2021

Understanding the Epigenetic Role of 8-oxoguanine and OGG1 in Non-Small Cell Lung Cancer

Kyrellos Ibrahim

Understanding the Epigenetic Role of 8-oxoguanine and OGG1 in Non-Small Cell Lung Cancer

A Thesis Presented
By
Kyrellos Ibrahim

To the Keck Science Department
Of Claremont McKenna, Pitzer, and Scripps Colleges
In partial fulfillment of
The degree of Bachelor of Arts

Senior Thesis in Biochemistry
28 January 2021

W.M. Keck Science Department
Claremont McKenna College • Pitzer College • Scripps College

Abstract

Oxidative damage to the genome can form 8-oxoguanine (oxoG), a premutagenic lesion suggested to play an epigenetic role in the regulation of various cellular pathways. Alongside oxoG in this regulation is the 8-oxoguanine DNA glycosylase (OGG1), which primarily functions to repair oxoG damage via base excision repair, but is also implicated in recruiting NFκB and impacting gene expression associated with cancer growth. This proposal aims to build genome-wide maps of oxoG occupancy, and indirectly OGG1 localization, in healthy lung cells and in non-small cell lung cancer adenocarcinoma cells in order to identify regulatory regions in the genome at which oxoG is prevalent at higher rates in cancer cells relative to healthy cells. Measuring the occupancy of oxoG will be accomplished using two relatively new methods known as OG-seq and enTRAP-seq, both of which improved upon the resolution, throughput, and the scope of previous mapping techniques. The transcriptional changes resulting from increased oxoG/OGG1 occupancy will be quantified using Real Time qRT-PCR. The results of these experiments will aid in characterizing the roles of oxoG and OGG1 in regulation of transcription associated with cancer growth and can inform effective oncological therapeutic development.

A. Specific Aims

This proposal will develop genome-wide maps of oxoG and OGG1 using the established OG-seq and enTRAP-seq methods in healthy human lung cells and in cancerous human lung cells. These maps will enable detection of significant differences in localization of oxoG between the cell types and can reveal correlations between the cancer phenotype and certain genes and their associated regulatory regions. Once identified, the expression of those genes will be investigated to further understand their relationship with oxoG, OGG1, and the proliferation of cancer cells. Inhibiting OGG1 while measuring changes in gene expression will aid in confirming that any alterations in expression are in fact associated with OGG1 occupancy and not as a result of some other source.

1. **Identify differences in oxoG enrichment in both quantity and location in the genome between healthy and cancerous lung cells.**
    a. **Genome-wide mapping of oxoG using OG-seq**
    b. **Genome-wide mapping of oxoG using enTRAP-seq**
2. **Investigate if oxoG enrichment and OGG1 occupancy are associated with the changes in expression level of key genes in both healthy lung epithelial cells and lung cancer cells.**
    a. **Healthy vs cancer**
    b. **Healthy + inhibition vs cancer + inhibition**

B. Background and Significance

I. OGG1 as a DNA Glycosylase

One of the most prominent forms of DNA damage comes in the form of oxidative damage inflicted on the genetic material by reactive oxygen species (ROS). Such damage contributes to the development of a myriad of degenerative diseases as the accumulation of oxidative damage takes a toll on the genomic integrity of the organism. Cardiovascular diseases, immune deficiencies, and cancerous cell growth can be tied to the store of mutations that result from the onslaught of both endogenous ROS from metabolism and exogenous ROS from UV radiation, pollution, and other sources.[1,2] Guanine serves as an electron rich target for these oxidative species, producing the mutagenic lesion 8-oxoguanine (oxoG). In this mutagenic form, oxoG can incorrectly base pair with an adenine base during replication at an equal rate to the correct cytosine pairing (oxoG:C).[3] This incorrect adenine pairing (oxoG:A) can result in a G:C to T:A mutation after two cycles of replication if not properly repaired. Thus, an efficient repair pathway is necessary to excise the damaged base before replication turns the mutagenic lesion into an undetectable mutation.

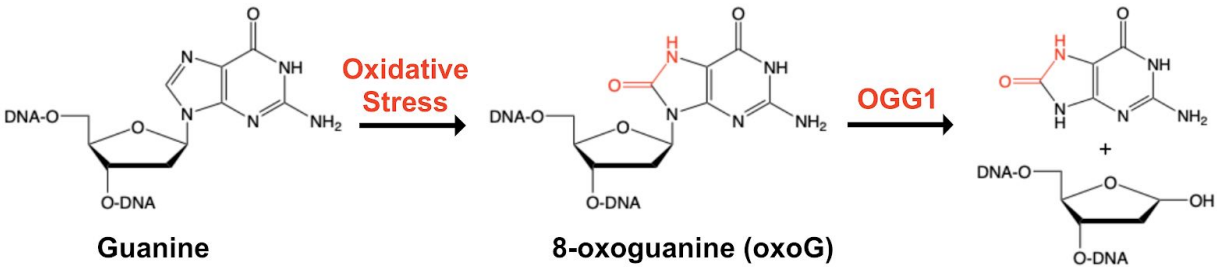

Figure 1. Oxidative stress in the form of ROS oxidize guanine to produce the mutagenic lesion 8-oxoguanine (oxoG). 8-oxoguanine DNA glycosylase (OGG1) excises the damaged base to produce an apurinic site and begin the base excision repair (BER) pathway.

The Base Excision Repair (BER) Pathway performs that precise repair action beginning with the removal of the damaged base using a DNA glycosylase, producing an apurinic site in the DNA (Figure 1) that is subsequently excised by an AP endonuclease (APE), and replaced by a DNA polymerase (DNAP) and ligase.[4–7] Specific to the oxoG lesion, 8-oxoguanine DNA glycosylase 1 (OGG1) functions as the initial repair protein that begins the BER pathway to correct this damage. OGG1 has been shown to only remove oxoG bases that are opposite a cytosine. Removing an oxoG that has incorrectly base-paired with an adenine would only guarantee the mutation will occur, as the machinery would identify an unpaired A and insert a T opposite, cementing the G:C to T:A transversion. To prevent this, an enzyme known as MutY in bacteria and MUTYH in humans removes any adenines paired with an oxoG to revert to an oxoG:C pairing that can be treated by OGG1.[4,7]

II. OGG1 as a Regulatory Element

While the pivotal role OGG1 plays in the maintenance of genomic integrity is well documented, recent studies have implicated human OGG1 (hOGG1) in a number of auxiliary pathways including the proinflammatory response pathway and innate immune responses.[8–10] Further investigations have pointed to the oxoG lesion serving as an epigenetic marker that specifically regulates gene expression driven by Nuclear Factor-κB (NFκB), a key transcription factor associated with cell viability, cytokine production, and immune response.[9,11] This makes OGG1 a promising candidate for therapeutic development in a number of disease pathologies and its potential role in cancer treatment is already being unveiled.

Nuclear Factor-κB (NFκB) can be activated by tumor necrosis factor alpha (TNFα), an inflammatory cytokine produced by immune cells during acute inflammation that enacts a myriad of signalling events within cells leading to necrosis or apoptosis.[12,13] As NFκB is implicated in the biological defense against tumorigenesis, among other diseases, its mechanistic integration with OGG1 points to OGG1 serving as a downstream target and/or an upstream initiator in innate immune pathways.[14–17] Oxidative stress induced by the proinflammatory response of TNFα and NFκB to combat infection or tumor growth causes reversible oxidation of the cysteine residues in OGG1, inhibiting the base excision function without compromising the enzyme's ability to bind to the DNA and flip the damaged base out of the helix.[18,19] This transient stalling is what enables OGG1 to recruit transcription factors such as NFκB directly to the chromatin in order to induce gene expression.[9,11,19,20] The proposed mechanism hypothesizes that oxidative stress recruits OGG1, then stalls OGG1 before repair in order to facilitate recruitment of transcription factors before OGG1 finally completes the repair (Figure 2).

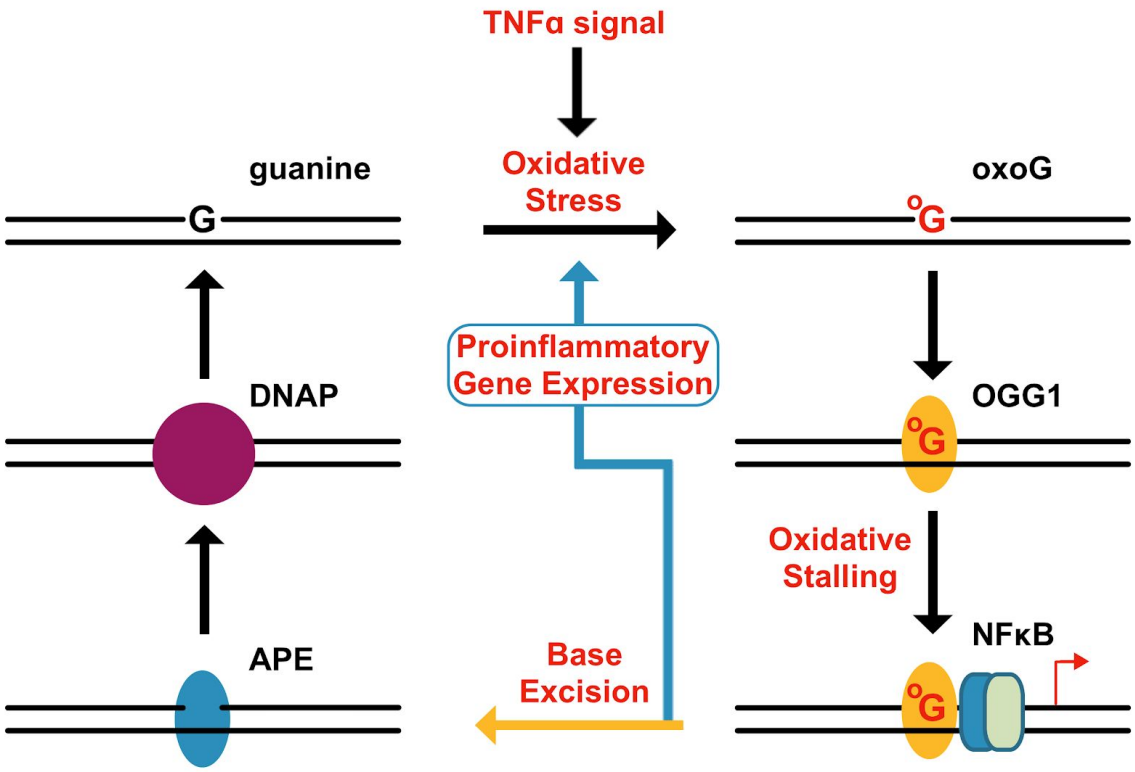

Figure 2. Oxidative stalling model presented by Vlahopoulos and colleagues.[19] Signals from TNFα induce oxidative stress and produce oxoG, initiating OGG1 occupancy at the damaged lesion. Remaining oxidative stress prevents immediate repair by temporarily oxidizing key cysteine residues in OGG1, allowing for localization of NFκB and subsequent gene expression. When the oxidation is reversed, the stalling ceases and the base excision repair (BER) pathway is completed.

III. OGG1 Inhibition in Cancer Treatment

Synthetic inhibitors of OGG1 were shown to decrease occupancy of NFκB on DNA containing oxoG in cells exposed to TNFα and the resultant oxidative stress.[21] Seeing that OGG1 may be deeply woven into cancer growth and the biological response, investigations into its potential as a therapeutic target are well warranted. Inhibition of OGG1 has potential to serve as a co-chemotherapeutic treatment in order to weaken the ability of targeted tumors to repair DNA under assault by chemotherapeutic agents. However, upregulation of OGG1 can be a potential preemptive or even early stage treatment method of enhancing the apoptotic pathway of TNFα and NFκB in hampering

tumorigenesis. Because oxoG may be serving as an epigenetic marker in cancer growth, a genomic map of these damaged guanines can serve as a guide as to whether these epigenetic markers are regulated by location: whether they are most prominently found in promoters, enhancers, or within the genes themselves. Both *in vitro* and *in vivo* studies found significantly higher levels of oxoG in open chromatin and regulatory elements, specifically in 5' untranslated regions (5'UTRs), promoter regions, and CpG islands, relative to an expected random dispersion of damage across the genome.[22–25] This is a promising indication that the enrichment of the oxidative damage is not randomly distributed, but is connected to regulatory mechanisms beyond simple damage-repair pathways. In part, this proposal aims to identify if the regulatory regions with higher incidence of oxoG lesions are connected to established proto-oncogenes and/or tumor suppressor genes. If such enrichment is uncovered, the implications of OGG1 in cancer treatment would be greatly expanded.

There are two primary methods for developing a genome-wide map of oxoG damage: OG-seq[22] and enTRAP-seq[23]. OG-seq labels the oxoG damage sites with biotin in order to enable affinity purification, replacing the previous standard of immunoprecipitation as the means of quantifying enrichment and identifying genomic location. On the other hand, enTRAP-seq uses an OGG1 with a K249Q point mutation to freeze the glycosylase at the oxoG damage without repairing it and without compromising substrate recognition. This mutation used in enTRAP-seq replicates the effect of the *in vitro* borohydride trapping reaction - which creates a covalent intermediate between oxoG and OGG1 - and the trapped complexes can also be precipitated using affinity purification.[23] Both of these methods will be valuable in identifying where the oxidation

occurs across the genome, indicating loci at which OGG1 engages with the genome and possibly recruits NFκB.

While these initial studies have expanded the possibilities of genome mapping and deepened the understanding of the epigenetic role of oxoG and OGG1, this proposal intends to dig into the specific genes impacted to determine any salient correlation between this damage-turned-signal and cancer growth. Additionally, the impact of OGG1 inhibitors on the expression of the genes of interest will be investigated as a manner of validating the reliance of the target gene expression on OGG1 occupancy, as well as the potential to take advantage of those findings in therapeutic development.

C. Research Design and Methods

The long term objective of this project is to understand if the proposed mechanistic interactions between OGG1, TNFα, and NFκB in gene regulation bears any significance in oncological therapeutic development. This proposal is distinct from the studies that established the methods that will be used for two primary reasons. First, as the broader goals of this investigation are tied to cancer, human lung cells will be used as opposed to the mouse embryonic fibroblasts in which these methods were developed. Non-cancerous cells will be from the BEAS-2B cell line (human lung epithelial cells) and the cancerous cells will be from the NCI-H1563 line (adenocarcinoma, non-small cell lung cancer) or the NCI-H1838 line (adenocarcinoma, non-small cell lung cancer). Each of these cultures will be grown under identical conditions. Second, these experiments seek to integrate these methods with an investigation into the relationship between oxoG, OGG1, cancer, and the resultant changes in gene expression. The previous studies were each isolated uses of the genome-wide mapping techniques and this

proposal intends to utilize the techniques to address the nature of the relationship between each of these elements.

The experiments proposed here will study the epigenetic behavior of the genomic 8-oxoguanine residue (oxoG) and the role of the glycosylase OGG1 as a regulatory element. As this epigenetic mechanism may be intrinsically tied to tumorigenesis and anticancer response pathways, identifying the implicated genes as proto-oncogenes or tumor suppressor genes will inform the prospects of OGG1 serving as a therapeutic target. Changes in gene expression between healthy cells, cancer cells, and cancer cells exposed to OGG1 inhibitors will further inform those prospects and uncover the relationship between TNFα, NFκB, and OGG1.

Aim 1: Identify genome-wide differences in oxoG enrichment between healthy and cancerous lung cells.

a. **Genome-wide mapping of oxoG using OG-seq**
b. **Genome-wide mapping of oxoG using enTRAP-seq**

Objective and Overview: We will establish the presence of any basal difference between the amount and location of oxoG damage in healthy and cancerous cells. This will be achieved by specifically mapping oxoG damage using both the OG-seq method and the enTRAP-seq method. Genomic localization of OGG1 is expected to match oxoG enrichment, which can be confirmed using ChIP-seq with anti-OGG1 antibodies. Any loci found to display differences in oxoG enrichment and/or OGG1 occupancy will be screened for specific enhancers, promoter regions, or genes documented as proto-oncogenic or tumor suppressors. If such differences are found, the impacted regions will be investigated during the following aim as to whether or not the gene

expression of those parts of the genome has changed at all. Any changes in gene expression can be targets for future investigations in supporting OGG1's potential as a therapeutic target.

Experimental design.

a. Genome-wide mapping of oxoG using OG-seq

First, the location and quantity of the oxoG lesion will be mapped using the OG-seq method developed by Ding and colleagues.[22] Essentially, OG-seq operates via biotinylation and affinity purification, as opposed to the antibody and immunoprecipitation (IP) utilized in previous methods. OG-seq has been developed as an alternative to those previously established methods of identifying oxoG damage, touting two main advantages: high resolution and whole-genome sequencing.

One promising attribute of OG-seq is the ability to maintain the integrity of the sequences containing oxoG at higher resolutions (approximately 0.15-kb resolution). Previous studies that used oxoG antibodies yielded genomic maps that were not resolute enough to determine precise genomic elements (ie. promoters, UTRs, etc.), producing a map at around 1000-kb resolution.[10,22,24,26,27] Additionally, the antibody method was hampered by DNA secondary structure, a challenge that OG-seq circumvents by labeling oxoG after fragmenting the genomic DNA as opposed to introducing oxoG antibodies for IP while the genomic DNA still has its secondary structure. The other primary method employed PCR primers to detect a G to T mutation that was assumed to be resultant from oxoG damage (following the G:C to T:A mutation previously described). During the replicative cycles of PCR, the oxoG damage would pair with an A, resulting in a G to T transversion as the sequence is amplified. Thus,

researchers could indicate that an oxoG lesion must have been present at that location in the sequence because of this G to T mutation. Unfortunately, this PCR method required targeting of specific sequences and was thus limited by low throughput from creating a full genome map.

The original use of the OG-seq method was in mouse embryonic fibroblasts, but this experiment will be executed in human lung epithelial cells, both non-cancerous and cancerous. OG-seq enables measuring the localization and enrichment of oxoG via the chemical labeling of the oxoG lesions using a commercially available amine-terminated biotin with a polyethylene glycol (PEG) linker (BTN). BTN is attached to oxoG through a one-electron oxidizing agent K_2IrBr_6 that is sufficiently reactive with oxoG without introducing any oxidative damage to the unaffected guanine residues (Figure 3).

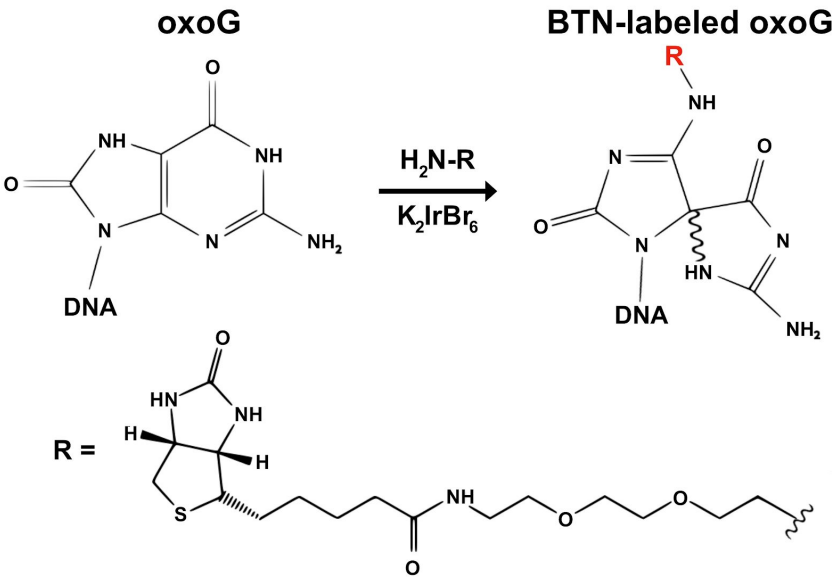

Figure 3. Covalent labeling of oxoG with BTN via selective oxidation using K_2IrBr_6. Method developed by Ding et al.[22]

BTN is introduced to the DNA after the genomic material is fragmented into segments of around 150 base pairs by sonication. The 150-mers are labeled with the amine adduct by incubating them with the oxidant K_2IrBr_6 and the BTN label.[22] Affinity purification (AP) using streptavidin (STP) coated agarose beads will extract the biotinylated strands, then the complementary strands will be released from the beads with 150 mM NaOH. The eluted strands will then be prepared with a single-stranded DNA adaptor kit before submission for Illumina Next Generation Sequencing (NGS) to determine the sequences of the enriched elements (Figure 4). We will compare the oxoG enrichment of the healthy lung epithelial BEAS-2B cells with the NCI-H1563 and the NCI-H1838 lung adenocarcinoma cells. To control for any sequence-dependent biases that may occur during the sequencing library preparation steps, an input control will be constructed by fragmenting a genome from the BEAS-2B cells without BTN labeling. Ding and colleagues produced an average depth of 3x genome coverage from their input control.

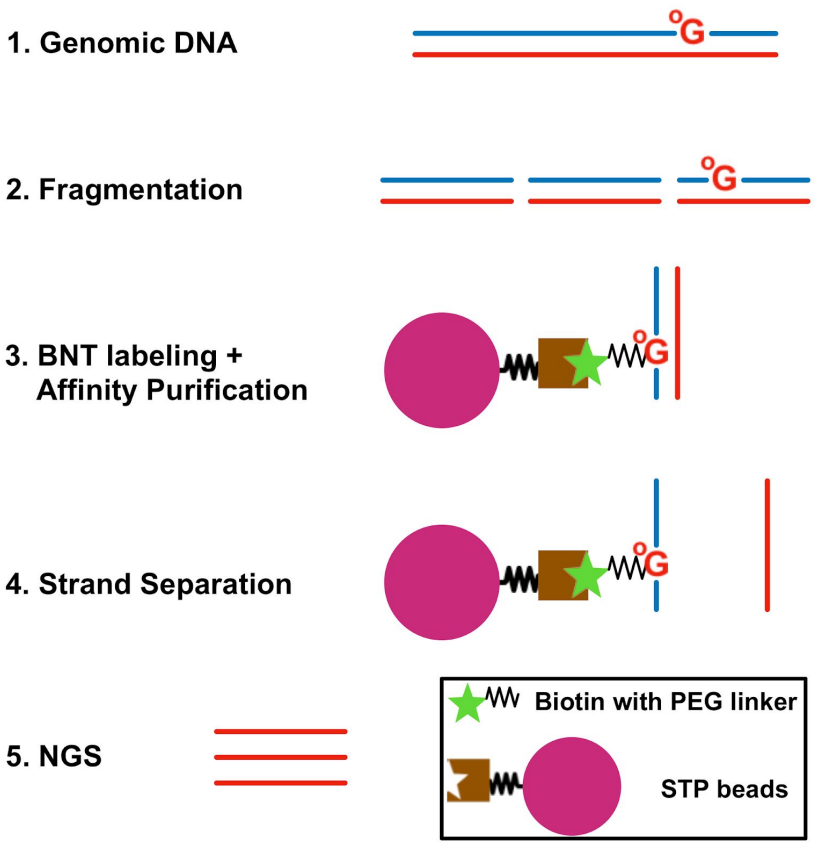

Figure 4. Enrichment of oxoG-containing strands from genomic DNA using the OG-seq method. *Adapted from Ding et al.*[22]

Enrichment will be determined relative to the results of the BEAS-2B OG-seq replicates. As certain regulatory regions are already established to have a higher incidence of oxoG, analysis will be focused on determining differences in prevalence between the healthy cells and cancer cells. This is not intended to explicitly confirm the enrichment results of previous studies, but to use the developed technology to delve into the specifics of that enrichment. Each OG-seq experiment will be performed with three replicates and enrichment cutoffs that will determine significant differences in oxoG localization will be at peaks of 3-fold enrichment and higher.[22] These experiments will yield a genome-wide map of oxoG enrichment variation between the lung cancer models and the healthy cell line.

b. Genome-wide mapping of oxoG using enTRAP-seq

The next method of oxoG mapping will be the enTRAP-seq method recently developed by Fang and Zou[23] that replicates the effects of *in vitro* borohydride trapping with an K249Q mutation in OGG1, followed by affinity purification and sequencing. When OGG1 interacts with the oxoG damage, it forms a Schiff base intermediate with the oxidized ring before excising the base. A common *in vitro* method of establishing a stable covalent attachment between the enzyme and oxoG lesion is to introduce sodium cyanoborohydride ($NaBH_3CN$) as a mild reducing agent to trap the enzyme onto the DNA strand (Figure 5). Sodium cyanoborohydride is a more optimal reducing agent than sodium borohydride ($NaBH_4$) because its capacity for reduction limits itself to the bond of interest without inducing any off-target reduction in the process. Alternatively, the enTRAP-seq method takes advantage of a mutated OGG1 protein. OGG1 Lys249 is essential for the glycosylase activity because it attacks the C1′ site of oxoG to form that

Schiff base intermediate. Mutating the Lysine residue to Glutamine (K249Q) has been reported to block the catalytic activity of the enzyme without compromising substrate recognition or specificity, meaning it also retains its ability to recognize the cytosine opposite the oxoG base.[23,28] Thus, the K249Q mutant non-covalently binds OGG1 to oxoG by rendering OGG1 catalytically inactive.

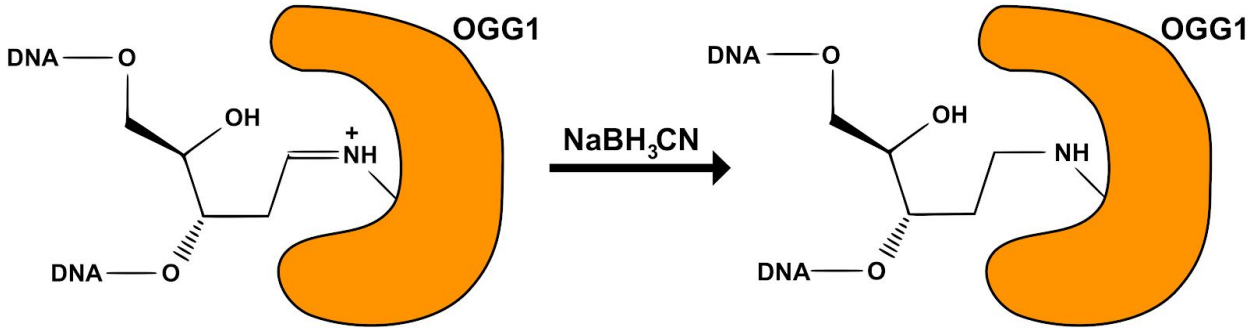

Figure 5. Borohydride trapping assay of oxoG with OGG1. Sodium cyanoborohydride ($NaBH_3CN$) selectively reduces the positively charged nitrogen of the Schiff base intermediate to form a stable covalent connection. The enTRAP-seq method mimics that connection with the K249Q point mutation, binding oxoG and OGG1 non-covalently by preventing catalysis. *Adapted from Fang et al*.[23]

The procedure for enTRAP-seq, established by Fang and colleagues[23], is relatively straightforward (Figure 6). Genomic DNA will be extracted from each cell line and fragmented using commercially available digestion methods. In order to prevent any further oxidation during the extraction and fragmentation process, desferal (deferoxamine) will be added during each step. Once fragmented, the DNA sample will be divided into a control group and an experimental group. The control group will be treated with Fpg (the bacterial functional analog of OGG1) to excise as much of the oxoG present as possible before introducing the K249Q OGG1. This will establish a baseline to which the enrichment peaks of the experimental sample can be compared. A true enrichment peak from the experimental group must be larger than that of the control group, thus preventing false positive peaks from influencing the data analysis. After the control treatment, both samples will be treated with the K249Q OGG1,

followed by affinity purification and elution of the glycosylase from the DNA samples. The resultant DNA will be prepared using Illumina's commercially available DNA Library Prep Kit and then submitted to Illumina for Next Generation Sequencing (NGS) and quantification.

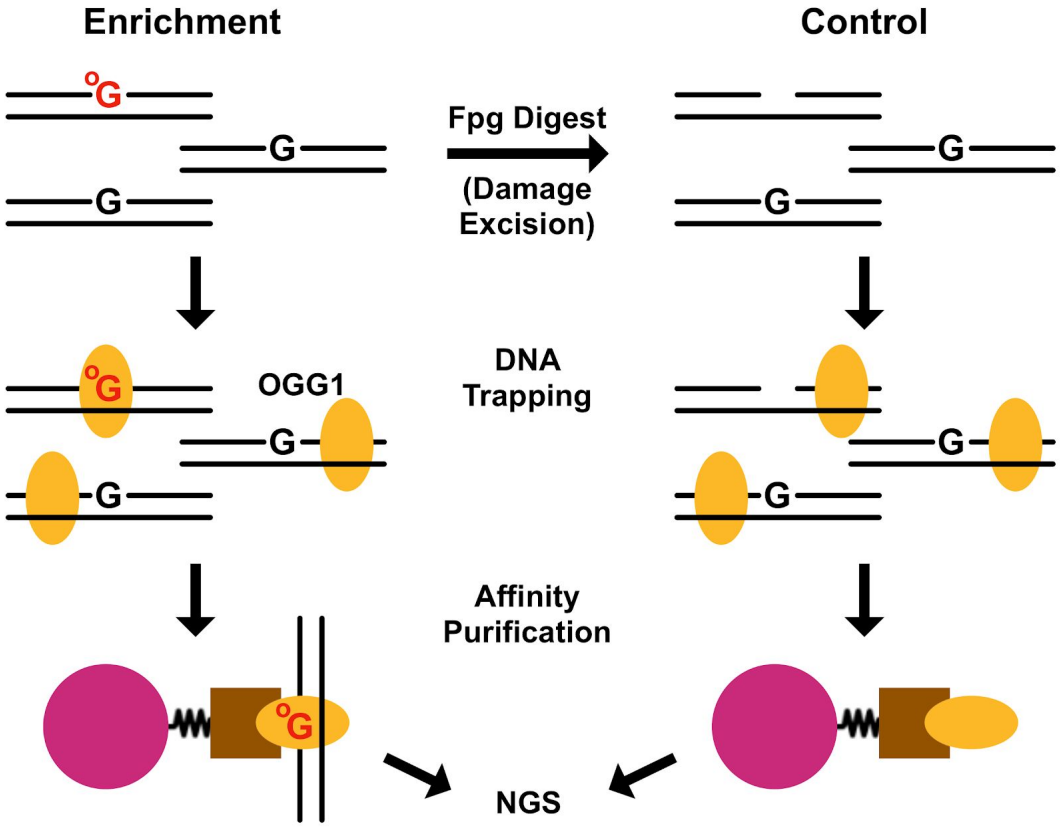

Figure 6. Scheme of the enTRAP-seq method developed by Fang et al[23] for enrichment of genomic DNA with OGG1 occupancy. NGS results from the control serve to establish the background signal against which the level of enrichment for each sequence will be measured.

This experiment will also be conducted on each cell line, contrasting the oxoG enrichment of the healthy lung epithelial BEAS-2B cells with the NCI-H1563 and the NCI-H1838 lung adenocarcinoma cells.

Auxiliary Approaches. In the event that no significant differences are found between the healthy and cancerous cell lines, then specific sequences can be targeted in place of the genome-wide mapping approach. These new target sequences would be

documented regulatory regions of known proto-oncogenes and tumor suppressors. As this approach has a more narrow focus and higher resolution than OG-seq and enTRAP-seq, it would yield promising results, though limited in scope, as the sequencing would be focused on a chosen set of genes instead of the full genome. The higher resolution is due to the nature of the PCR sequencing method previously described, as the location of the oxidative damage can be more accurately identified via this sequencing method. As lung cells are the model of this proposal, genes relevant to lung adenocarcinoma will be studied. For example, mutations in the TP53, EGFR, and KRAS genes are common in lung cancers and can provide the target genes, promoters, and UTRs for this auxiliary approach.[19,29,30]

Expected Results. The OG-seq experiments are expected to provide more than 90 million reads that will cover the human genome almost three times. Enrichment is expected to be found in promoter regions and UTRs, consistent with the findings of previous studies.[10,22,23,31] For the purposes of data analysis, promoter regions will be defined as 500 bp upstream and 100 bp downstream of the transcription start site and UTRs will be defined as beyond that interval.[23] The enTRAP-seq experiments are expected to yield on the order of 10,000 significant enrichment peaks, likely in the same promoter regions and UTRs found in the OG-seq experiments. OGG1 occupancy is expected to be nearly identical to the oxoG enrichment detected by both OG-seq and enTRAP-seq, but if we wish to verify this, we can perform a standard ChIP-seq assay using anti-OGG1 antibodies followed by NGS to confirm that assumed overlap.

Differences in quantity and location of oxoG between healthy cells and cancerous cells in the results of both experiments would point to an epigenetic and regulatory distinction between the cell types. The genes found to be associated with those differences will be of particular interest in the analysis of transcription that will be pursued in the second aim of this proposal.

Aim 2: Investigate if oxoG enrichment and OGG1 occupancy are associated with the changes in expression level of key genes in both healthy lung epithelial cells and lung cancer cells.

a. **Healthy vs cancer**
b. **Healthy + inhibition vs cancer + inhibition**

Objective and Overview. This aim is intended to determine if the transcriptional behavior of new target genes identified in Aim 1 also differ between healthy lung epithelial cells and lung cancer cells. In order to determine if the changes in gene expression that may be found are actually dependent on OGG1, the same transcriptional profiling will be performed with each cell line exposed to a series of OGG1 inhibitors. The inhibitors that will be employed are small molecules that have already been confirmed to be effective inhibitors specific to human OGG1 (hOGG1) and serve to replicate an OGG1 knockout strain, as there are not OGG1-deficient lung cells currently available. Future experiments could develop an OGG1 knockout line using the CRISPR-Cas9 system, but that has not yet been done. This proposal will use the following inhibitors: TH5487 (IC_{50} = 342 nM)[21], SU0268 (IC_{50} = 59 nM)[32], and D0167 (IC_{50} = 330 nM)[33] (Figure 7). Performing this experiment with three different inhibitors in parallel will help mitigate confounding effects from potential off-target interactions and account for distinct modes

of inhibition. The comparison of gene expression will be carried out using Real Time quantitative Reverse Transcription PCR (qRT-PCR) to measure the quantity of mRNA transcripts for each of the target genes in each cell line. Analysis will be carried out across a 3-by-4 matrix of experiments, comparing the healthy cells and cancer cells both with and without inhibitors (Table 1).

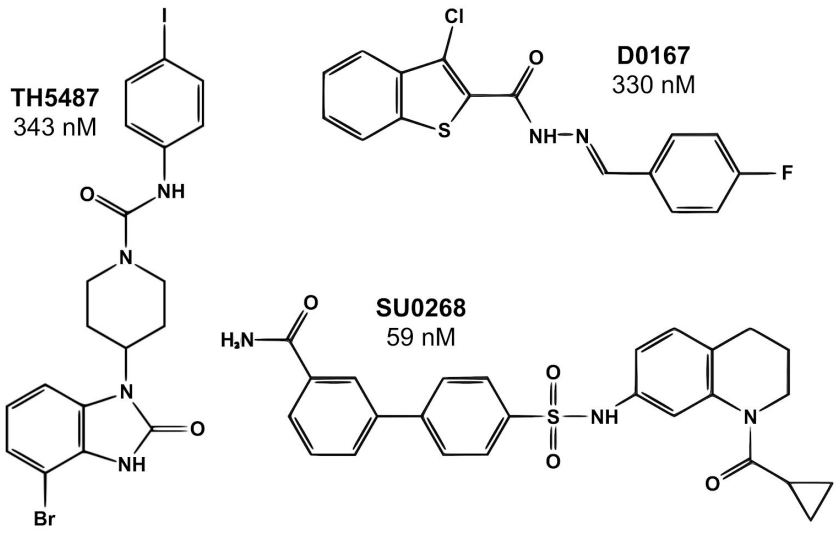

Figure 7. Structure of each inhibitor with the associated IC_{50} value.

Table 1. Comparison Matrix of Each Cell Line with Each OGG1 Inhibitor

BEAS-2B (healthy)	NCI-H1563 (cancer)	NCI-H1838 (cancer)
BEAS-2B **+ TH5487**	NCI-H1563 **+ TH5487**	NCI-H1838 **+ TH5487**
BEAS-2B **+ SU0268**	NCI-H1563 **+ SU0268**	NCI-H1838 **+ SU0268**
BEAS-2B **+ D0167**	NCI-H1563 **+ D0167**	NCI-H1838 **+ D0167**

Experimental design. Quantification of the mRNA transcripts in each of the experimental scenarios described in Table 1 will be accomplished using Real Time qRT-PCR. Real Time qRT-PCR combines the technologies of qPCR and reverse transcription with fluorescent probes that deliver quantitative data in real time during the replication process of PCR.[34,35] Because the starting material of this reaction will be mRNA transcripts after cell lysis and collection, reverse transcription is necessary to develop cDNA of the transcripts that can be amplified by PCR. After the cDNA creation, qPCR is

performed, amplifying and quantifying the amount of each transcript in the sample (Figure 8A). Running one-step qRT-PCR (performing reverse transcription and the Real Time qPCR at the same time) is the preferred method over the two-step method of reverse transcription and qPCR occurring separately, as it eliminates issues of experimental variation from the extra step.[34,36] However, because mRNA can be rapidly degraded if handled poorly, there will be three independent replicates of each cell culture from Table 1 as well as three replicates of each qRT-PCR reaction for each individual replicate. This series of replicates should properly safeguard against experimental errors as well as inconsistencies in the reaction efficiency of both the reverse transcription and qPCR.

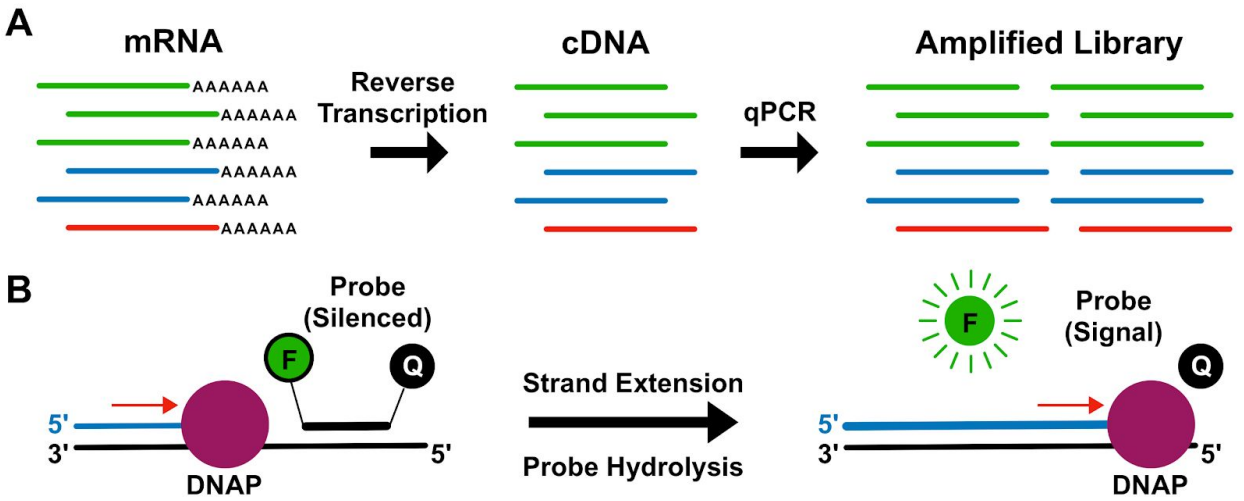

Figure 8. Real Time qRT-PCR. **A)** Conversion of mRNA extracted from the cells into cDNA via reverse transcription and quantification of the library via PCR amplification. **B)** Mechanism of the hydrolysis probe used for real time quantification. Hydrolysis occurs, releasing the fluorophore, when the DNA polymerase passes over the probe during strand extension.

Real Time qRT-PCR uses the standard components of RT-PCR - dNTPs, primers specific to the target transcripts, a reverse transcriptase, and the Taq DNA polymerase (Taq DNAP) - and a fluorescent probe. The fluorescent probe is the integral element of the real time data collection, pioneered as the TaqMan assay.[35] The fluorescent probe,

known as a hydrolysis probe, is a small oligonucleotide (~4 bases long) with a fluorophore at the 5' end and a quencher at the 3' end (Figure 8B). With such close proximity to the quencher, the fluorophore does not emit a signal when excited by light energy input. During the PCR cycles, this probe anneals to the target sequence downstream of the primer. As the Taq DNAP constructs the new strand and passes over the fluorophore, it releases the fluorophore, enabling a signal to be detected. This method is known as Fluorescence Resonance Energy Transfer (FRET) and is reviewed by Jares-Erijman and colleagues.[37] Because more than one transcript will be quantified in a single reaction, multiple probes with distinct sequences and distinct fluorophores will be used, depending on the target transcripts.

Analysis of the data focuses on the curve developed by the increasing fluorescent signal. These curves are divided into four phases: the linear ground phase, early exponential phase, exponential phase, and plateau phase. Baseline fluorescence is calculated during the linear ground phase as fluorescence has not yet risen above the background signal. During the early exponential phase, the signal reaches a significantly higher level (~10 times the standard deviation) of the background signal. The cycle at which this occurs is known as C_t or CP. That value is indicative of the starting number of transcripts in that sample.[34,38] Once the CP value is calculated for each transcript across all replicates and experimental conditions, comparisons will be made between each series of conditions using an ANOVA for the necessary matrix of comparisons to establish significant differences in the expression of the target genes.

Auxiliary Approaches. If the Real Time qRT-PCR does not yield results indicating significant differences in OGG1-mediated gene expression between the healthy and

cancer cell lines, another option for quantification is RNA-seq. RNA-seq would provide analysis for the entire transcriptome instead of targeting specific sequences. This method is a larger undertaking than qRT-PCR and would require much more extensive data analysis to glean any conclusions from the data. RNA-seq operates by extracting the complete set of mRNA from the cells, reverse transcribing the mRNA into a cDNA library, and fragmenting the library using sonication or other forms of digestion. The cDNA fragments are then fitted with sequencing adaptors, amplified by PCR, and submitted for Illumina NGS. All the sequenced reads are then mapped onto the human reference genome and the number of repeat matches on any one gene are quantified to determine the level of gene expression.[39,40] The technology for the next generation RNA sequencing, quantification, and the subsequent analysis is available through Illumina.

Expected Results. Considering the data from the Real Time qRT-PCR experiment, differential expressed transcripts will be defined as transcripts that have 3-fold or greater change from the point of comparison. The healthy uninhibited cells will be treated as a general indicator for the baseline, but as the data will be analysed using an ANOVA, each pairing may lead to interesting conclusions. The primary focus will be comparing the results of the healthy uninhibited cells to both cancerous uninhibited cell lines. If there is a reported difference in the gene expression of the experiment's target genes between the uninhibited cultures, this would indicate a series of notions. As the target genes would have been confirmed to be subject to regulation by OGG1 occupancy in Aim 1, significant results would support the hypothesis that the gene expression levels that are unique to the cancer phenotype are in fact associated with changes in OGG1 occupancy - OGG1 is present at the regulatory regions of those

genes when the changes in gene expression are induced. In order to further confirm this association, the inhibited cultures would necessarily yield opposite results: by removing OGG1, the changes in gene expression seen in the uninhibited cultures would be non-existent or even reversed. It should be noted that this expectation is reliant on the inhibitors interfering with OGG1 binding, not only with catalysis, which is in part why all three inhibitors are being used. A lack of this reversion in gene expression could indicate that either there is no correlation between the gene expression of the target genes and OGG1 occupancy or that there is a biological redundancy that compensates for the absence of OGG1.

D. Future Directions

Characterizing how oxoG and OGG1 are ingrained in the regulatory web of gene expression strengthens the current understanding of their epigenetic role and their impact on disease development. Identifying the role oxoG and OGG1 play in the regulation of promoters, transcription factors, and mediators associated with cancer enables therapeutic development to take advantage of this new insight in order to create more effective and more specific treatments. OGG1 has already been considered as a target for the development of co-chemotherapeutic treatments and the results of this proposal would inform the approach in crafting selective and effective interventions.

This proposal intends to identify specific genes whose expression are impacted by the epigenetic marker of oxoG and the resultant localization of OGG1 and to determine the role those genes play in cancer cells. As was noted, signaling elements and

transcription factors such as TNFα and NFκB, respectively, are thought to be mechanistically intertwined with OGG1. This relationship is a promising avenue for investigation, specifically in whether the oxidative-stress-recruitment model (Figure 2) holds merit. To determine a possible correlation with TNFα dependency and NFκB recruitment via OGG1, one could remove TNFα transgenically and expose the cells or organism to oxidative stress and measure with both OG-seq and enTRAP-seq how oxoG enrichment, and indirectly OGG1 localization, might be affected by those changes. The nuclear localization of NFκB could be measured simultaneously to detect any effects of the TNFα depletion farther downstream. Beyond simply observing NFκB localization, studies could be conducted into the expression of those pro-inflammatory genes known to rely on NFκB for transcriptional activation.

Along a different vein, the genes identified in this proposal can be studied more extensively. If any of those genes had not previously been associated with the cancer phenotype, there are myriads of opportunities to investigate how the altered expression of those genes contributes to tumorigenesis. The transcripts and peptide products of those genes could additionally serve as new therapeutic targets. The results of this proposal can also be replicated in cell types other than human lung tissue. Using other cell types could yield transcriptional changes unique to the type of cancer being studied, which would further expand the opportunity for more targeted and effective therapeutics. This proposal could open multiple new avenues of investigation into the mechanisms and associated players of cancer development as well as into the translational field in a manner that advances oncological research and treatments.

Acknowledgements

I would like to thank each of my thesis advisors for their guidance and unending patience throughout this process. I would like to thank Dr. Karl Haushalter, my research advisor at Harvey Mudd College, for engaging me as a young researcher and supporting me in producing this work from the ground up. I would also like to thank Dr. Kyle Jay and Dr. Ethan Van Arnam, my Keck Science faculty advisors, for stepping up during Dr. Haushalter's leave of absence and encouraging me through the challenges of this thesis. Additionally, I would like to thank the Keck Science Department and the Harvey Mudd Biology and Chemistry Departments for the opportunity to research and study. Lastly, I would like to thank my whole family, my friends, and my roommate for their continuing support throughout this project.

References

(1) Ames, B. N.; Shigenaga, M. K.; Hagen, T. M. Oxidants, Antioxidants, and the Degenerative Diseases of Aging. *Proc. Natl. Acad. Sci.* **1993**, *90* (17), 7915–7922. https://doi.org/10.1073/pnas.90.17.7915.

(2) Rosenquist, T. A.; Zharkov, D. O.; Grollman, A. P. Cloning and Characterization of a Mammalian 8-Oxoguanine DNA Glycosylase. *Proc. Natl. Acad. Sci.* **1997**, *94* (14), 7429–7434. https://doi.org/10.1073/pnas.94.14.7429.

(3) Grollmanand, A. P.; Moriya, M. Mutagenesis by 8-Oxoguanine: An Enemy Within. *Trends Genet.* **1993**, *9* (7), 246–249. https://doi.org/10.1016/0168-9525(93)90089-Z.

(4) Krokan, H. E.; Bjoras, M. Base Excision Repair. *Cold Spring Harb. Perspect. Biol.* **2013**, *5* (4), a012583–a012583. https://doi.org/10.1101/cshperspect.a012583.

(5) Lindahl, T. Quality Control by DNA Repair. *Science* **1999**, *286* (5446), 1897–1905. https://doi.org/10.1126/science.286.5446.1897.

(6) David, S. S.; Williams, S. D. Chemistry of Glycosylases and Endonucleases Involved in Base-Excision Repair. *Chem. Rev.* **1998**, *98* (3), 1221–1262. https://doi.org/10.1021/cr980321h.

(7) David, S. S.; O'Shea, V. L.; Kundu, S. Base-Excision Repair of Oxidative DNA Damage. *Nature* **2007**, *447* (7147), 941–950. https://doi.org/10.1038/nature05978.

(8) Aguilera-Aguirre, L.; Hosoki, K.; Bacsi, A.; Radák, Z.; Wood, T. G.; Widen, S. G.; Sur, S.; Ameredes, B. T.; Saavedra-Molina, A.; Brasier, A. R.; Ba, X.; Boldogh, I. Whole Transcriptome Analysis Reveals an 8-Oxoguanine DNA Glycosylase-1-Driven DNA Repair-Dependent Gene Expression Linked to Essential Biological Processes. *Free Radic. Biol. Med.* **2015**, *81*, 107–118. https://doi.org/10.1016/j.freeradbiomed.2015.01.004.

(9) Pan, L.; Zhu, B.; Hao, W.; Zeng, X.; Vlahopoulos, S. A.; Hazra, T. K.; Hegde, M. L.; Radak, Z.; Bacsi, A.; Brasier, A. R.; Ba, X.; Boldogh, I. Oxidized Guanine Base Lesions Function in 8-Oxoguanine DNA Glycosylase-1-Mediated Epigenetic Regulation of Nuclear Factor KB-Driven Gene Expression. *J. Biol. Chem.* **2016**, *291* (49), 25553–25566. https://doi.org/10.1074/jbc.M116.751453.

(10) Fleming, A. M.; Ding, Y.; Burrows, C. J. Oxidative DNA Damage Is Epigenetic by Regulating Gene Transcription via Base Excision Repair. *Proc. Natl. Acad. Sci.* **2017**, *114* (10), 2604–2609. https://doi.org/10.1073/pnas.1619809114.

(11) Pan, L.; Hao, W.; Zheng, X.; Zeng, X.; Ahmed Abbasi, A.; Boldogh, I.; Ba, X. OGG1-DNA Interactions Facilitate NF-KB Binding to DNA Targets. *Sci. Rep.* **2017**, *7* (1), 43297. https://doi.org/10.1038/srep43297.

(12) Idriss, H. T.; Naismith, J. H. TNFα and the TNF Receptor Superfamily: Structure-Function Relationship(s). *Microsc. Res. Tech.* **2000**, *50* (3), 184–195.

https://doi.org/10.1002/1097-0029(20000801)50:3<184::AID-JEMT2>3.0.CO;2-H.

(13) Locksley, R.; Killeen, N.; Michael, L. The TNF and TNF Receptor Superfamilies: Integrating Mammalian Biology. *Cell* **2001**, *104*, 487–501.

(14) Ghosh, G.; Wang, V. Y.-F.; Huang, D.-B.; Fusco, A. NF-KB Regulation: Lessons from Structures: NF-KB Regulation: Lessons from Structures. *Immunol. Rev.* **2012**, *246* (1), 36–58. https://doi.org/10.1111/j.1600-065X.2012.01097.x.

(15) Liu, T.; Zhang, L.; Joo, D.; Sun, S.-C. NF-KB Signaling in Inflammation. *Signal Transduct. Target. Ther.* **2017**, *2* (1), 17023. https://doi.org/10.1038/sigtrans.2017.23.

(16) Rayburn, E. R.; Ezell, S. J.; Zhang, R. Anti-Inflammatory Agents for Cancer Therapy. *Mol. Cell. Pharmacol.* **2009**, *1* (1), 29–43. https://doi.org/10.4255/mcpharmacol.09.05.

(17) Xia, L.; Tan, S.; Zhou, Y.; Lin, J.; Wang, H.; Oyang, L.; Tian, Y.; Liu, L.; Su, M.; Wang, H.; Cao, D.; Liao, Q. Role of the NFκB-Signaling Pathway in Cancer. *OncoTargets Ther.* **2018**, *11*, 2063–2073. https://doi.org/10.2147/OTT.S161109.

(18) Choudhary, S.; Boldogh, I.; Brasier, A. R. Inside-Out Signaling Pathways from Nuclear Reactive Oxygen Species Control Pulmonary Innate Immunity. *J. Innate Immun.* **2016**, *8* (2), 143–155. https://doi.org/10.1159/000442254.

(19) Vlahopoulos, S.; Adamaki, M.; Khoury, N.; Zoumpourlis, V.; Boldogh, I. Roles of DNA Repair Enzyme OGG1 in Innate Immunity and Its Significance for Lung Cancer. *Pharmacol. Ther.* **2019**, *194*, 59–72. https://doi.org/10.1016/j.pharmthera.2018.09.004.

(20) Hao, W.; Wang, J.; Zhang, Y.; Wang, C.; Xia, L.; Zhang, W.; Zafar, M.; Kang, J.; Wang, R.; Ali Bohio, A.; Pan, L.; Zeng, X.; Wei, M.; Boldogh, I.; Ba, X. Enzymatically Inactive OGG1 Binds to DNA and Steers Base Excision Repair toward Gene Transcription. *FASEB J.* **2020**, *34* (6), 7427–7441. https://doi.org/10.1096/fj.201902243R.

(21) Visnes, T.; Cázares-Körner, A.; Hao, W.; Wallner, O.; Masuyer, G.; Loseva, O.; Mortusewicz, O.; Wiita, E.; Sarno, A.; Manoilov, A.; Astorga-Wells, J.; Jemth, A.-S.; Pan, L.; Sanjiv, K.; Karsten, S.; Gokturk, C.; Grube, M.; Homan, E. J.; Hanna, B. M. F.; Paulin, C. B. J.; Pham, T.; Rasti, A.; Berglund, U. W.; Nicolai, C. von; Benitez-Buelga, C.; Koolmeister, T.; Ivanic, D.; Iliev, P.; Scobie, M.; Krokan, H. E.; Baranczewski, P.; Artursson, P.; Altun, M.; Jensen, A. J.; Kalderén, C.; Ba, X.; Zubarev, R. A.; Stenmark, P.; Boldogh, I.; Helleday, T. Small-Molecule Inhibitor of OGG1 Suppresses Proinflammatory Gene Expression and Inflammation. *Science* **2018**, *362* (6416), 834–839. https://doi.org/10.1126/science.aar8048.

(22) Ding, Y.; Fleming, A. M.; Burrows, C. J. Sequencing the Mouse Genome for the Oxidatively Modified Base 8-Oxo-7,8-Dihydroguanine by OG-Seq. *J. Am. Chem. Soc.* **2017**, *139* (7), 2569–2572. https://doi.org/10.1021/jacs.6b12604.

(23) Fang, Y.; Zou, P. Genome-Wide Mapping of Oxidative DNA Damage via Engineering of 8-Oxoguanine DNA Glycosylase. *Biochemistry* **2020**, *59* (1), 85–89. https://doi.org/10.1021/acs.biochem.9b00782.

(24) Yoshihara, M.; Jiang, L.; Akatsuka, S.; Suyama, M.; Toyokuni, S. Genome-Wide Profiling of 8-Oxoguanine Reveals Its Association with Spatial Positioning in Nucleus. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* **2014**, *21* (6), 603–612. https://doi.org/10.1093/dnares/dsu023.

(25) Gorini, F.; Scala, G.; Di Palo, G.; Dellino, G. I.; Cocozza, S.; Pelicci, P. G.; Lania, L.; Majello, B.; Amente, S. The Genomic Landscape of 8-OxodG Reveals Enrichment at Specific Inherently Fragile Promoters. *Nucleic Acids Res.* **2020**, *48* (8), 4309–4324. https://doi.org/10.1093/nar/gkaa175.

(26) Ohno, M. A Genome-Wide Distribution of 8-Oxoguanine Correlates with the Preferred Regions for Recombination and Single Nucleotide Polymorphism in the Human Genome. *Genome Res.* **2006**, *16* (5), 567–575. https://doi.org/10.1101/gr.4769606.

(27) Pastukh, V.; Roberts, J. T.; Clark, D. W.; Bardwell, G. C.; Patel, M.; Al-Mehdi, A.-B.; Borchert, G. M.; Gillespie, M. N. An Oxidative DNA "Damage" and Repair Mechanism Localized in the VEGF Promoter Is Important for Hypoxia-Induced VEGF MRNA Expression. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **2015**, *309* (11), L1367–L1375. https://doi.org/10.1152/ajplung.00236.2015.

(28) Bruner, S. D.; Norman, D. P. G.; Verdine, G. L. Structural Basis for Recognition and Repair of the Endogenous Mutagen 8-Oxoguanine in DNA. *Nature* **2000**, *403*, 859–866.

(29) Petitjean, A.; Achatz, M. I. W.; Borresen-Dale, A. L.; Hainaut, P.; Olivier, M. TP53 Mutations in Human Cancers: Functional Selection and Impact on Cancer Prognosis and Outcomes. *Oncogene* **2007**, *26* (15), 2157–2165. https://doi.org/10.1038/sj.onc.1210302.

(30) Kosaka, T.; Yatabe, Y.; Endoh, H.; Kuwano, H.; Takahashi, T.; Mitsudomi, T. Mutations of the Epidermal Growth Factor Receptor Gene in Lung Cancer: Biological and Clinical Implications. *Cancer Res.* **2004**, *64*, 8919–8923.

(31) Fleming, A. M.; Zhu, J.; Jara-Espejo, M.; Burrows, C. J. Cruciform DNA Sequences in Gene Promoters Can Impact Transcription upon Oxidative Modification of 2′-Deoxyguanosine. *Biochemistry* **2020**, *59* (28), 2616–2626. https://doi.org/10.1021/acs.biochem.0c00387.

(32) Tahara, Y.; Auld, D.; Ji, D.; Beharry, A. A.; Kietrys, A. M.; Wilson, D. L.; Jimenez, M.; King, D.; Nguyen, Z.; Kool, E. T. Potent and Selective Inhibitors of 8-Oxoguanine DNA Glycosylase. *J. Am. Chem. Soc.* **2018**, *140* (6), 2105–2114. https://doi.org/10.1021/jacs.7b09316.

(33) Donley, N.; Jaruga, P.; Coskun, E.; Dizdaroglu, M.; McCullough, A. K.; Lloyd, R. S. Small Molecule Inhibitors of 8-Oxoguanine DNA Glycosylase-1 (OGG1). *ACS Chem. Biol.* **2015**, *10* (10), 2334–2343. https://doi.org/10.1021/acschembio.5b00452.

(34) Wong, M. L.; Medrano, J. F. Real-Time PCR for MRNA Quantitation. *BioTechniques* **2005**, *39* (1), 75–85. https://doi.org/10.2144/05391RV01.

(35) Heid, C. A.; Stevens, J.; Livak, K. J.; Williams, P. M. Real Time Quantitative PCR. *Genome Res.* **1996**, *6*, 986–994.

(36) Battaglia, M.; Pedrazzoli, P.; Palermo, B.; Lanza, A.; Bertolini, F.; Gibelli, N.; Da Prada, G.; Zambelli, A.; Perotti, C.; Robustelli della Cuna, G. Epithelial Tumour Cell Detection and the Unsolved Problems of Nested RT-PCR: A New Sensitive One Step Method without False Positive Results. *Bone Marrow Transplant.* **1998**, *22* (7), 693–698. https://doi.org/10.1038/sj.bmt.1701405.

(37) Jares-Erijman, E. A.; Jovin, T. M. FRET Imaging. *Nat. Biotechnol.* **2003**, *21* (11), 1387–1395. https://doi.org/10.1038/nbt896.

(38) Tichopad, A. Standardized Determination of Real-Time PCR Efficiency from a Single Reaction Set-Up. *Nucleic Acids Res.* **2003**, *31* (20), 122e–1122. https://doi.org/10.1093/nar/gng122.

(39) Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A Revolutionary Tool for Transcriptomics. *Nat. Rev. Genet.* **2009**, *10* (1), 57–63. https://doi.org/10.1038/nrg2484.

(40) Aguilera-Aguirre, L.; Hosoki, K.; Bacsi, A.; Radák, Z.; Wood, T. G.; Widen, S. G.; Sur, S.; Ameredes, B. T.; Saavedra-Molina, A.; Brasier, A. R.; Ba, X.; Boldogh, I. Whole Transcriptome Analysis Reveals an 8-Oxoguanine DNA Glycosylase-1-Driven DNA Repair-Dependent Gene Expression Linked to Essential Biological Processes. *Free Radic. Biol. Med.* **2015**, *81*, 107–118. https://doi.org/10.1016/j.freeradbiomed.2015.01.004.