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

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

*Department of Chemistry, University of Utah, Salt Lake City, UT 84112-0850

Edited by Graham C. Walker, Massachusetts Institute of Technology, Cambridge, MA, and approved December 29, 2016 (received for review December 2, 2016)

Reactive oxygen species (ROS) have emerged as important cellular-signaling agents for cellular survival. Herein, we demonstrate that ROS-mediated oxidation of DNA to yield 8-oxo-7,8-dihydroguanine (OG) in gene promoters is a signaling agent for gene activation. Enhanced gene expression occurs when OG is formed in guanine-rich, potential G-quadruplex–forming sequences (PQS) in promoter-coding strands, initiating base excision repair (BER) by 8-oxoguanine DNA glycosylase (OGG1), yielding an abasic site (AP). The AP enables melting of the duplex to unmask the PQS, adopting a G-quadruplex fold in which apurinic/apyrimidinic endonuclease 1 (APE1) binds, but inefficiently cleaves, the AP for activation of vascular endothelial growth factor (VEGF) or endonuclease III-like protein 1 (NTHL1) genes. These details were mapped via synthesis of OG and AP analogs at single-nucleotide precision within the promoter of a lucerase reporter system. The reporters were analyzed in human and mouse cells while selectively knocking out or down critical BER proteins to identify the impact on lucerase expression. Identification of the oxidatively modified DNA base OG to guide BER activity in a gene promoter and impact cellular phenotype ascribes an epigenetic role to OG.

Significance

Damage to DNA bases due to oxidative stress is thought to be deleterious, leading to stalled replication forks and mutations. Similarly, folding of DNA strands into G-quadruplexes slows the progression of polymerases, requiring specialized helicases for unfolding before transcription. In the case of oxidative damage in a G-quadruplex-forming sequence of a promoter, we show that the presence of the DNA damage lesion 8-oxoguanine (OG) leads to an ∼300% increase in gene expression. This concept was demonstrated by chemical synthesis of a segment of the vascular endothelial growth factor (VEGF) or endonuclease III-like protein 1 (NTHL1) promoter with a site specifically incorporated lesion in a reporter plasmid. This observation is direct evidence that OG represents an epigenetic modification and G-quadruplex–forming sequences can serve as sensors of oxidative stress.

Author contributions: A.M.F. and C.J.B. designed research; A.M.F. and Y.D. performed research; A.M.F. and Y.D. analyzed data; and A.M.F. and C.J.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

To whom correspondence may be addressed. Email: burrows@chem.utah.edu or afleming@chem.utah.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619809114/-/DCSupplemental.
G-quadruplex structure solved by MEFs confirms OGG1 was critical to expression enhancement; however, more questions remain. In WT cells, after removal of OG by OGG1, an AP is present in DNA, and we questioned whether an AP could also impact transcription. Consequently, plasmids with APs (THF analog, F) at the reactive positions were introduced to establish whether gene induction still occurs. Transfection of the AP-containing plasmids into all three cell lines yielded Rluc expression enhancement at levels of 2.5- to 6.0-fold relative to the WT plasmid (Fig. 2C). These results identify a stronger sequence and cell line dependency in transcriptional amplification than observed with OG. The AP analog F is more stable than a natural AP, likely resulting in the increased luciferase expression detected. The enhanced expression observed with an abasic site, especially in the OGG1−/− MEFs, identifies APE1 as the possible BER enzyme responsible for enhancing gene expression in this promoter sequence.

To further confirm that APE1 induces transcription, we treated glioblastoma cells with siRNA to knock down APE1 and prevent further processing of the AP (Fig. 3A). Glioblastoma cells titrated with APE1-specific siRNA in the range of 1–50 nM yielded a dose–response in the extent of Rluc expression when AP-containing plasmids were transfected (Fig. 3B). At the highest siRNA concentration studied, Rluc expression for each position studied was similar to the WT plasmid (Fig. 3B). This initial observation adds further support for the importance of APE1 in gene activation when an AP is present in the VEGF promoter PQS. Next, we wanted to understand if the ability of APE1 to bind the AP or increased by 2.5- to 3.9-fold, whereas OG-containing plasmids in the OGG1−/− MEFs yielded essentially no change in the amounts of Rluc expression compared with the WT plasmid (Fig. 2B). Because OG was in the coding strand of the promoter, it does not interfere with advancement of RNA pol II on the promoter to the transcription start site in the OGG1−/− MEFs, and, therefore, the same gene output was observed with and without OG.

The observation that OG does not increase transcription in the OGG1−/− MEFs confirms OGG1 was critical to expression enhancement; however, more questions remain. In WT cells, after removal of OG by OGG1, an AP is present in DNA, and we questioned whether an AP could also impact transcription. Consequently, plasmids with APs (THF analog, F) at the reactive positions were introduced to establish whether gene induction still occurs. Transfection of the AP-containing plasmids into all three cell lines yielded Rluc expression enhancement at levels of 2.5- to 6.0-fold relative to the WT plasmid (Fig. 2C). These results identify a stronger sequence and cell line dependency in transcriptional amplification than observed with OG. The AP analog F is more stable than a natural AP, likely resulting in the increased luciferase expression detected. The enhanced expression observed with an abasic site, especially in the OGG1−/− MEFs, identifies APE1 as the possible BER enzyme responsible for enhancing gene expression in this promoter sequence.

To further confirm that APE1 induces transcription, we treated glioblastoma cells with siRNA to knock down APE1 and prevent further processing of the AP (Fig. 3A). Glioblastoma cells titrated with APE1-specific siRNA in the range of 1–50 nM yielded a dose–response in the extent of Rluc expression when AP-containing plasmids were transfected (Fig. 3B). At the highest siRNA concentration studied, Rluc expression for each position studied was similar to the WT plasmid (Fig. 3B). This initial observation adds further support for the importance of APE1 in gene activation when an AP is present in the VEGF promoter PQS. Next, we wanted to understand if the ability of APE1 to bind the AP or
cleave the site was responsible for gene activation (Fig. 3C). First, APE1 inhibitor III (Fig. 3D) was titrated at a concentration range of 100–1,000 nM into glioblastoma cells before transfection of the AP-containing plasmids. The inhibitor prevents cleavage of the backbone, whereas APE1 binding is not strongly impacted (Fig. 3C) (14). In these studies, Rluc expression increased up to >30-fold with a dose–response in the inhibitor concentration for the F-containing plasmid, whereas the WT plasmid remained the same throughout the titration (Fig. 3E).

Further verification that APE1 binding to APs without causing a strand break leads to gene induction was achieved by studying APs poorly cleaved by APE1 (Fig. 3C). Previous experiments found APE1 binds normally but poorly cleaves oligomers containing a phosphorothioate and 2′-OMe nucleotide 5′ to the AP analog F (2′-MeO-PS-F, Fig. 3D) (15). Therefore, 2′-MeO-PS-F modified plasmids were transfected into glioblastoma cells, and >10-fold Rluc expression was observed (Fig. 3F). To rule out a strand break leads to the increased expression observed, an OG-containing plasmid was treated with Fpg and APE1 to yield a strand break before transfection. Transfection of this plasmid did not yield an observable Rluc expression. This result is likely a consequence of the poor efficiency of transfecting strand-broken plasmids into cells. The combined results of these studies support OG inducing transcription via OGG1 generation of an AP in the VEGF PQS fold. (A) Expression observed when OG is in a G-quadruplex positive or negative folding sequence context. (B) Requirement of the fifth G run for maximal expression when OG is present. Error bars represent 95% CI on the basis of four or eight replicates. Significance values for each comparison were calculated by a Student’s t test. Significance at ***P < 0.0001 is indicated.

The next set of studies aimed to understand the possible involvement of G-quadruplex formation for induction of Rluc with the VEGF PQS promoter element. Previous experiments found the VEGF PQS to be bound by three equivalents of the SP1 transcription factor, and G4 formation is also involved in activation of this gene (8, 16). Decoupling of G-quadruplex formation from SP1 binding is complicated by the observation that this transcription factor binds both duplex and G-quadruplex folded DNA of the same sequence with similar affinity (17); additionally, OG in duplex DNA decreases SP1 binding by 70–90% depending on the position (2), and the impact in G-quadruplex DNA is not known. A sequence recognizable by SP1 but incapable of G-quadruplex formation (SI Appendix, Fig. S2) was designed for incorporation of OG (G-quadruplex negative SP1+; Fig. 4). This new OG-containing plasmid was transfected into glioblastoma cells yielding a similar expression as observed in the WT plasmid with the same G-quadruplex negative sequence (Fig. 4). This null result strongly supports the conclusion that the ability to adopt a G-quadruplex fold is important in gene activation when OG is present.

In another experiment, the VEGF sequence was judiciously mutated by removing a G run and converting one run to all T nucleotides to prevent G-quadruplex formation and SP1 binding. Transfection of this sequence with and without OG into glioblastoma...
cells also found OG did not impact transcription (SI Appendix, Fig. S3). These control experiments do not rule out all sequence context effects or the role of all protein activators and repressors but they do provide strong support for this sequence to adopt a G-quadruplex fold to induce the gene. The observation of transcriptional activation by G4 formation in the present study is consistent with a recent demonstration that these folds in promoters up-regulate transcription on the basis of G-quadruplex ChIP-Seq in tandem with RNA-Seq analysis (9).

We previously demonstrated the concept that the five-track VEGF G5 can switch structures to extrude damaged DNA bases to maintain the fold by replacing a damaged run with the fifth track (11). To verify if there is an essential role for the fifth G-track in gene activation, plasmons containing OG with only the four G runs required for folding (10) were transfected. The G4 samples gave significantly less gene expression than the G5 analogs, identifying the fifth G-track as essential to achieve the maximal amount of gene induction (Fig. 4B). Negative supercoiling in DNA was demonstrated by the Hurley laboratory to promote G4 formation (18); therefore, a comparison of Ruic expression in relaxed vs. supercoiled plasmids (plasmid state of all studies) with OG at position 12 of the VEGF G5 was made to find a similar result for both systems (SI Appendix, Fig. S4). This experiment supports the state of the plasmid leading to expression in the cell is similar for both supercoiled and relaxed plasmids. Furthermore, this observation hints that folding of VEGF to a G-quadruplex may be an essential step in the transcriptional activation process.

After establishing the essential role for a G-quadruplex in gene activation, we circled back to better understand how an AP site favors G-quadruplex formation and how APE1 functions in this context. When DNA is comprised of canonical Watson–Crick base pairs, the duplex context is thermodynamically favored, and even OG base paired with C does not significantly impact the duplex state (Fig. 5A); however, an AP site (i.e., F) significantly impacts the thermal stability of duplex DNA, leading to opening of the duplex, allowing the POS to adopt a G-quadruplex fold. The ability of an AP to unmask the G-quadruplex was demonstrated by conducting a series of thermal melting experiments (T_m). In the VEGF duplex an AP site decreased the T_m by ~20 °C at the positions studied (Fig. 5A). Evaluation of T_m values in the G4 context found loop APs decreased the T_m by ~20 °C; in contrast, when the fifth domain was present in the G5 context the T_m values remained nearly the same as the nonmodified sequence while maintaining a G-quadruplex structure (Fig. 5B and SI Appendix, Fig. S5). We propose that maintaining the thermal stability results from extrusion of the damaged G run and replacing it with the fifth G-track (Fig. 5B). This observation is consistent with our previous studies that found extrusion of a damaged G-track by the fifth domain yields a more stable fold for highly helix-distorting modified DNA bases (11). Furthermore, recruitment of the fifth domain was most efficient for APs in core positions because they caused the greatest disruption to the G4-fold. This observation supports a mechanism by which the AP can facilitate G-quadruplex formation by shifting the duplex-quadruplex equilibrium.

The interaction between APE1 and an AP was previously studied in the G4 context by the C.J.B. laboratory and others, and it was found that APE1 binds APs in G4s, but the cleavage rate was highly attenuated (12, 19). The attenuated cleavage by APE1 on the promoter may provide the trigger to increase the rate of transcription observed by allowing APE1 to idle on the G4/G5 structural motif. The Ref-1 domain of APE1 is known to interact with protein factors (e.g., HIF1-α, STAT3, and CBP/p300) that bind promoter sequences such as the VEGF POS and increase gene transcription (20); however, the detailed molecular interactions are not known at present. This proposed mechanism suggests any structure in DNA that prevents APE1 cleavage of an AP would lead to transcriptional activation. To test this hypothesis, we transfected plasmids into glioblastoma cells with noncleavable APs in both G-quadruplex negative sequences studied previously and found a significant increase of >1.5-fold in Ruic expression (SI Appendix, Fig. S6). The increase observed was less than that observed in the G-quadruplex context, suggesting the folded structure must play an additional role for induction of the gene that is not currently well understood. On the basis of these studies, gene induction by OG in the VEGF promoter POS occurs when OGG1 binds and cleaves OG from the duplex to yield an AP. The AP shifts the duplex-quadruplex equilibrium to the G-quadruplex fold followed by APE1 binding and gene activation via the gene regulatory Ref-1 domain of APE1 (Fig. 5C).

Activation of the VEGF promoter POS in the genomic context would require site-specific G oxidation to OG to initiate the BER process in a gene promoter. Oxidation can result from cellular oxidants (6), via long-range electron transfer (21), or active chromatin remodeling catalyzed by lysine demethylase 1A or 1B (LSD1 or LSD2) (22, 23). The latter pathway identified by Perillo et al. (22) occurs by the flavin-dependent LSD1 during demethylation of H3K9me2 to generate H2O2 near the DNA; LSD2 also catalyzes a similar reaction. The H2O2 so formed is most likely activated by the Fe(II)- or Cu(I)-mediated Fenton reaction to oxidize G to OG (and other base oxidation products) or an AP (24, 25). Perillo et al. demonstrated estrogen-induced activation of the BCL-2 gene occurs by this LSD1-mediated DNA oxidation mechanism in MCF7 cells, and they noted an essential role for OGG1 of BER in the activation process. Curiously, the region found to be oxidized in the BCL-2 promoter houses a POS hypothesized to be responsible for gene regulation (18). The present results support a proposal that oxidation of G to OG in the BCL-2 POS directs OGG1 to unmask the G4-fold to alter gene expression ultimately through APE1 (Fig. 5C).
Thus, G oxidation in a regulatory PQS may serve as an epigenetic DNA modification to alter cellular phenotype.

A recent report by Boldogh and coworkers (26) found cellular oxidative stress oxidized G most likely to OG in a G-rich region adjacent to an NF-kB binding site in the TNF-α promoter of HEK293 cells. They propose OG was then bound by OGG1 that recruited NF-kB to its binding site and increased transcription. In a study by Antonioli et al., gene activation by OG was suspected in SIRT1 expression under oxidative stress conditions in HeLa cells (27); their work highlighted the importance of OGG1 and APE1 in the gene activation process. Together, these three genome-level studies support the concept that the actions of BER in gene promoters guided by oxidative modification of G can increase transcription; however, they could not definitively claim that OG was the oxidative modification responsible for the observed phenotypic change, or if the presence of OG was a consequence of the change in cellular state. Our studies reported here provide this evidence.

In the present work, we addressed these questions by synthetically installing OG with single-nucleotide precision in a reporter system that allowed unambiguous demonstration of OG enhancing transcription. Furthermore, we confirmed OGG1 is required to initiate BER on OG, leading to increased transcription in the VEGF PQS promoter. The VEGF PQS can adopt a G4/G5-fold with an AP that APE1 binds yet poorly cleaves to increase transcription (Fig. 5C). Because the action of APE1 on an AP results in increased transcription, the present experiments address how oxidations yielding APs, such as LSD1-generated H2O2 followed by Fenton chemistry, can change the transcriptional output of a gene without requiring OG. However, other G oxidation products from Fenton chemistry, such as the hydantoin products (24, 25), are substrates for the NEIL1 and 2 glycosylases and do not require APE1 (28, 29) and, therefore, would not impact transcription via the mechanism outlined unless additional factors are involved. Thus, further questions remain concerning how an indiscriminate oxidant like H2O2 liberated by LSD1 or LSD2 would have evolved to write OG at critical locations for gene induction. Regardless, the present studies in tandem with the earlier reports (22, 26, 27) identify a common thread in which OGG1 activity on OG in gene promoters can regulate transcription.

The sequence of events described provides an alternative mechanistic paradigm for gene activation under oxidative stress. Is this mechanism specific to the VEGF PQS or can it be generalized to other genes? Because genomic OG concentrations are elevated in tandem with some DNA repair, cell cycle, and stress response genes during inflammation (6), we investigated DNA repair genes to identify another PQS with the potential to regulate transcription similarly to the VEGF PQS. The analysis found the endonuclease III-like protein 1 (NTHL1) gene harbors a coding strand PQS with five G runs (Fig. 6A). Additionally, the NTHL1 gene can be up-regulated under oxidative stress conditions (30). Structural studies (1H-NMR, CD, and Tm) found that the five-track NTHL1 PQS adopts a parallel-stranded G-quadruplex fold in vitro (SI Appendix, Fig. S7). Plasmids were prepared with the NTHL1 PQS with OG or an F located at a reactive loop or core position in the promoter of the Rluc reporter gene. Transfection of these plasmids into glioblastoma cells followed by a 48-h incubation found >fourfold increase in gene expression for all modifications studied (Fig. 6B). This second example provides some generality to the mechanism proposed in this work.

The alternative gene activation pathway identified requires base oxidation of G to OG in the PQS context to guide OGG1 for generation of an AP. The AP unmasks the G-quadruplex from the duplex state for prolonged binding by APE1 and associated factors (Fig. 5C). The function of oxidative base modifications in DNA to direct proteins to alter transcription ascribes an epigenetic role to OG. The protein readers and erasers are members of the BER pathway and, therefore, these proteins activate genes in addition to guarding the genome against insults such as oxidative stress. Coupling of DNA repair with transcription provides an efficient mechanism to complete two necessary cellular tasks during oxidative stress. An intertwining of these pathways is starting to emerge in other studies (22, 26, 27, 31). The present study demonstrates that cells can harness oxidized modifications of DNA bases for altering phenotype under oxidative stress and identifies a mechanism by which ROS are cellular-signaling agents, as previously hypothesized (32). Last, the observation that classically defined forms of DNA damage have an epigenetic role in the cell has surfaced with the recent demonstrations of 5-hydroxymethyluracil (33) and N4-methyladenine (34) also guiding cellular processes in higher eukaryotes. These studies extend the epigenetic landscape in DNA beyond methylation of cytosine and its oxidized derivatives (35–37).

Materials and Methods

Detailed materials and methods are described in SI Appendix, Materials and Methods. Characterization of the VEGF PQS with AP and structural characterization of the NTHL1 PQS can be found in the SI Appendix. The complete methods and data are located in the SI Appendix.

Plasmid Construction. The plasmids were constructed from the psiCHECK2 plasmid (Promega) that contains genes for the Renilla luciferase (Rluc) and firefly luciferase (luc) proteins. The luc gene is regulated by the HSV-TK promoter and used as the internal standard, whereas the Rluc gene was originally regulated by the SV40 early enhancer/promoter that was modified to include the PQS of interest. Additionally, the PQSs of interest were flanked by Nt.BspQI-nickingendonuclease recognition sequences. Insertion of the PQS and nicking endonuclease recognition sequences was achieved using restriction-free cloning, followed by transformation to competent Escherichia coli and isolation by miniprep kit (Qiagen), as described previously (38). The site-specific modifications were synthesized into short oligomers with the sequence between the two nicking endonuclease sites and they were inserted into the plasmid via literature methods (38, 39). Confirmation that the DNA modifications were introduced into the plasmid was achieved using a protocol established in the C.J.B. laboratory, in which the modification was removed by Fpg and APE1 to yield a ligatable gap (40). After ligation of the gap with T4-DNA ligase, Sanger sequencing provided a characteristic nucleotide loss at the modification site to confirm its presence (SI Appendix, Fig. S8).

Cell Culture Studies. All cells were grown in Dulbecco’s Modified Eagle Medium (DMEM) supplemented with FBS, gentamicin, glutamax, and nonessential amino acids. The WT MEF and OGG1−/− MEF cells were previously developed and reported on in the literature (41), and the glioblastoma cells (U87 MG) were purchased from ATCC. Transfection experiments were conducted in white, 96-well plates using X-tremeGene HP DNA transfection agent (Roche) with 200–750 ng of plasmid following the manufacturer’s protocol. The Dual-Glo luciferase (Promega) assay was conducted following the manufacturer’s protocol to monitor Rluc and luc expression levels. Each experiment was conducted in four or eight replicates, as recommended by the Dual-Glo luciferase assay. The APE1 inhibitor III studies were conducted by titration of APE1 inhibitor III from 0 to 1,000 nM from a DMSO stock solution to the cell culture media during...
transfection. Controls in which only DMSO was added to the cells found no change in expression level of the WT plasmid, ensuring the data obtained resulted from the inhibitor and not the DMSO. The APE1 siRNA knockdown studies were conducted with FlexiTube siRNAs (Qiagen) following the manufacturer’s protocol. The siRNAs were transfected 12 h before addition of the reporter plasmid in a concentration range from 0 to 50 nM.

Data Analysis. The data were analyzed by converting the luminescence measured into normalized relative response ratios (RRR), which is the luminescence of RLuc divided by the luminescence of Luc (i.e., RRR = RLuc/Luc). To obtain the normalized expression values reported, each RRR was divided by the RRR for the WT sequence in that data set. The error bars represent 95% confidence intervals obtained from the data. Statistical analysis was achieved using a two-tailed Student’s t test.

ACKNOWLEDGMENTS. The MEF cell lines were provided by Dr. Tomas Lindahl (Imperial Cancer Research Fund, UK). We thank the University of Utah core facilities for synthesizing the OG- and F-containing oligomers, as well as conducting Sanger sequencing of the plasmids. This work was supported by a National Cancer Institute Grant R01 CA090689.