Genome-Wide Mapping of Oxidative DNA Damage via Engineering of 8-Oxoguanine DNA Glycosylase

Yuxin Fang† and Peng Zou⁎†,‡,§

†College of Chemistry and Molecular Engineering, Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University, Beijing 100871, China
‡Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China
§PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China

Supporting Information

ABSTRACT: The occurrence of 8-oxo-7,8-dihydroguanine (OG) in the genome, as one of the major DNA oxidative damages, has been implicated in an array of biological processes, ranging from mutagenesis to transcriptional regulation. Genome-wide mapping of oxidative damages could shed light on the underlying cellular mechanism. In the present study, we engineered the hOGG1 enzyme, a primary 8-oxoguanine DNA glycosylase, into a guanine oxidation-profiling tool. Our method, called enTRAP-seq, successfully identified more than 1400 guanine oxidation sites in the mouse embryonic fibroblast genome. These OG peaks were enriched in open chromatin regions and regulatory elements, including promoters, S′ untranslated regions, and CpG islands. Collectively, we present a simple and generalizable approach for the genome-wide profiling of DNA damages with high sensitivity and specificity.

In living organisms, genomic DNA is constantly threatened by both endogenous and exogenous reactive oxygen species (ROS). Among the four DNA nucleobases, guanine is the most susceptible to oxidation due to its low redox potential, resulting in various guanine oxidative damages. Guanine oxidation is linked to mutagenesis and many diseases, including tumorigenesis, accelerated aging, and malfunction of the central nervous system. More recently, guanine oxidation has been shown to function as an epigenetic mark in the regulation of gene expression. For example, an increased level of 8-oxo-7,8-dihydroguanine (OG), a major form of guanine oxidation, was observed in the promoter regions of hypoxia-modulated genes, leading to the recruitment of repair enzymes and transcription factors. Similar models have also been proposed in other cases, including estrogen-modulated transcription and inflammatory response. Together, these studies have shed light on the regulatory role of guanine oxidation, revealing that the occurrence of oxidative damages in the genome may be context-dependent.

There is a growing interest in the genome-wide profiling of oxidative damages. Immunoprecipitation (IP) with OG-specific antibodies, followed by DNA microarray analysis or high-throughput sequencing, has enabled genome-wide mapping of OG sites in the human and rodent genomes. Since the binding affinity of antibodies was often impaired by the presence of DNA secondary structures, a chemical profiling approach was developed, where OG was oxidatively conjugated with a biotinylated amine probe, enriched via affinity purification, and sequenced (OG-Seq). Alternatively, oxidized DNA fragments could be enriched by a locator code inserted through a combination of in vitro base excision repair (BER) and click reaction (click-code-seq). While all of these techniques offer valuable information on the genomic distribution of DNA oxidative damages, some of their conclusions contradict with each other. For example, OG-Seq and OxiDIP-Seq (IP with a polyclonal antibody) studies revealed enrichment of OG in open chromatin regions, whereas click-code-seq and another IP study with a monoclonal antibody concluded that OG occurred more frequently within the heterochromatin, such as lamina-associated domains. Contradictions may arise from differences in the specificity of reagents, potential bias introduced in sequencing, and cell types under investigation (e.g., yeast in click-code-seq vs mammalian cells in OG-Seq and OxiDIP-Seq).

In this study, we present a novel tool for the genome-wide profiling of DNA oxidative damages. To achieve high sensitivity and accuracy in target identification, we employed DNA repair enzymes, which have evolved to recognize DNA damages with both high affinity and high specificity. We showed that by introducing a K249Q mutation to the substrate-binding pocket of DNA glycosylase hOGG1 (UniProt ID: O15527), the enzyme-DNA substrate intermediate could be trapped as a stable complex and enriched via affinity purification. We named this strategy as enTRAP-seq, for enzyme-mediated trapping and affinity precipitation of damaged DNA and sequencing (enTRAP-seq). We applied enTRAP-seq to analyze the guanine oxidation in the mouse embryonic fibroblast (MEF) genome, which revealed enrichment of OG in regulatory elements such as promoters, S′UTR, and CpG islands.

Special Issue: Future of Biochemistry: The Asia-Pacific Issue

Received: August 29, 2019
Revised: October 14, 2019
Published: October 16, 2019

DOI: 10.1021/acs.biochem.9b00782
Biochemistry 2020, 59, 85–89
The MutM (also known as Fpg) and the corresponding human enzyme OGG1 are primary glycosylases to remove OG. Both Fpg and OGG1 are bifunctional glycosylases, which catalyze not only the excision of the damaged base (glycosylase activity) to yield an apurinic/abasic (AP) site but also the subsequent cleavage of DNA strand (AP lyase activity). Repairing of OG undergoes a mechanism of covalent catalysis, in which a Schiff base is formed to displace the cleaved base and to promote the decomposition of the sugar moiety. Importantly, the Schiff base intermediate can be reduced by borohydride to form a stable adduct

We next applied this borohydride trapping strategy to enrich a 250-bp model DNA sequence containing a single OG (0.1% recovery yield vs 0.01% from control sequence) with an efficiency comparable to K249Q (9.3 ± 0.7% recovery yield vs 0.07 ± 0.01% from control sequence), while the efficiency of K249R (0.11 ± 0.01% recovery yield vs 0.01 ± 0.001% from control sequence) was approximately 80-fold weaker than that of K249Q (Figure S3).

The K249Q mutation has been reported to block the catalytic activity of the enzyme while retaining its substrate recognition, a feature that has assisted the cocystalization of hOGG1 and an OG:C-containing DNA. However, it has never been used in the context of the genomic profiling of OG damage sites. The noncovalent interaction between K249Q and OG-DNA was confirmed by our gel-shift assay (Figure S4). Isothermal titration calorimetry (ITC) measurement revealed a dissociation constant of 20.1 ± 10.2 nM (Figure S5). To avoid background signals caused by borohydride-induced nonspecific cross-linking between the glycosylase and DNA (Figure S6), we focused on the borohydride-free OG-DNA trapping methods with the K249Q hOGG1 mutant. We optimized the amount of blocking DNA and the salt concentration of the reaction buffer to finally achieve enrichment of 120-fold for our model DNA sequence (Figure S7). We also noted that this protocol could enrich a model DNA sequence containing a single AP:C site by 9-fold (Figure S8). As a product of OG removal by glycosylases, the average occurrence of AP:C sites in the genome (0.9 sites per million bases) is similar to OG:C (5.4 sites per million guanine). Taken together, the above data demonstrate that our method preferentially targets OG:C pairing, and is also capable of capturing repair intermediates.

We then applied our optimized protocol to the analysis of genomic DNA extracted from MEF cells and fragmented with a commercial enzymatic digestion method (see the supplementary methods in Supporting Information). As a control
also analyzed the enrichment of a heterochromatin marker and histone modification. We found that the enrichment of OG peaks was reduced to the background level after preincubation with Fpg, indicating good performance of damage excision. We further verified that Fpg treatment alone did not interfere with DNA sequencing at the genomic level (Figure S9). We moved on to profile the oxidative damages in the MEF genome with two biological replicates. By comparing signals between the enriched and the input samples, we obtained 9326 significantly enriched peaks in total. Among these, 1470 peaks (15.8%) were also sensitive to the Fpg treatment, which we defined as OG peaks (Figures S11 and S12 and Supporting Information). A representative OG peak is shown in Figure 3B.

To study the distribution of guanine oxidations, we calculated the relative enrichment of OG peaks in various genomic elements (Figure 4, Table S1). Within different functional regions, OG peaks occurred more frequently in regulatory elements such as promoters and 5′UTR, which was consistent with the data from OG-Seq. To examine the correlation between guanine oxidation and genome accessibility, we calculated the enrichment of OG peaks in open chromatin markers such as DNase I highly sensitive sites (DNase1_HS), RNA polymerase II occupied regions (PolII), and histone modification H3K27ac interacting regions. We also analyzed the enrichment of a heterochromatin marker H3K9me3. As shown in Figure 4, OG peaks were substantially enriched in open chromatin regions while slightly depleted in closed heterochromatin regions. Moreover, we discovered a significant enrichment of OG peaks in CpG islands, which has not been reported before.

To summarize, we have described a novel method, enTRAP-seq, for the genome-wide profiling of oxidative DNA damages. By introducing a mutation K249Q to stabilize the complex between hOGG1 and its DNA substrate, we have engineered hOGG1 into a useful tool for DNA OG damage detection. For the analysis of genomic DNA, we identified 1470 significantly enriched OG peaks in MEF cells across two independent biological replicate experiments. Our data revealed that OG peaks tend to occur in regulatory elements and transcriptionally active chromatin regions.

Our conclusion supports previously published results of OG-Seq and OxiDIP-Seq, while arguing against the result from click-code-seq, where the accumulation of OG in the heterochromatin rather than the euchromatin was observed. Moreover, our data revealed a strong enrichment of DNA oxidation in the CpG islands. Interestingly, there is increasing evidence suggesting that guanine oxidation can relieve the transcriptional repression imparted by cytosine methylation, either by interfering with the recognition of methyl-CpG-binding proteins (MBPs) or by promoting demethylation through the recruitment of OGG1 and Ten-eleven translocation (Tet) proteins. Our present work thus provides clues that guanines within CpGs are favorably targeted by oxidants.

A drawback of enTRAP-seq is its lack of single-base resolution at identifying the damaged sites. In principle, DNA damage could be detected as mutations introduced during PCR amplification. However, our SNPs analysis of the enriched data set only revealed 20 mutations among 1470 OG peaks (Table S4, and Supplementary Methods), likely due to low mutation frequency and limited sequencing depth.

It is worth noting that the enTRAP-seq strategy can be extended to profiling other DNA damages. According to a previous study by Bruner et al., the mutation R154H could increase glycosylase activity on OG with opposite bases other
than cytosine. We hypothesize that the double mutations of R154H and K249Q may extend targets of this method to OG: A base pairing. Covalent catalysis employed by OGG1 is a commonly observed mechanism for almost all DNA glycosylases. The Fpg/Nei family bonds to C1′ of deoxyribose with its N-terminal amino group of proline. Similarly, the HhH family, which OGG1 belongs to, typically utilizes a lysine as the nucleophile. Recently, MutY and MBD4 from the HhH family have been reported to form transient covalent intermediates using a conserved aspartic acid residue. We envision that introducing mutations in these enzymes that block the repair process may convert them into useful enTRAP-seq tools for targeting a battery of DNA damages.

**ASSOCIATED CONTENT**

### Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.biochem.9b00782.

Materials and methods, model DNA sequences, gel-shift assays, ITC binding isotherm, and qPCR and NGS data analysis (PDF)

Annotation of high-confidence peaks (XLSX)

### Accession Codes

UniProt ID: hOgg1, O15527; Fpg, P05523; scOgg1, P53397.

### AUTHOR INFORMATION

**Corresponding Author**

E-mail: zoupeng@pku.edu.cn.

**ORCID**

Peng Zou: 0000-0002-9798-5242

**Funding**

This work was supported by the Ministry of Science and Technology of China (2017YFA0503600 and 2018YFA0507600) and Li Ge-Zhao Ning Life Science Junior Research Fellowship. P.Z. was sponsored by the National Thousand Young Talents Award.

**Notes**

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

We thank Prof. Chengqi Yi (Peking University) for helpful discussions, H. Meng and M. Liu for advice on NGS data analysis, H. Li for assistance with ITC experiments, and the National Center for Protein Sciences at Peking University in Beijing, China, for assistance with Fragment Analyzer 12. The mouse embryonic fibroblast cell line was a kind gift from Prof. Chengqi Yi.

### REFERENCES


