propose there are similar mechanisms leading to the observed changes. A critical component of OG guiding mRNA synthesis in the PQS context was the structure-switching capability of the sequence.^{13,27} Therefore, the next studies conducted looked at the impact OG or F on Z-DNA structure to determine whether this sequence context is capable of structure switching when oxidatively modified. The initial structural experiments enabled additional studies to understand the process leading to gene modulation in the PZS context.

Z-DNA structure is impacted by OG or APs.

In model oligodeoxynucleotides comprised of the d(CG)_n repeat sequence, Z-DNA formation in solution is promoted by high monovalent salt concentrations (e.g., [NaCl] > 3 M), or high concentrations of multivalent metals (e.g., Mg²⁺, Ni²⁺, or Co³⁺).⁵⁶ Therefore, to interrogate the impact of OG or AP (i.e., F) on a Z-DNA duplex, we conducted solution studies on two different Z-DNA forming sequences from the literature that were modified with OG or F. The first sequence selected adopts a hairpin with a tetra-T loop (HP, Figure 4A),⁶⁰ and the second sequence is also a hairpin with a Z-DNA region adjacent to a B-DNA region allowing a study of how modifications can impact a sequence with a B-Z junction (BZ, Figure 4A).⁶¹ Hairpins that adopt Z-DNA were selected because they permitted the study of the modification's impact on structure while avoiding issues that would arise with competing uni- and bi-molecular duplexes that would occur with d(CG)_n strands annealed to adopt duplexes. Additionally, the hairpin systems selected allowed study of one modification per duplex in a manner similar to the cellular studies.

Native and modified Z-DNA strands were studied by titrating NaCl from 0.15 to 4.5 M while monitoring the B to Z transition via circular dichroism (CD) spectroscopy. The B-Z transition is easily observed by the direction and magnitude of the CD spectrum at 295 nm, in which B-DNA gives a positive signal and Z-DNA has a negative signal (Figure 4B).^{56,60,61} Thus, a plot of the intensity at 295 nm vs. a change in the [NaCl] yields a titration curve with a midpoint value, or [NaCl] transition midpoint (i.e., [NaCl]_T) that reports on the stability of the Z-DNA in an experimentally tractable system (Figure 4C). Lower [NaCl]_T values identify Z-DNA formation occurs more readily than higher [NaCl]_T values. In the Z-DNA hairpin HP selected for study, a central G in the stem was altered to OG or F, or a T in the loop was changed to F. The non-modified hairpin gave a [NaCl]_T = 3.3, inclusion of OG decreased the [NaCl]_T to 2.9, F in the stem increased the [NaCl]_T to 4.0, and F in the loop gave a similar [NaCl]_T = 3.3 as the native sequence (Figure 4D). In the B-Z hairpin system, the native sequence had a [NaCl]_T = 3.4, the presence of OG decreased the [NaCl]_T to 2.8, and F increased the [NaCl]_T to 3.6. These findings indicate OG favors Z-DNA formation and F inhibits Z-DNA when located in the stem or duplex of Z-DNA, while F in the loop of a Z-DNA forming hairpin had no impact on the [NaCl]_T value. These salt-dependent studies in oligodeoxynucleotides provide a quantitative measure of the relative differences in Z-DNA stabilities with and without chemical modifications, which is not achievable in plasmids; although, we note that the oligodeoxynucleotide studies may not provide a complete picture of the changes in the plasmid context.

Additional interrogation of the stability was performed by measuring the thermal stabilities (i.e., *T_m* values) for the HP sequence in the native or modified states. Determination of the *T_m* values for the sequences under high salt conditions to achieve Z-DNA was not possible because the *T_m* values were too high to be measured; to overcome this challenge, the salt concentration was dropped to 20 mM NaCl in 20 mM NaP_i (pH 7.4) buffer to obtain *T_m* values for the sequences. The native HP had a *T_m* value of 74.7 °C, OG in the stem caused a small decrease in *T_m* value to 72.1 °C, F in the stem of the duplex dramatically decreased the *T_m* value to 57.9 °C, and F in the loop did not impact the stability with a *T_m* value similar to the native sequence (75.3 °C; Figure 5). These findings provide some support for

the NaCl titration studies, in which F in a Z-DNA duplex is destabilizing and F in the loop of a Z-DNA forming hairpin has no impact on the stability. The results with OG between the T_m and salt-dependent studies cannot be compared because the T_m values were measured on B-form DNA while the $[\text{NaCl}]_T$ values monitor the B $\rightarrow$ Z transition.

The reason OG favors Z-DNA can be chemically rationalized by the difference in G•C base pairs between B- and Z-form duplexes (Figure 6A). In a B-form helix, the G nucleotides are in the *anti* conformation (Figure 6B),⁶² while in a Z-form helix, the G nucleotides are in the *syn* conformation (Figure 6C).⁶³ This is important because the carbonyl at C8 in OG introduces an electrostatic clash with the sugar-phosphate backbone⁶⁴ resulting in the OG nucleotide favoring the *syn* conformation that is found in Z-DNA (Figures 6B and 6C); further, Z-DNA has a larger major groove than B-DNA that will also favor OG. Prior studies identified OG can facilitate Z-DNA formation.⁶⁵ In contrast, the baseless site F is destabilizing to duplex nucleic acids as a result of the hole that disrupts π stacking between adjacent base pairs.

The studies suggest OG or F can have opposite impacts on Z-DNA stability as illustrated in Scheme 2. Oxidative modification of G to OG increases Z-DNA stability and likely shifts the equilibrium to the non-canonical fold; in contrast, an AP (i.e., F) destabilizes Z-DNA and shifts the equilibrium toward B-DNA. In B-DNA, AP cause significant destabilization of the duplex ($\sim 17^\circ\text{C}$ decrease in stability, Figure 5) that could provide the drive to an alternative structure for minimization of the impact of AP. On the basis of work conducted in plasmids, model oligonucleotides, and the present studies, we propose a possible model by which AP drives the formation of a non-canonical fold in DNA. The plasmid studies found PZSs under appropriate conditions can adopt a hairpin in each strand, or cruciform-like structure.⁶⁶ The Delaney laboratory demonstrated modifications in triplet-repeat expanded DNA hairpins can roll to adapt structures that place the modification in the loop of a hairpin.⁶⁷ Therefore, we propose the AP in the PZS context can shift the equilibrium to a hairpin, or cruciform-like structure, and extrude the AP into a hairpin loop (Scheme 2). Support for this proposal is derived from the finding that the HP sequence with an F in the loop was not destabilized by the modification (Figures 4A, 4D, and 5). This proposal allows a model to be proposed for how G oxidation in the PZS context in a promoter near the TSS can impact transcription.

The presence of an F yields a hairpin or cruciform-like structure presenting the baseless site to APE1 in a loop. Prior studies in the Tell laboratory found AP in the loop of a hairpin in the promoter of the *SIRT1* gene was bound by APE1 by not cleaved, resulting in the protein functioning as a *trans*-activator of transcription by recruitment of activating factors.²⁰ As a parallel to Tell's studies, we propose a similar pathway to gene activation when F is in the PZS in the coding strand of the promoter; while in the template strand of the promoter, the F in the non-canonical structure leads to gene inactivation by either blocking RNA pol II loading or activation of transcription-coupled repair. Both proposals are made on the basis of prior work in different structure-switching sequence contexts.^{20,27}

In a final study to support whether AP in the loop of a hairpin stalls APE1 endonuclease activity, we monitored the extent of APE1-mediated cleavage of F in a B-form duplex, the

stem of the BZ hairpin, or the loop of a well-defined hairpin (Figures 7A and S2). The B-form duplex was selected to be G-rich, similar to Z-DNA duplexes, and the well-defined hairpin had tails on the 5' and 3' sides that did not base pair to aid in blocking a bimolecular duplex to ensure a hairpin structure was studied. The amount of product was determined at 5- and 30-min reaction times on the three F-containing contexts that were allowed to be acted on by APE1. Further, the reactions were conducted in buffer suggested by the supplier of APE1 (NEB) that were of low ionic strength (50 mM KOAc). This means all duplexes were B-form. This points to a limitation in the enzyme studies being that the AP substrate could not be acted upon in a Z-DNA duplex because the high ionic strength required to achieve this fold prohibits enzyme activity. Nonetheless, studies were conducted to understand whether the structural context altered APE1 endonuclease activity on F-containing DNA. Cleavage yields of F in the B-DNA context were >75% at 5 min and increased slightly at 30 min (Figure 7B). The reaction yield for F in the stem of the BZ hairpin was ~10% at 5 min and increased to 65% at 30 min (Figure 7B). Lastly, when F was in the loop of a well-defined hairpin the yield at 5 min was ~2% and at 30 min the yield increased to ~6% (Figure 7B). These findings conclude that the sequence and structural context in which F resides influences the APE1 reaction yields significantly.

The observation that APE1 favors cleavage of F in B-form duplex DNA is consistent with prior studies.^{58,59} Noteworthy, is that the BZ hairpin existed as B-form under the ionic strength of the studies and that the APE1 reaction yield in this sequence was slower than the B-form duplex. The reason for this could be the location of F in the BZ hairpin eight base pairs from the end, or the G•C-rich nature of the hairpin stabilized the fold and attenuated APE1 activity. Studies to further understand the BZ hairpin were not pursued; although future work to understand sequence effects on repair could be valuable in understanding sequence dependency in DNA repair rates, which are important for mutagenesis.⁶⁸ Prior work found APE1 catalyzes F removal from single-stranded DNA nearly as efficiently as from a B-form duplex.⁶⁹ The surprising observation was the F in the loop of a hairpin, in which cleavage yields were <10%, even after a 30 min reaction, suggest APE1 is highly challenged by F in hairpin loops. Studies to understand whether APE1 could bind F in the loop of a hairpin were not pursued; however, previous reports have found APE1 binds AP in loops while poorly cleaving the strand.⁵⁸

Conclusion

The present experiments identify another possible structure-switching sequence context in the promoter of a gene near the TSS in which oxidative modification of a G nucleotide could be an on/off switch for transcription.^{20,22} The present findings are educational in providing scope to the plasticity of DNA under oxidative stress conditions; however, it is unlikely that cells utilize PZSs as redox-active sites for G oxidation leading to gene regulation. This is because the strand in which G oxidation occurs cannot be controlled within the d(CG)_n repeat sequence context as a result of the symmetry of the sequence in both strands. In contrast to PZSs, potential G-quadruplex forming sequences focus the highly reactive G-runs all on one strand wherein G oxidation occurs. The strand on which G oxidation occurs in a promoter will then determine whether oxidation up- or down-regulates transcription during oxidative stress.^{13,27} Moreover, sequencing studies have found >700,000 PQSs.⁷⁰

This number of PQSs is nearly 700-fold more than the number of PZSs found in the bioinformatic analysis reported herein (Figure 1). The fact that so many more PQSs exist in the human genome compared to PZSs, and that PQSs will direct the strand on which G oxidation occurs when they are in a promoter, argues that PQSs serve as redox switches for gene regulation, while PZSs probably do not. Even though this downplays the importance of the present results, the findings will be important for understanding sequencing data for OG. Now that methods have been developed to conduct next-generation sequencing for OG on eukaryotic genomes,^{71–73} understanding the implications of OG in a particular sequence context for phenotypic change will be important as more data become available. These studies highlight G oxidation in a PZS context can impact mRNA synthesis (Figure 2), but if OG is found in PZSs, it likely occurred by random G oxidation and did not occur with an evolved function.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by a National Cancer Institute Grant (R01 CA090689). The DNA strands and Sanger sequencing experiments were provided by the core facilities at the University of Utah, which are supported in part by an NCI Cancer Center Support Grant (P30 CA042014).

References

- [1]. Sohal RS, and Weindruch R (1996) Oxidative stress, caloric restriction, and aging. *Science* 273, 59–63. [PubMed: 8658196]
- [2]. Lonkar P, and Dedon PC (2011) Reactive species and DNA damage in chronic inflammation: reconciling chemical mechanisms and biological fates. *Int. J. Cancer* 128, 1999–2009. [PubMed: 21387284]
- [3]. Kurian GA, Rajagopal R, Vedantham S, and Rajesh M (2016) The role of oxidative stress in myocardial ischemia and reperfusion injury and remodeling: Revisited. *Oxid. Med. Cell Longev* 2016, 1656450. [PubMed: 27313825]
- [4]. Khatri N, Thakur M, Pareek V, Kumar S, Sharma S, and Datusalia AK (2018) Oxidative stress: Major threat in traumatic brain injury. *CNS Neurol. Disord. Drug Targets* 17, 689–695. [PubMed: 29952272]
- [5]. Giorgio M, Trinei M, Migliaccio E, and Pelicci PG (2007) Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? *Nat. Rev. Mol. Cell Biol* 8, 722–728. [PubMed: 17700625]
- [6]. Antoniali G, Malfatti MC, and Tell G (2017) Unveiling the non-repair face of the base excision repair pathway in RNA processing: A missing link between DNA repair and gene expression? *DNA Repair (Amst)* 56, 65–74. [PubMed: 28629776]
- [7]. Mittler R (2016) ROS are good. *Trends Plant Sci* 22, 11–19. [PubMed: 27666517]
- [8]. Visnes T, Cazares-Korner A, Hao W, Wallner O, Masuyer G, Loseva O, Mortusewicz O, Wiita E, Sarno A, Manoilov A, Astorga-Wells J, Jemth A-S, Pan L, Sanjiv K, Karsten S, Gokturk C, Grube M, Homan EJ, Hanna BMF, Paulin CBJ, Pham T, Rasti A, Berglund UW, von Nicolai C, Benitez-Buelga C, Koolmeister T, Ivanic D, Iliev P, Scobie M, Krokan HE, Baranczewski P, Artursson P, Altun M, Jensen AJ, Kalderen C, Ba X, Zubarev RA, Stenmark P, Boldogh I, and Helleday T (2018) Small-molecule inhibitor of OGG1 suppresses proinflammatory gene expression and inflammation. *Science* 362, 834–839. [PubMed: 30442810]
- [9]. Wood ZA, Poole LB, and Karplus PA (2003) Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* 300, 650–653. [PubMed: 12714747]

- [10]. Ba X, and Boldogh I (2018) 8-Oxoguanine DNA glycosylase 1: Beyond repair of the oxidatively modified base lesions. *Redox Biol* 14, 669–678. [PubMed: 29175754]
- [11]. Fleming AM, Zhu J, Ding Y, Visser JA, and Burrows CJ (2018) Human DNA repair genes possess potential G-quadruplex sequences in their promoters and 5'-untranslated regions. *Biochemistry* 57, 991–1002. [PubMed: 29320161]
- [12]. Cogoi S, Ferino A, Miglietta G, Pedersen EB, and Xodo LE (2018) The regulatory G4 motif of the Kirsten ras (KRAS) gene is sensitive to guanine oxidation: implications on transcription. *Nucleic Acids Res* 46, 661–676. [PubMed: 29165690]
- [13]. Fleming AM, Ding Y, and Burrows CJ (2017) Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. *Proc. Natl. Acad. Sci. U. S. A.*, 2604–2609. [PubMed: 28143930]
- [14]. Perillo B, Ombra MN, Bertoni A, Cuozzo C, Sacchetti S, Sasso A, Chiariotti L, Malorni A, Abbondanza C, and Avvedimento EV (2008) DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science* 319, 202–206. [PubMed: 18187655]
- [15]. Pastukh V, Roberts JT, Clark DW, Bardwell GC, Patel M, Al-Mehdi AB, Borchert GM, and Gillespie MN (2015) An oxidative DNA “damage” and repair mechanism localized in the VEGF promoter is important for hypoxia-induced VEGF mRNA expression. *Am. J. Physiol. Lung Cell Mol. Physiol* 309, L1367–1375. [PubMed: 26432868]
- [16]. Chen K, Zhao, Boxuan S, and He C (2016) Nucleic acid modifications in regulation of gene expression. *Cell Chem. Biol* 23, 74–85. [PubMed: 26933737]
- [17]. Hammoud SS, Low DH, Yi C, Carrell DT, Guccione E, and Cairns BR (2014) Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. *Cell Stem Cell* 15, 239–253. [PubMed: 24835570]
- [18]. Fong YW, Cattoglio C, and Tjian R (2013) The intertwined roles of transcription and repair proteins. *Mol. Cell* 52, 291–302. [PubMed: 24207023]
- [19]. Pan L, Zhu B, Hao W, Zeng X, Vlahopoulos SA, Hazra TK, Hegde ML, Radak Z, Bacci A, Brasier AR, Ba X, and Boldogh I (2016) Oxidized guanine base lesions function in 8-oxoguanine DNA glycosylase1-mediated epigenetic regulation of nuclear factor kappaB-driven gene expression. *J. Biol. Chem* 291, 25553–25566. [PubMed: 27756845]
- [20]. Antoniali G, Lirussi L, D'Ambrosio C, Dal Piaz F, Vascotto C, Casarano E, Marasco D, Scaloni A, Fogolari F, and Tell G (2014) SIRT1 gene expression upon genotoxic damage is regulated by APE1 through nCaRE-promoter elements. *Mol. Biol. Cell* 25, 532–547. [PubMed: 24356447]
- [21]. Wang R, Hao W, Pan L, Boldogh I, and Ba X (2018) The roles of base excision repair enzyme OGG1 in gene expression. *Cell. Mol. Life Sci* 75, 3741–3750. [PubMed: 30043138]
- [22]. Fleming AM, and Burrows CJ (2017) 8-Oxo-7,8-dihydroguanine, friend and foe: Epigenetic-like regulator versus initiator of mutagenesis. *DNA Repair (Amst)* 56, 75–83. [PubMed: 28629775]
- [23]. Seifermann M, and Epe B (2017) Oxidatively generated base modifications in DNA: Not only carcinogenic risk factor but also regulatory mark? *Free Radic. Biol. Med* 107, 258–265. [PubMed: 27871818]
- [24]. Zuchegna C, Aceto F, Bertoni A, Romano A, Perillo B, Laccetti P, Gottesman ME, Avvedimento EV, and Porcellini A (2014) Mechanism of retinoic acid-induced transcription: histone code, DNA oxidation and formation of chromatin loops. *Nucleic Acids Res* 42, 11040–11055. [PubMed: 25217584]
- [25]. Fleming AM, Zhou J, Wallace SS, and 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. [PubMed: 26405692]
- [26]. Bhakat KK, Mantha AK, and Mitra S (2009) Transcriptional regulatory functions of mammalian AP-endonuclease (APE1/Ref-1), an essential multifunctional protein. *Antioxid. Redox Signal* 11, 621–638. [PubMed: 18715144]
- [27]. Fleming AM, Zhu J, Ding Y, and Burrows CJ (2017) 8-Oxo-7,8-dihydroguanine in the context of a promoter G-quadruplex is an on-off switch for transcription. *ACS Chem. Biol* 12, 2417–2426. [PubMed: 28829124]

- [28]. Zhu J, Fleming AM, and Burrows CJ (2018) The RAD17 promoter sequence contains a potential tail-dependent G-quadruplex that downregulates gene expression upon oxidative modification. *ACS Chem. Biol* 13, 2577–2584. [PubMed: 30063821]
- [29]. Redstone SCJ, Fleming AM, and Burrows CJ (2019) Oxidative modification of the potential G-quadruplex sequence in the PCNA gene promoter can turn on transcription. *Chem. Res. Toxicol.* doi: 10.1021/acs.chemrestox.1028b00332.
- [30]. Saito I, Takayama M, Sugiyama H, Nakatani K, Tsuchida A, and Yamamoto M (1995) Photoinduced DNA cleavage via electron transfer: demonstration that guanine residues located 5' to guanine are the most electron-donating sites. *J. Am. Chem. Soc* 117, 6406–6407.
- [31]. Delaney S, and Barton JK (2003) Long-range DNA charge transport. *J. Org. Chem* 68, 6475–6483. [PubMed: 12919006]
- [32]. Schuster GB (2000) Long-range charge transfer in DNA: Transient structural distortions control the distance dependence. *Acc. Chem. Res* 33, 253–260. [PubMed: 10775318]
- [33]. Biffi G, Tannahill D, McCafferty J, and Balasubramanian S (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem* 5, 182–186. [PubMed: 23422559]
- [34]. Hansel-Hertsch R, Beraldi D, Lensing SV, Marsico G, Zyner K, Parry A, Di Antonio M, Pike J, Kimura H, Narita M, Tannahill D, and Balasubramanian S (2016) G-quadruplex structures mark human regulatory chromatin. *Nat. Genet* 48, 1267–1272. [PubMed: 27618450]
- [35]. Dzatko S, Krafcikova M, Hansel-Hertsch R, Fessl T, Fiala R, Loja T, Krafcik D, Mergny JL, Foldynova-Trantirkova S, and Trantirek L (2018) Evaluation of the stability of DNA i-motifs in the nuclei of living mammalian cells. *Angew. Chem. Int. Ed. Engl* 57, 2165–2169. [PubMed: 29266664]
- [36]. Zeraati M, Langley DB, Schofield P, Moye AL, Rouet R, Hughes WE, Bryan TM, Dinger ME, and Christ D (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem* 10, 631–637. [PubMed: 29686376]
- [37]. Brazda V, Laister RC, Jagelska EB, and Arrowsmith C (2011) Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Mol. Biol* 12, 33. [PubMed: 21816114]
- [38]. Kouzine F, Wojtowicz D, Baranello L, Yamane A, Nelson S, Resch W, Kieffer-Kwon KR, Benham CJ, Casellas R, Przytycka TM, and Levens D (2017) Permanganate/S1 nuclease footprinting reveals non-B DNA structures with regulatory potential across a mammalian genome. *Cell Syst* 4, 344–356. [PubMed: 28237796]
- [39]. Shin S-I, Ham S, Park J, Seo SH, Lim CH, Jeon H, Huh J, and Roh T-Y (2016) Z-DNA-forming sites identified by ChIP-Seq are associated with actively transcribed regions in the human genome. *DNA Res* 23, 477–486.
- [40]. Balasubramanian S, Hurley LH, and Neidle S (2011) Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? *Nat. Rev. Drug Discov* 10, 261–275. [PubMed: 21455236]
- [41]. Kang HJ, Kendrick S, Hecht SM, and Hurley LH (2014) The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc* 136, 4172–4185. [PubMed: 24559432]
- [42]. Holder IT, and Hartig JS (2014) A matter of location: influence of G-quadruplexes on *Escherichia coli* gene expression. *Chem. Biol* 21, 1511–1521. [PubMed: 25459072]
- [43]. Boži Abram S, Maruši M, Lenar i Živkovi M, Br i J, and Plavec J (2018) Non-canonical structures in promoter modulate gene expression in *Escherichia coli*. *Croatica Chemica Acta* 91, 163–170.
- [44]. Wang AH, Quigley GJ, Kolpak FJ, Crawford JL, van Boom JH, van der Marel G, and Rich A (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* 282, 680–686. [PubMed: 514347]
- [45]. Jovin TM, and Soumpasis DM (1987) The transition between B-DNA and Z-DNA. *Ann. Rev. Phys. Chem* 38, 521–560.
- [46]. Temiz NA, Donohue DE, Bacolla A, Luke BT, and Collins JR (2012) The role of methylation in the intrinsic dynamics of B- and Z-DNA. *PLoS One* 7, e35558. [PubMed: 22530050]

- [47]. Huppert JL, and Balasubramanian S (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* 33, 2908–2916. [PubMed: 15914667]
- [48]. Huang W, Loganantharaj R, Schroeder B, Fargo D, and Li L (2013) PAVIS: a tool for peak annotation and visualization. *Bioinformatics (Oxford, England)* 29, 3097–3099.
- [49]. Ding Y, Fleming AM, and Burrows CJ (2018) Case studies on potential G-quadruplex-forming sequences from the bacterial orders Deinococcales and Thermales derived from a survey of published genomes. *Sci. Rep* 8, 15679. [PubMed: 30356061]
- [50]. Riedl J, Fleming AM, and 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, 491–494.
- [51]. Kim J, Yang C, and DasSarma S (1996) Analysis of left-handed Z-DNA formation in short d(CG)n sequences in *Escherichia coli* and *Halobacterium halobium* plasmids. Stabilization by increasing repeat length and DNA supercoiling but not salinity. *J. Biol. Chem* 271, 9340–9346. [PubMed: 8621598]
- [52]. Zhao J, Bacolla A, Wang G, and Vasquez KM (2010) Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci* 67, 43–62. [PubMed: 19727556]
- [53]. Cer RZ, Bruce KH, Mudunuri US, Yi M, Volfovsky N, Luke BT, Bacolla A, Collins JR, and Stephens RM (2011) Non-B DB: a database of predicted non-B DNA-forming motifs in mammalian genomes. *Nucleic Acids Res* 39, D383–D391. [PubMed: 21097885]
- [54]. Tang N, Muller JG, Burrows CJ, and Rokita SE (1999) Nickel and cobalt reagents promote selective oxidation of Z-DNA. *Biochemistry* 38, 16648–16654. [PubMed: 10600127]
- [55]. Everett RD, Baty D, and Chambon P (1983) The repeated GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. *Nucleic Acids Res* 11, 2447–2464. [PubMed: 6304653]
- [56]. Peck LJ, Nordheim A, Rich A, and Wang JC (1982) Flipping of cloned d(pCpG)n.d(pCpG)n DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. *Proc. Natl. Acad. Sci. U. S. A* 79, 4560–4564. [PubMed: 6956879]
- [57]. Allgayer J, Kitsera N, Bartelt S, Epe B, and Khobta A (2016) Widespread transcriptional gene inactivation initiated by a repair intermediate of 8-oxoguanine. *Nucleic Acids Res* 44, 7267–7280. [PubMed: 27220469]
- [58]. Broxson C, Hayner JN, Beckett J, Bloom LB, and Tornaletti S (2014) Human AP endonuclease inefficiently removes abasic sites within G4 structures compared to duplex DNA. *Nucleic Acids Res* 42, 7708–7719. [PubMed: 24848015]
- [59]. Schermerhorn KM, and Delaney S (2013) Transient-state kinetics of apurinic/apyrimidinic (AP) endonuclease 1 acting on an authentic AP site and commonly used substrate analogs: the effect of diverse metal ions and base mismatches. *Biochemistry* 52, 7669–7677. [PubMed: 24079850]
- [60]. Benight AS, Wang YS, Amaratunga M, Chattopadhyaya R, Henderson J, Hanlon S, and Ikuta S (1989) Conformation and dynamics of a left-handed Z-DNA hairpin: studies of d(CGCGCGTTTTTCGCGCG) in solution. *Biochemistry* 28, 3323–3332. [PubMed: 2545256]
- [61]. Dumat B, Larsen AF, and Wilhelmsson LM (2016) Studying Z-DNA and B- to Z-DNA transitions using a cytosine analogue FRET-pair. *Nucleic Acids Res* 44, e101. [PubMed: 26896804]
- [62]. Drew HR, Wing RM, Takano T, Broka C, Tanaka S, Itakura K, and Dickerson RE (1981) Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 78, 2179–2183. [PubMed: 6941276]
- [63]. Luo Z, Dauter M, and Dauter Z (2014) Phosphates in the Z-DNA dodecamer are flexible, but their P-SAD signal is sufficient for structure solution. *Acta Crystallogr. D Biol. Crystallogr* 70, 1790–1800. [PubMed: 25004957]
- [64]. Lipscomb LA, Peek ME, Morningstar ML, Verghis SM, Miller EM, Rich A, Essigmann JM, and Williams LD (1995) X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Proc. Natl. Acad. Sci. U.S.A* 92, 719–723. [PubMed: 7846041]
- [65]. Wang J, Wang S, Zhong C, Tian T, and Zhou X (2015) Novel insights into a major DNA oxidative lesion: its effects on Z-DNA stabilization. *Org. Biomol. Chem* 13, 8996–8999. [PubMed: 26242474]

- [66]. McLean MJ, and Wells RD (1988) The role of DNA sequence in the formation of Z-DNA versus cruciforms in plasmids. *J. Biol. Chem* 263, 7370–7377. [PubMed: 2835374]
- [67]. Volle CB, Jarem DA, and Delaney S (2012) Trinucleotide repeat DNA alters structure to minimize the thermodynamic impact of 8-oxo-7,8-dihydroguanine. *Biochemistry* 51, 52–62. [PubMed: 22148399]
- [68]. Banda DM, Nunez NN, Burnside MA, Bradshaw KM, and David SS (2017) Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: Mechanism, metals and medicine. *Free Radic. Biol. Med* 107, 202–215. [PubMed: 28087410]
- [69]. Marenstein DR, Wilson DM 3rd, and Teebor GW (2004) Human AP endonuclease (APE1) demonstrates endonucleolytic activity against AP sites in single-stranded DNA. *DNA Repair (Amst)* 3, 527–533. [PubMed: 15084314]
- [70]. Chambers VS, Marsico G, Boutell JM, Di Antonio M, Smith GP, and Balasubramanian S (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol* 33, 877–881. [PubMed: 26192317]
- [71]. Ding Y, Fleming AM, and Burrows CJ (2017) Sequencing the mouse genome for the oxidatively modified base 8-oxo-7,8-dihydroguanine by OG-Seq. *J. Am. Chem. Soc* 139, 2569–2572. [PubMed: 28150947]
- [72]. Wu J, McKeague M, and Sturla SJ (2018) Nucleotide-resolution genome-wide mapping of oxidative DNA damage by click-code-seq. *J. Am. Chem. Soc* 140, 9783–9787. [PubMed: 29944356]
- [73]. Amente S, Di Palo G, Scala G, Castrignanò T, Gorini F, Coccozza S, Moresano A, Pucci P, Ma B, Stepanov I, Lania L, Pelicci PG, Dellino GI, and Majello B (2019) Genome-wide mapping of 8-oxo-7,8-dihydro-2'-deoxyguanosine reveals accumulation of oxidatively-generated damage at DNA replication origins within transcribed long genes of mammalian cells. *Nucleic Acids Res* 47, 221–236. [PubMed: 30462294]

Figure 1.

Length and location analysis for PZSs in the human genome. (A) The number of PZSs identified of lengths 10–26 nt (i.e., d(CG)_n where n = 5, 6, 7...13). The inset to panel A in the figure provides the number of expected PZSs with lengths 10, 12, 14, and 16 nt. (B) Percentage of PZSs found in defined genomic elements. (C) Table of percentage of various regions distributed in the human genome and the relative distribution of PZSs in these regions. (D) Distribution of PZSs around human TSSs.

Figure 2.

The impact of OG on transcription when located in various structure-switching sequence contexts in the promoter of the Rluc gene in a reporter plasmid transfected in U87 glioblastoma cells. The normalized RRRs observed for OG in the (A) coding or (B) template strand of the promoter in the context of a Z-DNA, G4, or iM structural switches compared to a sequence incapable of structure switching (i.e., B-DNA). The data for the G4, iM, and B-DNA were previously reported and are provided for comparison.^{13,27} The sequence selected as a B-DNA structure as a negative control was previously shown to be void of alternative structure by ¹H-NMR and CD analysis.¹³ The data were obtained by transfection of the reporter plasmids into U87 cells followed by a 48 h incubation prior to conducting a dual-glo luciferase assay. Significance was determined by a Student's t-test where * = $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

Figure 3.

An F analog of an AP in the coding vs. template strand in the PZS context of the Rluc promoter near the TSS impacts transcription similar to OG. The data were obtained by transfection of the reporter plasmids into U87 cells followed by a 48-h incubation prior to conducting a dual-glo luciferase assay. Significance was determined by a Student's t-test where * = $P < 0.05$ and ** $P < 0.01$.

Figure 4.

Analysis of the impact of OG or F on two model Z-DNA forming sequences. (A) The two model sequences selected for study. (B) Monitoring the B to Z transition for the model sequences by CD spectroscopy as the [NaCl] was increased from 0.15 to 4.5 M. The example provided was derived from a study on HP-OG. (C) Plot of CD signal at 295 nm vs. [NaCl] to determine the [NaCl]_T. (D) Plot of [NaCl]_T values measured for the sequences. Significance was determined by a Student's t-test where * $P < 0.05$.

Figure 5.

The T_m values determined for the potential Z-DNA-forming hairpin HP in the native or modified state with OG or F. The values were determined in 20 mM NaPi (pH 7.4) with 20 mM NaCl by following the melting process via the change in absorbance at 260 nm. Under the low-ionic strength conditions of the analysis, the sequences were B-form duplexes. Significance was determined by a Student's t-test where ** $P < 0.01$.

Figure 6. Comparison of B- and Z-form duplexes and base pairs between G or OG and C in these helices. (A) Top and side views of B-form (pdb 1BNA)⁶² and Z-form (pdb 4OCB)⁶³ duplexes. The G•C and OG•C base pairs in a (B) B-form vs. (C) Z-form duplex to illustrate the *syn* vs. *anti* conformation of the nucleotides in these two duplexes.

Figure 7. Efficiency of APE1-mediated cleavage of an F site in different structural contexts. **(A)** Sequences for the three contexts studied, and **(B)** the reaction yields at 5 and 30 min after initiation.

Scheme 1.

Oxidative modification of G to OG in a promoter PQS is a transcriptional on/off switch. (A) Oxidation of the G heterocycle by two electrons yields OG and BER yields an abasic site (AP) in the *VEGF* potential G-quadruplex forming sequence (PQS) context. (B) Oxidation of the *VEGF* promoter PQS results in modulation of gene expression with the up or down levels of gene expression determined by the coding vs. template strand of occupancy and the PQS location in the promoter.

Scheme 2.

Oxidative modification of G to OG or OG release to AP (or F) impacts Z-DNA structure that can influence gene expression depending on the strand in which the initial oxidation occurs.

Table 1.

The structure-switching sequences studied in the Rluc promoter in either the coding or template strand.

Identifier	Sequence	Strand (Coding or Template)
Z-DNA	5'-CGCGCGCGCGCGCGCGCGCGCGCGCG-3'	Coding or Template
<i>VEGFG4</i>	5'-CT <u>GGG</u> C <u>GGG</u> CC <u>GGGGG</u> C <u>GGGG</u> TCCGGC <u>GGGG</u> - CC-3'	Coding * or Template
<i>VEGFiM</i>	5'-GG <u>CCCC</u> GCCGGA <u>CCCC</u> G <u>CCCCC</u> GC <u>CCC</u> G <u>CCC</u> AG-3'	Coding or Template *
Neg. Control	5'CT GGG G GAA CC TTGGG C GGAA TCCGGC TGGG CC-3'	Coding or Template

* The native strand of occupancy for the *VEGFG4* or *VEGFiM* sequences.