Published in final edited form as:

J Am Chem Soc. 2020 January 22; 142(3): 1115–1136. doi:10.1021/jacs.9b11050.

Interplay of guanine oxidation and G-quadruplex folding in gene promoters

Aaron M. Fleming*, Cynthia J. Burrows*

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

Abstract

Living in an oxygen atmosphere demands an ability to thrive in the presence of reactive oxygen species (ROS). Aerobic organisms have successfully found solutions to the oxidative threats imposed by ROS by evolving an elaborate detoxification system, upregulating ROS during inflammation, and utilizing ROS as messenger molecules. In this Perspective, recent studies are discussed that demonstrate ROS as signaling molecules for gene regulation by combining two emergent properties of the guanine (G) heterocycle in DNA, namely, oxidation sensitivity and a propensity for G-quadruplex (G4) folding, both of which depend upon sequence context. In human gene promoters, this results from an elevated 5'-GG-3' dinucleotide frequency and GC enrichment near transcription start sites. Oxidation of DNA by ROS drives conversion of G to 8-oxo-7,8-dihydroguanine (OG) to mark target promoters for base excision repair initiated by OG-glycosylase I (OGG1). Sequence-dependent mechanisms for gene activation are available to OGG1 to induce transcription. Either OGG1 releases OG to yield an abasic site driving formation of a non-canonical fold, such as a G4, to be displayed to apurinic/apyrimidinic 1 (APE1) and stalling on the fold to recruit activating factors, or OGG1 binds OG and facilitates activator protein recruitment. The mechanisms described drive induction of stress response, DNA repair, or estrogen-induced genes, and these pathways are novel potential anticancer targets for therapeutic intervention. Chemical concepts provide a framework to discuss the regulatory or possible epigenetic potential of the OG modification in DNA, in which DNA "damage" and non-canonical folds collaborate to turn on or off gene expression. The next steps for scientific discovery in this growing field are discussed.

Graphical Abstract

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

Introduction.

The origin of biomolecules is a fascinating question that has led to a rich set of hypotheses regarding the chemistry and macromolecular assembly of biological systems in prebiotic settings.¹ After establishment of the nucleotide monomers in DNA and RNA, an equally fascinating question asks what were the drivers of the primary sequence of these polymers. Our interests in this question are focused on the chemical and biological properties of the DNA monomer 2'-deoxyguanosine (dG) that drove sequence selection in modern genomes. The two unique properties of dG that stand out relative to the other nucleotides are the sensitivity of the guanine (G) heterocycle to oxidation, particularly in G runs, and the ability of certain G-rich sequences to adopt G-quadruplex (G4) folds.^{2,3} Recent discoveries have identified that an interplay of these two emergent properties of G can regulate transcription during oxidative and inflammatory stress.^{4,5}

Aerobic organisms carry out oxidative phosphorylation to funnel electrons from food to O_2 for reduction to H_2O establishing a proton gradient to drive ATP formation. There exist three important points regarding the utilization of O_2 in biology that pertain to the present discussion. One is that the conversion of $O_2 \rightarrow H_2O$ is not quantitative, and partially reduced oxygen species such as superoxide ($O_2^{\bullet-}$), one member of a family of reactive oxygen species (ROS),⁶ are formed.⁷ Intracellular bursts of ROS are well established to cause oxidation of all classes of biomolecules.⁸⁻¹⁰ The second observation is that nature actively produces ROS during inflammation when exposed to unwanted microorganisms or pollen, as well as during physical trauma.¹¹ The third observation is that cellular pathways harness ROS as signaling agents to chemically modify proteins to direct specific cellular tasks.^{12,13} Even under normal cellular conditions, ROS serve a vital role to maintain a healthy cellular state. In this Perspective, chemical concepts offer a backdrop to outline the interplay of G oxidation in DNA by ROS as the input signal to unmask a G4 fold during times of oxidative and inflammatory stress and high metabolic demand for modulation of cellular phenotype. These studies broaden the signaling capabilities of ROS to include oxidative modification of DNA for gene regulation.

Emergent properties of G-rich DNA sequences.

Since completion of the Human Genome Project, inspection of the sequence has uncovered many interesting features. For instance, a frequency count of dinucleotides in the human genome identified 5'-GG-3' to be the most commonly observed of the possible combinations (Figure 1A).^{14,15} In comparison to a randomized genome, the human genome has 5'-GG-3' dinucleotides at a 1.23-fold increased abundance. This significant enrichment of 5'-GG-3' sequences is highlighted by a comparison to the *Caenorhabditis elegans* (*C. elegans*) genome that is similar in G content (human ~ 20% vs. worm ~ 18%) yet possesses this dinucleotide at a frequency close to expected (1.06-fold; Figure 1A).¹⁴ Another feature of the human genome sequence is the presence of spikes in the GC content occurring around transcription start sites (TSSs) of genes.¹⁶ The overabundance of GC base pairs near TSSs and 5'-GG-3' dinucleotides in the human genome results in new emergent chemical features in the genome. In this context, “emergent” properties are properties in DNA that are different from the sum of the nucleotide monomers.

The first emergent property that G-rich regions, and particularly 5'-GG-3' dinucleotides, bring to DNA is a greater sensitivity to oxidation.^{2,17,18} Of the four canonical bases in DNA, the G heterocycle is the most electron rich, and therefore, the most prone to oxidation (Figure 1B).¹⁹ When G bases π stack upon one another, the ionization potential of the 5'-most G is lowered leading to a greater sensitivity to oxidation at a 5'-GG-3' site (Figure 1B).² The π -stacking effect becomes more dramatic as the length of the G run increases from two to four on the basis of theoretical calculations by Saito and co-workers.² Another feature of duplex DNA is the ability for an electron hole (i.e., oxidation site) to migrate through the π -stacked network to effect an oxidative chemical reaction at the most prone site, such as the 5' G in 5'-GG-3' steps (Figure 1C; electron hole = h^+).^{17,18,20,21} The ability of duplex DNA to funnel electron holes to oxidation-susceptible sites has been experimentally validated numerous times.^{17,18,20-25} This property of duplex DNA led Heller to hypothesize that the overabundance of Gs in introns near boundaries serves as “cathodic protection”,²⁶ reasoning that electron holes generated by oxidation funnel damage to introns to complete the reaction so that the chemically modified nucleotide would not impact the coding potential of exons.^{26,27} More recent studies are finding that oxidation of Gs in regulatory regions, such as gene promoters where they are also overly abundant, can function as chemical marks for modulation of gene expression.²⁸⁻³⁰ In some respects, this is similar to Heller's hypothesis in that oxidation of the DNA outside of protein coding regions of the genome is a desirable genomic feature derived from selection of the G-rich sequence.

A second emergent property that occurs with an increased frequency of 5'-GG-3' dinucleotides is a greater number of sequences with the potential to adopt G4 folds.^{3,31-33} Potential G-quadruplex sequences (PQSs) occur in DNA when there is a high local density of G runs. G-Quadruplexes are non-canonical DNA or RNA folds that differ in their structural and physical properties when compared to duplexes, and they can be bound by an alternative set of regulatory DNA-binding proteins. As described in greater detail below, these folds can function as structural switches to guide cellular physiology, particularly at the point of regulating when and to what extent mRNAs are synthesized. The redox

sensitivity of G-rich sequences and the propensity for some of these to adopt G4 folds for gene regulation appear to be advantageously selected in mammals.

Guanine oxidation is inevitable when living in an oxygen atmosphere.

The concentration of O_2 in buffer that is equilibrated with the atmosphere is $\sim 200 \mu M$; this value is lower in vivo under physiological conditions (i.e., normoxia or physoxia), and ranges from ~ 20 - $100 \mu M$ O_2 depending on the tissue type.³⁴ To meet metabolic demands, the flux of O_2 into a cell can be very high. Hegg and co-workers found $\sim 70\%$ of intracellular H_2O in *E. coli* during log-phase growth was generated from O_2 by the electron transport chain.³⁵ While the electron transport chain is highly efficient, release of partially reduced O_2 in the form of $O_2^{\bullet -}$ occurs with a yield of 0.1-2% depending on the cellular state (Figure 2).⁷ With the high flux of O_2 through mitochondria, even at low yields, $O_2^{\bullet -}$ is generated in substantial amounts. Superoxide has a short cellular half-life ($\sim 1 \mu s$) and undergoes disproportionation catalyzed by superoxide dismutase to furnish hydrogen peroxide (H_2O_2),¹² an ROS with a longer half-life ($\sim 10 \mu s$) prior to catalase disproportionation of two H_2O_2 equivalents to yield two equivalents of H_2O and one of O_2 .¹² The steady-state concentration of H_2O_2 is ~ 200 nM and can increase many fold under high metabolic demand or cellular disease.¹² These numbers highlight primary cellular tasks that can generate significant concentrations of ROS inside cells. The oxidizing power of ROS is balanced by a competing system of enzymes and small molecule reductants to quench inappropriate oxidation of biomolecules.³⁶ During times of stress, high metabolic demand, such as during intense exercise, or disease, the ROS load can overburden these defense systems.^{11,37,38}

Nature has found ways to harness ROS in cellular processes via natural selection. Hydrogen peroxide is a signaling agent for immune cell activation and vascular remodeling.³⁹ Additionally, studies identified that the antioxidant proteins thioredoxin, nucleoredoxin, and peroxiredoxin sense H_2O_2 via redox-active cysteine residues for induction of cell apoptosis, modulation of the androgen receptor, and function in circadian rhythms.^{40,41} A chemical side effect of H_2O_2 is the Fenton reaction that occurs when an iron or a copper ion reacts with H_2O_2 to yield the potent one-electron oxidant hydroxyl radical ($HO^\bullet$; $E_{red} = 2.31$ V vs. NHE; Figure 2) for indiscriminate oxidation of all biomolecules.⁹ Oxidation of DNA and G nucleotides by $HO^\bullet$ yields many products resulting from initial oxidation at the heterocyclic bases or the sugar-phosphate backbone.⁴²⁻⁴⁵ In duplex DNA, the $HO^\bullet$ -mediated oxidation is favored on the base, while the ultimate reaction site leading to products is highly dependent on the sequence context, O_2 concentration, and presence of relevant concentrations of reductant.^{44,46} However, the amount of DNA damage assignable to $HO^\bullet$ may be less than previously suspected; Meyerstein and coworkers recently showed that in the presence of physiological bicarbonate (~ 40 mM), the Fenton reaction principally yields carbonate radicals, $CO_3^{\bullet -}$ (see below).⁴⁷ Thus, iron, copper, and ROS, such as H_2O_2 , are tightly regulated in cells, and the latter can become unregulated to reach elevated concentrations during oxidative and inflammatory stress leading to carbonate radical chemistry with DNA.^{12,39,48}

When biological systems must combat threats from harmful sources such as bacteria, viruses, pollen, trauma, or toxins, they release bursts of ROS and reactive nitrogen species

(RNS) in an inflammatory response.^{9,11,49} Beyond the overproduction of $O_2^{\bullet-}$ by various oxidases, the free radical nitric oxide ($^{\bullet}NO$) is released by nitric oxide synthase, allowing reaction of these two radicals in a diffusion-limited step to generate peroxynitrite anion ($ONOO^-$).⁸ Peroxynitrite reacts with cellular CO_2 to yield short-lived nitrosoperoxy carbonate anion ($ONOOCO_2^-$). Homolytic degradation of $ONOOCO_2^-$ yields the one-electron oxidants carbonate radical anion ($CO_3^{\bullet-}$; $E_{red} = 1.59$ V vs. NHE) and nitrogen dioxide radical ($^{\bullet}NO_2$; $E_{red} = 1.04$ V vs. NHE) in $\sim 70\%$ yield (Figure 2).⁵⁰ The Geacintov and Shafirovich team found that $CO_3^{\bullet-}$ specifically oxidizes the G heterocycle to furnish OG in high yields.⁵¹ These inflammation-derived ROS and RNS target invading agents to minimize their threat via oxidation, and collectively, they oxidize all biomolecular components including the DNA nucleotides.^{8,50}

Of the four canonical nucleotides, the G heterocycle is the most electron rich resulting in G being the most susceptible to oxidative modification (Figure 1B; $E_{red} = 1.29$ V vs. NHE).¹⁹ Products of G oxidation by ROS were reported by our laboratory and others.^{24,52-61} A consistent set of products has been observed that include the two-electron oxidation products 8-oxo-7,8-dihydroguanine (OG) and 2-iminohydantoin (2Ih), as well as the four-electron oxidation products spiroiminodihydantoin (Sp), 5-guanidinohydantoin (Gh), and imidazolone/oxazolone (Iz/Z; Figure 2). At present, the results of studies in model cells or organisms living under aerobic conditions have found OG in the highest yields within genomes at ~ 1 in 10^6 Gs followed by Z, Sp, and Gh at levels around two orders of magnitude lower than OG.⁶²⁻⁶⁴ Oxidative and inflammatory stress increase the levels of oxidized nucleotides in the genome.⁶² The changes in OG concentration are routinely used as a biomarker for quantification of the impact of stress on cells.⁶³

Direct oxidation of G to OG via oxidation can occur as previously described to “write” OG into a genome. Alternatively, indiscriminate one-electron oxidation of duplex DNA can yield an electron hole that can migrate through unperturbed duplexes allowing oxidation at a low ionization potential site such as a 5'-GG-3' step (Figure 1B).¹⁷ A characteristic signature of hole migration leading to G oxidation is enhanced reaction at the 5' most G of the G run as a consequence of the π -stacking effect (Figure 1C).² The Barton laboratory recently demonstrated that charge transport (i.e., electron hole migration) can occur over thousands of base pairs in a redox exchange between DNA repair proteins.⁶⁵ In model duplex DNA, remote ROS-mediated oxidation followed by hole migration to a susceptible 5'-GG-3' for termination by a chemical reaction is mapped by hot piperidine cleavage and gel analysis.⁶⁶ The 5' G oxidation effect diagnostic of hole migration occurs during photoredox catalysis (e.g., riboflavin + $h\nu$), $CO_3^{\bullet-}$ -mediated oxidations, and was modestly observed in $HO^{\bullet}$ -mediated oxidations.⁶⁷ The Barton laboratory treated HeLa chromatin with a rhodium photooxidant and then mapped the oxidation sites in the G-rich promoter of the transcriptionally active phosphoglycerate kinase gene. Using a PCR-based approach to identify the oxidized Gs, the team found that the relative reactivity decreased from 5'-GGG-3' > 5'-GG-3' > 5'-GA-3' verifying electron hole migration can occur in chromatin.⁶⁸ Recently, the Greenberg laboratory found electron hole migration in nucleosomes shows dependency on the sequence and the site of hole injection relative to specific histone tyrosine interactions with the DNA.⁶⁹ Nonetheless, oxidation of DNA results in favorable oxidation of G runs like those found in PQSSs.

A recent test to understand whether oxidation of G favors 5'-GG-3' sites in genomic DNA was provided by the Sturla laboratory when they sequenced the yeast genome for OG at single-nucleotide resolution.⁷⁰ The sequencing work found >50% of the sequenced OG was located 5' to a G nucleotide, consistent with oxidation occurring remotely and the corresponding hole migrating to an easily oxidized G run. Lastly, the favorability of G oxidation occurring within PQS contexts is gleaned from OG sequencing at >100-nt resolution by OG-seq in our laboratory, OxiDIP-seq by Amente et al., OG ChIP-seq by the Gillespie laboratory, and enTRAP-seq by Fang et al., in which OG-enriched regions in mammalian DNA were favorably biased to reside in gene promoters and within PQSs;⁷¹⁻⁷⁴ a limitation of these sequencing studies as a result of the low resolution for locating OG is not knowing the exact location in the sequenced fragment in which the modification resides. Park et al. utilized a PCR-based approach to identify OG is formed in specific genes in mammalian cells experiencing oxidative stress in tandem with upregulated expression.⁷⁵ These sequencing results likely provide a lower bound on the number of PQSs that harbor OG because it is well established that G-rich regions poorly PCR amplify during library preparation for sequencing.⁷⁶ Collectively, the OG-sequencing reports highlight sites of G oxidation to OG can occur at specific low-energy sequences, such as G-rich PQSs, and this may occur with gene selectivity.

G-Quadruplex folds in the human genome: How, where, and why?

A region of the genome is a PQS when there are at least four G runs with at least three Gs per run in close proximity to one another (Figure 3A; ~50-nucleotide window with loop regions of ~12).⁷⁷⁻⁷⁹ These sequences enable one G from each run to embrace one another in Hoogsteen base pairing to form a G-tetrad (Figure 3B), and because there are at least three Gs per run, at a minimum, three stacked G-tetrads are formed.^{3,31-33} Assembly of G-tetrad stacks orient the unbonded electron pair on O⁶ of each G to be positioned in a way that forms a channel with dimensions appropriate for coordination of potassium ions to stabilize the structure. Potassium ions are preferentially coordinated because they exist in the highest intracellular concentration (~140 mM) and provide the best size match for the channel generated by the G-tetrad stack in cells.⁸⁰ The core of G-tetrads is maintained by the intervening nucleotides that are the linkers or loops holding the structure together. There exist two global G4 types on the basis of the direction in which the subsections of the sequence are oriented. Parallel G4 folds have each G-track oriented in the same 5' to 3' direction, and antiparallel folds mix the 5' to 3' orientation of the G runs (Figure 3C). Specific nomenclature exists for the G4 folds in the telomere repeat sequence, in which the parallel fold is called propeller, and the antiparallel fold coordinated to potassium ions is called hybrid and exists in two different conformations.³³ The glycosidic bond angle of G (*syn* vs. *anti*) is a key driver for the fold type in which parallel G4s have all *anti* Gs, and antiparallel-stranded G4s contain a mixture of *syn* and *anti* Gs.³³ Additionally, the fold type directs the loop configuration between adjacent G runs. In parallel folds, the loops are double-chain reversals, and in antiparallel folds the loops are diagonal, edgewise, and/or double-chain reversals (Figure 3C).³¹ These G4 qualities combine to yield a characteristic structure with fine features determined by the primary sequence presenting different motifs for protein recognition.

When there are more Gs in a run than required for core G-tetrad formation, or there are more than four G runs, dynamic structures are inevitable. In DNA G4s, three tetrads are the usual core size—but exceptions do exist—and when a run has four or more Gs, selection of three to form the core occurs leaving the others to occupy loop positions.^{3,31} This scenario sets up the possibility of a heterogeneous population of folds in solution by selection of different Gs to participate in the core. Many biological sequences fall into this category by having more Gs than strictly required for G4 formation.⁸¹⁻⁸⁴ Another observation from our laboratory is that many human PQSs, in addition to the telomere repeat sequence, possess more than the necessary four G runs required for folding, with an extra track on one side or the other of the principal G4-folding region (Figure 3A).^{82,84} As described below, we hypothesize these additional G runs provide a backup system to ensure G4 folding when the sequence is chemically modified in a way that negatively impacts the principal G4 structure. Said another way, relevant PQSs harness excess Gs for greater plasticity to ensure folding when constrained by other competing factors, such as incompatible base modifications, in the genomic context.

Global identification of PQSs in the human genome was approached with bioinformatic tools inspecting for the classical consensus sequence by assuming a limit of up to seven-nucleotide loop lengths (Figure 3A). The seven-nucleotide loop length set a lower limit of >375,000 PQSs in the human genome.^{77,78} Experimentally, the number of G4 folds found by the Balasubramanian laboratory using G4-Seq was >700,000.⁸⁵ The experimental versus theoretical difference observed resulted from the presence of G4s with longer loops than seven nucleotides or bulges intervening in the G runs in the experimental cohort. Mapping these PQSs to the genomic features in which they occur has identified an abundance of them in important gene regulatory regions.⁷⁹ Enrichment of PQSs occurs in gene promoters and 5'-untranslated regions (5'-UTRs), and predominantly in the first intron of many genes; additionally, PQSs are found in G-rich repetitive telomeres and satellite DNA (Figure 4A).^{33,79,86-88} Closer inspection of PQS distributions around TSSs of human genes found enrichment of these sequences strongly suggesting they were selected for gene regulation (Figure 4B). The presence of PQSs in regulatory regions positions them in prime genomic locations to impact RNA biogenesis if and when they fold.

The ability of PQSs to adopt G4 folds in human chromatin was demonstrated by immunofluorescent visualization in cell nuclei.^{90,91} Beyond the presence of G4s in the single-stranded telomere regions where they are free to fold, they were also found in chromosome interiors in which the C-rich complementary strand provides a barrier to folding by favoring the duplex state; thus, cellular mechanisms must exist to overcome this energetic barrier. A final observation was that G4s maximally fold during the replication phase of the cell cycle (S phase).⁹⁰ During S phase, active replication and transcription occur that open the duplex and relieve the thermodynamic constraint for G4 formation; however, this is not a complete picture because folded G4s in chromosome interiors were also found in the quiescent cell state (G₀). Lastly, the Balasubramanian laboratory developed G4-ChIP-Seq that was initially applied to human cells under normal growth conditions to identify ~10,000 folded G4s in the genome.⁹² These were biased toward gene promoters and 5'-UTRs in which they appear to function as up or down regulatory elements of transcription. Our laboratory inspected the pool of folded G4 sequences to find ~60% had five or more G

tracks, while ~40% of all PQSs found computationally had 5 or more G tracks that suggests PQSs that fold to G4s in cells favor the flexibility in structure afforded by the additional G run.⁸⁹ The number of folded G4s found by G4-ChIP-Seq was nearly two-orders of magnitude less than the folded G4s found by G4-Seq;⁸⁵ therefore, a large genomic reserve of untapped PQSs for G4 folding exists for gene regulation when cells experience stress.

Studies targeting specific G4s in the genome suggest a wide variety of functions for these non-canonical structures. On one hand, G4 folds are an impediment to polymerase bypass on template strands that require suppression of folding by specialized helicases.^{93,94} On the other hand, G4s were demonstrated to function in gene regulation. The Hurley laboratory conducted a thorough analysis of the *c-MYC* promoter to identify negative supercoiling sensory input for switching topological forms from duplex to a non-canonical structure, leading to differential binding of effector proteins.⁹⁵ The work also highlighted a key point that will be addressed later, which is that folding of the C-rich complementary strand to an i-motif fold can occur as the regulatory non-canonical structural unit.⁹⁶ Many other proto-oncogenes (e.g., *VEGF-A*, *SRC*, *c-KIT*, etc.) contain promoter PQSs that can fold to G4s and have been identified for regulation of gene expression.⁹⁷⁻⁹⁹ There is mounting evidence for G4s folding at origins of replication to aid in activation of replication.¹⁰⁰ In the genome, folded G4s are associated with decreased writing of the epigenetic modification 5-methyl-2'-deoxycytosine (5mC) into DNA and can impact writing of chemical modifications on histones when the G4s are ligand stabilized.¹⁰¹⁻¹⁰³ Our studies highlight that the G-richness of PQSs render them sensors of oxidative stress to initiate G4 folding as a structural switch for gene regulation.⁴ G-Quadruplexes in DNA appear to have a wide variety of functions in channeling cellular activity, and the list will likely continue to grow.

G-Quadruplex structural plasticity accommodates base modifications.

One consequence of oxidatively modifying PQSs is that the resulting products can perturb the structural integrity of the corresponding G4 fold. The structural impact of a G to OG conversion on a G4 fold was showcased in the four-track human telomere sequence that utilizes all Gs in G-tetrad formation.¹⁰⁴⁻¹⁰⁷ Thermal melting analysis identified a >15 °C loss in stability when OG was present. From a molecular recognition perspective, converting the N7 hydrogen bond acceptor on G to a hydrogen bond donor at the same position in OG, and introduction of a more sterically demanding *O*⁸ carbonyl, lead to the thermal changes observed (Figure 5A). Work in our laboratory showed that with more complete models of the human telomere with a fifth G track present, the additional G track could exchange with the OG-bearing track to maintain the stability of the fold.^{82,84} This result suggests that the plasticity of native G4s can accommodate structurally perturbing oxidative chemical modifications. A similar structural impact was observed in the human telomere sequence when inosine,⁸³ an abasic site, or Gh was placed at a core site in a G4 with >4 G tracks.^{104,108}

The plasticity of promoter G4 folds is dependent on the specific sequence. In general, human regulatory promoter G4s favor shorter loops; however, it is common that they have a run of Gs longer than needed for folding, and about half have five or more G runs.⁸¹ Promoter G4 dynamics is evident in NMR analysis of the native sequences yielding poorly defined imino

spectra in which the polymorphism is typically minimized to a single species upon judicious mutation of excess Gs to Ts in G runs >3 nucleotides long and removal of extra G runs;^{109,110} this approach aids in identification of principal G4 folds. Analysis of the *VEGF* promoter G4 in our laboratory and the *KRAS* promoter G4 in the Xodo laboratory, which both have excess Gs in the principal G runs and five or more G runs total, identified that OG is structurally well accommodated as a result of the spare Gs (Figure 5B)^{5,82} enabling the modified nucleotide to reside in a loop position that does not impact the core of Gs involved in G-tetrad formation. We hypothesize the overabundance of G runs needed for principal G4 folding provides a selective advantage to serve as a “spare tire” or backup to preserve the G4 fold when oxidatively modified.⁸² Engagement of a spare tire domain was documented for the G4 destabilizing modifications OG, abasic sites, or Gh in human promoter PQSs;^{82,108,111} additionally, recovery of an OG-destabilized promoter G4 can occur via introduction of a pyrene-conjugated G-rich oligonucleotide.¹¹² In comparison, OG is well tolerated in duplex DNA while abasic sites and Gh/Sp significantly impact the helix stability.¹¹³⁻¹¹⁶ In summary, a small number of chemically modified G sites studied to date indicate they can significantly impact duplex stability while the flexibility of G4 folds enables residence of these modifications in a loop without globally impacting the fold.

Historically, OG is exclusively a mutagenic lesion in the genome.

The mutagenic nature of OG has emerged from a number of in vitro and in vivo studies. As our arguments develop, we support all of these observations; however, as will be pointed out below, it appears as though these studies represent only one facet of its cellular role. In duplex DNA, OG base paired with C minimally perturbs the structure via a slight steric clash between the *O*⁸ carbonyl and the phosphodiester backbone (Figure 6). The impact is slight in thermal melting studies (−1 °C),¹¹³ is silent in NMR analysis,¹¹⁷ and observable in high-resolution crystallographic structures.¹¹⁸ The steric clash between OG and the backbone is notable because events that open the duplex, such as replication or transcription, allow the OG heterocycle to undergo an *anti* to *syn* conformational change relative to the sugar to relieve the steric repulsion. Consequently, the *syn* conformation of OG presents the Hoogsteen face of the purine for complementary hydrogen bonding with the Watson-Crick face of A leading to a stable OG:A base pair (Figure 6).¹¹⁹ The OG:A base pair causes a small (~3 °C) decrease in thermal stability and has minimal impact on duplex structure;¹¹³ however, during replication, a second polymerase bypass allows the template A to direct T insertion in the strand originally occupied by OG. This leads to a G→T transversion mutation that is a signature of a G oxidation to OG at a given site (Figure 6).¹²⁰

Molecular changes associated with an OG:C compared to a G:C base pair occur in the major groove of duplex DNA. The difference in donor-acceptor pattern on the imidazole ring of G versus OG impacts reader protein binding. In model duplex DNA, targets with OG synthesized within the known consensus sequences for the transcription factors SP1,¹²¹ NF-κB,¹²² or CREB,¹²³ and the methyl-CpG binding domain MBD of methyl-CpG binding protein 2 (MeCP2)¹²⁴ found decreased association constants of 4- to 100-fold depending on the sequence and protein. A conclusion from the Strauss laboratory with CREB was that OG could be epigenetic by functioning as a transcriptional repressor.¹²³ The Sugden laboratory studied NF-κB binding and concluded OG is either beneficial or aberrant depending on the

gene product that is formed from the OG-containing promoter.¹²² The Cadet laboratory studied SP1 binding and concluded that OG may alter gene expression.¹²¹ Lastly, when OG resides in transcribed regions of DNA, the transcriptional impact is dependent on the coding versus template strand of OG occupancy, the cell type studied, and whether the DNA repair status of the cell was functioning normally, for which OG can stall or abolish mRNA synthesis.¹²⁵⁻¹²⁸

The biological significance of OG is amplified by evolution of two DNA repair glycosylases (i.e., reader proteins) that specifically recognize this oxidized heterocycle: OG-glycosylase 1 (OGG1) reads the OG:C base pair and MutY homolog (MUTYH) reads the OG:A base pair.¹²⁹ In Ogg1^{-/-} mouse models, the genomic OG levels increased as aging progressed (7-fold increase at 14 weeks old),¹³⁰ but Ogg1^{-/-} mice successfully developed with no significant change in lifespan or increase in tumor frequency. However, they showed an attenuated response to inflammatory agents.²⁹ On the other hand, Mutyh^{-/-} mice are more prone to tumor formation, and oxidative stress enhances the tumor count formed in these mice relative to wild-type controls.¹³¹ Humans with a biallelic germ-line mutation in the MUTYH gene harbor an elevated level of G→T transversions and greater colorectal cancer incidence.¹³² Thus, OGG1 and MUTYH have different impacts on cells, in which the OG reader OGG1 has no clear-cut ties with cancer, but appears to function in the cellular response to inflammation; however, inappropriate base pairing of OG with A is correlated with cancer as seen in the MUTYH mutant studies. In summary, OG can be mutagenic under certain circumstances when base paired with A, is benign in some contexts, or can be impactful on mRNA expression when located in transcription factor binding sites.

There exist two key insights derived from the knowledge of OG mutagenesis and its repair in cells. (1) The mutation frequency of OG under normal physiological conditions is very low (~0.1% in U2OS human osteosarcoma cells¹³³ and <2% in *E. coli*¹³⁴) as a result of faithful DNA repair pathways that avoid mutagenesis.¹²⁹ *Thus, OG is not the “highly mutagenic” nucleotide that has dominated the historical literature regarding this DNA modification.* (2) Over the course of evolution, G has not been systematically removed from the human genome via G→T transversions;¹³⁵ if anything, G has been maintained with significant enrichment as noted by the increased frequency of oxidation-prone 5'-GG-3' dinucleotides (Figure 1A).^{14,15} This final point will reappear below.

Faithful base excision repair of OG provides entry for transcription proteins.

A commonality shared by DNA repair and transcriptional proteins is intimate interaction with DNA, and therefore, a blurring of the functions for these proteins is not surprising. The intertwined roles of the major DNA repair pathways and transcription have been reviewed;¹³⁶ here we focus on the points of base excision repair (BER) of OG at which transcriptional regulation may occur. The glycosylase OGG1 only removes the OG heterocycle from an OG:C base pair, while MUTYH only removes A from an OG:A base pair in order to provide a polymerase an opportunity to insert C opposite yielding the OG:C base pair that is an OGG1 substrate (Figure 7).^{129,137,138} Interestingly, MUTYH has some unusual implications

in cancer suggesting a broader role in DNA damage response that was recently reviewed.¹³⁹ The following discussion will focus only on the OG:C base pair and repair by OGG1.

Release of the OG heterocycle mediated by OGG1 in vitro is described as bifunctional resulting in a strand break in the DNA, but whether the second step occurs in cells is not well established.¹⁴⁰ In cellulo, mechanistic examination suggests OGG1 is monofunctional yielding an abasic site (AP) that is handed off to apurinic-apyrimidinic endonuclease 1 (APE1; Figure 7) for strand cleavage.¹²⁵ Studies in vitro have recreated the stimulation of OGG1 by APE1 suggesting these proteins can function together,^{141,142} although the complete mechanistic details of OGG1 remain an open question. Nonetheless, after OG release yields an AP or strand break, both are substrates for APE1. For the intact AP, APE1 hydrolyzes the 5'-phosphodiester bond of the AP to yield a 3' hydroxyl on the 5' side of the nick and a 5'-deoxyribosephosphate (5'-dRP) on the 3' side of the nick (Figure 7).¹⁴³ Next, polymerase beta (Polβ) has two functions.¹⁴³ First, the lyase activity removes the 5'-deoxyribosephosphate, and then in the second step with the assistance of XRCC1, the single-nucleotide gap is filled with dGTP via the polymerase activity (Figure 7). In the final step, DNA ligase III (LIG3) seals the nick to yield a repaired strand (Figure 7).¹⁴³ Each of these points provides a substrate, protein member, or both for cross-interaction with transcription initiation.

Epigenetic-like OG in gene promoters allows interplay between DNA repair and transcription initiation.

The initial cellular observation of controlled G oxidation in a gene promoter driving productive transcription was reported by the Avvedimento laboratory (Figure 8A).¹⁴⁴ They found estrogen-responsive genes were subject to promoter DNA oxidation delivered during receptor-targeted demethylation of H3K9me2 by lysine-specific histone demethylase 1A (LSD1 also KDM1A). The chromatin remodeler LSD1 is a flavin-dependent monoamine oxidase that generates H₂O₂ from O₂ to reestablish the flavin catalyst during the demethylation process.¹⁴⁵ The locally formed H₂O₂ drives G oxidation in discrete genomic loci as determined by immunofluorescence in cellulo. The oxidation events lead to recruitment of OGG1 and topoisomerase II beta (TOP2B; Figure 8A) to promoters for induction of transcription. Localized OGG1-mediated release of OG in the promoter was proposed to provide an entry point for TOP2B for gene activation (Figure 7). A recent review highlighted studies suggesting transcriptional activation by TOP2B may occur via interactions with CCCTC-binding factor (CTCF) and cohesin;¹⁴⁶ however, more work is needed to better understand the TOP2B interactome. Avvedimento's studies support OG as the DNA oxidation product because OGG1 recruitment to promoters was essential for gene activation, in which *BCL2* was one gene found upregulated by this mechanism.¹⁴⁴ Their subsequent work identified retinoic acid-induced genes follow a similar induction via DNA oxidation driven by the Jumonji C domain-containing histone demethylases.¹⁴⁷

Our inspection of the data found that the *BCL2* PQS is in the vicinity of the oxidation event, suggesting there may be a role for OG formation and BER processing in this G-rich region (Figure 2).⁴ Prior work reported that the *BCL2* PQS and potential i-motif sequence can

adopt non-canonical folds, and studies have shown the i-motif is key to guiding transcription of this gene.^{96,150,151} Thus, oxidation may provide the signal for the structural switch (duplex→G4 and/or i-motif) to induce transcription, as we discuss in more detail below. Future studies are needed to clarify this point.

The finding of LSD1-mediated DNA oxidation begs many questions. Why would evolution have selected a chromatin remodeler to deliver a ROS in close vicinity to the genetic material to promote DNA oxidation? Does this provide access to the emergent properties of G nucleotides, such as G4 folding, for regulation of cellular processes? Further, there exist many examples of oxidases functioning on or in close proximity to the genome, and enzymes harboring Fe-S clusters capable of conducting redox chemistry near DNA. Examples of proteins containing a Fe-S cluster that function on DNA include MUTYH, XPD, and DNA polymerases (α and δ families),¹⁵² and these enzymes can be damaged (i.e., oxidized) by $O_2^{\bullet-}$ to liberate Fe(II) near DNA.¹⁵³ Additionally, the Fe(II)- and α -ketoglutarate-dependent Jumonji C domain-containing histone demethylases and the ten-eleven translocation methylcytosine dioxygenases (TET) that oxidize 5mC all function near or on DNA. A ready source for Fe(II) release near DNA can drive the Fenton reaction to release $HO^{\bullet}$ and oxidize G nucleotides;⁴⁴ however, the indiscriminate oxidation of DNA by $HO^{\bullet}$ yields such a diverse set of damage products that it would be nearly impossible to regulate a coherent cellular response.^{43,44,129,154} If the major ROS from the cellular Fenton reaction is $CO_3^{\bullet-}$,⁴⁷ then G^{++} formation and long-range charge transport would focus the modification at a PQS. Alternatively, reactive intermediates in Fe(II)- α -ketoglutarate enzymes could conduct one-electron oxidation of DNA to yield a specific product at a specific site to impact gene expression. Future studies are needed to address these questions.

Boldogh and Ba explored DNA oxidation in gene promoters followed by OGG1-mediated induction of inflammatory response genes in cellulo and in mouse models (Figure 8B). They stimulate an inflammatory response of ROS (e.g., $CO_3^{\bullet-}$) in mammalian cells with the cytokine tumor necrosis factor alpha (TNF α).^{29,148,155} Oxidized regions in the genome were then mapped using qPCR before and after Fpg treatment (a BER enzyme with a large substrate scope for G oxidation products)¹⁴³ to find the proinflammatory gene promoters for TNF, CCL20, and CXCL1 as example sites of oxidation. Using PCR-based methods, they demonstrated that induction occurred via OGG1 binding to the oxidized promoters prior to NF- κ B binding to its consensus sequence at a site adjacent to bound OGG1. Additional studies found OGG1 is coupled to the DNA occupancy of NF- κ B that suggested to them an epigenetic function of OG and OGG1. A key component of their pathway was oxidation of a cysteine residue in OGG1 (OGG1^{ox}) that attenuated the glycosylase activity while the DNA binding ability remained; this posttranslational modification allowed stalling of the glycosylase on DNA for recruitment of NF- κ B and gene induction. Prior work identified cysteine oxidation in OGG1 modulates repair capacity in cellulo.¹⁵⁶ The importance of cysteine oxidation as a modification in DNA repair proteins that alters their activity was recently reviewed to illustrate the broad scope of posttranslational modifications impacting DNA repair proteins.¹⁵⁷ The work of Boldogh and Ba demonstrated that the susceptibility of G in a gene promoter toward oxidation can modulate gene expression via the BER glycosylase OGG1 providing an entry for site-specific recruitment of the activating transcription factor NF- κ B.

In the third set of studies, induction of the *VEGF* gene under oxidative stress nicely illustrates the two emergent properties of G coming together as a regulatory switch in the genome. Initial studies by the Gillespie laboratory found using a Fpg-coupled qPCR approach that oxidative stress induced by hypoxia in rat pulmonary cells caused DNA oxidation in G-rich gene promoters such as the *Vegf* gene.^{73,158} In sync with the promoter oxidation was recruitment of the DNA repair proteins Ogg1, Ape1, and the transcription factor Hif-1 α leading to a 300% increase in *Vegf* mRNA. Their work suggested the BER pathway links hypoxia-induced promoter oxidation and gene activation. Inspired by their biological studies, we used chemical tools to explore the molecular details driving *VEGF* gene induction in human cells via oxidation of G in the promoter in cellulo.

In the human *VEGF* gene promoter, there is a PQS positioned at -50 in the coding strand; a similar sequence exists in the rat gene. Prior work in the Hurley laboratory demonstrated that the PQS can fold in the plasmid context, and G4-binding ligands added to cell culture impact *VEGF* expression levels.¹⁵⁹ Further, the Maizels laboratory reported that the G4-specific helicase XPB produced a positive ChIP-Seq signal for the *VEGF* promoter.¹⁶⁰ These studies support the proposal that the *VEGF*PQS can adopt a G4 fold in plasmids and chromatin. Thus, we first studied the oxidation chemistry of the *VEGF*PQS to locate the hotspot sites of oxidation at the 5' Gs in G runs.⁸² Yields of OG were significant during oxidation reactions conducted under conditions that modeled the reducing nature of the cellular context. This knowledge enabled us to synthesize OG site specifically at oxidation prone positions in the context of the *VEGF*PQS in a promoter driving a luciferase expression vector.⁴ Transfection of the OG-containing vector into mammalian cells led, after incubation, to a 300% increase in expression levels compared to cells expressing the vector without OG. This study was the first to demonstrate that OG at a specific location in a gene promoter can induce transcription (Figure 8C) and the first to link G oxidation in a promoter G4 as a key event in gene induction. Additionally, our findings in a chemically-defined system reproduced the biological results reported by Gillespie's laboratory.¹⁵⁸ We continued the studies to verify G oxidation to OG at other hotspot sites in the *VEGF*PQS led to the same induction of the reporter gene.⁴ This suggests that the PQS functions as a unit, and oxidation can occur anywhere within the G-rich sequence to induce gene expression.

Utilizing the chemically-modified vectors in cellulo and synthetic models in vitro, we proposed a BER-dependent pathway harnessing a G4 fold for gene induction (Figures 8C and 9).⁴ The promoter DNA exists as a duplex in the resting state. Upon oxidation of duplex DNA, the electron hole generated migrates through the π -stacked DNA base pairs to a low energy G-run that is part of a PQS. A G run in the PQS is oxidized at a 5' G to yield OG that slightly destabilizes the duplex. Promoter OG provides the signal for recruitment of OGG1 to bind and release the oxidized heterocycle yielding an AP in the duplex (Figures 7 and 9). The AP provides the thermodynamic drive enabling the duplex $\rightarrow$ G4 structural switch because an AP significantly impacts duplex stability (>15 °C), while as a result of G4 plasticity (Figures 5 and 9), the AP resides in a loop position and does not decrease the thermal stability of the non-canonical fold. This duplex $\rightarrow$ G4 structural switch displays the AP to APE1 for binding, but the structural context attenuates endonuclease activity of this protein.^{108,161} Prior work demonstrated APE1 via its redox-effector factor 1 (Ref-1) function interacts with HIF1- α or AP-1 for gene activation,^{162,163} in which the importance

of HIF1- α was also noted in the work by Gillespie's laboratory.¹⁵⁸ Thus, we propose stalling of APE1 on the AP-bearing G4 provides a platform for HIF1- α recruitment and gene activation; alternatively, APE1 may bind the folded G4 preloaded with HIF1- α , for example, leading to gene activation.⁴ Future studies are needed to address the final choreography of events regarding APE1 interaction with activating protein transcription factors leading to gene induction.

Cellular support for the stalling of APE1 on the genome was found when a plasmid bearing a non-cleavable AP analog in a sequence that could still adopt a G4 fold was transfected in the cells; we found a 1500% greater expression level in this study that strongly implicates APE1 stalling on the DNA as a key factor leading to enhanced gene expression.⁴ After induction, the repair process eventually goes to completion, and gene expression returns to basal levels.¹⁶⁴ A key control was to repeat the work and synthesize OG into a promoter that retained the native transcription factor consensus sequence but was incapable of G4 formation; this experiment led to a null result. This final observation supports the role of a G4 fold in the gene activation process.^{4,165} Finally, when the spare tire domain was omitted from the sequence, gene expression was attenuated, supporting the need for the 5th domain for maximal gene expression with OG.⁴ The *VEGF* work demonstrates that both emergent properties of G—sensitivity to oxidation and propensity to adopt G4 folds—collaborate to induce transcription under oxidative stress conditions. Furthermore, APE1 is the central hub protein to bridge a non-canonical DNA fold and activating transcription factors together.¹³

Continued efforts with OG-modified luciferase reporters transfected into mammalian cells provided us with the opportunity to explore with greater detail the spatial and PQS constraints on gene expression. Next, we placed the *VEGFPQS* and OG in the non-transcribed template strand and found the presence of OG downregulated luciferase expression with dependency on Cockayne syndrome B protein (CSB or ERCC6).¹⁶⁴ The observation of a role for CSB suggests OG embedded in the template strand can initiate transcription-coupled nucleotide excision repair because this protein is fundamental for activation of this pathway. Further, these results are consistent with prior work showing OG on template strands and folded G4s in this location negatively impact transcription.^{111,125} Movement of the PQS bearing OG further from the TSS showed that up- or downregulation could occur out to ~200 nucleotides with dependency on occupancy of the coding versus template strand.¹⁶⁵ Beyond the ~200-nt window, the impact of OG in the PQS context on transcription was minimal. These studies provide support for a hypothesis that G oxidation to OG is the stimulus for G4 folding to function as an on/off switch of gene expression leading to the claim OG has epigenetic-like qualities.²⁸ The emergent properties of G allow a dormant region of the genome under normal conditions to signal alteration of gene expression under oxidative and inflammatory stress via a duplex→G4 switch mediated by G oxidation to OG.

The Xodo laboratory demonstrated oxidation of G in the *KRAS* promoter drives a non-canonical fold leading to upregulation of transcription (Figure 8D).⁵ The *KRAS* gene has a PQS –116 nucleotides upstream of the TSS in the promoter on the template strand.¹⁶⁶ Further, the *KRAS* PQS has excess G nucleotides in the G runs and extra G runs (i.e., spare tires). Xodo, et al. demonstrated that the G4 fold is maintained upon G oxidation to OG for

enhanced recruitment of the transcription factors MYC-associated zinc-finger protein (MAZ) and heterogeneous nuclear ribonucleoprotein A1 (hnRNP-A1 or A1) for gene activation. In their proposal, the BER pathway activates transcription after returning the sequence back to the wild-type state. The overall findings of DNA oxidation in the *KRAS* promoter to induce transcription are fundamentally the same as observed in the *VEGF* system; however, the fine details regarding the protein players differ.

The Tell laboratory provides a final example of promoter G oxidation driving a non-canonical fold in the genome, in which they focused their analysis on the sirtuin1 (SIRT1) gene promoter (Figure 8E).¹⁴⁹ Treatment of HeLa cells with H₂O₂ drove G oxidation at the negative calcium responsive element (nCaRE) in the SIRT1 gene promoter to guide location-specific BER initiated by OGG1 at this site as determined by ChIP analysis. They proposed OG release occurs in the nCaRE sequence that can adopt a hairpin (i.e., cruciform) with the AP in the loop region. The hairpin displays the AP to APE1 for binding resulting in the buildup of a multiprotein complex forming between APE1, Ku70/80, and RNA polymerase II for induction of transcription. Tell's studies provide an example that broadens the scope of non-canonical DNA secondary structures as switches for transcription under oxidative stress conditions. Similar to the *VEGF* studies, APE1 was an essential hub protein driving the interplay of DNA repair and transcriptional activation.

Cellular responses to oxidative stress set in motion by emergent properties of G.

During oxidative stress, expression of DNA repair proteins needs to be modulated. Thus, we asked whether promoter PQSs in DNA repair genes could signal transcriptional changes via G oxidation to OG to focus BER and unmask a G4 fold as a regulatory structure for activating protein factor binding. To this end, we first used bioinformatic tools to determine that DNA repair genes harbor PQSs in promoters at a two-fold greater frequency compared to all other genes.⁸⁹ Interestingly, these PQSs are distributed nearly equally on the coding and template strand, as well as throughout the promoter region within a 200-nucleotide window (Figure 4B). Additionally, more than half of these PQSs include a 5th track (or spare tire).⁸¹ Thus, up- and downregulation of DNA repair protein transcription levels would likely occur via the pathways we and others have described.^{4,5} Upregulation of DNA repair genes allows more efficient repair of the DNA damage inflicted by ROS; however, downregulation of DNA repair genes is not an intuitive result. As an example, the BER pathway has 11 glycosylases for initiation of repair and many of these have overlapping functions.¹⁶⁷ Some of these glycosylases are tightly coupled to replication (e.g., NEIL1), during which oxidative stress stops replication, and those genes coupled to this pathway are not needed; hence, lower expression levels would help conserve energy.^{9,89}

A feedback loop between DNA damage and repair may exist (Figure 10). Steps to showcase this genomic feature driving upregulated DNA repair gene expression come from the *NTHL1*, *NEIL3*, and *PCNA* genes that harbor a PQS in the coding strand of their promoters.^{4,108,168} Biophysical studies determined that the three PQSs adopt stable G4 folds in vitro, and each possesses excess Gs and G runs to maintain the fold when oxidatively modified.

The *NEIL3* system was explored with the greatest detail.¹⁰⁸ We placed the wild-type (i.e., all G) *NEIL3* PQS into the promoter of a luciferase gene and found it could be activated upon treatment of transfected cells with TNF α to induce inflammation. Importantly, treatment with TNF α mimicked the impact of the synthetic installation of OG in the same promoter. A similar set of chemical and biological experiments were conducted to verify OG, BER, and APE1 stalling on a G4 fold were critical for gene activation (Figure 8C). This proposed pathway is consistent with *NEIL3* upregulation during oxidative stress induced by hypoxia-ischemia in neurogenesis,¹⁶⁹ cardiac arrest,¹⁷⁰ or cancer.^{108,169,171} Our model studies were continued with the *NTHL1* and *PCNA* gene promoter PQSs,^{4,168} in which the findings in our chemically defined systems are in line with cell studies showing upregulation of these genes upon oxidative stress.^{4,168} As for PQS oxidation mediating downregulation, the sensor protein RAD17 has a PQS in the template strand of the gene promoter.¹⁷² Upon oxidation of this G-rich region, expression was downregulated, which is consistent with downregulated *RAD17* expression in cancer cells experiencing high ROS loads.¹⁷² Consistent with our hypothesis is a recent report from the Pei and Strauss laboratories who demonstrated in zebrafish that Apex1 (i.e., Ape1), oxidative modification of G to OG in DNA, and the Creb1 transcription factor are all essential for BER protein expression during neural development.¹⁷³ Thus, we propose a feedback loop exists in cells to set DNA repair proteins at appropriate levels for maintaining genomic integrity during oxidative and inflammatory stress. Additionally, high metabolic demand, such as long-distance running, is known to elevate OG levels and increase expression of stress response genes.^{174,175} The proposed feedback loop would keep these essential repair proteins upregulated leading to health benefits of exercise.

On the other hand, cancer cells commonly have DNA repair systems upregulated that cause DNA damaging anticancer drugs to fail because the lesions formed by the drug are rapidly repaired. The possibility of promoter PQSs in DNA repair proteins as essential sequences for regulation of these genes render the folded G4s as targets for cancer therapeutics. Targeting gene promoter G4s with ligands to regulate gene expression was proposed and is actively being pursued by many laboratories.^{176,177} The best studied to date is the *c-MYC* oncogene promoter PQS that when targeted with a G4 ligand, the overexpression of c-MYC in the cancer cells is highly attenuated.¹⁷⁸ A similar strategy can be employed to target the over expressed DNA repair genes in cancer cells to attenuate expression of these genes that cause these cells to elude traditional cancer therapies.

Why OG? Mutagenesis vs. Epigenetics.

The case for OG as an epigenetic mark is multifaceted and allows parallels to be drawn to the established 5mC epigenetic system. The base modification 5mC retains Watson-Crick base pairing while altering the protein readout in the major groove of duplex DNA. It also has a clear protein system for reading, writing, and erasing the epigenetic mark, and the mark is heritable to daughter generations.¹⁷⁹ The OG and 5mC heterocycles are both chemically stable; both maintain strong Watson-Crick base pairing to their complementary partner that is not diminished by competing tautomeric forms of the base (Figure 11A).¹⁸⁰ A drawback to OG is the *anti-syn* rotation to allow OG base pairing with A. As described above (Figure 6), it appears as though the OG epigenetic system is employed for quick

changes to cellular phenotype (i.e., stress or estrogen/retinoic acid mediated changes); therefore, a replication event enabling the OG:A base pair to form is less likely. Moreover, if the OG:A base pair does form, there is a well-established repair system to suppress the mutagenesis of this inappropriate base pair to low levels.¹²⁹ Analogous to this minor drawback is the natural ability of G to wobble base pair with T that is suppressed to negligible amounts by mismatch repair,¹⁸¹ thus, non-canonical base pairing is not an issue with the genetic and epigenetic systems.

In the human genome, 5mC is primarily found in CpG dinucleotide contexts, in which nearly all CpGs are methylated with the exception of CpG islands that are located in ~70% of promoters.¹⁸² 5-Methylcytosine is written into DNA at specific sites by DNA methyltransferase proteins (DNMTs).¹⁸² As for OG, its formation is favorable in GpG (i.e., 5'-GG-3') dinucleotide contexts, and recent sequencing efforts support OG formation is biased to GpG sites and PQS contexts;^{70-72,74} however, the OG modification is maintained at a much lower frequency in a genome compared to 5mC (5mC:OG ~1,500:1).¹⁸³ At present, a direct protein writer of OG has been elusive. Currently, the most compelling study for site-specific OG formation in DNA was reported by Avvedimento and co-workers when they identified the chromatin remodelers LSD1 and Jumonji C domain-containing histone demethylases actively formed discrete OG foci in genomic DNA.¹⁴⁴ Like all other reports of OG formation in gene promoters, the oxidation relied on a diffusible oxidant system (e.g., H₂O₂ or CO₃^{•-}).^{4,5,148,149,158} Alternatively, identification of an oxidase for selective modification of the G heterocycle to OG is not farfetched because the TET enzymes perform a similar task by oxidation of 5mC to yield 5-hydroxycytosine (5hmC), and other downstream products.¹⁸³ Future work in this domain is needed.

Simultaneously OG and 5mC alter the hydrogen bond donor-accepter pattern in the major groove of duplex DNA that can alter the protein readers of the sequence (Figure 11B). The presence of 5mC has a repulsive effect on transcription factors such as c-MYC and CRE, while attracting methyl CpG-binding proteins (MBDs).^{179,182} Similarly, OG repels SP1,¹²¹ NF-κB,¹²² and CREB,¹²³ while attracting OGG1 binding.¹⁴⁸ The modulation of protein readers in regulatory regions by these two epigenetic modifications provides the chemical signal to change cellular phenotype. Future work is needed to address whether other OG reader proteins exist.

The erasing mechanisms for OG and 5mC both utilize the BER pathway. The dormant state for each system is the wild-type G-rich PQS for OG or the 5mC-modified sequence. Interestingly, it is oxidation of G to OG, and oxidation of 5mC to downstream products that poise the gene for activation, and the products are BER substrates.²⁸ For G oxidation to OG, OGG1 is the gatekeeper to an AP, while oxidation of 5mC by at least four electrons to 5-formylcytosine (5fC) or six electrons to 5-carboxylcytosine (5caC) are substrates for thymine DNA glycosylase (TDG) to release the base and yield an AP (Figure 12).¹⁸⁴ In both situations, completion of the removal process starts with APE1 to return the sequence back to the non-modified state. Because OG-mediated gene activation was found to be APE1 dependent in our studies and in Tell's work,^{4,149} it is tempting to conclude that 5mC removal via oxidation and BER may function in a similar mode, as we have described (Figures 8C and 8E). However, Balasubramanian's laboratory demonstrated PQSs block writing of 5mC,

and therefore, prevent any of these reactions from occurring under normal conditions.¹⁰¹ Oxidation of G in PQSs may also collaborate with the 5mC epigenetic writing and/or erasing pathways leading to an additional layer of base modifications that regulate transcription. Future work is needed to address whether the 5mC and OG systems are functional together in gene regulation. Additional examination of bisulfite sequencing data is needed to address whether 5mC is favorably installed in other potential non-canonical folds in which the proposed APE1-dependent mode of gene activation could occur.

One defining feature of the established epigenetic system is heritability from mother to daughter generations. This is easy to rationalize chemically because 5mC is installed on both strands in the CpG context. After replication, the cell can read the hemimethylation status of the replicons and return them to the bismethylated state. In contrast, OG formation in a GpG context cannot follow this mechanism. On the basis of the studies so far, there is possibly a good reason for this. The established system using 5mC is utilized for long-term information storage in the genome; in contrast, the OG system appears to function in the day-to-day management of stress. Because of the vastly different timeframes under which these two systems are employed, 5mC needs a robust heritability mechanism, while the OG system functions without being heritable through generations; however, the G-rich PQSs in which oxidation occurs are heritable resulting in reproducible formation of OG in specific genomic regions for gene regulation. There exists an extreme importance of epigenetic 5mC during development that is essential for cellular differentiation.¹⁷⁹ During development, OG could be essential for managing the ROS formed during the high metabolic demand that occurs while a fetus is growing and developing, and hints in the literature suggest this may occur;¹⁷³ however, future studies are needed to better understand the role of OG in the genome during development.

The mutagenic nature of OG should designate this DNA modification as a lesion. This point is mute when looking at the consequence of utilizing 5mC as an epigenetic mark. Methylation of the cytosine heterocycle increases the rate of deamination to yield the thymine base by 2.2-fold, that after replication creates a C→T transition mutation.¹⁸⁵ Sites containing 5mC in the human genome account for ~30% of all germline and somatic mutations.¹⁸⁵ The burden of using this epigenetic system is so large that it has fundamentally altered the dinucleotide frequency of the human genome (Figure 1A).^{14,15} The CpG dinucleotide frequency in humans is ~20% of the expected value because of utilizing 5mC as an epigenetic mark; in contrast, *C. elegans* that does not use 5mC has the CpG dinucleotide at the expected frequency (Figure 1A). The reason for the low abundance of CpG sites in humans is proposed to be the enhanced deamination rate of 5mC.^{14,15} As stated above, OG is favorably found in GpG sites, and this dinucleotide exists in higher frequency in the human genome than expected. Retention of the high 5'-GG-3' dinucleotide frequency occurs because OG has a low mutation rate under normal growth conditions.¹³³ Even in cancer genomes in which the BER proteins are not mutated, the G→T mutation signature diagnostic of OG is not generally observed at substantial levels.¹⁸⁶ Interestingly, *C. elegans* has a 5'-GG-3' frequency close to expected that may suggest this organism has not evolved to use OG as an epigenetic mark, similar to its avoidance of 5mC.

The static amount of OG in the mammalian genome is $\sim 10^3$ and can increase by an order of magnitude under oxidative or inflammatory stress conditions.^{62,183,187} To put this number in perspective, the levels of modified bases in mouse embryonic stem cells descends in the order $5\text{mC} \gg 5\text{hmC} \gg 5\text{fC} \sim \text{OG} \sim 5\text{fU} > 5\text{caC} \sim 5\text{hmU}$, and this changes order after cell differentiation in the 3-month-old mouse cortex tissue to $5\text{mC} \gg 5\text{hmC} \gg \text{OG} \sim 5\text{fU} > 5\text{fC} > 5\text{hmU} \gg 5\text{caC}$.¹⁸³ The amounts of OG and 5fC are similar prior to differentiation and after differentiation OG is present at about one order of magnitude greater concentration than 5fC in the genome. The low levels of 5fC resulting in an impact on cellular phenotype are compensated for by its enrichment at key genomic sites. 5-Formylcytosine is favorably located in genomic sites associated with enhanced gene expression as identified by colocalization with the active histone marks H3K27ac and H3K4me1.¹⁸⁸⁻¹⁹⁰ The molecular pathway regarding how 5fC causes an increase in gene expression remains an open question. Similarly, the OG-Seq data from our laboratory was further analyzed to find OG enriched sites in the mouse genome are colocalized with the active H3K27ac histone mark.^{71,191} Molecular details regarding possible pathways by which OG mediates gene activation have been proposed (Figure 8).^{4,144,148,149,166} Lastly, OG levels are dynamic with $\sim 3,000$ writing events occurring per day, and the half-life of OG in a genome is short (< 30 min) reflecting efficient and active removal via BER.^{186,192} These findings support OG levels, albeit low, are likely sufficiently written and erased in critical regulatory locations at a frequency that can impact cellular processes.

Why OG? There are a few key reasons OG may have been selected over other modifications regarded as DNA lesions. At present, there are plausible mechanisms for site- or region-specific introduction of OG into the genome (i.e., electron transfer or chromatin remodeling).^{17,144} Once formed, OG like 5mC only impacts the genetic code in the major groove of DNA where protein readers decipher the message to provide the dynamic repulsive and attractive changes needed for selection of specific protein readers to alter transcription (Figure 11B). To reiterate, there is a well evolved repair system targeting the oxidized G base OG if it falls out of line and base pairs with A.¹²⁹ The presence of OG in the genome directs the BER process initiated by OGG1 that hands off the AP product to APE1 for gene activation via the Ref-1 activity of this protein.^{129,162,193} This is possible because OG exists in the genome at sufficient levels and is enriched in gene promoters on the basis of initial OG sequencing results.^{70-72,74} As described next, these features are not present with most other DNA “lesions,” or they have another cellular purposes.

Other “lesions” in DNA to consider include cytosine deamination to uracil, heterocycle alkylation, or other base oxidation products. At present, we know activation-induced cytidine deaminase (AID) catalyzes C deamination for immunoglobulin diversification (i.e., somatic hypermutation, class switch recombination and gene conversion).¹⁹⁴ This DNA “lesion” appears to have already been selected for other purposes. Base alkylation to guide where DNA repair occurs at first glance seems promising; however, nearly all alkylators, endogenous and exogenous, react at more than one nucleotide. As an example, Lindahl showed that non-enzymatic *S*-adenosyl-*L*-methionine when allowed to react with DNA yields N7-methylguanine, *O*⁶-methylguanine, and N3-methyladenine.¹⁹⁵ Furthermore, alkylation does not allow reaction in one site to funnel to a regulatory site as does base

oxidation via long-range charge transfer.¹⁷ Also, base adducts can reside in either the major or minor groove of DNA or the Watson-Crick face of the modified base, in which those in the minor groove or on the Watson-Crick face will be ineffective at modulating protein readout in the major groove. Lastly, alkylation activates BER or nucleotide excision repair, and the pathway dominating the repair process shows dependency on the adduct size and shape.^{154,196} Because of the diversity in alkylation products, the impact on DNA structure, and their variable repair mechanisms, it would be nearly impossible for a cell to respond reproducibly to this type of DNA modification. Other oxidation products such as the hydantoins Sp/Gh/2Ih may function, but Sp and Gh concentrations are very low, less than 5caC, and 2Ih has yet to be found in cellulose.^{62,183} The T oxidation product 5-hydroxymethyluracil (5hmU) was found by the Carell laboratory to be a product of TET-mediated oxidation, although future work is needed to address the biological significance of this finding beyond stimulation of DNA repair.¹⁸³ Thus, OG may not be the only base modification with epigenetic potential; however, at present, it seems to be the oxidatively modified base well-suited for the job.

Are PQSs the only non-canonical fold available for gene regulation?

The studies in our laboratory and Xodo's laboratory have identified G oxidation to OG in the PQS context can function as the initiating event for G4 folding and gene activation, although the fine details differ in how this occurs.^{4,5} Nonetheless, the G4 fold appears to be functional. These observations lead to the question of other non-canonical folds in DNA. First, Tell's laboratory found oxidative modification of a hairpin-forming sequence can drive transcription,¹⁴⁹ and we further demonstrated OG in a potential Z-DNA forming sequence context can do the same thing (Figures 13A and 13B).¹⁹⁷ The Z-DNA results are instructive but perhaps not biologically relevant because oxidation-prone Gs reside equally on the coding versus template strand that would result in an inconsistent response to oxidation, i.e., up- and downregulation at the same time.

A role for i-motif folds in gene regulation upon G oxidation.

The other non-canonical fold capable of driving gene regulation is the C-rich potential i-motifs that are complementary to PQSs.¹⁵¹ The Hurley laboratory has conducted the most thorough studies to support folded i-motifs can regulate gene expression, in which ligand stabilization of the *BCL2* i-motif leads to upregulated transcription.⁹⁶ This leads to the question, do i-motif folds function when G oxidation occurs in a gene promoter to regulate transcription? Folded i-motifs can play a role either as the site of G oxidation, if a G appears in a loop region between the C runs, or their folding can direct binding of gene activating proteins.¹⁵¹ In the first scenario, we have contributed to understanding whether OG in an i-motif context can impact gene expression by synthesizing this modification in a vector for study in cellulose (Figures 13A and 13B).¹⁶⁴ The studies found that OG in the i-motif context can modulate gene expression with the same strand dependency as found with the G4 fold. A conclusion drawn from this observation is that i-motifs can regulate gene expression; however, we cannot say whether the i-motif exclusively folded or folded in tandem with the G4. This observation is likely not biologically relevant because G-runs dominate the location of G oxidation resulting in an overwhelming favorability of oxidation in the G-rich strand

rather than the C-rich strand, and this determines the strand occupancy of APE1 for increasing gene expression.^{2,17,24}

The Hurley laboratory has proposed a model of mutual exclusivity for native G4 and i-motif folding in which the non-canonical G4 fold functions as a repressor and the i-motif sequence functions as an activator or repressor depending on its folding status and the gene promoter.^{151,198,199} In contrast, when oxidation of the potential non-canonical structural unit occurs, a different class of proteins is recruited to the site (i.e., BER) driving the gene modulation process when compared to the native transcriptional structures in the wild-type state. Thus, native and oxidatively modified non-canonical folds function differently in the context of the cell, and findings from one study are likely incongruent with those from the other study. One lingering open question regarding G4 and i-motif folds is what is the role of the DNA-RNA heteroduplex leaving an unpaired DNA strand, or R loop, in the folding and regulatory process?^{200,201} Recent reports are suggesting that R-loop structures may have a critical role in gene activation when G4 folds occur,²⁰² and future studies geared toward understanding them in the context of oxidatively modified promoters are needed.

Future perspective and concluding remarks.

A captivating mystery of the human genome sequence is why has it evolved to have an overabundance of 5'-GG-3' dinucleotides and PQSs around transcription start sites (Figures 1A and 4B). Fundamentally, why would thermodynamically stable G-rich sequences be enriched in regions that must be actively opened for mRNA synthesis? We propose the G richness around TSSs brings about two new emergent properties to DNA. First, as a consequence of the GC-richness around TSSs there exists an enrichment of non-canonical G4s, i-motifs, and Z-DNA.^{79,197} Second, the G runs associated with PQSs render these sites highly sensitive to oxidative modification as a result of π -stacking of adjacent Gs lowering the ionization potential of the 5' most G in the run (Figure 1B).² These two features serve as a regulatory system during oxidative stress as the input signal to drive a duplex→G4 switch that can overcome the thermodynamic barrier to control gene expression.⁴

Fine tuning of a genomic sequence to have new chemical properties for biological function is growing in awareness.^{13,28-30} The ability of oxidation at a distance in DNA to generate an electron hole that funnels to key sites before completing the reaction and yielding an oxidatively modified base is fascinating. Beyond the work in the Barton laboratory with Fe-S containing proteins,⁶⁵ biology may have found other ways to use emergent properties of DNA for gene regulation. Many G4-specific helicases also possess redox-active Fe-S clusters, and G4s attenuate electron hole migration,^{152,203} thus leading to a possibility that G4s are found by a redox relay system similar to that for DNA damage. Future experiments will shed some light on these questions.

A final point is the possible existence of OG during natural selection to arrive at modern genomes that have highly selected sequences, sequence-specific protein binding sites, and chemical or epigenetic marks to modulate protein binding. Under prebiotic conditions, a plausible synthetic pathway for 8-oxopurines, such as OG, has been described.²⁰⁴ Further, UV radiation like that found on prebiotic Earth drives the facile conversion of G to OG

suggesting that this purine has been with biology from the beginning.²⁰⁵ Apparently, OG was deselected as a storage of genetic information because in the absence of a developed repair system, the *anti-syn* conversion of OG causes low fidelity replication.²⁰⁶ The recent studies described in this Perspective suggest biology has found another way to utilize the facile formation of OG.

Acknowledgements

We are grateful to the National Cancer Institute for funding our work on this topic (R01 CA090689). The work described in our laboratory could not have happened without the expert help of Yun Ding and Judy Zhu (University of Utah).

References

- (1). Joyce GF; Szostak JW Protocells and RNA self-replication. Cold Spring Harb. Perspect. Biol 2018, 10, 1–22.
- (2). Saito I; Takayama M; Sugiyama H; Nakatani K; Tsuchida A; Yamamoto M Photoinduced DNA cleavage via electron transfer: demonstration that guanine residues located 5' to guanine are the most electron-donating sites. J. Am. Chem. Soc 1995, 117, 6406–6407.
- (3). Lane AN; Chaires JB; Gray RD; Trent JO Stability and kinetics of G-quadruplex structures. Nucleic Acids Res. 2008, 36, 5482–5515. [PubMed: 18718931]
- (4). Fleming AM; Ding Y; Burrows CJ Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. Proc. Natl. Acad. Sci. U.S.A 2017, 114, 2604–2609. [PubMed: 28143930]
- (5). Cogoi S; Ferino A; Miglietta G; Pedersen EB; Xodo LE The regulatory G4 motif of the Kirsten ras (KRAS) gene is sensitive to guanine oxidation: implications on transcription. Nucleic Acids Res. 2018, 46, 661–676. [PubMed: 29165690]
- (6). Winterbourn CC Reconciling the chemistry and biology of reactive oxygen species. Nat. Chem. Biol 2008, 4, 278–286. [PubMed: 18421291]
- (7). Murphy MP How mitochondria produce reactive oxygen species. Biochem. J 2009, 417, 1–13. [PubMed: 19061483]
- (8). Lonkar P; Dedon PC Reactive species and DNA damage in chronic inflammation: reconciling chemical mechanisms and biological fates. Int. J. Cancer 2011, 128, 1999–2009. [PubMed: 21387284]
- (9). Fleming AM; Burrows CJ Formation and processing of DNA damage substrates for the hNEIL enzymes. Free Radic. Biol. Med 2017, 107, 35–52. [PubMed: 27880870]
- (10). Berlett BS; Stadtman ER Protein oxidation in aging, disease, and oxidative stress. J. Biol. Chem 1997, 272, 20313–20316. [PubMed: 9252331]
- (11). Dedon PC; Tannenbaum SR Reactive nitrogen species in the chemical biology of inflammation. Arch. Biochem. Biophys 2004, 423, 12–22. [PubMed: 14989259]
- (12). Giorgio M; Trinei M; Migliaccio E; Pelicci PG Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? Nat. Rev. Mol. Cell Biol 2007, 8, 722–728. [PubMed: 17700625]
- (13). Antoniali G; Malfatti MC; Tell G 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) 2017, 56, 65–74. [PubMed: 28629776]
- (14). Simmen MW Genome-scale relationships between cytosine methylation and dinucleotide abundances in animals. Genomics 2008, 92, 33–40. [PubMed: 18485662]
- (15). Gentles AJ; Karlin S Genome-scale compositional comparisons in eukaryotes. Genome Res. 2001, 11, 540–546. [PubMed: 11282969]
- (16). Zhang L; Kasif S; Cantor CR; Broude NE GC/AT-content spikes as genomic punctuation marks. Proc. Natl. Acad. Sci. U.S.A 2004, 101, 16855–16860. [PubMed: 15548610]

- (17). Genereux JC; Barton JK Mechanisms for DNA charge transport. *Chem. Rev* 2010, 110, 1642–1662. [PubMed: 20214403]
- (18). Schuster GB Long-range charge transfer in DNA: Transient structural distortions control the distance dependence. *Acc. Chem. Res* 2000, 33, 253–260. [PubMed: 10775318]
- (19). Steenken S; Jovanovic SV How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution. *J. Am. Chem. Soc* 1997, 119, 617–618.
- (20). Arnold AR; Grodick MA; Barton JK DNA charge transport: from chemical principles to the cell. *Cell Chem. Biol* 2016, 23, 183–197. [PubMed: 26933744]
- (21). Kawai K; Majima T Hole transfer kinetics of DNA. *Acc. Chem. Res* 2013, 46, 2616–2625. [PubMed: 23805774]
- (22). Delaney S; Barton JK Long-range DNA charge transport. *J. Org. Chem* 2003, 68, 6475–6483. [PubMed: 12919006]
- (23). Barnett RN; Bongiorno A; Cleveland CL; Joy A; Landman U; Schuster GB Oxidative damage to DNA: Counterion-assisted addition of water to ionized DNA. *J. Am. Chem. Soc* 2006, 128, 10795–10800. [PubMed: 16910674]
- (24). Fleming AM; Burrows CJ G-Quadruplex folds of the human telomere sequence alter the site reactivity and reaction pathway of guanine oxidation compared to duplex DNA. *Chem. Res. Toxicol* 2013, 26, 593–607. [PubMed: 23438298]
- (25). Liu Y; Liu Z; Geacintov NE; Shafirovich V Proton-coupled hole hopping in nucleosomal and free DNA initiated by site-specific hole injection. *Phys. Chem. Chem. Phys* 2012, 14, 7400–7410. [PubMed: 22526555]
- (26). Heller A Spiers Memorial Lecture. On the hypothesis of cathodic protection of genes. *Faraday Discuss.* 2000, 116, 1–13.
- (27). Friedman KA; Heller A Guanosine distribution and oxidation resistance in eight eukaryotic genomes. *J. Am. Chem. Soc* 2004, 126, 2368–2371. [PubMed: 14982441]
- (28). Fleming AM; Burrows CJ 8-Oxo-7,8-dihydroguanine, friend and foe: Epigenetic-like regulator versus initiator of mutagenesis. *DNA Repair (Amst)* 2017, 56, 75–83. [PubMed: 28629775]
- (29). Hao W; Qi T; Pan L; Wang R; Zhu B; Aguilera-Aguirre L; Radak Z; Hazra TK; Vlahopoulos SA; Bacsí A; Brasier AR; Ba X; Boldogh I Effects of the stimuli-dependent enrichment of 8-oxoguanine DNA glycosylase I on chromatinized DNA. *Redox Biol.* 2018, 18, 43–53. [PubMed: 29940424]
- (30). Seifermann M; Epe B Oxidatively generated base modifications in DNA: Not only carcinogenic risk factor but also regulatory mark? *Free Radic. Biol. Med* 2017, 107, 258–265. [PubMed: 27871818]
- (31). Patel DJ; Phan AT; Kuryavyi V Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.* 2007, 35, 7429–7455. [PubMed: 17913750]
- (32). Mergny JL; Sen D DNA quadruple helices in nanotechnology. *Chem. Rev* 2019, 119, 6290–6325. [PubMed: 30605316]
- (33). Xu Y Chemistry in human telomere biology: structure, function and targeting of telomere DNA/RNA. *Chem. Soc. Rev* 2011, 40, 2719–2740. [PubMed: 21301727]
- (34). Edwald E; Stone MB; Gray EM; Wu J; Veatch SL Oxygen depletion speeds and simplifies diffusion in HeLa cells. *Biophys. J* 2014, 107, 1873–1884. [PubMed: 25418168]
- (35). Kreuzer-Martin HW; Ehleringer JR; Hegg EL Oxygen isotopes indicate most intracellular water in log-phase *Escherichia coli* is derived from metabolism. *Proc. Natl. Acad. Sci. U.S.A* 2005, 102, 17337–17341. [PubMed: 16301541]
- (36). Lü J-M; Lin PH; Yao Q; Chen C Chemical and molecular mechanisms of antioxidants: experimental approaches and model systems. *J. Cell. Mol. Med* 2010, 14, 840–860. [PubMed: 19754673]
- (37). Villano D; Vilaplana C; Medina S; Cejuela-Anta R; Martinez-Sanz JM; Gil P; Genieser HG; Ferreres F; Gil-Izquierdo A Effect of elite physical exercise by triathletes on seven catabolites of DNA oxidation. *Free Radic. Res* 2015, 49, 973–983. [PubMed: 25786325]
- (38). Van Houten B; Santa-Gonzalez GA; Camargo M DNA repair after oxidative stress: current challenges. *Curr. Opin. Toxicol* 2018, 7, 9–16. [PubMed: 29159324]

- (39). Veal EA; Day AM; Morgan BA Hydrogen peroxide sensing and signaling. *Mol. Cell* 2007, 26, 1–14. [PubMed: 17434122]
- (40). Finkel T Signal transduction by reactive oxygen species. *J. Cell Biol* 2011, 194, 7–15. [PubMed: 21746850]
- (41). Saitoh M; Nishitoh H; Fujii M; Takeda K; Tobiume K; Sawada Y; Kawabata M; Miyazono K; Ichijo H Mammalian thioredoxin is a direct inhibitor of apoptosis signal-regulating kinase (ASK) 1. *EMBO J.* 1998, 17, 2596–2606. [PubMed: 9564042]
- (42). Cadet J; Wagner JR; Shafirovich V; Geacintov NE One-electron oxidation reactions of purine and pyrimidine bases in cellular DNA. *Int. J. Radiat. Biol* 2014, 90, 423–432. [PubMed: 24369822]
- (43). Fleming AM; Muller JG; Ji I; Burrows CJ Characterization of 2'-deoxyguanosine oxidation products observed in the Fenton-like system Cu(II)/H₂O₂/reductant in nucleoside and oligodeoxynucleotide contexts. *Org. Biomol. Chem* 2011, 9, 3338–3348. [PubMed: 21445431]
- (44). Alshykhly OR; Fleming AM; Burrows CJ 5-Carboxamido-5-formamido-2-iminohydantoin, in addition to 8-oxo-7,8-dihydroguanine, is the major product of the iron-Fenton or X-ray radiation-induced oxidation of guanine under aerobic reducing conditions in nucleoside and DNA contexts. *J. Org. Chem* 2015, 80, 6996–7007. [PubMed: 26092110]
- (45). Zheng L; Greenberg MM DNA damage emanating from a neutral purine radical reveals the sequence dependent convergence of the direct and indirect effects of gamma-radiolysis. *J. Am. Chem. Soc* 2017, 139, 17751–17754. [PubMed: 29190086]
- (46). Chatgililoglu C; Ferreri C; Geacintov NE; Krokidis MG; Liu Y; Masi A; Shafirovich V; Terzidis MA; Tsegay PS 5',8-Cyclopurine lesions in DNA damage: Chemical, analytical, biological, and diagnostic significance. *Cells* 2019, 8, 513.
- (47). Illes E; Mizrahi A; Marks V; Meyerstein D Carbonate-radical-anions, and not hydroxyl radicals, are the products of the Fenton reaction in neutral solutions containing bicarbonate. *Free Radical Biol. Med* 2019, 131, 1–6. [PubMed: 30458276]
- (48). Kozłowski H; Janicka-Kłos A; Brasun J; Gaggelli E; Valensin D; Valensin G Copper, iron, and zinc ions homeostasis and their role in neurodegenerative disorders (metal uptake, transport, distribution and regulation). *Coord. Chem. Rev* 2009, 253, 2665–2685.
- (49). Aguilera-Aguirre L; Hosoki K; Bacsí A; Radak Z; Sur S; Hegde ML; Tian B; Saavedra-Molina A; Brasier AR; Ba X; Boldogh I Whole transcriptome analysis reveals a role for OGG1-initiated DNA repair signaling in airway remodeling. *Free Radic. Biol. Med* 2015, 89, 20–33. [PubMed: 26187872]
- (50). Joffe A; Geacintov NE; Shafirovich V DNA lesions derived from the site selective oxidation of guanine by carbonate radical anions. *Chem. Res. Toxicol* 2003, 16, 1528–1538. [PubMed: 14680366]
- (51). Crean C; Geacintov NE; Shafirovich V Oxidation of guanine and 8-oxo-7,8-dihydroguanine by carbonate radical anions: Insight from oxygen-18 labeling experiments. *Angew. Chem., Int. Ed* 2005, 44, 5057–5060.
- (52). Luo W; Muller JG; Rachlin EM; Burrows CJ Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-oxo-7,8-dihydroguanosine. *Org. Lett* 2000, 2, 613–616. [PubMed: 10814391]
- (53). Luo W; Muller JG; Rachlin EM; Burrows CJ Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model. *Chem. Res. Toxicol* 2001, 14, 927–938. [PubMed: 11453741]
- (54). Cadet J; Douki T; Ravanat J-L Oxidatively generated damage to the guanine moiety of DNA: Mechanistic aspects and formation in cells. *Acc. Chem. Res* 2008, 41, 1075–1083. [PubMed: 18666785]
- (55). Cui L; Ye W; Prestwich EG; Wishnok JS; Taghizadeh K; Dedon PC; Tannenbaum SR Comparative analysis of four oxidized guanine lesions from reactions of DNA with peroxynitrite, singlet oxygen, and gamma-radiation. *Chem. Res. Toxicol* 2013, 26, 195–202. [PubMed: 23140136]
- (56). Delaney S; Delaney JC; Essigmann JM Chemical-biological fingerprinting: probing the properties of DNA lesions formed by peroxynitrite. *Chem. Res. Toxicol* 2007, 20, 1718–1729. [PubMed: 17941698]

- (57). Rokhlenko Y; Cadet J; Geacintov NE; Shafirovich V Mechanistic aspects of hydration of guanine radical cations in DNA. *J. Am. Chem. Soc* 2014, 136, 5956–5962. [PubMed: 24689701]
- (58). Hong H; Cao H; Wang Y Identification and quantification of a guanine-thymine intrastrand cross-link lesion induced by Cu(II)/H₂O₂/ascorbate. *Chem. Res. Toxicol* 2006, 19, 614–621. [PubMed: 16696563]
- (59). Matter B; Seiler CL; Murphy K; Ming X; Zhao J; Lindgren B; Jones R; Tretyakova N Mapping three guanine oxidation products along DNA following exposure to three types of reactive oxygen species. *Free Radic. Biol. Med* 2018, 121, 180–189. [PubMed: 29702150]
- (60). Kino K; Saito I; Sugiyama H Product analysis of GG-specific photooxidation of DNA via electron transfer: 2-Aminoimidazolone as a major guanine oxidation product. *J. Am. Chem. Soc* 1998, 120, 7373–7374.
- (61). Minko IG; Vartanian VL; Tozaki NN; Coskun E; Coskun SH; Jaruga P; Yeo J; David SS; Stone MP; Egli M; Dizdaroglu M; McCullough AK; Lloyd RS Recognition of DNA adducts by edited and unedited forms of DNA glycosylase NEIL1. *DNA Repair (Amst)* 2019, 85, 102741. [PubMed: 31733589]
- (62). Mangerich A; Knutson CG; Parry NM; Muthupalani S; Ye W; Prestwich E; Cui L; McFaline JL; Mobley M; Ge Z; Taghizadeh K; Wishnok JS; Wogan GN; Fox JG; Tannenbaum SR; Dedon PC Infection-induced colitis in mice causes dynamic and tissue-specific changes in stress response and DNA damage leading to colon cancer. *Proc. Natl. Acad. Sci. U.S.A* 2012, 109, E1820–E1829. [PubMed: 22689960]
- (63). Gedik CM; Collins A Establishing the background level of base oxidation in human lymphocyte DNA: results of an interlaboratory validation study. *FASEB J.* 2005, 19, 82–84. [PubMed: 15533950]
- (64). Matter B; Malejka-Giganti D; Csallany AS; Tretyakova N Quantitative analysis of the oxidative DNA lesion, 2,2-diamino-4-(2-deoxy-β-D-erythro-pentofuranosyl)amino]-5(2*H*)-oxazolone (oxazolone), in vitro and in vivo by isotope dilution-capillary HPLC-ESI-MS/MS. *Nucleic Acids Res.* 2006, 34, 5449–5460. [PubMed: 17020926]
- (65). Tse ECM; Zwang TJ; Bedoya S; Barton JK Effective distance for DNA-mediated charge transport between repair proteins. *ACS Cent. Sci* 2019, 5, 65–72. [PubMed: 30693326]
- (66). Fleming AM; Alshykhly O; Zhu J; Muller JG; Burrows CJ Rates of chemical cleavage of DNA and RNA oligomers containing guanine oxidation products. *Chem. Res. Toxicol* 2015, 28, 1292–1300. [PubMed: 25853314]
- (67). Margolin Y; Shafirovich V; Geacintov NE; DeMott MS; Dedon PC DNA sequence context as a determinant of the quantity and chemistry of guanine oxidation produced by hydroxyl radicals and one-electron oxidants. *J. Biol. Chem* 2008, 283, 35569–35578. [PubMed: 18948263]
- (68). Nunez ME; Holmquist GP; Barton JK Evidence for DNA charge transport in the nucleus. *Biochemistry* 2001, 40, 12465–12471. [PubMed: 11601969]
- (69). Sun H; Zheng L; Yang K; Greenberg MM Positional dependence of DNA hole transfer efficiency in nucleosome core particles. *J. Am. Chem. Soc* 2019, 141, 10154–10158. [PubMed: 31244168]
- (70). Wu J; McKeague M; Sturla SJ Nucleotide-resolution genome-wide mapping of oxidative DNA damage by click-code-seq. *J. Am. Chem. Soc* 2018, 140, 9783–9787. [PubMed: 29944356]
- (71). Ding Y; Fleming AM; Burrows CJ Sequencing the mouse genome for the oxidatively modified base 8-oxo-7,8-dihydroguanine by OG-Seq. *J. Am. Chem. Soc* 2017, 139, 2569–2572. [PubMed: 28150947]
- (72). Amente S; Di Palo G; Scala G; Castrignanò T; Gorini F; Cocozza S; Moresano A; Pucci P; Ma B; Stepanov I; Lania L; Pelicci PG; Dellino GI; Majello B 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.* 2019, 47, 221–236. [PubMed: 30462294]
- (73). Clark DW; Phang T; Edwards MG; Geraci MW; Gillespie MN Promoter G-quadruplex sequences are targets for base oxidation and strand cleavage during hypoxia-induced transcription. *Free Radical Biol. Med* 2012, 53, 51–59. [PubMed: 22583700]
- (74). Fang Y; Zou P Genome-wide mapping of oxidative DNA damage via engineering of 8-oxoguanine DNA glycosylase. *Biochemistry* 2019, doi: 10.1021/acs.biochem.1029b00782.

- (75). Park J; Park JW; Oh H; Maria FS; Kang J; Tian X Gene-specific assessment of guanine oxidation as an epigenetic modulator for cardiac specification of mouse embryonic stem cells. *PLoS One* 2016, 11, e0155792. [PubMed: 27249188]
- (76). Green MR; Sambrook J Polymerase Chain Reaction (PCR) amplification of GC-rich templates. *Cold Spring Harb. Protoc* 2019, 2019, doi:10.1101/pdb.prot095141.
- (77). Huppert JL; Balasubramanian S Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* 2005, 33, 2908–2916. [PubMed: 15914667]
- (78). Todd AK; Johnston M; Neidle S Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* 2005, 33, 2901–2907. [PubMed: 15914666]
- (79). Maizels N; Gray LT The G4 genome. *PLoS Genet.* 2013, 9, e1003468. [PubMed: 23637633]
- (80). Engelhart AE; Plavec J; Persil Ö; Hud NV: Metal ion interactions with G-quadruplex structures In *Nucleic Acid-Metal Ion Interactions*; The Royal Society of Chemistry, 2009; pp 118–153.
- (81). Ding Y; Fleming AM; Burrows CJ Case studies on potential G-quadruplex-forming sequences from the bacterial orders Deinococcales and Thermales derived from a survey of published genomes. *Sci. Rep* 2018, 8, 15679. [PubMed: 30356061]
- (82). Fleming AM; Zhou J; Wallace SS; Burrows CJ 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* 2015, 1, 226–233. [PubMed: 26405692]
- (83). Yue DJ; Lim KW; Phan AT Formation of (3+1) G-quadruplexes with a long loop by human telomeric DNA spanning five or more repeats. *J. Am. Chem. Soc* 2011, 133, 11462–11465. [PubMed: 21702440]
- (84). An N; Fleming AM; Burrows CJ Human telomere G-quadruplexes with five repeats accommodate 8-oxo-7,8-dihydroguanine by looping out the DNA damage. *ACS Chem. Biol* 2016, 11, 500–507. [PubMed: 26686913]
- (85). Chambers VS; Marsico G; Boutell JM; Di Antonio M; Smith GP; Balasubramanian S High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol* 2015, 33, 877–881. [PubMed: 26192317]
- (86). Puig Lombardi E; Holmes A; Verga D; Teulade-Fichou M-P; Nicolas A; Londoño-Vallejo A Thermodynamically stable and genetically unstable G-quadruplexes are depleted in genomes across species. *Nucleic Acids Res.* 2019, 47, 6098–6113. [PubMed: 31114920]
- (87). Eddy J; Maizels N Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.* 2006, 34, 3887–3896. [PubMed: 16914419]
- (88). Piazza A; Adrian M; Samazan F; Heddi B; Hamon F; Serero A; Lopes J; Teulade-Fichou MP; Phan AT; Nicolas A Short loop length and high thermal stability determine genomic instability induced by G-quadruplex-forming minisatellites. *EMBO J.* 2015, 34, 1718–1734. [PubMed: 25956747]
- (89). Fleming AM; Zhu J; Ding Y; Visser JA; Zhu J; Burrows CJ Human DNA repair genes possess potential G-quadruplex sequences in their promoters and 5'-untranslated regions. *Biochemistry* 2018, 57, 991–1002. [PubMed: 29320161]
- (90). Biffi G; Tannahill D; McCafferty J; Balasubramanian S Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem* 2013, 5, 182–186. [PubMed: 23422559]
- (91). Henderson A; Wu Y; Huang YC; Chavez EA; Platt J; Johnson FB; Brosh RM Jr.; Sen D; Lansdorp PM Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* 2014, 42, 860–869. [PubMed: 24163102]
- (92). Hansel-Hertsch R; Beraldi D; Lensing SV; Marsico G; Zyner K; Parry A; Di Antonio M; Pike J; Kimura H; Narita M; Tannahill D; Balasubramanian S G-quadruplex structures mark human regulatory chromatin. *Nat. Genet* 2016, 48, 1267–1272. [PubMed: 27618450]
- (93). Mohaghegh P; Karow JK; Brosh RM Jr.; Bohr VA; Hickson ID The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.* 2001, 29, 2843–2849. [PubMed: 11433031]
- (94). Maizels N Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat. Struct. Mol. Biol* 2006, 13, 1055–1059. [PubMed: 17146462]
- (95). Brooks TA; Kendrick S; Hurley L Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* 2010, 277, 3459–3469. [PubMed: 20670278]

- (96). Kendrick S; Kang HJ; Alam MP; Madathil MM; Agrawal P; Gokhale V; Yang D; Hecht SM; Hurley LH The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc* 2014, 136, 4161–4171. [PubMed: 24559410]
- (97). Guo K; Gokhale V; Hurley LH; Sun D Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.* 2008, 36, 4598–4608. [PubMed: 18614607]
- (98). Wei D; Parkinson GN; Reszka AP; Neidle S Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation. *Nucleic Acids Res.* 2012, 40, 4691–4700. [PubMed: 22287624]
- (99). Rodriguez R; Miller KM; Forment JV; Bradshaw CR; Nikan M; Britton S; Oelschlaegel T; Xhemalce B; Balasubramanian S; Jackson SP Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol* 2012, 8, 301–310. [PubMed: 22306580]
- (100). Prorok P; Artufel M; Aze A; Coulombe P; Peiffer I; Lacroix L; Guedin A; Mergny JL; Damaschke J; Schepers A; Ballester B; Mechali M Involvement of G-quadruplex regions in mammalian replication origin activity. *Nat. Commun* 2019, 10, 3274. [PubMed: 31332171]
- (101). Mao SQ; Ghanbarian AT; Spiegel J; Martinez Cuesta S; Beraldi D; Di Antonio M; Marsico G; Hansel-Hertsch R; Tannahill D; Balasubramanian S DNA G-quadruplex structures mold the DNA methylome. *Nat. Struct. Mol. Biol* 2018, 25, 951–957. [PubMed: 30275516]
- (102). Guilbaud G; Murat P; Recolin B; Campbell BC; Maiter A; Sale JE; Balasubramanian S Local epigenetic reprogramming induced by G-quadruplex ligands. *Nat. Chem* 2017, 9, 1110–1117. [PubMed: 29064488]
- (103). Jara-Espejo M; Peres Line SR DNA G-quadruplex stability, position and chromatin accessibility are associated with CpG island methylation. *FEBS J.* 2019, doi:10.1111/febs.15065.
- (104). Zhou J; Fleming AM; Averill AM; Burrows CJ; Wallace SS The NEIL glycosylases remove oxidized guanine lesions from telomeric and promoter quadruplex DNA structures. *Nucleic Acids Res.* 2015, 43, 4039–4054. [PubMed: 25813041]
- (105). Vorlíčková M; Tomasko M; Sagi AJ; Bednarova K; Sagi J 8-Oxoguanine in a quadruplex of the human telomere DNA sequence. *FEBS J.* 2012, 279, 29–39. [PubMed: 22008383]
- (106). Lech CJ; Cheow Lim JK; Wen Lim JM; Amrane S; Heddi B; Phan AT Effects of site-specific guanine C8-modifications on an intramolecular DNA G-quadruplex. *Biophys. J* 2011, 101, 1987–1998. [PubMed: 22004753]
- (107). Bielskute S; Plavec J; Podbevsek P Impact of oxidative lesions on the human telomeric G-quadruplex. *J. Am. Chem. Soc* 2019, 141, 2594–2603. [PubMed: 30657306]
- (108). Fleming AM; Zhu J; Howpay Manage SA; Burrows CJ Human NEIL3 gene expression is regulated by epigenetic-like oxidative DNA modification. *J. Am. Chem. Soc* 2019, 141, 11036–11046. [PubMed: 31241930]
- (109). Agrawal P; Hatzakis E; Guo K; Carver M; Yang D Solution structure of the major G-quadruplex formed in the human VEGF promoter in K^+ : insights into loop interactions of the parallel G-quadruplexes. *Nucleic Acids Res.* 2013, 41, 10584–10592. [PubMed: 24005038]
- (110). Phan AT; Modi YS; Patel DJ Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc* 2004, 126, 8710–8716. [PubMed: 15250723]
- (111). Omega CA; Fleming AM; Burrows CJ The fifth domain in the G-quadruplex-forming sequence of the human NEIL3 promoter locks DNA folding in response to oxidative damage. *Biochemistry* 2018, 57, 2958–2970. [PubMed: 29718661]
- (112). Takahashi S; Kim KT; Podbevsek P; Plavec J; Kim BH; Sugimoto N Recovery of the formation and function of oxidized G-quadruplexes by a pyrene-modified guanine tract. *J. Am. Chem. Soc* 2018, 140, 5774–5783. [PubMed: 29608858]
- (113). Plum GE; Grollman AP; Johnson F; Breslauer KJ Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex. *Biochemistry* 1995, 34, 16148–16160. [PubMed: 8519772]

- (114). Fleming AM; Burrows CJ 8-Oxo-7,8-dihydro-2'-deoxyguanosine and abasic site tandem lesions are oxidation prone yielding hydantoin products that strongly destabilize duplex DNA. *Org. Biomol. Chem* 2017, 15, 8341–8353. [PubMed: 28936535]
- (115). Jia L; Shafirovich V; Shapiro R; Geacintov NE; Broyde S Structural and thermodynamic features of spiroiminodihydantoin damaged DNA duplexes. *Biochemistry* 2005, 44, 13342–13353. [PubMed: 16201759]
- (116). Khutsishvili I; Zhang N; Marky LA; Crean C; Patel DJ; Geacintov NE; Shafirovich V Thermodynamic profiles and nuclear magnetic resonance studies of oligonucleotide duplexes containing single diastereomeric spiroiminodihydantoin lesions. *Biochemistry* 2013, 52, 1354–1363. [PubMed: 23360616]
- (117). Oda Y; Uesugi S; Ikehara M; Nishimura S; Kawase Y; Ishikawa H; Inoue H; Ohtsuka E NMR studies of a DNA containing 8-hydroxydeoxyguanosine. *Nucleic Acids Res.* 1991, 19, 1407–1412. [PubMed: 2027747]
- (118). Lipscomb LA; Peek ME; Morningstar ML; Verghis SM; Miller EM; Rich A; Essigmann JM; Williams LD X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Proc. Natl. Acad. Sci. U.S.A* 1995, 92, 719–723. [PubMed: 7846041]
- (119). Kouchakdjian M; Bodepudi V; Shibutani S; Eisenberg M; Johnson F; Grollman AP; Patel DJ NMR structural studies of the ionizing radiation adduct 7-hydro-8-oxodeoxyguanosine (8-oxo-7H-dG) opposite deoxyadenosine in a DNA duplex. 8-Oxo-7H-dG(syn) dA(anti) alignment at lesion site. *Biochemistry* 1991, 30, 1403–1412. [PubMed: 1991121]
- (120). Shibutani S; Takeshita M; Grollman AP Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* 1991, 349, 431–434. [PubMed: 1992344]
- (121). Ramon O; Sauvaigo S; Gasparutto D; Faure P; Favier A; Cadet J Effects of 8-oxo-7,8-dihydro-2'-deoxyguanosine on the binding of the transcription factor Sp1 to its cognate target DNA sequence (GC box). *Free Radical Res.* 1999, 31, 217–229. [PubMed: 10499779]
- (122). Hailer-Morrison MK; Kotler JM; Martin BD; Sugden KD Oxidized guanine lesions as modulators of gene transcription. Altered p50 binding affinity and repair shielding by 7,8-dihydro-8-oxo-2'-deoxyguanosine lesions in the NF-kappaB promoter element. *Biochemistry* 2003, 42, 9761–9770. [PubMed: 12911319]
- (123). Moore SP; Toomire KJ; Strauss PR DNA modifications repaired by base excision repair are epigenetic. *DNA Repair (Amst)* 2013, 12, 1152–1158. [PubMed: 24216087]
- (124). Valinluck V; Tsai HH; Rogstad DK; Burdzy A; Bird A; Sowers LC Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res.* 2004, 32, 4100–4108. [PubMed: 15302911]
- (125). Allgayer J; Kitsera N; Bartelt S; Epe B; Khobta A Widespread transcriptional gene inactivation initiated by a repair intermediate of 8-oxoguanine. *Nucleic Acids Res.* 2016, 44, 7267–7280. [PubMed: 27220469]
- (126). Tornaletti S; Maeda LS; Kolodner RD; Hanawalt PC Effect of 8-oxoguanine on transcription elongation by T7 RNA polymerase and mammalian RNA polymerase II. *DNA Repair (Amst)* 2004, 3, 483–494. [PubMed: 15084310]
- (127). Le Page F; Klungland A; Barnes DE; Sarasin A; Boiteux S Transcription coupled repair of 8-oxoguanine in murine cells: the Ogg1 protein is required for repair in nontranscribed sequences but not in transcribed sequences. *Proc. Natl. Acad. Sci. U.S.A* 2000, 97, 8397–8402. [PubMed: 10890888]
- (128). Pastoriza-Gallego M; Armier J; Sarasin A Transcription through 8-oxoguanine in DNA repair-proficient and Csb(-)/Ogg1(-) DNA repair-deficient mouse embryonic fibroblasts is dependent upon promoter strength and sequence context. *Mutagenesis* 2007, 22, 343–351. [PubMed: 17630408]
- (129). David SS; O'Shea VL; Kundu S Base-excision repair of oxidative DNA damage. *Nature* 2007, 447, 941–950. [PubMed: 17581577]
- (130). Minowa O; Arai T; Hirano M; Monden Y; Nakai S; Fukuda M; Itoh M; Takano H; Hippou Y; Aburatani H; Masumura K; Nohmi T; Nishimura S; Noda T Mmh/Ogg1 gene inactivation results

in accumulation of 8-hydroxyguanine in mice. *Proc. Natl. Acad. Sci. U.S.A* 2000, 97, 4156–4161. [PubMed: 10725358]

- (131). Sakamoto K; Tominaga Y; Yamauchi K; Nakatsu Y; Sakumi K; Yoshiyama K; Egashira A; Kura S; Yao T; Tsuneyoshi M; Maki H; Nakabeppu Y; Tsuzuki T MUTYH-null mice are susceptible to spontaneous and oxidative stress induced intestinal tumorigenesis. *Cancer Res.* 2007, 67, 6599–6604. [PubMed: 17638869]
- (132). Al-Tassan N; Chmiel NH; Maynard J; Fleming N; Livingston AL; Williams GT; Hodges AK; Davies DR; David SS; Sampson JR; Cheadle JP Inherited variants of MYH associated with somatic G:C→T:A mutations in colorectal tumors. *Nat. Genet* 2002, 30, 227–232. [PubMed: 11818965]
- (133). Kamiya H; Kurokawa M; Makino T; Kobayashi M; Matsuoka I Induction of action-at-a-distance mutagenesis by 8-oxo-7,8-dihydroguanine in DNA pol lambda-knockdown cells. *Genes Environ.* 2015, 37, 10. [PubMed: 27350807]
- (134). Neeley WL; Delaney S; Alekseyev YO; Jarosz DF; Delaney JC; Walker GC; Essigmann JM DNA Polymerase V allows bypass of toxic guanine oxidation products in vivo. *J. Biol. Chem* 2007, 282, 12741–12748. [PubMed: 17322566]
- (135). D'Onofrio G; Jabbari K; Musto H; Alvarez-Valin F; Cruveiller S; Bernardi G Evolutionary genomics of vertebrates and its implications. *Ann. N.Y. Acad. Sci* 1999, 870, 81–94. [PubMed: 10415475]
- (136). Fong YW; Cattoglio C; Tjian R The intertwined roles of transcription and repair proteins. *Mol. Cell* 2013, 52, 291–302. [PubMed: 24207023]
- (137). Banda DM; Nunez NN; Burnside MA; Bradshaw KM; David SS Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: Mechanism, metals and medicine. *Free Radic. Biol. Med* 2017, 107, 202–215. [PubMed: 28087410]
- (138). Leipold MD; Workman H; Muller JG; Burrows CJ; David SS Recognition and removal of oxidized guanines in duplex DNA by the base excision repair enzymes hOGG1, yOGG1, and yOGG2 *Biochemistry* 2003, 42, 11373–11381. [PubMed: 14503888]
- (139). Raetz AG; David SS When you're strange: Unusual features of the MUTYH glycosylase and implications in cancer. *DNA Repair (Amst)* 2019, 80, 16–25. [PubMed: 31203172]
- (140). Nash HM; Lu R; Lane WS; Verdine GL The critical active-site amine of the human 8-oxoguanine DNA glycosylase, hOgg1: direct identification, ablation and chemical reconstitution. *Chem. Biol* 1997, 4, 693–702. [PubMed: 9331411]
- (141). Esadze A; Rodriguez G; Cravens SL; Stivers JT AP-endonuclease 1 accelerates turnover of human 8-oxoguanine DNA glycosylase by preventing retrograde binding to the abasic-site product. *Biochemistry* 2017, 56, 1974–1986. [PubMed: 28345889]
- (142). Hill JW; Hazra TK; Izumi T; Mitra S Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic Acids Res.* 2001, 29, 430–438. [PubMed: 11139613]
- (143). David SS; Williams SD Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chem. Rev* 1998, 98, 1221–1262. [PubMed: 11848931]
- (144). Perillo B; Ombra MN; Bertoni A; Cuzzo C; Sacchetti S; Sasso A; Chiariotti L; Malorni A; Abbondanza C; Avvedimento EV DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science* 2008, 319, 202–206. [PubMed: 18187655]
- (145). Culhane JC; Cole PA LSD1 and the chemistry of histone demethylation. *Curr. Opin. Chem. Biol* 2007, 11, 561–568. [PubMed: 17851108]
- (146). Austin CA; Lee KC; Swan RL; Khazeem MM; Manville CM; Cridland P; Treumann A; Porter A; Morris NJ; Cowell IG TOP2B: The first thirty years. *Int. J. Mol. Sci* 2018, 19.
- (147). Zuchegna C; Aceto F; Bertoni A; Romano A; Perillo B; Laccetti P; Gottesman ME; Avvedimento EV; Porcellini A Mechanism of retinoic acid-induced transcription: histone code, DNA oxidation and formation of chromatin loops. *Nucleic Acids Res.* 2014, 42, 11040–11055. [PubMed: 25217584]
- (148). Pan L; Zhu B; Hao W; Zeng X; Vlahopoulos SA; Hazra TK; Hegde ML; Radak Z; Bacsí A; Brasier AR; Ba X; Boldogh I Oxidized guanine base lesions function in 8-oxoguanine DNA

glycosylase1-mediated epigenetic regulation of nuclear factor kappaB-driven gene expression. *J. Biol. Chem* 2016, 291, 25553–25566. [PubMed: 27756845]

- (149). Antoniali G; Lirussi L; D'Ambrosio C; Dal Piaz F; Vascotto C; Casarano E; Marasco D; Scaloni A; Fogolari F; Tell G SIRT1 gene expression upon genotoxic damage is regulated by APE1 through nCaRE-promoter elements. *Mol. Biol. Cell* 2014, 25, 532–547. [PubMed: 24356447]
- (150). Dai J; Chen D; Jones RA; Hurley LH; Yang D NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res.* 2006, 34, 5133–5144. [PubMed: 16998187]
- (151). Kang HJ; Kendrick S; Hecht SM; Hurley LH 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* 2014, 136, 4172–4185. [PubMed: 24559432]
- (152). Fuss JO; Tsai C-L; Ishida JP; Tainer JA Emerging critical roles of Fe-S clusters in DNA replication and repair. *Biochem. Biophys. Acta* 2015, 1853, 1253–1271. [PubMed: 25655665]
- (153). Imlay JA Iron-sulphur clusters and the problem with oxygen. *Mol. Microbiol* 2006, 59, 1073–1082. [PubMed: 16430685]
- (154). Shafirovich V; Geacintov NE Removal of oxidatively generated DNA damage by overlapping repair pathways. *Free Radic. Biol. Med* 2017, 107, 53–61. [PubMed: 27818219]
- (155). 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 AS; 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; Helleday T Small-molecule inhibitor of OGG1 suppresses proinflammatory gene expression and inflammation. *Science* 2018, 362, 834–839. [PubMed: 30442810]
- (156). Bravard A; Vacher M; Moritz E; Vaslin L; Hall J; Epe B; Radicella JP Oxidation status of human OGG1-S326C polymorphic variant determines cellular DNA repair capacity. *Cancer Res.* 2009, 69, 3642–3649. [PubMed: 19351836]
- (157). Alnajjar KS; Sweasy JB A new perspective on oxidation of DNA repair proteins and cancer. *DNA Repair (Amst)* 2019, 76, 60–69. [PubMed: 30818170]
- (158). Pastukh V; Roberts JT; Clark DW; Bardwell GC; Patel M; Al-Mehdi AB; Borchert GM; Gillespie MN 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* 2015, 309, L1367–1375. [PubMed: 26432868]
- (159). Sun D; Liu WJ; Guo K; Rusche JJ; Ebbinghaus S; Gokhale V; Hurley LH The proximal promoter region of the human vascular endothelial growth factor gene has a G-quadruplex structure that can be targeted by G-quadruplex-interactive agents. *Mol. Cancer Ther* 2008, 7, 880–889. [PubMed: 18413801]
- (160). Gray LT; Vallur AC; Eddy J; Maizels N G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nat. Chem. Biol* 2014, 10, 313–318. [PubMed: 24609361]
- (161). Broxson C; Hayner JN; Beckett J; Bloom LB; Tornaletti S Human AP endonuclease inefficiently removes abasic sites within G4 structures compared to duplex DNA. *Nucleic Acids Res.* 2014, 42, 7708–7719. [PubMed: 24848015]
- (162). Bhakat KK; Mantha AK; Mitra S Transcriptional regulatory functions of mammalian AP-endonuclease (APE1/Ref-1), an essential multifunctional protein. *Antioxid. Redox Signal* 2009, 11, 621–638. [PubMed: 18715144]
- (163). Burra S; Marasco D; Malfatti MC; Antoniali G; Virgilio A; Esposito V; Demple B; Galeone A; Tell G Human AP-endonuclease (Ape1) activity on telomeric G4 structures is modulated by acetyltable lysine residues in the N-terminal sequence. *DNA Repair (Amst)* 2019, 73, 129–143. [PubMed: 30509560]
- (164). Fleming AM; Zhu J; Ding Y; Burrows CJ 8-Oxo-7,8-dihydroguanine in the context of a promoter G-quadruplex is an on-off switch for transcription. *ACS Chem. Biol* 2017, 12, 2417–2426. [PubMed: 28829124]

- (165). Fleming AM; Zhu J; Ding Y; Burrows CJ Location dependence of the transcriptional response of a potential G-quadruplex in gene promoters under oxidative stress. *Nucleic Acids Res.* 2019, 47, 5049–5060. [PubMed: 30916339]
- (166). Cogoi S; Xodo LE G-quadruplex formation within the promoter of the *KRAS* proto-oncogene and its effect on transcription. *Nucleic Acids Res.* 2006, 34, 2536–2549. [PubMed: 16687659]
- (167). Wallace SS Base excision repair: A critical player in many games. *DNA Repair* 2014, 19, 14–26. [PubMed: 24780558]
- (168). Redstone SCJ; Fleming AM; Burrows CJ Oxidative modification of the potential G-quadruplex sequence in the PCNA gene promoter can turn on transcription. *Chem. Res. Toxicol* 2019, 32, 437–446. [PubMed: 30604962]
- (169). Sejersted Y; Hildrestrand GA; Kunke D; Rolseth V; Krokeide SZ; Neuraater CG; Suganthan R; Atneosen-Asegg M; Fleming AM; Saugstad OD; Burrows CJ; Luna L; Bjoras M Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia. *Proc. Natl. Acad. Sci. U.S.A* 2011, 108, 18802–18807. [PubMed: 22065741]
- (170). Wang J; Wang Q; Watson LJ; Jones SP; Epstein PN Cardiac overexpression of 8-oxoguanine DNA glycosylase 1 protects mitochondrial DNA and reduces cardiac fibrosis following transaortic constriction. *Am. J. Physiol. Heart Circ. Physiol* 2011, 301, H2073–H2080. [PubMed: 21873502]
- (171). Shinmura K; Kato H; Kawanishi Y; Igarashi H; Goto M; Tao H; Inoue Y; Nakamura S; Misawa K; Mineta H; Sugimura H Abnormal expressions of DNA glycosylase genes NEIL1, NEIL2, and NEIL3 are associated with somatic mutation loads in human cancer. *Oxid. Med. Cell. Longev* 2016, 2016, 1546392. [PubMed: 27042257]
- (172). Zhu J; Fleming AM; Burrows CJ The RAD17 promoter sequence contains a potential tail-dependent G-quadruplex that downregulates gene expression upon oxidative modification. *ACS Chem. Biol* 2018, 13, 2577–2584. [PubMed: 30063821]
- (173). Pei D-S; Jia P-P; Luo J-J; Liu W; Strauss PR AP endonuclease 1 (Apex1) influences brain development linking oxidative stress and DNA repair. *Cell Death Dis.* 2019, 10, 348. [PubMed: 31024003]
- (174). Okamura K; Doi T; Hamada K; Sakurai M; Yoshioka Y; Mitsuzono R; Migita T; Sumida S; Sugawa-Katayama Y Effect of repeated exercise on urinary 8-hydroxy-deoxyguanosine excretion in humans. *Free Radic. Res* 1997, 26, 507–514. [PubMed: 9212344]
- (175). Marfe G; Tafani M; Pucci B; Di Stefano C; Indelicato M; Andreoli A; Russo MA; Sinibaldi-Salimei P; Manzi V The effect of marathon on mRNA expression of anti-apoptotic and pro-apoptotic proteins and sirtuins family in male recreational long-distance runners. *BMC Physiol.* 2010, 10, 7. [PubMed: 20462402]
- (176). Balasubramanian S; Hurley LH; Neidle S Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? *Nat. Rev. Drug Discov* 2011, 10, 261–275. [PubMed: 21455236]
- (177). Asamitsu S; Obata S; Yu Z; Bando T; Sugiyama H Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules (Basel, Switzerland)* 2019, 24, 10.3390/molecules24030429.
- (178). Brooks TA; Hurley LH The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat. Rev. Cancer* 2009, 9, 849–861. [PubMed: 19907434]
- (179). Chen K; Zhao Boxuan S.; He C Nucleic acid modifications in regulation of gene expression. *Cell Chem. Biol* 2016, 23, 74–85. [PubMed: 26933737]
- (180). Culp SJ; Cho BP; Kadlubar FF; Evans FE Structural and conformational analyses of 8-hydroxy-2'-deoxyguanosine. *Chem. Res. Toxicol* 1989, 2, 416–422. [PubMed: 2519731]
- (181). Drake JW The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. *Ann. N.Y. Acad. Sci* 1999, 870, 100–107. [PubMed: 10415476]
- (182). Hudson NO; Buck-Koehntop BA Zinc finger readers of methylated DNA. *Molecules (Basel, Switzerland)* 2018, 23, 2555.
- (183). Pfaffeneder T; Spada F; Wagner M; Brandmayr C; Laube SK; Eisen D; Truss M; Steinbacher J; Hackner B; Kotljarova O; Schuermann D; Michalakakis S; Kosmatchev O; Schiesser S; Steigenberger B; Raddaoui N; Kashiwazaki G; Muller U; Spruijt CG; Vermeulen M; Leonhardt

- H; Schar P; Muller M; Carell T Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol* 2014, 10, 574–581. [PubMed: 24838012]
- (184). Carell T; Kurz MQ; Muller M; Rossa M; Spada F Non-canonical bases in the genome: The regulatory information layer in DNA. *Angew. Chem. Int. Ed. Engl* 2018, 57, 4296–4312. [PubMed: 28941008]
- (185). Shen JC; Rideout WM 3rd; Jones PA The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.* 1994, 22, 972–976. [PubMed: 8152929]
- (186). Tubbs A; Nussenzweig A Endogenous DNA damage as a source of genomic instability in cancer. *Cell* 2017, 168, 644–656. [PubMed: 28187286]
- (187). Arai T; Kelly VP; Minowa O; Noda T; Nishimura S High accumulation of oxidative DNA damage, 8-hydroxyguanine, in Mmh/Ogg1 deficient mice by chronic oxidative stress. *Carcinogenesis* 2002, 23, 2005–2010. [PubMed: 12507922]
- (188). Song CX; Szulwach KE; Dai Q; Fu Y; Mao SQ; Lin L; Street C; Li Y; Poidevin M; Wu H; Gao J; Liu P; Li L; Xu GL; Jin P; He C Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* 2013, 153, 678–691. [PubMed: 23602153]
- (189). Raiber E-A; Beraldi D; Ficiz G; Burgess HE; Branco MR; Murat P; Oxley D; Booth MJ; Reik W; Balasubramanian S Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol.* 2012, 13, R69. [PubMed: 22902005]
- (190). Zhang Y; Zhou C Formation and biological consequences of 5-Formylcytosine in genomic DNA. *DNA Repair (Amst)* 2019, 81, 102649. [PubMed: 31303545]
- (191). Xue G; Chu XY; Zhang HY Oxidative DNA damage is associated more with genome accessibility than spatial positioning in the nucleus. *J. Biomol. Struct. Dyn* 2019, 37, 1857–1862. [PubMed: 29663850]
- (192). Hamilton ML; Guo Z; Fuller CD; Van Remmen H; Ward WF; Austad SN; Troyer DA; Thompson I; Richardson A A reliable assessment of 8-oxo-2-deoxyguanosine levels in nuclear and mitochondrial DNA using the sodium iodide method to isolate DNA. *Nucleic Acids Res.* 2001, 29, 2117–2126. [PubMed: 11353081]
- (193). Rahimoff R; Kosmatchev O; Kirchner A; Pfaffeneder T; Spada F; Brantl V; Muller M; Carell T 5-Formyl- and 5-carboxydeoxycytidines do not cause accumulation of harmful repair intermediates in stem cells. *J. Am. Chem. Soc* 2017, 139, 10359–10364. [PubMed: 28715893]
- (194). Meng FL; Du Z; Federation A; Hu J; Wang Q; Kieffer-Kwon KR; Meyers RM; Amor C; Wasserman CR; Neuberg D; Casellas R; Nussenzweig MC; Bradner JE; Liu XS; Alt FW Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* 2014, 159, 1538–1548. [PubMed: 25483776]
- (195). Rydberg B; Lindahl T Nonenzymatic methylation of DNA by the intracellular methyl group donor *S*-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.* 1982, 1, 211–216. [PubMed: 7188181]
- (196). McKibbin PL; Fleming AM; Towheed MA; Van Houten B; Burrows CJ; David SS Repair of hydantoin lesions and their amine adducts in DNA by base and nucleotide excision repair. *J. Am. Chem. Soc* 2013, 135, 13851–13861. [PubMed: 23930966]
- (197). Fleming AM; Zhu J; Ding Y; Esders S; Burrows CJ Oxidative modification of guanine in a potential Z-DNA-forming sequence of a gene promoter impacts gene expression. *Chem. Res. Toxicol* 2019, 32, 899–909. [PubMed: 30821442]
- (198). Kaiser CE; Van Ert NA; Agrawal P; Chawla R; Yang D; Hurley LH Insight into the complexity of the i-motif and G-quadruplex DNA structures formed in the KRAS promoter and subsequent drug-induced gene repression. *J. Am. Chem. Soc* 2017, 139, 8522–8536. [PubMed: 28570076]
- (199). Brown RV; Wang T; Chappeta VR; Wu G; Onel B; Chawla R; Quijada H; Camp SM; Chiang ET; Lassiter QR; Lee C; Phanse S; Turnidge MA; Zhao P; Garcia JGN; Gokhale V; Yang D; Hurley LH The consequences of overlapping G-quadruplexes and i-motifs in the platelet-derived growth factor receptor beta core promoter nuclease hypersensitive element can explain the unexpected effects of mutations and provide opportunities for selective targeting of both structures by small molecules to downregulate gene expression. *J. Am. Chem. Soc* 2017, 139, 7456–7475. [PubMed: 28471683]

- (200). Belotserkovskii BP; Tornaletti S; D'Souza AD; Hanawalt PC R-loop generation during transcription: Formation, processing and cellular outcomes. *DNA Repair (Amst)* 2018, 71, 69–81. [PubMed: 30190235]
- (201). Kuznetsov VA; Bondarenko V; Wongsurawat T; Yenamandra SP; Jenjaroenpun P Toward predictive R-loop computational biology: genome-scale prediction of R-loops reveals their association with complex promoter structures, G-quadruplexes and transcriptionally active enhancers. *Nucleic Acids Res.* 2018, 46, 7566–7585. [PubMed: 29945198]
- (202). Zheng KW; He YD; Liu HH; Li XM; Hao YH; Tan Z Superhelicity constrains a localized and R-loop-dependent formation of G-quadruplexes at the upstream region of transcription. *ACS Chem. Biol* 2017, 12, 2609–2618. [PubMed: 28846373]
- (203). Delaney S; Barton JK Charge Transport in DNA Duplex/Quadruplex Conjugates *Biochemistry* 2003, 42, 14159–14165. [PubMed: 14640683]
- (204). Stairs S; Nikmal A; Bucar DK; Zheng SL; Szostak JW; Powner MW Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides. *Nat. Commun* 2017, 8, 15270. [PubMed: 28524845]
- (205). Nguyen KV; Burrows CJ Whence flavins? Redox-active ribonucleotides link metabolism and genome repair to the RNA world. *Acc. Chem. Res* 2012, 45, 2151–2159. [PubMed: 23054469]
- (206). Kim SC; O'Flaherty DK; Zhou L; Lelyveld VS; Szostak JW Inosine, but none of the 8-oxo-purines, is a plausible component of a primordial version of RNA. *Proc. Natl. Acad. Sci. U.S.A* 2018, 115, 13318–13323. [PubMed: 30509978]

Figure 1.

Sequences of DNA abundant in 5'-GG-3' dinucleotides are oxidation prone. (A) Comparison of dinucleotide frequencies between *H. sapiens* and *C. elegans* that have a similar G abundance but different dinucleotide frequencies. The data are replotted from the literature.¹⁴ Because of sequence complementarity in duplex DNA, not all possible dinucleotide combinations are shown (i.e., GG = CC and only GG is shown). (B) Plot of ionization potentials from Saito et al. illustrating G has the lowest potential, and runs of G render these sequences even more oxidation prone.² (C) In duplex DNA, oxidation at a remote site can induce a chemical reaction at a distal 5'-GG-3' via generation of a migratory electron hole (h^+).

Figure 2. Metabolism and inflammation yield ROS that induce oxidative modification of the G heterocycle.

Figure 3.

Certain G-rich sequences can adopt G4 folds. (A) The consensus sequence for a PQS (2 examples with a 5th G track on either the 5' or 3' side of the principal G4 folding sequence). (B) Hoogsteen base pairing in a G-tetrad. (C) Cartoon representations of the global G4 folds to depict the location of the Gs, *syn* (red) vs. *anti* (blue) conformations of the Gs, and loop nomenclature.

Figure 4.

Potential G-quadruplex forming sequences show enrichment in the human genome in regulatory elements. (A) Regions in which PQSs are found enriched in the human genome, and (B) histograms of PQS distributions around the TSSs of human genes in the template (left) or coding (right) strands. The complete distributions of PQSs around TSSs (gray left or red right) are also compared to PQSs specific to human DNA repair genes (blue lines). The data in panel B are replotted from our original report on these distributions.⁸⁹

Figure 5.

The structural plasticity of native G4 folds can accommodate the structurally destabilizing modification OG. (A) G-Tetrad hydrogen bonding is compromised by OG. (B) In the human *VEGF* promoter G4, the presence of OG engages the 5th G run or “spare tire” to maintain the G4 fold.⁸²

Figure 6.

Oxidative modification of G to OG yields a Janus-faced purine that base pairs with C on the Watson-Crick face or A on the Hoogsteen face leading to a G→T transversion mutation in the absence of faithful DNA repair.

Figure 7.

The base excision repair of an OG:C base pair initiated by OGG1. *The proposed mechanism assumes OGG1 is a monofunctional glycosylase as supported by recent cellular studies.¹²⁵

Figure 8.

Proposed pathways by which G oxidation to OG in a gene promoter activates transcription.

(A) Avvendimento's proposed pathway by estrogen-receptor mediated gene activation for driving OG formation in a promoter and OGG1 activity leading to an entry point for TOP2B-mediated gene activation.¹⁴⁴ (B) Boldogh/Ba and co-workers found Cys oxidation in OGG1 (OGG1^{ox}) attenuates the glycosylase activity to stall the protein for active recruitment of NF-κB to OG-bearing promoters for gene activation during inflammation.¹⁴⁸

(C) Our laboratory found OGG1 release of OG in a promoter provides the drive for a duplex→G4 switch to display an AP to APE1 with stalled endonuclease activity on the G4 allowing recruitment of activating transcription factors via the Ref-1 properties during oxidative/inflammatory stress.⁴ (D) Xodo and co-workers found OG stabilized G4s were binding substrates for MAZ and A1 that led to resolving the non-canonical fold for BER-mediated activation during oxidative stress.⁵ (E) Tell and co-workers found oxidation in the nCaRE hairpin forming sequence provided a substrate for APE1 to interact with Ku70/80 and RNA pol II for gene activation under oxidative stress.¹⁴⁹

Figure 9.
Proposed relative energy-level diagram for gene regulation via G oxidation in a promoter PQS.

Figure 10.
DNA repair gene promoter oxidation is the stimulus for upregulated transcription.

Figure 11. Structural comparisons of OG and 5mC. (A) The dominant tautomers of OG and 5mC favor Watson-Crick base pairing. (B) The OG:C and 5mC:G base pairs alter the major groove hydrogen bond donor and acceptor patterns.

Figure 12.

The BER process returns the oxidation products of G and 5mC back to the native state.

Figure 13.

The impact of G oxidation to OG in a gene promoter sequence that can switch structure to a non-canonical fold. (A) In the coding strand of a promoter, OG in sequences that can adopt G4s, i-motifs (iM), Z-DNA enhance gene expression; in contrast, OG in a sequence only capable of B-DNA formation has no impact on transcription. (B) In the template strand of the promoter, OG in all four sequence contexts caused a significant decrease in gene expression. Significance was determined by a Student's t-test where $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$. This figure is adapted from our prior publication on this topic that provides all the experimental details for collection of the data.¹⁹⁷