SURVEY AND SUMMARY

Oxidative stress-mediated epigenetic regulation by G-quadruplexes

Aaron M. Fleming[*] and Cynthia J. Burrows [iD][*]

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

ABSTRACT

Many cancer-associated genes are regulated by guanine (G)-rich sequences that are capable of refolding from the canonical duplex structure to an intrastrand G-quadruplex. These same sequences are sensitive to oxidative damage that is repaired by the base excision repair glycosylases OGG1 and NEIL1–3. We describe studies indicating that oxidation of a guanosine base in a gene promoter G-quadruplex can lead to up- and downregulation of gene expression that is location dependent and involves the base excision repair pathway in which the first intermediate, an apurinic (AP) site, plays a key role mediated by AP endonuclease 1 (APE1/REF1). The nuclease activity of APE1 is paused at a G-quadruplex, while the REF1 capacity of this protein engages activating transcription factors such as HIF-1α, AP-1 and p53. The mechanism has been probed by *in vitro* biophysical studies, whole-genome approaches and reporter plasmids *in cellulo*. Replacement of promoter elements by a G-quadruplex sequence usually led to upregulation, but depending on the strand and precise location, examples of downregulation were also found. The impact of oxidative stress-mediated lesions in the G-rich sequence enhanced the effect, whether it was positive or negative.

GRAPHICAL ABSTRACT
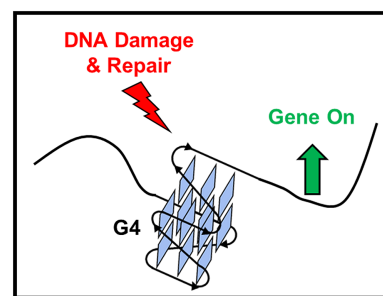
INTRODUCTION

Many cancer-associated genes are regulated by guanine (G)-rich sequences in their gene promoters through the formation of G-quadruplex (G4) secondary structures [1]. These sites also represent hotspots for oxidative damage [2], and oxidative stress is a feature common to many cancer types [3]. Thus, the intersection of G oxidation and G4 folding in promoter regions can have a major impact on gene expression and consequently on the progression of cancer [4]. Here, we summarize the major pathways of G oxidation, its impact on G4 folding and what we know so far about the role of DNA repair in modulating gene expression when DNA damage is present in a potential G4 sequence (PQS).

G4s can form from DNA or RNA sequences containing four or more closely spaced G tracks comprised of $G_{n\geq 3}$ for DNA or $G_{n\geq 2}$ for RNA in which one G from each of four tracks is hydrogen bonded in Hoogsteen fashion to create a flat G tetrad; these then assemble via π stacking with K^+ ions bound in central positions in between layers (Figure 1A and B) [5]. Although the human telomere sequence (hTelo) exhibits a more complex fold, adopting parallel, antiparallel and hybrid folds with cation dependence, most G4 structures in DNA gene promoters and in RNA are parallel stranded with all of the G tracks oriented in the same direction yielding a compact structure with extended

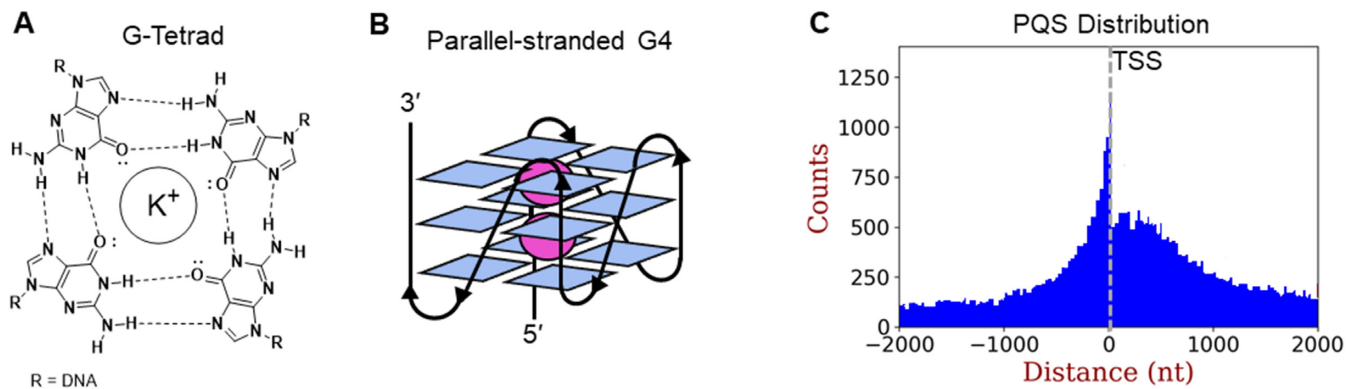**Figure 1.** (**A**) The G-tetrad forms one layer of a folded G4. (**B**) A parallel G4. (**C**) Distribution of PQSs with respect to the TSS (11).

flat surfaces above and below the stacked quartets (6). The core is held together by loops, comprised of typically one to three nucleotides, that bridge from the top of the stack to the bottom. Interestingly, this is most commonly achieved with only one nucleotide in the loop. Longer loops can favor antiparallel arrangements of the G tracks and may also form hairpins or more complex structures (7,8). Myriad examples have been analyzed by NMR spectroscopy that include bulged nucleotides or other unusual secondary structures (7,8). Generally speaking, cationic small molecules with extended aromatic rings are likely to be candidates for G4 binding, and the diversity of 3D structures provided by the loops and flanking sequences lends hope to the idea of G4-specific targeting for therapeutic applications (9,10).

PQSs exist in hundreds of thousands of sites in the human genome, but their distribution is non-random (12). Apart from telomeres at the ends of chromosomes, PQSs are common in regulatory regions in mammalian and plant genomes but less common in bacteria and yeast (13,14). In mammals, the intron-1/exon-1 boundary has an unusually high frequency of PQSs, where the sequence may serve as a site for N^6-methylation of adenosine in pre-mRNA and regulation of mRNA splicing (15,16). Of interest to our laboratory is the high frequency of PQSs in gene promoters, just in advance of the transcription start site (TSS) (17,18). Pertinent to this discussion, we found that many of the 191 genes involved in mammalian DNA repair have PQSs in their promoters, and that their distribution was similar to the overall set of ~10 000 genes that harbor G4s (Figure 1C) (17,19). Notably, many oncogenes and other cancer-associated genes (e.g. *c-MYC*, *VEGF*; see Table 1) have G4 sequences that appear to regulate their expression making these sites targets for pharmaceutical intervention (1,20). For DNA repair genes, the prevalence of G4s makes sense if the G4 structure is sensitive to the presence of chemical modifications from oxidation, alkylation, deamination or strand breaks. For oxidation, G4 folding during endogenous oxidative stress can upregulate the repair genes needed for correcting lesions in the genome. Additionally, 5'-untranslated regions (5'-UTRs) just downstream of the TSS often contain PQSs, and if present in the non-template strand of DNA, they would be transcribed to form PQSs in the 5'-UTR of mRNA, presenting two opportunities to regulate gene expression either in DNA or in RNA (17,18).

Furthermore, mounting evidence suggests that G4s near the TSS, which can be a difficult site to pinpoint, can participate in R-loop formation wherein the DNA–RNA hybrid plays a role in gene initiation (21,22). In this situation, the presence of folded G4s in the non-template strand helps keep the two strands separated and may assist in the recruitment of factors needed for RNA synthesis and elongation (23).

DNA damage and repair as a trigger for G4 folding

Hypoxia, metabolic levels and inflammation intertwine in cells to generate oxidative stress—an imbalance of oxidants (O_2) and reductants [thiols and NAD(P)H] (34). Reactive oxygen species (ROS) are generated from both very low [O_2] and high [O_2] conditions and can take many forms. Although the Fenton reaction is invoked as a cellular source of hydroxyl radical (HO$^\bullet$), this reaction is intercepted by even low concentrations of bicarbonate, a major buffering component of mammalian cells, so that Fe(II) complexes reacting with H_2O_2 produce $CO_3^{\bullet-}$ as a major ROS species instead of HO$^\bullet$ (35,36). Carbonate radical anion is also formed during inflammation in which peroxynitrite (from upregulated iNOS—inducible nitric oxide synthase) reacts with dissolved CO_2 to yield ultimately $CO_3^{\bullet-}$ again as the principal ROS (37). The distinction between HO$^\bullet$ and $CO_3^{\bullet-}$ is important because these two radicals react very differently with DNA. Hydroxyl radical is highly reactive and abstracts H$^\bullet$ indiscriminately from all sites in 2-deoxyribose as well as forming adducts with all DNA and RNA bases (38). Ionizing radiation can also directly form base radicals by electron ejection as well as forming HO$^\bullet$ (39); however, we have proposed that *endogenous* pathways, via peroxynitrite- or Fe(II)-mediated oxidation of HCO_3^-/CO_2, tend to favor the formation of $CO_3^{\bullet-}$ as the major ROS attacking DNA in aerobic organisms (35,36).

In contrast to HO$^\bullet$, $CO_3^{\bullet-}$ acts as a one-electron oxidant and is more selective in its reactions (40). Because G is the most electron-rich of the nucleobases, and correspondingly has the lowest reduction potential (Figure 2), it is the most common nucleotide to be modified during oxidative stress (40). The immediate product of electron loss is a guanosine radical cation, G$^{\bullet+}$, which represents an electron hole. In DNA that is perfectly base paired, the double helix can easily transport electron holes along the π-stacked bases until the most stable site is found (41,42). Following this

Table 1. Examples of G4-forming sequences in the human genome and their locations

Gene/name	Strand	Location w.r.t. TSS	Sequence (5′ to 3′)	Reference
hTelo	n.a.	n.a.	TA**GGG**TTA**GGG**TTA**GGG**TTA**GGG**TTA**GGG**TT…	(24)
VEGF	Non-template	−50	CC**GGGG**C**GGG**CC**GGGG**C**GGGG**TCCC**GGG**C**GGGG**TC	(25)
c-MYC	Template	−262	CC**GGG**A**GGGG**CGCTTATG**GGG**A**GGG**T**GG**GGA**GGG**T**GGG**GAA**GG**T**GGGG**AG	(26)
KRAS	Template	−104	GA**GGG**AGC**GG**CTGA**GGG**C**GG**TGT**GGG**AAGA**GG**GAAGA**GGGG**GAG	(27)
HIF-1α	Template	−48	TC**GGG**CGCG**CGGGG**A**GGGG**AGA**GGGG**C**GG**GAGC	
HSP90	Non-template	−77	GA**GGG**C**GGG**CCAAA**GGG**AA**GGGG**T**GGG**CT	(28)
TERT	Template	−18	AA**GGGG**A**GGGG**CT**GGG**A**GGG**CCC**GG**A**GGGG**GCT**GGG**CC**GGGG**ACCC**GG**GA**GGGG**TC**GG**GAC**GGGG**C**GGGG**TC	(29)
BCL2	Template	−1386	GA**GGGG**C**GGG**CGC**GGG**A**GG**AA**GGGG**C**GGG**AGC**GGGG**CT	(7)
NEIL3	Non-template	−3	TA**GGG**TGCTGTTT**GGG**C**GGGG**CCT**GGG**C**GG**GGCC	(30)
NTHL1	Non-template	−279	TC**GGG**TTGCAGT**GGG**CGC**GGG**TGA**GGG**CCC**GG**GAC	(31)
PCNA	Non-template	−126	CA**GGG**A**GG**CA**GGG**CGAC**GGGGG**C**GGGG**C**GG**GGCG	(32)
RAD17	Template	−18	CC**GGG**A**GGG**ACT**GGG**CT**GGGG**CA**GG**CT**GG**GGCG	(33)
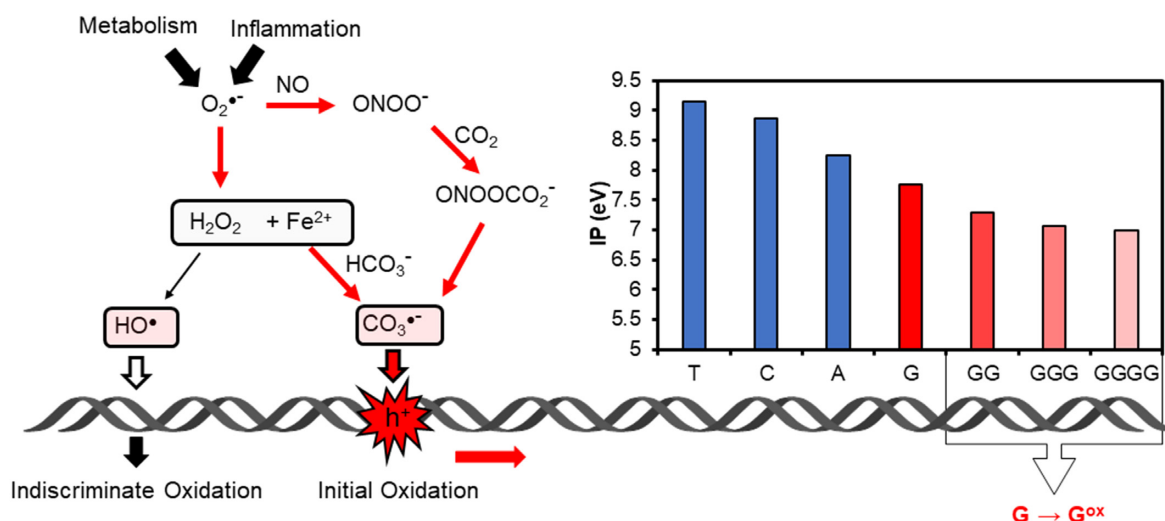
Figure 2. ROS, principally $CO_3^{\bullet-}$ from endogenous oxidative stress, generates electron holes in DNA that travel the duplex looking for a low ionization potential site such as a G track of a PQS (4).

rapid charge transport, which has been demonstrated to occur over hundreds or even thousands of base pairs (43), a slower chemical step produces the oxidized base, typically 8-oxo-7,8-dihydro-2′-deoxyguanosine (OG, Figure 3) (38). The most stable site for the $G^{\bullet+}$ is one in which one or more G bases are stacked on its 3′ side. Thus, the G tracks of a PQS, i.e. 5′-GGG-3′, in its duplex state, are the best sites for one-electron oxidation by $CO_3^{\bullet-}$ directly, or indirectly via charge transport (41,42). In such a sequence, any G followed by a 3′ G will be a good site for oxidation; thus, underlined Gs in the sequence above would be most susceptible to form OG.

Folded G4s are also targets for oxidation, but the pattern of reactivity is somewhat different. Among the π-stacked Gs of a 5′-GGG-3′ track, there is a distinct preference for oxidation of the 5′-most G (44). We speculate that the central G, which would normally follow the rule of being highly reactive in a duplex because of the presence of the 3′ G, is less reactive in a folded G4 because of the two K^+ ions closely associated with the central tetrad that might disfavor formation of $G^{\bullet+}$ at this site. Importantly, one should note that any Gs that are extruded from the double helix or the quadruplex are particularly reactive toward oxidants because of solvent accessibility (38,45). Many promoter G4s have an excess number of Gs; that is, some but not all tracks have four or five Gs instead of three, such that they will reside in loops of the folded G4 (Table 1) (45). For example, observation of an oxidation pattern of 5′-GGGG-3′ would suggest that the first three Gs are part of a G4 and the 3′-most G is looped out and reactive. Other ROS, such as singlet oxygen and ozone, do not follow a sequence pattern but rely entirely on the accessibility of the π face of G toward re-
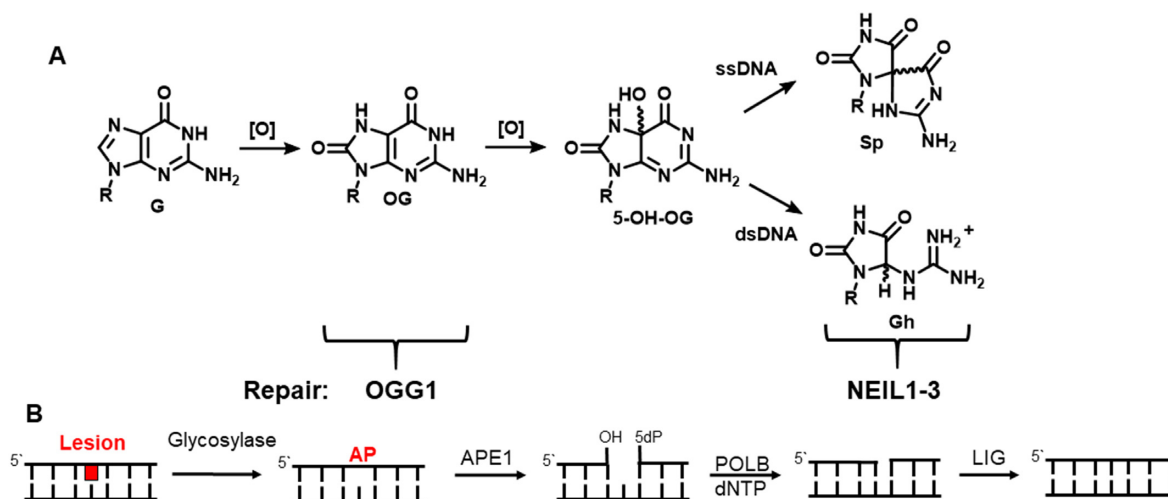
Figure 3. (**A**) Oxidation products of G include OG, the most common, and its overoxidation products Sp and Gh. (**B**) These lesions are repaired by the BER pathway using OGG1 for OG and NEIL glycosylases for Sp and Gh.

action (46). Indeed, we used this feature with G4s of known structure to suggest the absolute configuration of chiral oxidation products, such as the hydantoins Sp and Gh (see Figure 3) (47). Sp and Gh are further oxidation products of OG (48,49), present in lower amounts than OG in cells undergoing oxidative stress, that arise because OG is thousands of times more reactive than G toward one-electron oxidation and many other ROS as well (50). Many other oxidation products exist besides OG, Sp and Gh, and these have been reviewed (39,51); however, we will focus on OG for the remainder of this discussion.

Oxidized base lesions such as OG are typically repaired by the base excision repair (BER) pathway (Figure 3B), and for OG this is a multilayer process (52). OG itself is not very mutagenic because it is well accommodated in the B-form helix and base pairs well with C (53). However, its potential mutagenicity lies with the ability of OG in single-stranded DNA (ssDNA) to rotate about the glycosidic bond from the *anti* to the *syn* conformation, alleviating some unfavorable interactions between the C8 carbonyl and oxygen atoms of the deoxyribose–phosphate backbone. OG_{syn} forms a good base pair with A during replication, which would lead to a G→T transversion mutation if not repaired (54). Thus, the BER glycosylases that break the glycosidic bond between sugar and base are highly specific in the case of OG for double-stranded (dsDNA) DNA substrates. OGG1 is a glycosylase that recognizes and removes OG when paired with C; MUTYH correspondingly recognizes OG paired with A and removes the unmodified, but incorrect, A (52,54). After repairing the OG:A mispair, the OG:C base pair can then undergo normal repair by OGG1 in a second round. Accordingly, cell lines or mice with Ogg1 knocked out are viable and do not have a prominent change in phenotype, but are somewhat more susceptible to inflammatory stress (55). In contrast, mutations in MUTYH can lead to colon cancer because the OG:A mispair is poised for mutagenesis (56).

Excision of incorrect nucleobases by BER glycosylases leads to the formation of an abasic site, or apurinic (AP) site in the cases of OGG1 and MUTYH (52,54,57). Some glycosylases can also remove the sugar fragment by β- and δ-elimination reactions, but for OGG1 this reaction is slow; instead, the very toxic abasic site is likely handed off *in vivo* to AP endonuclease 1 (APE1) for strand scission at the 5′ phosphodiester bond adjacent to the AP (58). APE1 is a much more abundant protein in cells than BER glycosylases, and it serves multiple functions, as discussed below. Repair is completed after removal of the sugar fragment by dNTP insertion into the gap of the duplex by polymerase β (POLB) and ligation of the nicked strand.

Other oxidized lesions such as the hydantoins Sp and Gh can follow a similar repair pathway, although the glycosylases involved are the NEIL1–3 family that can process a wide variety of oxidized base lesions other than OG (40). Hydantoins are the best substrates studied so far, but unlike OGG1, there is no strong preference for the base opposite, or even a requirement that the lesion be part of a duplex (59–62). Interestingly, Sp and Gh in the context of a G4 are particularly good substrates for NEIL3, a monofunctional glycosylase that also requires APE1 to complete strand scission (45,61,63). Unlike OG, the hydantoins are all duplex-destabilizing lesions, as is the AP site (53). This has broader implications when the lesion is part of a PQS, and in fact, the process of DNA damage and repair might explain an energy conundrum with promoter G4s. These GC-rich sequences should be very stable as Watson–Crick base-paired duplexes. What is it that triggers strand separation and refolding to a G4 and perhaps also an i-motif (1)? Typically, i-motif folding may require a drop in pH below 7 (64), but why would a G4 want to fold? In terms of the duplex–quadruplex equilibrium, the duplex should always be more stable because of a larger number of base pairs and strong π stacking, although supercoiling and protein binding could impact the equilibrium (20).

Consider the energy diagram of Figure 4A. The G:C duplex represents a PQS present in a double-helical context that is presumably of very high stability because it is GC rich, and this is typically measured in oligonucleotides as a thermal melting temperature, T_m (45,53). Introduction
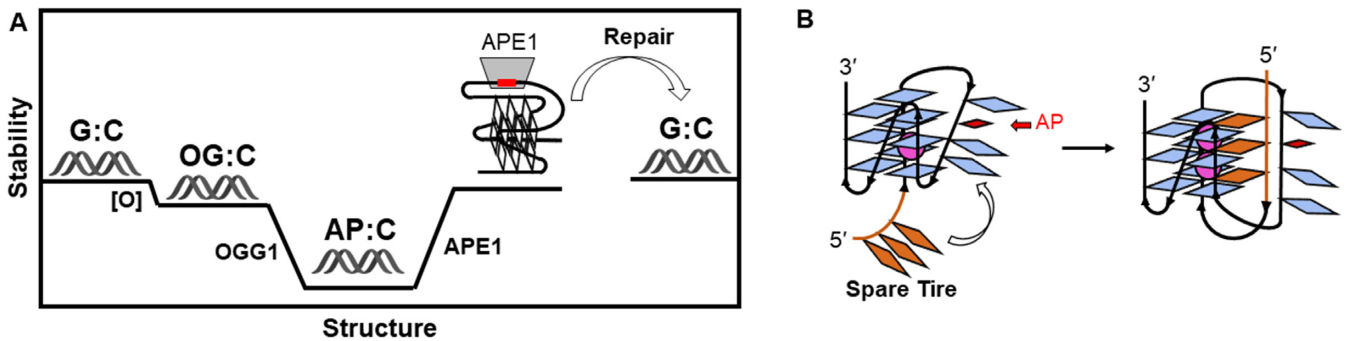
Figure 4. (**A**) Schematic drawing of stabilities, roughly equating to T_m measurements, for G:C base pairs versus lesions in various contexts. The AP site has low stability in a duplex context but can be refolded to a stable G4 if the AP site is looped out. (**B**) A fifth track, or 'spare tire', can be swapped for the damaged track to generate a stably folded G4.

of OG into the sequence will only destabilize the duplex slightly, by $\sim$2–3°C in T_m. The DNA repair pathway initiated by OGG1 leads to an AP that is highly destabilizing to the duplex. In general, either OG or AP would also be destabilizing to a quadruplex structure (31,45). However, as noted earlier, essentially all promoter PQSs have excess Gs. Either there are four or more Gs in some of the tracks or, more commonly, there are more than four G tracks. In the case of a base lesion, the G4 can still assemble by either sliding a track to a GGG run or by looping out the damaged track and swapping in the fifth track, or 'spare tire' (Figure 4B) (45). Essentially every promoter G4 that has been characterized, as well as the hTelo (65), has the ability to accommodate a damaged base either by sliding or by swapping.

Evidence that the 'spare tire' mechanism works has been seen for several G4s (30,45,66). NMR characterization of these dynamic processes in *c-MYC* indicates that refolding of a stable G4 is quite slow, on the order of many minutes to hours (67). Thermal denaturation studies of AP-containing G4s show that wild-type stabilities are achieved when a fifth track is engaged and the AP is looped out (30,31). If the AP is forced to reside in a core position of the G4, T_m values are 15–20°C lower, which is about the same degree of destabilization as an AP in a duplex. When repair is complete, the stability of the duplex returns to its original high level.

This picture helps explain how DNA damage and repair could collaborate with the G-rich sequence to facilitate G4 folding. This would imply that G oxidation and repair by OGG1 plus APE1 provides a mechanism to reveal the folded G4, dissociating the template and non-template strands, and assisting with transcription initiation. As we detail below, both OGG1 and APE1 provide additional benefits to gene induction by serving as sites for binding of activating transcription factors, including NF-κB, HIF-1α, AP-1 and others.

G4s, oxidative stress and gene expression

From the point of view of DNA damage and genome integrity, placement in key regulatory sites of poly-G tracks that have a high potential for oxidative damage and mutations is a poor evolutionary choice. However, if we assert

that OG is not highly mutagenic because DNA repair is efficient, then use of OG as a transient epigenetic mark in a PQS makes sense (4). The refolding of the promoter PQS upon BER processing of OG to AP provides a switching mechanism to recruit new proteins to the folded G4. A recent analysis described G4s as 'hubs' for binding of transcription factors (68); they might also be thought of as bus stops where DNA processes pause to let different passenger proteins on and off the bus.

Evidence of a correlation between oxidation in promoters and increased gene expression was found more than a decade ago when Avvedimento and colleagues found increases in OG (via ChIP-PCR studies for OGG1) during estrogen- or retinoic acid-induced expression of BCL2 (69,70). Interestingly, their studies pointed to H_2O_2-induced oxidation of G during LSD1-mediated demethylation of histone lysines. Independently, Gillespie and coworkers showed that the G4-containing promoter of *VEGF* was oxidized during oxidative stress-induced upregulation of the gene (71,72). The proximity of the G4 in *VEGF* to the TSS suggested a role for OG in this sequence that was known to impact gene induction, although the sequencing studies for OG were performed at 0.5 kb resolution.

We therefore asked whether we could recapitulate the increase in gene expression observed with G oxidation in a molecularly defined reporter system in cells (31). We chose to manipulate DNA plasmids synthetically so that we could control the precise location of DNA modifications and their sequence location at the same time as monitoring gene expression under different cellular conditions. The plasmid system we chose was a dual-luciferase plasmid with the SV40 promoter controlling the *Renilla* luciferase (*Rluc*) gene and the HSV-TK promoter ahead of the firefly luciferase gene (*luc*, Figure 5). Both promoters have TATA boxes at $\sim$25 nt upstream of the TSS in the non-template (or coding) strand. We replaced one of the TATA boxes with a PQS leaving the other promoter as 'wild type' to serve as an internal standard. Next, after growing sufficient quantities of plasmid, we used nicking endonucleases to remove a portion of the G-rich strand containing the PQS and to ligate in the same oligomer but with a single modification, for example either OG or F at a precise location. The latter, F, is an AP analog that is more chemically stable than AP but still a good substrate for APE1 during BER (73).
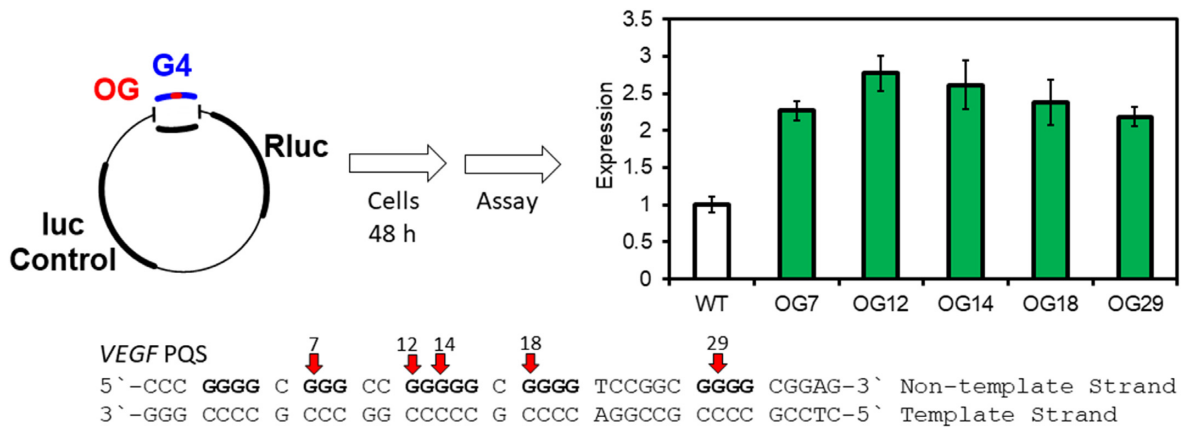
Figure 5. A dual-luciferase plasmid was manipulated to have an OG-containing PQS replacing the TATA box of the promoter. Transfection into cells and analysis after 48 h showed robust increases in gene expression (31).

In the first study, we placed an OG or F modification at one of five different positions in the *VEGF* G4 sequence; some of these positions were in presumed loops of a folded G4 and others were in core positions that would require the engagement of the fifth track to form a stable quadruplex; thus, both four-track and five-track versions of the *VEGF* G4 sequence were studied (31). Transfection into mouse embryonic fibroblasts (MEFs) of the OG-modified plasmid compared to normal G-containing plasmids could be compared in WT MEFs, OGG1$^{-/-}$ MEFs and WT MEFs with APE1 siRNA knockdown. Typical results are shown in Figure 5.

Several conclusions can be drawn from these and subsequent studies performed with the G4 sequences from promoters in *NTHL1* and *NEIL3* (BER glycosylases) (30,31), *PCNA* (repair processing) (32) and *RAD17* (damage response) (33): (1) A G4 sequence containing DNA damage in the non-template strand of the promoter increases gene expression by ∼3-fold for OG or ∼5-fold for AP. (2) The exact position of the damage in the G4 is not very important. (3) The fifth track enhances the effect, either by ensuring G4 folding or by assisting dynamics to find the best fold. (4) The BER pathway is required—OGG1 is needed to convert OG to AP and APE1 is absolutely required to see the increase in gene induction. (5) The magnitude of the effect is somewhat cell line dependent; in addition to MEFs, several human cancer cell lines were studied (11). (6) Even without OG or AP, a PQS in the promoter could increase gene expression under conditions of oxidative stress induced by addition of tumor necrosis factor α (TNFα) to the cell medium (30). TNFα leads to higher expression of iNOS, which generates peroxynitrite and ultimately carbonate radical anion (74). The finding of increased gene expression is consistent with the conclusion that oxidized lesions generated during endogenous oxidative stress are giving the same response as those specifically inserted by oligonucleotide synthesis. An overall mechanism is proposed in Figure 6.

As discussed in the earlier sections, endogenous oxidative stress generates $CO_3^{\bullet-}$ that leads to selective reaction at poly-G sequences such as a PQS. These sites become a focal point for DNA damage in which long-range charge transport will relocate an electron hole to these sites (4,41).

If a promoter PQS is oxidized at any one of multiple sites in the sequence, the DNA duplex is still stable (53). However, initiation of BER by the action of OGG1 leads to the destabilizing lesion AP (52,53). Like many proteins in the repair pathway, OGG1 binds tightly to the AP product, but then hands off to APE1 for the next step, strand cleavage (58). APE1 is the centerpiece of many studies on the impact of DNA repair on gene expression because it serves multiple functions: BER nuclease, transcription factor recruiter and G4 binder (75–77). When APE1 binds to AP in a G4 context, the nuclease function is paused, and this pausing allows the redox effector factor function of APE1 to recruit activating transcription factors, including HIF-1α, AP-1 and others (30,72,75).

Overall, the mechanism in Figure 6 proposes that G4 sequences in promoters are sensors of oxidative stress that act as an antenna for DNA oxidation. When OG is formed, the BER pathway begins with OGG1 creating the duplex-destabilizing abasic site AP. The AP site may be better accommodated in a folded G4 in which the AP is extruded into a loop; this is the switching point from repair to gene induction, because APE1 binds to G4s but its nuclease activity pauses. Once the repair bus has stopped at the G4 bus stop, APE1 facilitates gene expression through transcription factor binding (4).

This proposed mechanism is not the only one involving G oxidation and gene expression. In related work, Boldogh and coworkers showed that the transcription factor NF-κB was recruited to OGG1 sites during oxidative stress (78,79). Xodo and coworkers found *KRAS* upregulation after oxidation of G to OG in a promoter PQS was acted upon by the BER machinery to recruit activating protein factors (66). Tell and coworkers found that transcription was initiated when OGG1 sites were processed in the SIRT1 promoter to allow buildup on APE1, Ku70/80 and RNA pol II for gene induction (80). Hanawalt (22), Myong (23) and others point to modulation of gene expression by R loops, and the ability to form G4s impacts the stability of the R loop. Furthermore, it is well known that DNA damage on the template strand can impact gene expression because the DNA repair machinery should be recruited to TSSs in advance of RNA synthesis, i.e. transcription-coupled repair
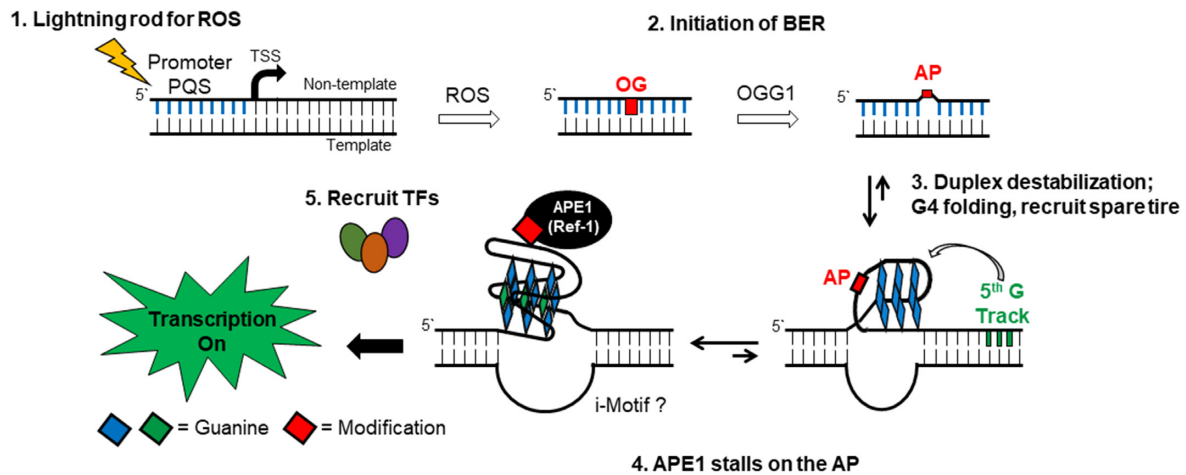
Figure 6. Proposed mechanism for upregulation of genes with oxidative damage in their promoter G4. The PQS acts as a sensor of oxidative stress; subsequent BER refolds the sequence to a G4 at which APE1 is bound and can recruit activating transcription factors.

(81). However, it is the gene activation mechanism involving DNA damage on the non-template strand that has intrigued many laboratories in the past few years. Central to this appears to be the ability of BER proteins to recruit transcription factors as passengers on the bus to gene expression (4,79,82).

APE1 drives the bus

What chemical and biological features of APE1 enable this protein to drive the bus that coordinates gene regulation under oxidative stress conditions? Mammalian cell studies identified APE1 as a high copy number protein ($\sim 10^4$–10^5 copies/cell) with a long half-life (~ 9 h) found in nearly all cell types that is essential for cellular survival (77,83,84). The biological importance of APE1 was established when embryonic lethality was observed in $Ape1^{-/-}$ knocked out mice, while heterozygous mice ($Ape1^{+/-}$) were found to thrive early on but displayed an increased apoptotic response to oxidative stress, and a high incidence of cancers later in life (85–87). The protein is found in the nucleus, cytosol and mitochondria, for which the distribution can change in some cancers (88–90). The human APE1 protein is 318 amino acids long, has a nuclease domain from positions 60 to 318 homologous to *Escherichia coli* exonuclease III and has an N-terminal region of ~ 60 amino acids (Figure 7A) (77,91). The first ~ 40 amino acids are unstructured, and the amino acids between ~ 40 and 60 are structured in solution (92). In humans, APE1 possesses a variety of functions on DNA that include an AP endonuclease, 3′,5′ exonuclease, 3′ phosphatase and a nucleotide incision repair catalyst; in RNA, APE1 has RNase H activity and functions in RNA metabolism; and lastly, the protein is a redox effector factor (REF1; positions 30–127, Figure 7A) for regulation of gene expression (76,93–95). The latter property of APE1 is so vital to cellular processes that the complete descriptive name for the protein is APE1/REF1 (95). Thus, APE1 is critical to cellular survival, and how this protein achieves this wide variety of functions has been a long-standing research question that remains an active area of study (76).

How APE1 can guide protein binding on DNA during oxidative stress in which both the BER and gene regulatory features of the protein are needed requires closer inspection of the chemistry, which will be the focus of this section. For readers interested in the other activities of APE1 that include its exonuclease and RNA processing properties, there exist excellent reports on these topics (94,97,98). Abasic site substrates for APE1 are provided by monofunctional glycosylases that hydrolyze the glycosidic bond of chemically modified nucleotides to yield the AP (Figure 3B) (52). As an endonuclease, APE1 is proposed to scan the DNA by diffusion along the strand with a residence time of 2–3 μs per base pair until it binds the AP in DNA (99); then, APE1 hydrolyzes the 5′ phosphodiester of the AP site yielding a 5′ fragment with a 3′ hydroxyl group, and a 3′ fragment with a 5′-deoxyribose phosphate (5dP; Figure 3B) (96). Structurally, APE1 engages an AP by flipping it into an active site binding pocket that kinks the DNA by $\sim 35°$, leaving an orphan base in the site opposite the AP in the complementary strand that is stabilized by protein interactions. There are many important active site residues identified by biochemical, structural and computational analyses (86,100–102). Noteworthy residues are Asp210 that is essential for activation of the H_2O that attacks the phosphodiester to yield the pentavalent intermediate, as well as Asp70 and Glu96 involved in the coordination of the Mg^{2+} cofactor needed for substrate binding, transition state stabilization and product binding (Figure 7B). Arginine 177 stabilizes the APE1–AP DNA-bound structure via intercalation to slow product release and allow the next enzyme to complete the repair process (83). In duplex DNA, APE1 binds the product to allow a handoff to POLB as the next member in the BER process that prevents the release of the more genotoxic lesion, a single-strand break (73,83,103). POLB has both lyase activity to remove the 5dP and polymerase activity to select the correct dNTP directed by the nucleotide opposite the AP site (104). Finally, ligase seals the nick to yield the repaired product (52). This provides a clear picture of how the enzyme is proposed to function as an endonuclease on a canonical substrate in short-patch BER.
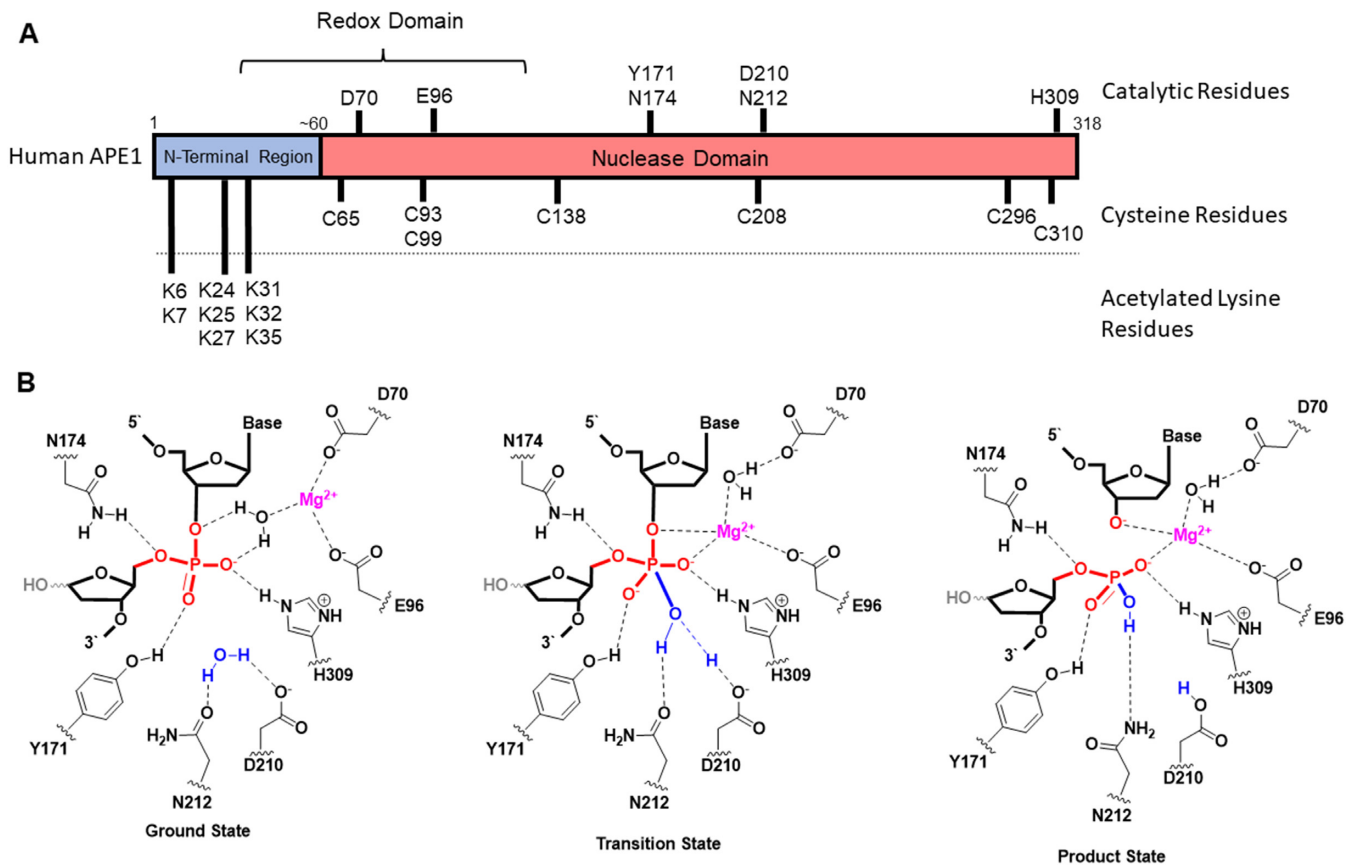
Figure 7. Key amino acids in the APE1 protein essential for its endonuclease activity, REF abilities and post-translational regulation. (**A**) Schematic of human APE1 to draw attention to key catalytic amino acids based on a solved structure (96), as well as cysteine and acetylatable lysine residues. (**B**) Proposed mechanism for APE1-catalyzed hydrolysis of a phosphodiester bond 5′ to an AP site as previously reported (96). The structure was studied with the tetrahydrofuran (F) analog of the AP (96). For clarity, the 1′-OH on the baseless sugar was added to the mechanism in light gray to display the *in vivo* substrate.

Enzyme activity assays for APE1 have found that the cleavage is most robust for AP-containing dsDNA and shows product inhibition as discussed earlier. In solution, APE1 favors endonuclease activity at slightly alkaline pH (7.4–7.8) and near-physiological concentrations of K^+ (50–140 mM) and Mg^{2+} ions (>2 mM) (105,106). Initial studies found APE1 can cleave an AP in ssDNA, small dsDNA bulges, model replication forks, transcription bubbles, R loops and DNA/RNA hybrid duplexes (106,107). The activity of APE1 toward AP-containing G4s found in the *c-MYC* (108), hTelo (61,75,109), *VEGF* (110) and *NEIL3* (30) sequence contexts has been analyzed to find reduced activity compared to dsDNA substrates. In the hTelo, *VEGF* and *NEIL3* G4 studies, the position of the lesion impacted the endonuclease cleavage yield; additionally, in the *VEGF* and *NEIL3* G4 studies whether the fifth G run was present or not impacted the APE1 cleavage yields. The efficiency and rate of APE1 cleavage of an AP in G4 contexts display differences in the dependence on the concentration of K^+ ions compared to dsDNA (61,75). In general, the enzymatic activity was nearly abolished as the concentrations of K^+ approach physiological concentrations. This may have an important role in preventing APE1 from cleaving an AP when a complementary strand is not present that would be necessary to permit the handoff to POLB. In the proposal

that G4 folds serve as genomic stops for APE1 to coordinate movement of protein transcription factors, the reduced activity supports this claim, but binding must be retained at the higher ionic strength.

The binding of APE1 to AP or an F analog in dsDNA has been analyzed by a variety of methods, under similar but not identical conditions, and on different sequence contexts and nucleotides opposite the F site (111–114). The key observation is that APE1 binds an AP (F)-containing dsDNA with a low nanomolar dissociation constant (K_D). The N-terminal region of APE1 when removed has a minor impact on the binding interaction with dsDNA; however, the activity of APE1 increases when the N-terminal region is removed as a result of decreased product inhibition suggesting the region is involved in the transfer of the product to the next BER member POLB (115–117). In G4 contexts, the catalytically inactive D210A APE1 mutant was first demonstrated to bind the *c-MYC* parallel-stranded G4 in 50 mM K^+ based on a supershift on a native gel; a binding constant was not determined in this study (108). The binding of wild-type APE1 to the wild-type hTelo sequence in the presence of 150 mM Na^+ ion studied by SPR (surface plasmon resonance) provided a 30 nM K_D value (75). It is worth noting that Na^+-bound hTelo G4s adopt basket (antiparallel) topologies (118). Further, the observation of APE1 binding

to the Na$^+$-bound hTelo G4 was the first to identify APE1 binds non-lesion-containing G4 folds (75). A recent study confirmed wild-type APE1 binds native and F-containing hTelo G4s in 140 mM K$^+$ ion-containing solutions (109), where the native G4 adopts a hybrid fold and the lesion-bearing G4s display folds not fully characterized (61,109); furthermore, the binding was location dependent (109). In our studies, wild-type APE1 bound the four- and five-track *VEGF* G4s that favorably adopt parallel folds in K$^+$ solution, and there was dependence in the K_D value on the location of F and whether the fifth G track was present or not (119). Lastly, the binding of APE1 to hTelo and *VEGF* G4s was found to be highly dependent on the presence of the N-terminal region (75,119), a sharp deviation from ds-DNA binding by this protein. This final observation suggests that G4 binding through the disordered N-terminal region may be strong. Analysis of tryptophan fluorescence changes when APE1 engages its substrate found large structural deviations for binding to F-containing dsDNA and small deviations for F-containing G4s (109), which might be explained by binding to the non-canonical folds via the disordered N-terminal region where tryptophan is not found in the sequence.

A few critical conclusions can be drawn from the binding studies: (1) Wild-type APE1 binds lesion-containing G4s coordinated to K$^+$ ions. (2) Binding of G4s by the N-terminal region of APE1 helps address the poor enzyme activity observed for AP-containing folds. (3) Wild-type APE1 binds native G4 folds adopting basket, hybrid or parallel topologies. This final point is interesting because folded genomic G4s were recently found to be hubs for transcription factor binding essential for gene regulation (68). Thus, these studies demonstrate that APE1 can stop the bus (i.e. bind) at G4 hubs to allow the exchange of proteins to coordinate gene expression. Whether the AP lesion is needed for APE1 binding *in cellulo* is not currently known.

One more unique property of APE1 is its REF activity to regulate downstream transcriptional activity by controlling the binding of critical transcription factors (120–122). This feature of APE1 is essential to cell growth (76) and is a promising drug target for the treatment of many different types of cancer (122). Transcription factors controlled by APE1 required for the response to oxidative stress include AP-1 (123), HIF-1α (72), NF-κB (78), p53 (124) and STAT3 (125). The N-terminal region of APE1 is critical for redox regulation, wherein C65 is the most important residue for the regulation (Figure 7A) (126). One proposal is that C65 and C93/99 function in the thiol–disulfide exchange reaction to reduce target disulfide-containing dormant transcription factors to sulfhydryls to activate them for gene regulation (126,127). Two challenges to this proposal exist. First, the cysteine residues (C65 and C93 or C99) are too far apart in the crystal structure to rationalize chemically a disulfide bond in APE1 (83,96); second, only C99 and C138 are solvent exposed, while the critical C65 is buried within the protein (96). A solution to this chemical challenge is the observation that APE1 is highly dynamic in solution allowing disulfide bond formation to occur (128,129). Similarly, APE1 could form an intermolecular disulfide bond with the transcription factor to aid in DNA binding and gene regulation, thus avoiding the complication of disul-

fide bond formation within APE1. Alternatively, a recent report proposed that APE1 induces a change in the target DNA conformation to enable transcription factor loading, and therefore does not invoke the REF1 domain in APE1 for disulfide-exchange chemistry (130). More studies are needed to better address whether these represent similar or distinct phenomena for gene activation.

Gene regulation during oxidative stress via G4 hubs and APE1 binding as a means to select transcription factor partners to induce mRNA synthesis are supported by prior work. This occurs as a result of preferential G oxidation in G-rich human promoters to localize the BER proteins, in which APE1 can bind the lesion-containing G4 folds using the N-terminal region of the protein. Transcription factors will then be recruited to turn transcription on, and then eventually the complex will dissociate to allow completion of the DNA repair process, returning the promoter to the native dsDNA state. The finding that APE1 can bind G4s without a lesion may provide an additional mechanism for gene induction facilitated by APE1. Future work is needed to address this point. The N-terminal region of APE1 possesses critical lysine residues that are post-translationally modified with acetyl groups (Figure 7A). The acetylation status of the protein is another layer of modifications used to fine-tune the protein–protein interactions of APE1 (116,131), and this has been proposed as a component of the APE1–G4 interaction responsible for gene regulation (132). Studies with other DNA repair proteins such as OGG1 have identified cysteine oxidation as an additional post-translational modification used by cells for regulation when the protein switches between DNA repair and gene regulation (79,133). Interestingly, APE1 has seven cysteine residues that can be oxidized (Figure 7A), a few of which were found to be oxidized in cells (134), and how this alters the APE1–G4 interactions has yet to be addressed (Figure 7A). The knowledge that APE1 binds G4s requiring the N-terminal region, but the catalytic activity is attenuated, suggests a possibility that the protein can bind the G4 through an alternative conformation. Future experiments to understand the full scope of post-translational modifications to APE1 responsible for switching the protein from repair to regulation and structural analysis of APE1 binding G4s are needed.

Location, location, location

The discovery of gene regulation guided by an APE1–promoter G4 interaction was made on a sequence found close to the TSS on the non-template strand (Figure 8A) (31). Human promoters are known to have significant GC skew correlated with elevated gene expression (135). As the GC skew increases, the probability of non-canonical structures such as G4s and the propensity for R-loop formation increase (4,21). The non-canonical structures such as R loops and possibly G4-forming sequences have been proposed to play a key role in gene regulation (4,21,135). Research in our laboratory in the model SV40 promoter with an embedded *VEGF* PQS without a lesion studied in mammalian cells found that the presence of the G-rich sequence increased the promoter strength (Figure 8B and C, bars labeled G) (11). The magnitude of enhancement was de-
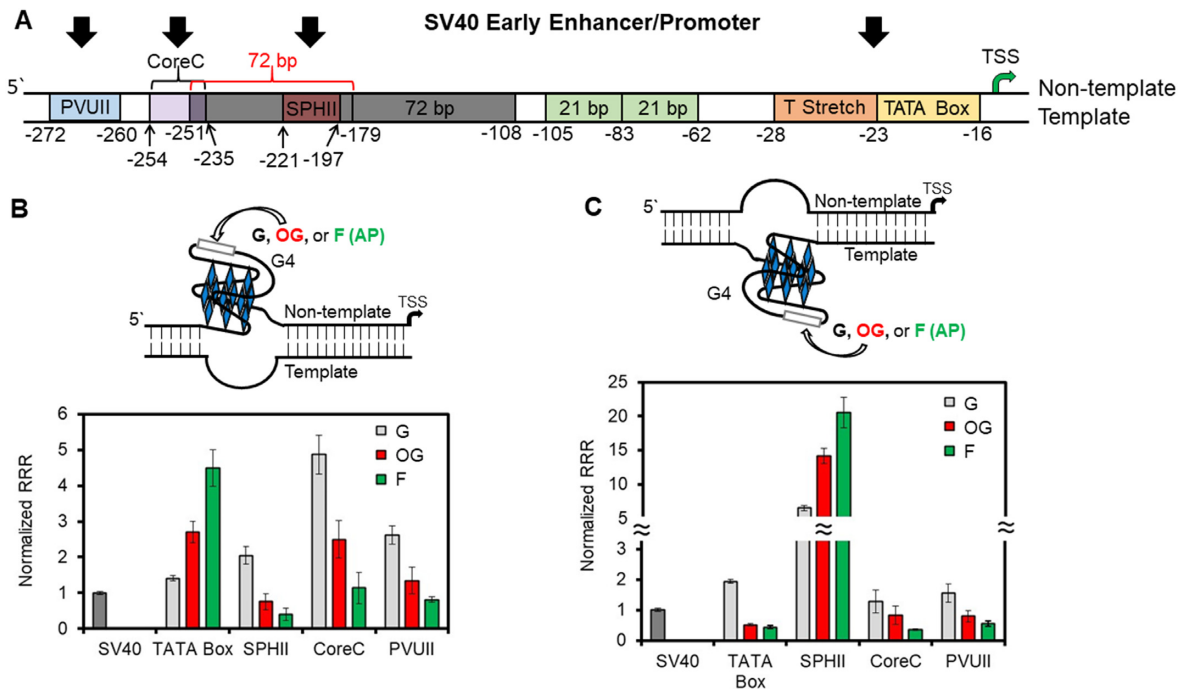
Figure 8. The *VEGF* PQS was moved through the SV40 promoter on both strands with the native G or synthetically installed OG or F to analyze the impact on gene expression. (**A**) Schematic of the SV40 promoter and key elements. Gene expression changes were observed when the G-rich strands with and without OG or F were synthetically installed on the (**B**) non-template or (**C**) template strand at locations indicated by arrows in panel (A) (11). In panels (B) and (C), the *x*-axis refers to the promoter element replaced with the PQS and the *y*-axis is the relative response ratio (RRR = Rluc/luc) normalized to the SV40 promoter expression. The Rluc promoter was the one replaced with the PQS.

pendent on the location or distance from the TSS and the non-template versus template strand of occupancy for the G-rich sequence. Generalizations to human promoters are not direct because other factors in the genome are at play; nonetheless, this observation demonstrates that a PQS can influence gene expression likely by an increase in GC skew for recruitment of activating transcription factors.

As described earlier, amplification of this activation effect by oxidative modification of a G nucleotide to OG or its repair intermediate AP (F) was demonstrated for the *VEGF*, *NEIL3*, *NTHL1* and *PCNA* PQSs near the TSS on the non-template strand with dependence on APE1 (Figure 8B) (30–32). Does the same impact on gene expression occur when the construct is flipped over to the template strand of the promoter at the same site? Studies with the *VEGF* PQS containing OG or F at this site revealed that transcription was turned down 3-fold (Figure 8C) (136). A similar observation was found for the *RAD17* PQS containing OG near the TSS on the template strand (33). Through the utility of siRNA knockdown experiments, the gene deactivation was found to be dependent on the Cockayne syndrome B protein (CSB, aka ERCC6) (136); this is consistent with a prior study that found CSB is needed for resolving folded G4s to allow advancement of the transcriptional machinery (137). The CSB protein interacts with the RNA pol II complex to aid in sensing and resolving DNA damage via the transcription-coupled repair pathway (138), which suggests the presence of a damaged PQS could fold to a G4 to engage CSB and stall transcription. In contrast, when a non-modified PQS was in the same location gene expression

was not CSB dependent, and therefore likely the sequence failed to adopt a G4 fold, and transcription continued unabated (136). Many mysteries still exist regarding the role of CSB in gene deactivation. (1) The CSB strongly binds intermolecular G4s but has a low affinity for intramolecular G4 folds (139), in which the latter is the likely fold in the studies described. (2) Is the function of CSB to resolve the lesion-containing G4 fold, recruit the transcription-coupled repair machinery or is CSB needed to achieve both functions?

Computational inspection of the human genome for PQSs found that they are enriched in promoters with the distribution centered at the TSS and expanding outward in both directions by ∼1000 nt (17,18). With this knowledge, our laboratory moved the OG-containing *VEGF* PQS to other promoter positions in either strand to observe the impact on gene expression (11). In the positions studied, gene activation was observed for OG in the PQS context at two locations: the first near the TSS on the non-template strand as previously described, and the second on the template strand near position −200 (Figure 8B and C). All other positions found OG in the PQS context turned transcription off. The most important takeaway from these studies is that promoters are like real estate, and it is all about location, location, location! Experimentally, this means that studies to understand which oxidation events in a promoter PQS result in activation versus those that lead to deactivation will need to be on a case-by-case basis.

The other genomic element around TSSs where PQSs are enriched is in 5′-UTRs (17,18). Examination of how oxidative modification of G in the PQS in 5′-UTRs impacts

gene expression has yet to be examined. The 5′-UTR differs from the promoter in that the non-template strand PQS in the genome is transcribed into the mRNA transcript. For non-oxidized PQSs, studies have found that when these G-rich sequences reside in the non-template strand of the 5′-UTR, the transcriptional output of the gene is enhanced (23). The proposal for the increased expression is the formation of a transcription-induced R loop stabilized by a G4 fold that results in successive rounds of efficient transcription. The regulation of transcription by R loops has been reviewed (21,22). Whether oxidation of G in the non-template strand PQS when located in the 5′-UTR will have an impact on gene expression is not known, although binding of the AP-containing G4 by APE1 could significantly stabilize the G4 fold and impact expression. Oxidation of G to OG in the non-template strand (i.e. coding strand) of a 5′-UTR outside of a PQS context drove downregulated transcription (57). As for the template strand, G oxidation to OG in a non-PQS context stalls transcription because of initiation of transcription-coupled repair (140,141), and therefore, it is anticipated that G oxidation in a PQS context in this strand will also downregulate gene expression. In situations where RNA pol II can read through the template OG-containing PQS, the OG heterocycle can be read as a T nucleotide via the Hoogsteen face to direct misincorporation of A in the transcript (142).

Critical ties to cancer

Critical components of cancer are DNA damage, DNA repair and aberrant gene expression. A common example of DNA damage and its ties to cancer is G oxidation to OG. The OG base when located in a DNA template strand can direct insertion of dCTP opposite via pairing with the Watson–Crick face or dATP by pairing on the Hoogsteen face during replication (143). In the event of dATP insertion and no repair, a second replication event would result in a T:A base pair in one daughter strand where there was originally a G:C base pair (i.e. G→T transversion). The frequency of OG causing G→T transversions is low in DNA repair-competent cells ($\sim$0.1%) (144). Hyperoxidation of G to Sp or Gh in template strands yields a mixture of G→T and G→C transversion signatures as a consequence of the hydantoins pairing with either purine 2′-deoxynucleotide triphosphate leading to nearly 100% mutation (145,146). Long-term exposure to oxidative and inflammatory stress has shown significant levels of G→T and G→C transversions in the genome in support of DNA damage as a mechanism for cancer initiation and progression (3). Definitive demonstration that OG, Sp/Gh or any other damaged G nucleotide is the origin of these mutations has yet to be reported.

BER pathways for DNA damage associated with cancer are dependent on the protein. The glycosylase MUYTH is responsible for initiating repair of an OG:A base pair, and when the enzyme is mutated to an inactive state, a high incidence of colon cancer is observed (54,56). Additionally, loss of APE1 is embryonic lethal; however, misregulation of APE1 and altered cell localization impact its repair and redox signaling capabilities in many cancers (86,87,122). The APEX1 gene is relatively small and is not highly mutation prone like some genes, but a few mutations have been noted in cancers that result in persistent genomic stress (147). One of many cellular pathways impacted by aberrant APE1 in cancer is the ability to respond to stress via binding to regulatory G4s as described earlier. An active area of cancer research is to drug APE1 via its nuclease abilities to increase the success of DNA-damaging chemotherapeutics, or alternatively to target the redox signaling of APE1 to interfere with gene expression in diseased cells (122). The situation is different for OGG1 and the NEILs. Mice that are Ogg1$^{-/-}$ are overall normal, except they have 2-fold higher levels of genomic OG and are more sensitive to oxidative and inflammatory stress compared to wild-type mice (55,148). The stress sensitivity observed likely results from an inadequate stress response by the mechanisms outlined herein. Lastly, the Neil1/2/3 glycosylases have been knocked out of mice to find no predisposition for cancer and no increased frequency of mutations, but they may have some metabolic disorders (149,150). This suggests the NEILs are not backup enzymes for one another and have distinct cellular functions beyond canonical DNA repair. These observations highlight that BER proteins differ widely as to their impact on cancer predisposition.

Genome stability can be negatively impacted by folded G4s. On template strands, G4 folds can cause stalled replication progress causing replication-fork collapse resulting in double-strand breaks that are highly mutagenic and specifically repaired by non-homologous DNA end joining (151). Sequences in the human genome that can adopt highly thermodynamically stable G4s are sites of known somatic mutations and recurrent mutations in cancer genomes (152). Sequencing for G4s has exposed these sites to be associated with gene amplification events found in cancers (12,19). Targeting cellular G4s with small molecules is an approach proposed for the treatment of cancer by changing the expressional output of oncogenes (9,10). This approach to downregulate an oncogene has been best studied for the *c-MYC* promoter G4, in which the c-MYC protein is overexpressed in a majority of human cancers (20). However, small molecules targeting G4s usually bind indiscriminately in cells and could target the $\sim$10 000 G4s that are known to fold, as well as induce folding of some of the other 700 000 PQSs that reside in the human genome (12,19). Ligand-bound G4s represent roadblocks to polymerases, they disrupt protein binding to G4s and they can alter chromatin structure resulting in many off-target impacts to the cells treated with these compounds. A solution is to target G4s with sequence specificity to avoid, or at least minimize, off-target impacts when treating cancer via regulatory G4s (153). The function of oxidized G4s for stress response regulation in cancers is not known, and whether targeting this pathway will be successful awaits more research discoveries.

CONCLUSIONS

At first glance, the collaboration of G4 folding, which inhibits the progression of polymerases, with DNA damage and repair pathways that may be mutagenic, to upregulate gene expression, appears to be an unlikely one—a sort of double negative. Yet, multiple lines of evidence suggest

that Nature is using these two key properties of G-rich sequences, namely their sensitivity to oxidation and their propensity to fold to stable G4s, in a productive way (4). Formation of OG in a promoter PQS is likely if long-range charge transport can conduct the electron hole to these sensitive sites. Refolding of the sequence to a G4 is then facilitated by the repair process in which APE1 plays the critical role—pausing at the G4 to switch to a gene regulatory pathway versus completion of repair. In this sense, OG serves a temporary epigenetic function as a DNA base modification capable of up- or downregulation of gene expression via its protein readers APE1 and CSB (136). OG as an epigenetic mark is not passed on to future generations of cells, but the potential quadruplex sequence in a regulatory site is (154). From viruses to humans, there are many examples of G-rich sequences being highly conserved, despite susceptibility to G→T mutations if not repaired, underscoring the important roles played by G-rich sequences (154). The participation of APE1/REF1 in both avoidance of mutagenesis and upregulation of oxidative stress-related genes mirrors the importance of both pathways, mutagenesis and aberrant gene expression, in cancer (4,77,121,122). Genomic hubs comprised of G4s could be more versatile than interactions with BER proteins, and crosstalk with other DNA repair pathways for gene regulation (155).

The study of G4s and DNA oxidation and repair in gene expression is being carried out from many different perspectives, to the ultimate benefit of science. On the one hand, biophysical chemists and structural biologists are characterizing how G4s fold, how specific proteins interact with DNA and the enzymology of DNA repair (17,27,59,63,75,83,96,109). All of these are *in vitro* studies, and while they provide detailed pictures of a given biomolecule, said molecule is not in the cellular setting. For whole-genome, potentially live-cell approaches, scientists use both low-resolution methods such as fluorescence microscopy or ChIP-seq and related methods to image and sequence sites in chromatin and targeted methods that deliver designed small molecules to folded G4s (10,19,69,71). This latter approach provides information about biomolecules in their native setting, but can alter the natural status of the cell by shifting the duplex–quadruplex equilibrium or by displacing G4-binding proteins with tight-binding small molecules; thus, there are uncertainties in these approaches also. We and others have found that some of the middle ground between *in vitro* and *in cellulo* DNA biochemistry can be discovered by the use of plasmids from the chemical biology toolbox (20,31,156). These are typically ~7000 base-pair circular DNA duplexes that can be site-specifically altered to contain essentially any promoter sequence and any chemically stable base modification at a defined site. Plasmids containing G4 sequences transfect well into many cell lines, and upon entry, they are packaged with histones and transported to the nucleus where they experience the same suite of DNA-interacting proteins as chromatin (157). A further advantage of plasmid studies is that hundreds of copies are introduced to the cell, providing a robust readout for gene expression studies. There are disadvantages, however; these include (1) the fact that the histone modifications in plasmid nucleosomes may not be native, (2) supercoiling could impact the formation of G4s differently

in plasmids and (3) long-range interactions, such as those mediated by large loop formation with CTCF, would likely be missing. In sum, all of the different approaches help us paint the complex picture of gene expression and the role of DNA repair and G4s in this important process.

REFERENCES

1. Brooks,T.A., Kendrick,S. and Hurley,L. (2010) Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.*, **277**, 3459–3469.
2. Ding,Y., Fleming,A.M. and Burrows,C.J. (2017) Sequencing the mouse genome for the oxidatively modified base 8-oxo-7,8-dihydroguanine by OG-Seq. *J. Am. Chem. Soc.*, **139**, 2569–2572.
3. Pfeifer,G. and Besaratinia,A. (2009) Mutational spectra of human cancer. *Hum. Genet.*, **125**, 493–506.
4. Fleming,A.M. and Burrows,C.J. (2020) Interplay of guanine oxidation and G-quadruplex folding in gene promoters. *J. Am. Chem. Soc.*, **142**, 1115–1136.
5. Yang,D. (2019) G-quadruplex DNA and RNA. *Methods Mol. Biol.*, **2035**, 1–24.
6. Patel,D.J., Phan,A.T. and Kuryavyi,V. (2007) Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.*, **35**, 7429–7455.
7. Onel,B., Carver,M., Wu,G., Timonina,D., Kalarn,S., Larriva,M. and Yang,D. (2016) A new G-quadruplex with hairpin loop immediately upstream of the human BCL2 P1 promoter modulates transcription. *J. Am. Chem. Soc.*, **138**, 2563–2570.
8. Ngoc Nguyen,T.Q., Lim,K.W. and Phan,A.T. (2020) Duplex formation in a G-quadruplex bulge. *Nucleic Acids Res.*, **48**, 10567–10575.
9. Asamitsu,S., Obata,S., Yu,Z., Bando,T. and Sugiyama,H. (2019) Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules*, **24**, E429.
10. Kosiol,N., Juranek,S., Brossart,P., Heine,A. and Paeschke,K. (2021) G-quadruplexes: a promising target for cancer therapy. *Mol. Cancer*, **20**, 40.
11. Fleming,A.M., Zhu,J., Ding,Y. and Burrows,C.J. (2019) Location dependence of the transcriptional response of a potential G-quadruplex in gene promoters under oxidative stress. *Nucleic Acids Res.*, **47**, 5049–5060.
12. Chambers,V.S., Marsico,G., Boutell,J.M., Di Antonio,M., Smith,G.P. and Balasubramanian,S. (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.*, **33**, 877–881.
13. Marsico,G., Chambers,V.S., Sahakyan,A.B., McCauley,P., Boutell,J.M., Antonio,M.D. and Balasubramanian,S. (2019) Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res.*, **47**, 3862–3874.
14. Ding,Y., Fleming,A.M. and Burrows,C.J. (2018) Case studies on potential G-quadruplex-forming sequences from the bacterial orders Deinococcales and Thermales derived from a survey of published genomes. *Sci. Rep.*, **8**, 15679.
15. Jara-Espejo,M., Fleming,A.M. and Burrows,C.J. (2020) Potential G-quadruplex forming sequences and N^6-methyladenosine colocalize at human pre-mRNA intron splice sites. *ACS Chem. Biol.*, **15**, 1292–1300.
16. Maizels,N. and Gray,L.T. (2013) The G4 genome. *PLoS Genet.*, **9**, e1003468.
17. Fleming,A.M., Zhu,J., Ding,Y., Visser,J.A., Zhu,J. and Burrows,C.J. (2018) Human DNA repair genes possess potential G-quadruplex sequences in their promoters and 5′-untranslated regions. *Biochemistry*, **57**, 991–1002.

18. Du,Z., Zhao,Y. and Li,N. (2008) Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription. *Genome Res.*, **18**, 233–241.

19. Hansel-Hertsch,R., Beraldi,D., Lensing,S.V., Marsico,G., Zyner,K., Parry,A., Di Antonio,M., Pike,J., Kimura,H., Narita,M. *et al.* (2016) G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.*, **48**, 1267–1272.

20. Brooks,T.A. and Hurley,L.H. (2009) The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat. Rev. Cancer*, **9**, 849–861.

21. Sollier,J. and Cimprich,K.A. (2015) Breaking bad: R-loops and genome integrity. *Trends Cell Biol.*, **25**, 514–522.

22. Belotserkovskii,B.P., Tornaletti,S., D'Souza,A.D. and Hanawalt,P.C. (2018) R-loop generation during transcription: formation, processing and cellular outcomes. *DNA Repair*, **71**, 69–81.

23. Lee,C.Y., McNerney,C., Ma,K., Zhao,W., Wang,A. and Myong,S. (2020) R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nat. Commun.*, **11**, 3392.

24. Xu,Y. (2011) Chemistry in human telomere biology: structure, function and targeting of telomere DNA/RNA. *Chem. Soc. Rev.*, **40**, 2719–2740.

25. Agrawal,P., Hatzakis,E., Guo,K., Carver,M. and Yang,D. (2013) 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.*, **41**, 10584–10592.

26. Phan,A.T., Modi,Y.S. and Patel,D.J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc.*, **126**, 8710–8716.

27. Cogoi,S. and Xodo,L.E. (2006) G-quadruplex formation within the promoter of the *KRAS* proto-oncogene and its effect on transcription. *Nucleic Acids Res.*, **34**, 2536–2549.

28. Ohnmacht,S.A., Micco,M., Petrucci,V., Todd,A.K., Reszka,A.P., Gunaratnam,M., Carvalho,M.A., Zloh,M. and Neidle,S. (2012) Sequences in the HSP90 promoter form G-quadruplex structures with selectivity for disubstituted phenyl bis-oxazole derivatives. *Bioorg. Med. Chem. Lett.*, **22**, 5930–5935.

29. Chaires,J.B., Trent,J.O., Gray,R.D., Dean,W.L., Buscaglia,R., Thomas,S.D. and Miller,D.M. (2014) An improved model for the hTERT promoter quadruplex. *PLoS One*, **9**, e115580.

30. Fleming,A.M., Zhu,J., Howpay Manage,S.A. and Burrows,C.J. (2019) Human NEIL3 gene expression is regulated by epigenetic-like oxidative DNA modification. *J. Am. Chem. Soc.*, **141**, 11036–11046.

31. Fleming,A.M., Ding,Y. and Burrows,C.J. (2017) Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. *Proc. Natl Acad. Sci. U.S.A.*, **114**, 2604–2609.

32. Redstone,S.C.J., Fleming,A.M. and Burrows,C.J. (2019) Oxidative modification of the potential G-quadruplex sequence in the PCNA gene promoter can turn on transcription. *Chem. Res. Toxicol.*, **32**, 437–446.

33. Zhu,J., Fleming,A.M. and Burrows,C.J. (2018) The RAD17 promoter sequence contains a potential tail-dependent G-quadruplex that downregulates gene expression upon oxidative modification. *ACS Chem. Biol.*, **13**, 2577–2584.

34. Forrester,S.J., Kikuchi,D.S., Hernandes,M.S., Xu,Q. and Griendling,K.K. (2018) Reactive oxygen species in metabolic and inflammatory signaling. *Circ. Res.*, **122**, 877–902.

35. Fleming,A.M. and Burrows,C.J. (2020) On the irrelevancy of hydroxyl radical to DNA damage from oxidative stress and implications for epigenetics. *Chem. Soc. Rev.*, **49**, 6524–6528.

36. Fleming,A.M. and Burrows,C.J. (2020) Iron Fenton oxidation of 2′-deoxyguanosine in physiological bicarbonate buffer yields products consistent with the reactive oxygen species carbonate radical anion not the hydroxyl radical. *Chem. Commun.*, **56**, 9779–9782.

37. Lonkar,P. and Dedon,P.C. (2011) Reactive species and DNA damage in chronic inflammation: reconciling chemical mechanisms and biological fates. *Int. J. Cancer*, **128**, 1999–2009.

38. Burrows,C.J. and Muller,J.G. (1998) Oxidative nucleobase modifications leading to strand scission. *Chem. Rev.*, **98**, 1109–1152.

39. Chatgilialoglu,C., Ferreri,C., Krokidis,M.G., Masi,A. and Terzidis,M.A. (2021) On the relevance of hydroxyl radical to purine DNA damage. *Free Radic. Res.*, https://doi.org/10.1080/10715762.2021.1876855.

40. Fleming,A.M. and Burrows,C.J. (2017) Formation and processing of DNA damage substrates for the hNEIL enzymes. *Free Radic. Biol. Med.*, **107**, 35–52.

41. Delaney,S. and Barton,J.K. (2003) Long-range DNA charge transport. *J. Org. Chem.*, **68**, 6475–6483.

42. Saito,I., Takayama,M., Sugiyama,H., Nakatani,K., Tsuchida,A. and Yamamoto,M. (1995) Photoinduced DNA cleavage via electron transfer: demonstration that guanine residues located 5′ to guanine are the most electron-donating sites. *J. Am. Chem. Soc.*, **117**, 6406–6407.

43. Tse,E.C.M., Zwang,T.J., Bedoya,S. and Barton,J.K. (2019) Effective distance for DNA-mediated charge transport between repair proteins. *ACS Cent. Sci.*, **5**, 65–72.

44. Fleming,A.M. and Burrows,C.J. (2013) 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.*, **26**, 593–607.

45. Fleming,A.M., Zhou,J., Wallace,S.S. and Burrows,C.J. (2015) A role for the fifth G-track in G-quadruplex forming oncogene promoter sequences during oxidative stress: do these "spare tires" have an evolved function? *ACS Cent. Sci.*, **1**, 226–233.

46. Di Mascio,P., Martinez,G.R., Miyamoto,S., Ronsein,G.E., Medeiros,M.H.G. and Cadet,J. (2019) Singlet molecular oxygen reactions with nucleic acids, lipids, and proteins. *Chem. Rev.*, **119**, 2043–2086.

47. Fleming,A.M., Orendt,A.M., He,Y., Zhu,J., Dukor,R.K. and Burrows,C.J. (2013) Reconciliation of chemical, enzymatic, spectroscopic and computational data to assign the absolute configuration of the DNA base lesion spiroiminodihydantoin. *J. Am. Chem. Soc.*, **135**, 18191–18204.

48. Luo,W., Muller,J.G., Rachlin,E.M. and Burrows,C.J. (2000) Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-oxo-7,8-dihydroguanosine. *Org. Lett.*, **2**, 613–616.

49. Luo,W., Muller,J.G., Rachlin,E.M. and Burrows,C.J. (2001) Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model. *Chem. Res. Toxicol.*, **14**, 927–938.

50. Mangerich,A., Knutson,C.G., Parry,N.M., Muthupalani,S., Ye,W., Prestwich,E., Cui,L., McFaline,J.L., Mobley,M., Ge,Z. *et al.* (2012) 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.*, **109**, E1820–E1829.

51. Cadet,J., Wagner,R.J., Shafirovich,V. and Geacintov,N.E. (2014) One-electron oxidation reactions of purine and pyrimidine bases in cellular DNA. *Int. J. Radiat. Biol.*, **90**, 423–432.

52. David,S.S. and Williams,S.D. (1998) Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chem. Rev.*, **98**, 1221–1262.

53. Fleming,A.M. and Burrows,C.J. (2017) 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.*, **15**, 8341–8353.

54. Banda,D.M., Nunez,N.N., Burnside,M.A., Bradshaw,K.M. and David,S.S. (2017) Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: mechanism, metals and medicine. *Free Radic. Biol. Med.*, **107**, 202–215.

55. Minowa,O., Arai,T., Hirano,M., Monden,Y., Nakai,S., Fukuda,M., Itoh,M., Takano,H., Hippou,Y., Aburatani,H. *et al.* (2000) Mmh/Ogg1 gene inactivation results in accumulation of 8-hydroxyguanine in mice. *Proc. Natl Acad. Sci. U.S.A.*, **97**, 4156–4161.

56. Al-Tassan,N., Chmiel,N.H., Maynard,J., Fleming,N., Livingston,A.L., Williams,G.T., Hodges,A.K., Davies,D.R., David,S.S., Sampson,J.R. *et al.* (2002) Inherited variants of MYH associated with somatic G:C→T:A mutations in colorectal tumors. *Nat. Genet.*, **30**, 227–232.

57. Allgayer,J., Kitsera,N., Bartelt,S., Epe,B. and Khobta,A. (2016) Widespread transcriptional gene inactivation initiated by a repair intermediate of 8-oxoguanine. *Nucleic Acids Res.*, **44**, 7267–7280.

58. Hill,J.W., Hazra,T.K., Izumi,T. and Mitra,S. (2001) Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic Acids Res.*, **29**, 430–438.

59. Krishnamurthy,N., Zhao,X., Burrows,C.J. and David,S.S. (2008) Superior removal of hydantoin lesions relative to other oxidized bases by the human DNA glycosylase hNEIL1. *Biochemistry*, **47**, 7137–7146.

60. Zhao,X., Krishnamurthy,N., Burrows,C.J. and David,S.S. (2010) Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts. *Biochemistry*, **49**, 1658–1666.

61. Zhou,J., Fleming,A.M., Averill,A.M., Burrows,C.J. and Wallace,S.S. (2015) The NEIL glycosylases remove oxidized guanine lesions from telomeric and promoter quadruplex DNA structures. *Nucleic Acids Res.*, **43**, 4039–4054.

62. Zhou,J., Liu,M., Fleming,A.M., Burrows,C.J. and Wallace,S.S. (2013) Neil3 and NEIL1 DNA glycosylases remove oxidative damages from quadruplex DNA and exhibit preferences for lesions in the telomeric sequence context. *J. Biol. Chem.*, **288**, 27263–27272.

63. Krokeide,S.Z., Laerdahl,J.K., Salah,M., Luna,L., Cederkvist,F.H., Fleming,A.M., Burrows,C.J., Dalhus,B. and Bjoras,M. (2013) Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimindiohydantoin and guanidinohydantoin. *DNA Repair*, **12**, 1159–1164.

64. Day,H.A., Pavlou,P. and Waller,Z.A. (2014) i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.*, **22**, 4407–4418.

65. An,N., Fleming,A.M. and Burrows,C.J. (2016) Human telomere G-quadruplexes with five repeats accommodate 8-oxo-7,8-dihydroguanine by looping out the DNA damage. *ACS Chem. Biol.*, **11**, 500–507.

66. Cogoi,S., Ferino,A., Miglietta,G., Pedersen,E.B. and Xodo,L.E. (2018) The regulatory G4 motif of the Kirsten ras (KRAS) gene is sensitive to guanine oxidation: implications on transcription. *Nucleic Acids Res.*, **46**, 661–676.

67. Grün,J.T., Hennecker,C., Klötzner,D.-P., Harkness,R.W., Bessi,I., Heckel,A., Mittermaier,A.K. and Schwalbe,H. (2020) Conformational dynamics of strand register shifts in DNA G-quadruplexes. *J. Am. Chem. Soc.*, **142**, 264–273.

68. Spiegel,J., Cuesta,S.M., Adhikari,S., Hänsel-Hertsch,R., Tannahill,D. and Balasubramanian,S. (2021) G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol.*, **22**, 117.

69. Perillo,B., Ombra,M.N., Bertoni,A., Cuozzo,C., Sacchetti,S., Sasso,A., Chiariotti,L., Malorni,A., Abbondanza,C. and Avvedimento,E.V. (2008) DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science*, **319**, 202–206.

70. Zuchegna,C., Aceto,F., Bertoni,A., Romano,A., Perillo,B., Laccetti,P., Gottesman,M.E., Avvedimento,E.V. and Porcellini,A. (2014) Mechanism of retinoic acid-induced transcription: histone code, DNA oxidation and formation of chromatin loops. *Nucleic Acids Res.*, **42**, 11040–11055.

71. Clark,D.W., Phang,T., Edwards,M.G., Geraci,M.W. and Gillespie,M.N. (2012) Promoter G-quadruplex sequences are targets for base oxidation and strand cleavage during hypoxia-induced transcription. *Free Radic. Biol. Med.*, **53**, 51–59.

72. Pastukh,V., Roberts,J.T., Clark,D.W., Bardwell,G.C., Patel,M., Al-Mehdi,A.B., Borchert,G.M. and Gillespie,M.N. (2015) An oxidative DNA "damage" and repair mechanism localized in the VEGF promoter is important for hypoxia-induced VEGF mRNA expression. *Am. J. Physiol. Lung Cell Mol. Physiol.*, **309**, L1367–L1375.

73. Schermerhorn,K.M. and Delaney,S. (2013) Transient-state kinetics of apurinic/apyrimidinic (AP) endonuclease 1 acting on an authentic AP site and commonly used substrate analogs: the effect of diverse metal ions and base mismatches. *Biochemistry*, **52**, 7669–7677.

74. Fulton,A. and Chong,Y. (1992) The role of macrophage-derived TNFα in the induction of sublethal tumor cell DNA damage. *Carcinogenesis*, **13**, 77–81.

75. Burra,S., Marasco,D., Malfatti,M.C., Antoniali,G., Virgilio,A., Esposito,V., Demple,B., Galeone,A. and Tell,G. (2019) Human AP-endonuclease (Ape1) activity on telomeric G4 structures is modulated by acetylatable lysine residues in the N-terminal sequence. *DNA Repair*, **73**, 129–143.

76. McNeill,D.R., Whitaker,A.M., Stark,W.J., Illuzzi,J.L., McKinnon,P.J., Freudenthal,B.D. and Wilson,D.M. (2020) Functions of the major abasic endonuclease (APE1) in cell viability and genotoxin resistance. *Mutagenesis*, **35**, 27–38.

77. Tell,G., Quadrifoglio,F., Tiribelli,C. and Kelley,M.R. (2009) The many functions of APE1/Ref-1: not only a DNA repair enzyme. *Antioxid. Redox Signal.*, **11**, 601–620.

78. Pan,L., Zhu,B., Hao,W., Zeng,X., Vlahopoulos,S.A., Hazra,T.K., Hegde,M.L., Radak,Z., Bacsi,A., Brasier,A.R. *et al.* (2016) Oxidized guanine base lesions function in 8-oxoguanine DNA glycosylase1-mediated epigenetic regulation of nuclear factor κB-driven gene expression. *J. Biol. Chem.*, **291**, 25553–25566.

79. Wang,R., Hao,W., Pan,L., Boldogh,I. and Ba,X. (2018) The roles of base excision repair enzyme OGG1 in gene expression. *Cell. Mol. Life Sci.*, **75**, 3741–3750.

80. Antoniali,G., Lirussi,L., D'Ambrosio,C., Dal Piaz,F., Vascotto,C., Casarano,E., Marasco,D., Scaloni,A., Fogolari,F. and Tell,G. (2014) SIRT1 gene expression upon genotoxic damage is regulated by APE1 through nCaRE-promoter elements. *Mol. Biol. Cell*, **25**, 532–547.

81. Hanawalt,P.C. and Spivak,G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.*, **9**, 958–970.

82. Seifermann,M. and Epe,B. (2017) Oxidatively generated base modifications in DNA: not only carcinogenic risk factor but also regulatory mark? *Free Radic. Biol. Med.*, **107**, 258–265.

83. Mol,C.D., Izumi,T., Mitra,S. and Tainer,J.A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination. *Nature*, **403**, 451–456.

84. Tell,G., Piazza,M.D., Kamocka,M.M. and Vascotto,C. (2013) Combining RNAi and *in vivo* confocal microscopy analysis of the photoconvertible fluorescent protein Dendra2 to study a DNA repair protein. *BioTechniques*, **55**, 198–203.

85. Unnikrishnan,A., Raffoul,J.J., Patel,H.V., Prychitko,T.M., Anyangwe,N., Meira,L.B., Friedberg,E.C., Cabelof,D.C. and Heydari,A.R. (2009) Oxidative stress alters base excision repair pathway and increases apoptotic response in apurinic/apyrimidinic endonuclease 1/redox factor-1 haploinsufficient mice. *Free Radic. Biol. Med.*, **46**, 1488–1499.

86. Xanthoudakis,S., Smeyne,R.J., Wallace,J.D. and Curran,T. (1996) The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice. *Proc. Natl Acad. Sci. U.S.A.*, **93**, 8919–8923.

87. Huamani,J., McMahan,C.A., Herbert,D.C., Reddick,R., McCarrey,J.R., MacInnes,M.I., Chen,D.J. and Walter,C.A. (2004) Spontaneous mutagenesis is enhanced in Apex heterozygous mice. *Mol. Cell Biol.*, **24**, 8145–8153.

88. Tell,G., Damante,G., Caldwell,D. and Kelley,M.R. (2005) The intracellular localization of APE1/Ref-1: more than a passive phenomenon? *Antioxid. Redox Signal.*, **7**, 367–384.

89. Di Maso,V., Avellini,C., Crocè,L.S., Rosso,N., Quadrifoglio,F., Cesaratto,L., Codarin,E., Bedogni,G., Beltrami,C.A., Tell,G. *et al.* (2007) Subcellular localization of APE1/Ref-1 in human hepatocellular carcinoma: possible prognostic significance. *Mol. Med.*, **13**, 89–96.

90. McIlwain,D.W., Fishel,M.L., Boos,A., Kelley,M.R. and Jerde,T.J. (2018) APE1/Ref-1 redox-specific inhibition decreases survivin protein levels and induces cell cycle arrest in prostate cancer cells. *Oncotarget*, **9**, 10962–10977.

91. Hadi,M.Z., Ginalski,K., Nguyen,L.H. and Wilson,D.M. 3rd (2002) Determinants in nuclease specificity of Ape1 and Ape2, human homologues of *Escherichia coli* exonuclease III. *J. Mol. Biol.*, **316**, 853–866.

92. Manvilla,B.A., Varney,K.M. and Drohat,A.C. (2010) Chemical shift assignments for human apurinic/apyrimidinic endonuclease 1. *Biomol. NMR Assign.*, **4**, 5–8.

93. Li,M. and Wilson,D.M. 3rd (2014) Human apurinic/apyrimidinic endonuclease 1. *Antioxid. Redox Signal.*, **20**, 678–707.

94. Antoniali,G., Serra,F., Lirussi,L., Tanaka,M., D'Ambrosio,C., Zhang,S., Radovic,S., Dalla,E., Ciani,Y., Scaloni,A. *et al.* (2017) Mammalian APE1 controls miRNA processing and its interactome is linked to cancer RNA metabolism. *Nat. Commun.*, **8**, 797–797.

95. Gampala,S., Caston,R.A., Fishel,M.L. and Kelley,M.R. (2021) In: Dizdaroglu,M.L. and Lloyd,R.S. (eds). *DNA Damage, DNA Repair, and Disease*. Vol. 2. Royal Society of Chemistry, Cambridge, UK, pp. 286–318.

96. Freudenthal,B.D., Beard,W.A., Cuneo,M.J., Dyrkheeva,N.S. and Wilson,S.H. (2015) Capturing snapshots of APE1 processing DNA damage. *Nat. Struct. Mol. Biol.*, **22**, 924–931.

97. Whitaker,A.M., Flynn,T.S. and Freudenthal,B.D. (2018) Molecular snapshots of APE1 proofreading mismatches and removing DNA damage. *Nat. Commun.*, **9**, 399.

98. Liu,T.-C., Lin,C.-T., Chang,K.-C., Guo,K.-W., Wang,S., Chu,J.-W. and Hsiao,Y.-Y. (2021) APE1 distinguishes DNA substrates in exonucleolytic cleavage by induced space-filling. *Nat. Commun.*, **12**, 601.

99. Liu,L., Kong,M., Gassman,N.R., Freudenthal,B.D., Prasad,R., Zhen,S., Watkins,S.C., Wilson,S.H. and Van Houten,B. (2017) PARP1 changes from three-dimensional DNA damage searching to one-dimensional diffusion after auto-PARylation or in the presence of APE1. *Nucleic Acids Res.*, **45**, 12834–12847.

100. Aboelnga,M.M. and Wetmore,S.D. (2019) Unveiling a single-metal-mediated phosphodiester bond cleavage mechanism for nucleic acids: a multiscale computational investigation of a human DNA repair enzyme. *J. Am. Chem. Soc.*, **141**, 8646–8656.

101. Erzberger,J.P. and Wilson,D.M. 3rd (1999) The role of Mg^{2+} and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease: new insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis. *J. Mol. Biol.*, **290**, 447–457.

102. Kanazhevskaya,L.Y., Koval,V.V., Lomzov,A.A. and Fedorova,O.S. (2014) The role of Asn-212 in the catalytic mechanism of human endonuclease APE1: stopped-flow kinetic study of incision activity on a natural AP site and a tetrahydrofuran analogue. *DNA Repair*, **21**, 43–54.

103. Maher,R.L. and Bloom,L.B. (2007) Pre-steady-state kinetic characterization of the AP endonuclease activity of human AP endonuclease 1. *J. Biol. Chem.*, **282**, 30577–30585.

104. Ray,S., Menezes,M.R., Senejani,A. and Sweasy,J.B. (2013) Cellular roles of DNA polymerase beta. *Yale J. Biol. Med.*, **86**, 463–469.

105. Gros,L., Ishchenko,A.A., Ide,H., Elder,R.H. and Saparbaev,M.K. (2004) The major human AP endonuclease (Ape1) is involved in the nucleotide incision repair pathway. *Nucleic Acids Res.*, **32**, 73–81.

106. Wilson,D.M. 3rd (2005) Ape1 abasic endonuclease activity is regulated by magnesium and potassium concentrations and is robust on alternative DNA structures. *J. Mol. Biol.*, **345**, 1003–1014.

107. Berquist,B.R., McNeill,D.R. and Wilson,D.M. 3rd (2008) Characterization of abasic endonuclease activity of human Ape1 on alternative substrates, as well as effects of ATP and sequence context on AP site incision. *J. Mol. Biol.*, **379**, 17–27.

108. Broxson,C., Hayner,J.N., Beckett,J., Bloom,L.B. and Tornaletti,S. (2014) Human AP endonuclease inefficiently removes abasic sites within G4 structures compared to duplex DNA. *Nucleic Acids Res.*, **42**, 7708–7719.

109. Davletgildeeva,A.T., Kuznetsova,A.A., Fedorova,O.S. and Kuznetsov,N.A. (2020) Activity of human apurinic/apyrimidinic endonuclease APE1 toward damaged DNA and native RNA with non-canonical structures. *Front. Cell Dev. Biol.*, **8**, 590848.

110. Omaga,C.A. (2018) In: *Analysis of the effect of the G-quadruplex fifth domain on DNA folding in response to oxidative and abasic lesions, Ph.D. thesis, University of Utah, Salt Lake City, UT*. https://collections.lib.utah.edu/ark:/87278/s6839rjq.

111. Lucas,J.A., Masuda,Y., Bennett,R.A., Strauss,N.S. and Strauss,P.R. (1999) Single-turnover analysis of mutant human apurinic/apyrimidinic endonuclease. *Biochemistry*, **38**, 4958–4964.

112. Whitaker,A.M., Stark,W.J., Flynn,T.S. and Freudenthal,B.D. (2020) Molecular and structural characterization of disease-associated APE1 polymorphisms. *DNA Repair*, **91–92**, 102867.

113. Hadi,M.Z., Coleman,M.A., Fidelis,K., Mohrenweiser,H.W. and Wilson,D.M. 3rd (2000) Functional characterization of Ape1 variants identified in the human population. *Nucleic Acids Res.*, **28**, 3871–3879.

114. Miroshnikova,A.D., Kuznetsova,A.A., Vorobjev,Y.N., Kuznetsov,N.A. and Fedorova,O.S. (2016) Effects of mono- and divalent metal ions on DNA binding and catalysis of human apurinic/apyrimidinic endonuclease 1. *Mol. Biosyst.*, **12**, 1527–1539.

115. Moor,N., Vasil'eva,I. and Lavrik,O. (2020) Functional role of N-terminal extension of human AP endonuclease 1 in coordination of base excision DNA repair via protein–protein interactions. *Int. J. Mol. Sci.*, **21**, 3122.

116. Fantini,D., Vascotto,C., Marasco,D., D'Ambrosio,C., Romanello,M., Vitagliano,L., Pedone,C., Poletto,M., Cesaratto,L., Quadrifoglio,F. *et al.* (2010) Critical lysine residues within the overlooked N-terminal domain of human APE1 regulate its biological functions. *Nucleic Acids Res.*, **38**, 8239–8256.

117. Chattopadhyay,R., Wiederhold,L., Szczesny,B., Boldogh,I., Hazra,T.K., Izumi,T. and Mitra,S. (2006) Identification and characterization of mitochondrial abasic (AP)-endonuclease in mammalian cells. *Nucleic Acids Res.*, **34**, 2067–2076.

118. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat d[AG$_3$(T$_2$AG$_3$)$_3$] G-tetraplex. *Structure*, **1**, 263–282.

119. Fleming,A.M., Howpay Manage,S.A. and Burrows,C.J. (2021) Binding of AP endonuclease-1 to G-quadruplex DNA depends on the N-terminal domain, Mg^{2+} and ionic strength. bioRxiv doi: https://doi.org/10.1101/2021.08.25.457676, 25 August 2021, preprint: not peer reviewed.

120. Jayaraman,L., Murthy,K.G., Zhu,C., Curran,T., Xanthoudakis,S. and Prives,C. (1997) Identification of redox/repair protein Ref-1 as a potent activator of p53. *Genes Dev.*, **11**, 558–570.

121. Malfatti,M.C., Antoniali,G., Codrich,M., Burra,S., Mangiapane,G., Dalla,E. and Tell,G. (2020) New perspectives in cancer biology from a study of canonical and non-canonical functions of base excision repair proteins with a focus on early steps. *Mutagenesis*, **35**, 129–149.

122. Caston,R.A., Gampala,S., Armstrong,L., Messmann,R.A., Fishel,M.L. and Kelley,M.R. (2021) The multifunctional APE1 DNA repair–redox signaling protein as a drug target in human disease. *Drug Discov. Today*, **26**, 218–228.

123. Chiu,W.C., Chen,C.J., Lee,T.S., Chen,Z.J., Ke,P.H. and Chiang,A.N. (2010) Oxidative stress enhances AP-1 and NF-κB-mediated regulation of β$_2$-glycoprotein I gene expression in hepatoma cells. *J. Cell Biochem.*, **111**, 988–998.

124. Liu,D. and Xu,Y. (2011) p53, oxidative stress, and aging. *Antioxid. Redox Signal.*, **15**, 1669–1678.

125. Barry,S.P., Townsend,P.A., McCormick,J., Knight,R.A., Scarabelli,T.M., Latchman,D.S. and Stephanou,A. (2009) STAT3 deletion sensitizes cells to oxidative stress. *Biochem. Biophys. Res. Commun.*, **385**, 324–329.

126. Georgiadis,M.M., Luo,M., Gaur,R.K., Delaplane,S., Li,X. and Kelley,M.R. (2008) Evolution of the redox function in mammalian apurinic/apyrimidinic endonuclease. *Mutat. Res.*, **643**, 54–63.

127. Walker,L.J., Robson,C.N., Black,E., Gillespie,D. and Hickson,I.D. (1993) Identification of residues in the human DNA repair enzyme HAP1 (Ref-1) that are essential for redox regulation of Jun DNA binding. *Mol. Cell Biol.*, **13**, 5370–5376.

128. Luo,M., Zhang,J., He,H., Su,D., Chen,Q., Gross,M.L., Kelley,M.R. and Georgiadis,M.M. (2012) Characterization of the redox activity and disulfide bond formation in apurinic/apyrimidinic endonuclease. *Biochemistry*, **51**, 695–705.

129. Su,D., Delaplane,S., Luo,M., Rempel,D.L., Vu,B., Kelley,M.R., Gross,M.L. and Georgiadis,M.M. (2011) Interactions of apurinic/apyrimidinic endonuclease with a redox inhibitor: evidence for an alternate conformation of the enzyme. *Biochemistry*, **50**, 82–92.

130. Bazlekowa-Karaban,M., Prorok,P., Baconnais,S., Taipakova,S., Akishev,Z., Zembrzuska,D., Popov,A.V., Endutkin,A.V., Groisman,R., Ishchenko,A.A. *et al.* (2019) Mechanism of stimulation of DNA binding of the transcription factors by human apurinic/apyrimidinic endonuclease 1, APE1. *DNA Repair*, **82**, 102698.

131. López,D.J., Rodríguez,J.A. and Bañuelos,S. (2021) Molecular mechanisms regulating the DNA repair protein APE1: a focus on its flexible N-terminal tail domain. *Int. J. Mol. Sci.*, **22**, 6308.

132. Roychoudhury,S., Pramanik,S., Harris,H.L., Tarpley,M., Sarkar,A., Spagnol,G., Sorgen,P.L., Chowdhury,D., Band,V., Klinkebiel,D. *et al.* (2020) Endogenous oxidized DNA bases and APE1 regulate the formation of G-quadruplex structures in the genome. *Proc. Natl Acad. Sci. U.S.A.*, **117**, 11409–11420.

133. Wang,K., Maayah,M., Sweasy,J.B. and Alnajjar,K.S. (2021) The role of cysteines in the structure and function of OGG1. *J. Biol. Chem.*, **296**, 100093.

134. van der Reest,J., Lilla,S., Zheng,L., Zanivan,S. and Gottlieb,E. (2018) Proteome-wide analysis of cysteine oxidation reveals metabolic sensitivity to redox stress. *Nat. Commun.*, **9**, 1581.

135. Ginno,P.A., Lott,P.L., Christensen,H.C., Korf,I. and Chédin,F. (2012) R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell*, **45**, 814–825.

136. Fleming,A.M., Zhu,J., Ding,Y. and Burrows,C.J. (2017) 8-Oxo-7,8-dihydroguanine in the context of a promoter G-quadruplex is an on–off switch for transcription. *ACS Chem. Biol.*, **12**, 2417–2426.

137. Scheibye-Knudsen,M., Tseng,A., Borch Jensen,M., Scheibye-Alsing,K., Fang,E.F., Iyama,T., Bharti,S.K., Marosi,K., Froetscher,L., Kassahun,H. *et al.* (2016) Cockayne syndrome group A and B proteins converge on transcription-linked resolution of non-B DNA. *Proc. Natl Acad. Sci. U.S.A.*, **113**, 12502–12507.

138. van Gool,A.J., Citterio,E., Rademakers,S., van Os,R., Vermeulen,W., Constantinou,A., Egly,J.M., Bootsma,D. and Hoeijmakers,J.H. (1997) The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerase II-containing complex. *EMBO J.*, **16**, 5955–5965.

139. Liano,D. and Antonio,M.D. (2021) Cockayne Syndrome B protein selectively interacts and resolves intermolecular DNA G-quadruplex structures. bioRxiv doi: http://doi.org/10.1101/2021.03.25.436565, 25 March 2021, preprint: not peer reviewed.

140. Guo,J., Hanawalt,P.C. and Spivak,G. (2013) Comet-FISH with strand-specific probes reveals transcription-coupled repair of 8-oxoguanine in human cells. *Nucleic Acids Res.*, **41**, 7700–7712.

141. Allgayer,J., Kitsera,N., von der Lippen,C., Epe,B. and Khobta,A. (2013) Modulation of base excision repair of 8-oxoguanine by the nucleotide sequence. *Nucleic Acids Res.*, **41**, 8559–8571.

142. Oh,J., Fleming,A.M., Xu,J., Chong,J., Burrows,C.J. and Wang,D. (2020) RNA polymerase II stalls on oxidative DNA damage via a torsion-latch mechanism involving lone pair–π and CH–π interactions. *Proc. Natl Acad. Sci. U.S.A.*, **117**, 9338–9348.

143. Shibutani,S., Takeshita,M. and Grollman,A.P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature*, **349**, 431–434.

144. Kamiya,H., Kurokawa,M., Makino,T., Kobayashi,M. and Matsuoka,I. (2015) Induction of action-at-a-distance mutagenesis by 8-oxo-7,8-dihydroguanine in DNA pol lambda-knockdown cells. *Genes Environ.*, **37**, 10.

145. Kornyushyna,O., Berges,A.M., Muller,J.G. and Burrows,C.J. (2002) *In vitro* nucleotide misinsertion opposite the oxidized guanosine lesions spiroiminodihydantoin and guanidinohydantoin and DNA synthesis past the lesions using *Escherichia coli* DNA polymerase I (Klenow fragment). *Biochemistry*, **41**, 15304–15314.

146. Henderson,P.T., Delaney,J.C., Muller,J.G., Neeley,W.L., Tannenbaum,S.R., Burrows,C.J. and Essigmann,J.M. (2003) The hydantoin lesions formed from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors *in vivo*. *Biochemistry*, **42**, 9257–9262.

147. Lirussi,L., Antoniali,G., D'Ambrosio,C., Scaloni,A., Nilsen,H. and Tell,G. (2016) APE1 polymorphic variants cause persistent genomic stress and affect cancer cell proliferation. *Oncotarget*, **7**, 26293–26306.

148. Hao,W., Qi,T., Pan,L., Wang,R., Zhu,B., Aguilera-Aguirre,L., Radak,Z., Hazra,T.K., Vlahopoulos,S.A., Bacsi,A. *et al.* (2018) Effects of the stimuli-dependent enrichment of 8-oxoguanine DNA glycosylase1 on chromatinized DNA. *Redox Biol.*, **18**, 43–53.

149. Rolseth,V., Luna,L., Olsen,A.K., Suganthan,R., Scheffler,K., Neurauter,C.G., Esbensen,Y., Kusnierczyk,A., Hildrestrand,G.A., Graupner,A. *et al.* (2017) No cancer predisposition or increased spontaneous mutation frequencies in NEIL DNA glycosylases-deficient mice. *Sci. Rep.*, **7**, 4384.

150. Vartanian,V., Lowell,B., Minko,I.G., Wood,T.G., Ceci,J.D., George,S., Ballinger,S.W., Corless,C.L., McCullough,A.K. and Lloyd,R.S. (2006) The metabolic syndrome resulting from a knockout of the NEIL1 DNA glycosylase. *Proc. Natl Acad. Sci. U.S.A.*, **103**, 1864–1869.

151. Puget,N., Miller,K.M. and Legube,G. (2019) Non-canonical DNA/RNA structures during transcription-coupled double-strand break repair: roadblocks or bona fide repair intermediates? *DNA Repair*, **81**, 102661.

152. Georgakopoulos-Soares,I., Morganella,S., Jain,N., Hemberg,M. and Nik-Zainal,S. (2018) Noncanonical secondary structures arising from non-B DNA motifs are determinants of mutagenesis. *Genome Res.*, **28**, 1264–1271.

153. Cadoni,E., De Paepe,L., Manicardi,A. and Madder,A. (2021) Beyond small molecules: targeting G-quadruplex structures with oligonucleotides and their analogues. *Nucleic Acids Res.*, **49**, 6638–6659.

154. Fleming,A.M. and Burrows,C.J. (2017) 8-Oxo-7,8-dihydroguanine, friend and foe: epigenetic-like regulator versus initiator of mutagenesis. *DNA Repair*, **56**, 75–83.

155. Limpose,K.L., Corbett,A.H. and Doetsch,P.W. (2017) BERing the burden of damage: pathway crosstalk and posttranslational modification of base excision repair proteins regulate DNA damage management. *DNA Repair*, **56**, 51–64.

156. You,C., Dai,X., Yuan,B., Wang,J., Brooks,P.J., Niedernhofer,L.J. and Wang,Y. (2012) A quantitative assay for assessing the effects of DNA lesions on transcription. *Nat. Chem. Biol.*, **8**, 817–822.

157. Mladenova,V., Mladenov,E. and Russev,G. (2009) Organization of plasmid DNA into nucleosome-like structures after transfection in eukaryotic cells. *Biotechnol. Biotechnol. Equip.*, **23**, 1044–1047.