PNAS
www.pnas.org

Supplementary Information for

Endogenous oxidized DNA bases and APE1 regulate the formation of G-quadruplex structures in the genome

Shrabasti Roychoudhury[1], Suravi Pramanik[1], Hannah L. Harris[1], Mason Tarpley[1], Aniruddha Sarkar[1], Gaelle Spagnol[2], Paul L. Sorgen[2], Dipanjan Chowdhury[3], Vimla Band[1&4], David Klinkebiel[2&4] and Kishor K. Bhakat[1&4]

To whom the correspondence should be addressed: Kishor K. Bhakat, Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, NE, 68198, USA, Phone: 402-559-8467, Fax: 402-559-7328, E-mail: kishor.bhakat@unmc.edu

This PDF file includes:

Supplementary Materials and method
Figures S1 to S6
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 and S2

www.pnas.org/cgi/doi/10.1073/pnas.1912355117

Supplementary Information

SI MATERIALS AND METHODS:

Cell culture, plasmids, reagents: The human lung epithelial adenocarcinoma A549 (ATCC # CCL-185), human lung fibroblast IMR-90 (ATCC #CCL-186), human embryonic kidney HEK-293T (ATCC # CRL-3216), and human pancreatic ductal adenocarcinoma PANC-1 (ATCC #CRL-1469) cell lines were maintained in high glucose Dulbecco's Modified Eagle's Medium (DMEM) (Thermo Fisher Scientific). Wild-type and *OGG1* null mouse embryonic fibroblast (MEF) cells (kindly provided by Dr. Istvan Boldogh, University of Texas Medical Branch, Galveston) were maintained in DMEM: Nutrient Mixture F-12 (DMEM/F-12) medium (Thermo Fisher Scientific). The human colon cancer HCT116 cell line (ATCC #CCL-247) was grown in McCoy 5A medium (Thermo Fisher Scientific). All media were supplemented with 10% fetal bovine serum (FBS; Sigma) and an antibiotic mixture of 100 U/mL penicillin and 100 µg/mL streptomycin (Gibco-BRL). All cell lines were authenticated by STR DNA profiling by Genetica DNA laboratories, Burlington, NC. For APE1 knockdown studies, HCT116 cells stably expressing APE1-shRNA or control shRNA (1) were maintained in 10% FBS supplemented McCoy's 5A medium with 1µg/ml puromycin. Transient APE1 knockdown was achieved using APE1siRNA (Sigma #SAS1-Hs01-00027147) transfected to the indicated cell lines for 72 hours. To generate the doxycycline-inducible stable expression APE1 shRNA and control shRNA A549, HCT116, HEK-293T, and PANC-1 cells, three different doxycycline-inducible human APE1-shRNA constructs (shRNA #V3IHSHEG_5634292, #V3IHSHEG_6377584, and #V3IHSHEG_7228555 named as 1, 2, and 3, respectively; Dharmacon) and a non-targeting control (NTC) shRNA lentiviral

SMARTvector constructs with GFP were used. Lentiviral supernatants were generated by individually transfecting the shRNA lentiviral SMARTvector constructs into HEK-293T cells with packaging plasmids using X-tremeGENE HP DNA transfection reagent (Sigma). Lentiviral supernatants were transduced into the aforementioned cell lines and selected with 1µg/ml puromycin. All doxycycline-inducible APE1-shRNA and NTC-shRNA stable cell lines were maintained in the respective media supplemented with tetracycline free 10% FBS (Atlanta biologicals) and 1 µg/mL puromycin. For rescue experiments, APE1-shRNA and NTC-shRNA HEK-293T and HCTT16 cells were treated with Doxycycline (2 µg/mL) for 3 days and subsequently transfected in serum-free Opti-MEM media (Thermo Fisher Scientific) with FLAG-tagged WT, acetylation defective K5R (lysine 6, 7, 27, 31, and 32 mutated to arginine), repair defective H309A (histidine 309 to alanine), or redox defective C65/99S (cysteine 65 and 99 to serine) APE1 plasmid constructs (pFLAG-CMV-5.1) using Lipofectamine 3000 (Invitrogen). Methoxyamine, Glucose Oxidase, Hydrogen Peroxide, Triptolide, Actinomycin D, and Aldehyde Reactive Probe (ARP) were obtained from Sigma.

Immunofluorescence analysis: Cells were cultured on coverslips (Fisherbrand) and fixed in 4% paraformaldehyde (PFA; Sigma) in PBS for 30 mins at room temperature (RT). Cells were subsequently permeabilized and blocked for 1 hour at RT using a blocking buffer containing 0.5% Triton X-100 (Sigma), 10% goat serum (Thermo-Fisher #50062Z), glycine, and sodium azide in PBS. After permeabilization and blocking, cells were incubated in blocking buffer with primary antibodies overnight at 4° C and subsequently the corresponding secondary antibodies for 1 hour at RT. Cells were then washed in PBS and mounted using mounting media with DAPI (Vector Laboratories –Item # VV- 93952-

27). For G quadruplex staining, an addition permeabilization step was performed after formaldehyde fixation by incubating cells in PBS containing 0.5% Tween 20 for 20 mins at 37° C. Following Tween 20 permeabilization, cells were treated with 0.04 µg/µL RNase A (Invitrogen) and subsequently blocked and processed as previously described (2). For the DNase experiments, cells were fixed in 4% PFA in PBS for 30 mins, permeabilized with 0.2% Triton X-100 in PBS for 1 min, and washed in PBS at RT. Cells were then incubated for 2 hours at 37 °C in DNase reaction buffer with or without 0.06 U/µl of DNase I (RQ1 DNase, Promega). The working DNase reaction buffer contained 40 mM Tris-HCl (pH 8), 5 mM $CaCl_2$, 2 mM $MgCl_2$, and 100 µg/ml BSA. Following DNase digestion, the cells were washed in PBS and stained as previously described. All immunofluorescence images were captured by confocal microscopy or super-resolution structured illumination microscopy (SIM) where indicated. Confocal microscopy images were captured using the Zeiss LSM 800 with Airyscan Microscope in the Advanced Microscopy Core Facility (AMCF) at the University of Nebraska Medical Center (UNMC). Three-dimensional (3D) SIM images were collected with the Zeiss ELYRA PS.1 Super-Resolution Microscope in the AMCF at UNMC. Pearson correlation coefficients were calculated using the ImageJ software JACoP co-localization analysis module. Quantification of co-localization was determined by establishing a threshold using the JACoP threshold optimizer followed by calculation of correlation coefficients. Primary antibodies used for the immunofluorescence studies were mouse monoclonal anti-APE1 (1:100; Novus Biologicals; # NB100-116), anti-AcAPE1 (1:50; (3)), Anti-AcOGG1 (1:50; (4)), Anti-G4-1H6 (1:50; Millipore), Anti-ECD (1:100; (5)) Anti-H3K27 acetylated histone (1:100; Millipore; #05-1334), Anti- H3K4Me3 (1:100; Millipore) and anti-c-Jun (1:100; Abcam).

Secondary antibodies used for the immunofluorescence studies were Alexa Fluor 488 anti-mouse IgG (Life Technologies, 1:500) or Alexa Fluor 594 anti-rabbit IgG (Life Technologies; 1:500).

Chromatin Immunoprecipitation (ChIP): Cells were plated in a 150 mm cell culture dish. The following day, cells were incubated with 1% formaldehyde in PBS for 15 mins at RT to crosslink protein-DNA complexes and subsequently washed in PBS. Cells were collected in PBS containing 1X protease inhibitor cocktail (Roche cOmplete™) with a plastic cell scraper and pelleted at 1000 rpm for 10 mins at 4º C. The pelleted cells were lysed and incubated on ice in SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8) for 10 minutes and subsequently subjected to sonication (Misonix sonicator 3000). Sonication was performed on ice with 8 rounds of 15-second pulses with 15-second pause between each pulse. The sonicated lysate was centrifuged at 14000 rpm for 15 min at 4º C, and the supernatant containing the clear sheared chromatin lysate was collected. The collected lysate was diluted 1:10 to a total volume of 2 mL with ChIP dilution buffer containing 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl pH 8.1, and 167 mM NaCl. Immunoprecipitation (IP) was performed by incubating 5 µg of the respective antibody (control IgG included in a separate IP) with the diluted lysate overnight at 4º C with constant nutating rocking. The next day, 25 µl of Protein A/ G Dynabeads™ (Invitrogen) was added to each reaction and incubated for 2 hours at 4° C with constant nutating rocking. The IPs were washed sequentially with low salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8, 150 mM NaCl), high salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8, 500 mM NaCl), LiCl immune complex wash buffer (0.25 M LiCl, 1%

NP40, 1% Na-deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8), and TE buffer (10 mM Tris-HCl, 1 mM EDTA pH 8). The protein-DNA complexes were eluted in ChIP elution buffer (1% SDS, 0.1 M $NaHCO_3$) and reverse cross-linked by incubating the elutes in 200 mM NaCl overnight at 65ºC. ChIP DNA was subjected to RNAse and proteinase K digestion. ChIP DNA was precipitated by following standard protocol of phenol-chloroform extraction, 100% ethanol precipitation, and a 70% ethanol wash. The purified ChIP DNA was dissolved in ultrapure water or TE buffer. Immunoprecipitation (IP) was performed with rabbit α-APE1, rabbit α-AcAPE1 (3), rabbit α-AcOGG1 (4) and mouse α-BG4 (Absolute biology).

AP-seq: For mapping AP sites, 1 mM biotin-conjugated ARP (Life Technologies, #A10550) was fed to cells for 3 hours. Sonicated lysates were harvested and diluted as described for ChIP. Biotin-conjugated ARP–tagged chromatin fragments from sonicated lysates were pulldown with 20 µl of streptavidin-conjugated MyOne Dynabeads (Life Technologies, # 65601) in a total volume of 2 ml for 2 hours. The streptavidin-conjugated MyOne Dynabeads were pre-washed in ChIP dilution buffer. Following incubation with the sonicated lysate, the beads were washed sequentially with low and high salt buffer as described for ChIP. DNA was purified and resuspended in 20 µl of TE for library preparation. Cells that were not fed biotinylated-ARP were used as negative control.

ChIP-seq and AP-seq data sets: Result for each ChIP-seq and AP-seq experiment with a given antibody was obtained from multiple biological replicates, AP-seq (n=2), AcAPE1 (n=3), BG4 (n=3), APE1 (n=2), and AcOGG1 (n=2). Replicates were merged, and the merged data files were utilized for statistical analysis. The numbers of replicate for each ChIP-seq and AP-seq experiment are shown in the table below.

ChIP seq	No. of biological replicate
AP seq_A549 cells	2
APE1 seq_A549 cells	2
AcAPE1 seq_A549 cells	3
AcOGG1 seq_A549 cells	2
BG4 seq_A549 cells	3
BG4 seq_APE1shRNA A549 cells	2
Actinomycin D_APE1_A549 cells	1
Actinomycin D_AcAPE1_A549 cells	1
Actinomycin D_BG4_A549 cells	1
Actinomycin D_ AcOGG1_A549 cells	1
AP seq_HCT116 cells	1
AcAPE1 seq_HCT116 cells	2
AcAPE1 seq_APE1shRNA HCT116 cells	1

ChIP-Sequencing analysis: The purified ChIP DNA was provided to the University of Nebraska Epigenomics Core. ChIP DNA was quantified by Qubit (Invitrogen). The library was made using New England Biolabs NEB Next Ultra II DNA Library Prep Kit. High-throughput sequencing was conducted at the UNMC Sequencing Core, using an Illumina

HiSeq 2500 Genome Analyzer. FASTQ files were generated after adaptor sequences, and low quality (Phred score < 20) ends were trimmed from the ChIP sequences using the Trim Galore software package (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). FASTQ files were aligned to the human genome (GRCh37/hg19) using the sequence aligner Bowtie2 (version 2.2.3) (6). The software package Picard routine Mark Duplicates (http://broadinstitute.github.io/picard/) was used to remove sequence duplications. Biological replicates were then merged using SAMtools. The merged files were used to call peaks using MACS2 peak caller (7) software (version 2.1.1). A peak from a given antibody was considered significant if it had a fold change ≥2 over IgG, ≥15 reads, and a p-value<0.001. Both analyses programs scored read-density across genomic DNA and defined peak regions where a bimodal enrichment of reads has occurred. All bigWig files generated using the deepTools bamCoverage routine (https://deeptools.readthedocs.io/en/develop/) were uploaded and displayed on the UCSC Genome browser with a cut off at 5 reads. Metaprofiles of coding genes were compiled using the UCSC transcript annotation. ChIPseeker package in R was used to analyze the peaks with respect to the TSS +/- 2 kb, with 95% confidence levels (8).

AP-seq analysis: Library prep and read processing was completed in the same manner as described above for ChIP-seq data. Individual biological replicates were then merged into a single file. All analyses were performed using the average relative enrichment compared to a non-ARP treated negative control, in appropriate bin sizes tiled across the genome or covering genomic elements. The bins were normalized using reads per million mapped

reads (RPM) and plotted in R. For a large-scale overview, a bin size of 100-kb was chosen for comparability with other ChIP-seq studies.

Genome-wide correlation and statistical analysis: Genome-wide correlation between AP sites, APE1, AcAPE1 binding, and G4 was determined using the *StereoGene* software with default parameters (9). *Stereogene* is based on Kernel correlation (KC), which provides an estimate of local genomic position correlation between two genomic features. Foreground KC was calculated from paired windows of the same genomic position for two profiles. Background KC was determined using a null distribution generated by a shuffling procedure that randomly matches windows between two profiles. The statistical significance of correlations were evaluated by a permutation-based Mann-Whitney test which compares the difference between foreground and background KC. The false discovery rate (FDR) for correlation was estimated using the background distribution as null distribution and the foreground as the signal. Correlations between two genomic features were also calculated using spearman's correlation with deepTools2 plotCorrelation (10).

Analysis of PQS scores: PQS scores were obtained from Marsico et al (11). G4 and AcAPE1 bam files were filtered based on reading strand (+ or -) using bedTools and then compared with PQS regions on the same strand using intersectBed. Regions that had a PQS score ≥20, ≥2-fold coverage over IgG, and ≥15 total reads were considered significant. Both plus and minus strands were combined to compare overlapping regions between G4 and AcAPE1 ChIP seq. IntersectBed was used to determine overlapping regions between AcOGG1, AcAPE1, and G4. Data are represented as Venn diagrams.

Differential BG4 enrichment analysis. Differential BG4 ChIP-seq peaks between control and APE1 knockdown cells were determined using THOR peak caller based on an FDR adjusted p-value <0.05 (12).

ChIP qPCR analysis: ChIP DNA and 1% input DNA were subjected to SYBR GREEN-based Real-Time PCR (7500 Real-Time PCR System; Applied Biosystems) with primers (Table below). The Real-Time PCR cycle threshold (CT) values were used to determine % ChIP enrichment normalized to input, which was calculated by (2^(adjusted input CT-ChIP CT))*100. Immunoprecipitation (IP) was performed with rabbit α-APE1, rabbit α-AcAPE1 (3), rabbit α-AcOGG1 (4), mouse α-BG4 (Absolute biology), mouse α -MAZ (Santa Cruz), or control IgG (Santa Cruz) primary antibodies.

Primer name	Forward sequence	Reverse sequence
KRAS G4	GTACGCCCGTCTGAAGAAGA	GAGCACACCGATGAGTTCGG
Neg G4	CTCCGACTCTCAGGCTCAAG	CAGCACTTTGGGAGGCTTAG
MYC	CAGGCAGACACATCTCAGGG	CGTATACTTGGAGAGCGCGT
P21	CAGGCTGTGGCTCTGATTGG	TTCAGAGTAACAGGCTAAGG

Western blotting: Western blotting was performed using standard methods, as described previously (13). Primary antibodies used in the study includes mouse monoclonal anti-APE1 (1:5000; Novus Biologicals; # NB100-116), mouse anti-HSC70 (1:10000; Santa Cruz), anti-ECD (1:1000; (5)) , and anti-beta-actin (1:1000) antibodies.

RNA-Seq and quantitative RT-PCR: Doxycycline inducible APE1-shRNA and NTC-shRNA A549 cells were treated with Doxycycline (2 µg/mL) for 3 days. For RNA-seq, total RNA was purified using a Qiagen kit (#74104) and sequenced using Illumina Hi-seq. For quantitative Real-Time PCR (qRT-PCR), total RNA was isolated from cells using the standard TRIZOL protocol. Synthesis of cDNA using 1 µg of total RNA was performed with the MulV RT kit (Invitrogen) using random hexamer primers, and 1/50th of each reaction was used in qRT-PCR. SYBR green (Applied Biosystems) was used for detection with an Applied Biosystems StepOne Plus system. Fold change in expression was calculated by specific gene $2^{\wedge\text{-(target-reference)}}$/*GAPDH* $2^{\wedge\text{-(target-reference)}}$. The following gene-specific primers were used:

Primer name	Forward sequence	Reverse sequence
KRAS	TCTTGCCTCCCTACCTTCCACAT	CTGTCAGATTCTCTTGAGCCCTG
GAPDH	TGGGCTACACTGGAGCACCAG	GGGTGTCGCTGTTGAAGTCA

Site-directed mutagenesis and Luciferase assay: The Del4 luciferase reporter plasmid, harboring the 22-mer c-*MYC* G4 forming sequence in the P1 promoter upstream of the luciferase reporter (Addgene plasmid # 16604) was mutated using the QuickChange Site-

directed Mutagenesis kit (Agilent Technologies, CA, USA) to generate a mutated G4-forming sequence.

- Del4: 5′-G4AG3TG4AG3TG4-3′
- G12A: 5′-G4AG3TGAG2AG3TG4-3′ (Kindly provided by Dr. Jyoti Dash; Indian Association for the cultivation of science; India,(14))
- G18A: 5′-G4AG3TG4AGAGTG4-3′

The pRL-TK renilla luciferase (Addgene plasmid #E2241) control reporter vector was used for measuring transfection efficiency and as a non-G4 sequence control. Cells were plated in a 6-well cell culture plate and transfected with 1 µg of Del4 plasmid with 100 ng of pRL -TK plasmid using Lipofectamine 3000 as per manufacture's protocol. After 48 hours, cells were lysed, and the Dual-luciferase reporter assay (Promega) was performed according to manufacturer protocol in Glow-max (Promega) system.

Preparation of G4 DNA-containing templates: Synthetic single-stranded DNA templates and complementary strands were purchased from Midland-certified reagents. The oligomer sequences are listed in the table below. Synthetic double-stranded DNA templates were generated by incubating 10 µM of each strand with annealing buffer with a final volume of 100 µL for 5 mins at 95° C, followed by slow cooling from 95° C to 37 °C. G4 DNA formation was induced by adding 100 mM KCl to the annealing reaction. Oligos were stored at −20° C.

Oligo name	Sequence

MYC-AP 75 mer template strand	(Cy5)ATAAGCTTCCCGGGGTCGACCACGTCTGGGGAGGGTG(abasic)GGAGGGTGGGG AAGGTCTAGATCTGGTACCGAATTCT
Complimentary strand	TATTCGAAGGGCCCCAGCTGGTGCAGACCCCTCCCACTCCTCC CACCCCTTCCAGATCTAG ACCATGGCTTAAGA
*MYC*WT	(Cy5)ATAAGCTTCCCGGGGTCGACCACGTCTGGGGAGGGTGGG GAGGGTGGGG AAGGTCTAGATCTGGTACCGAATTCT

CD spectroscopy: Circular Dichroism (CD) spectra were obtained on a JASCO J-810 spectropolarimeter, equipped with a thermostatted cell holder. CD readings were obtained with 1 µM oligonucleotides solutions in 50 mM HEPES pH 7.5, 10 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 1 mM DTT, and 0.1 mg/ml BSA. Purification of Wild-Type and mutant APE1 proteins were performed according to previously published protocol (15).

The spectra were recorded in a 10 mm quartz cuvette at 20 °C. Scans were performed at 20°C over a wavelength range of 210-330 nm with a response time of 0.5 s, 1 nm pitch, and 1 nm bandwidth. Blank spectra of samples containing buffer were subtracted from DNA samples. Spectra were reported as ellipticity (mdeg) versus wavelength (nm). Each spectrum was recorded five times and smoothed and subtracted to the baseline.

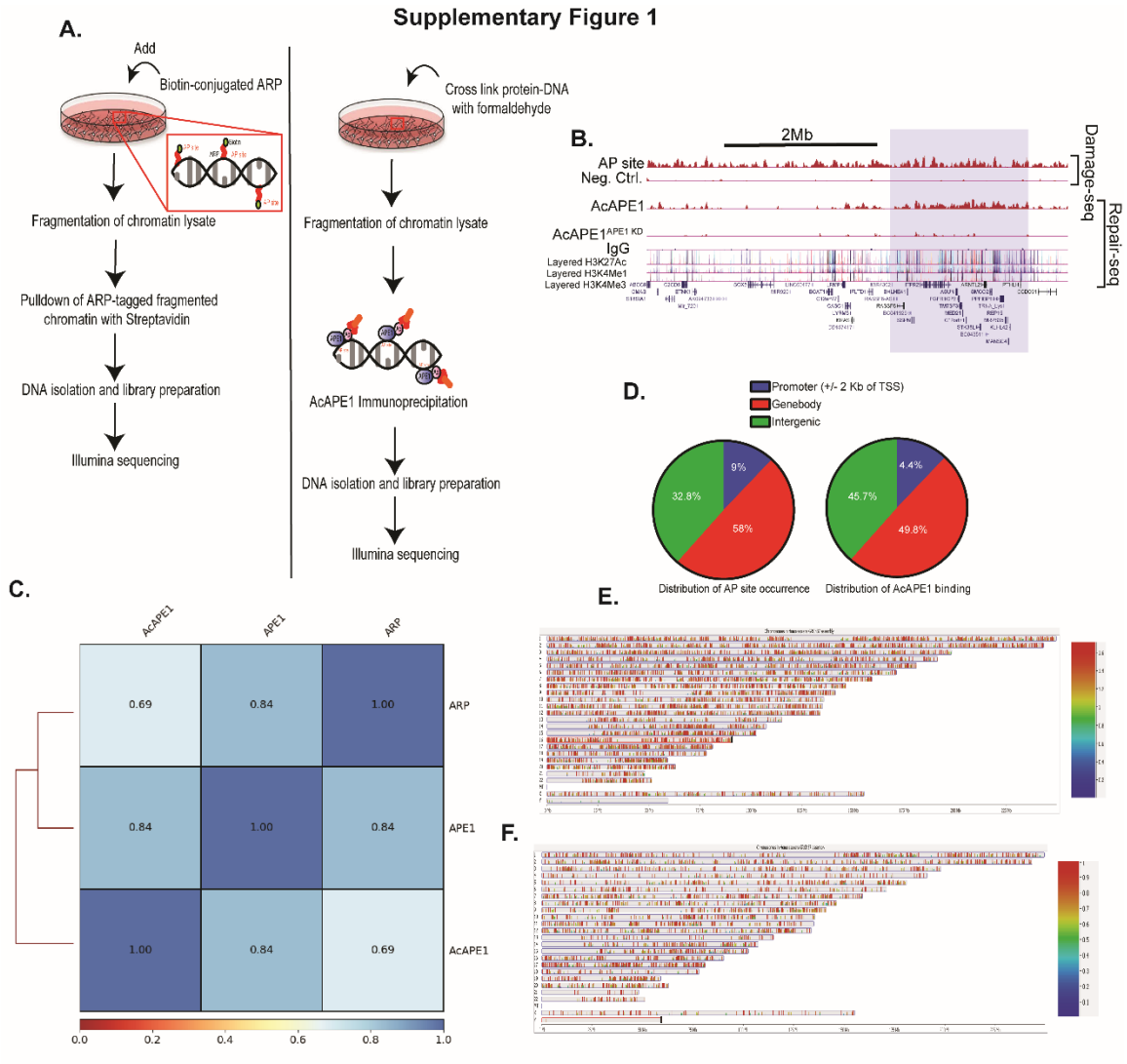
EMSA: A 75-mer c-*MYC* G4-forming sequence with or without tetrahydrofuran (THF), an AP site analog, at position 12 was labeled at the 5' end with Cy5 and annealed to a complementary strand as described above. The anneals oligos (30 nM) were independently

mixed with an increasing amount of recombinant APE1 (6.4, 64 and 640 ng) at room temperature in electrophoretic mobility shift assay buffer containing 50 mM HEPES pH 7.5, 10 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 1 mM DTT, 1 mM EDTA, and 0.1 mg/ mL BSA. One µg of poly-deoxy-inosinic-deoxy-cytidylic acid (Poly [dI-dC]) was used in each reaction as non-specific competitor. After 10 minutes, the EMSA binding reactions were kept on ice and loaded into a native gel. The protein-bound DNA was separated from the unbound (free) substrate by electrophoresis on a 10% non-denaturing polyacrylamide gel at 100 V/cm^2 for 60 mins. EMSA gels were visualized using the Odyssey Li-COR machine.

Fluorescence Recovery After Photobleaching (FRAP): A small region in the nucleus of a cell was photobleached with three iterations at 100% laser intensity and recovery of fluorescence was monitored 200 times every 20 ms at 1% laser intensity (25 mW Argon laser, 514 nm line) to allow the measured fluorescence to reach a steady-state level. The shown FRAP data were corrected for background noise and normalized to pre-bleach values.

Apurinic/apyrimidinic (AP) endonuclease activity assay: A 75-mer oligonucleotide containing a THF at nucleotide 37 (Midland Corp) was labeled at the 5´ end with Cy5. The oligo was annealed to the complementary strand with a G base opposite THF. Endonuclease reactions were performed by incubating the annealed duplex with recombinant APE1 or in vitro acetylated APE1 in endonuclease reaction buffer with a final reaction volume of 15 µL at 37 º C for increasing time points within the linear range (15). The working endonuclease reaction buffer contained 50 mM Tris-HCl pH 8.5, 50 mM KCl, 1mM $MgCl_2$, 1 mM DTT, 0.1 mM EDTA, and 100 µg/mL BSA. The endonuclease reaction was stopped by adding 10 µL of 80% formamide/40 mM NaOH containing 0.05%

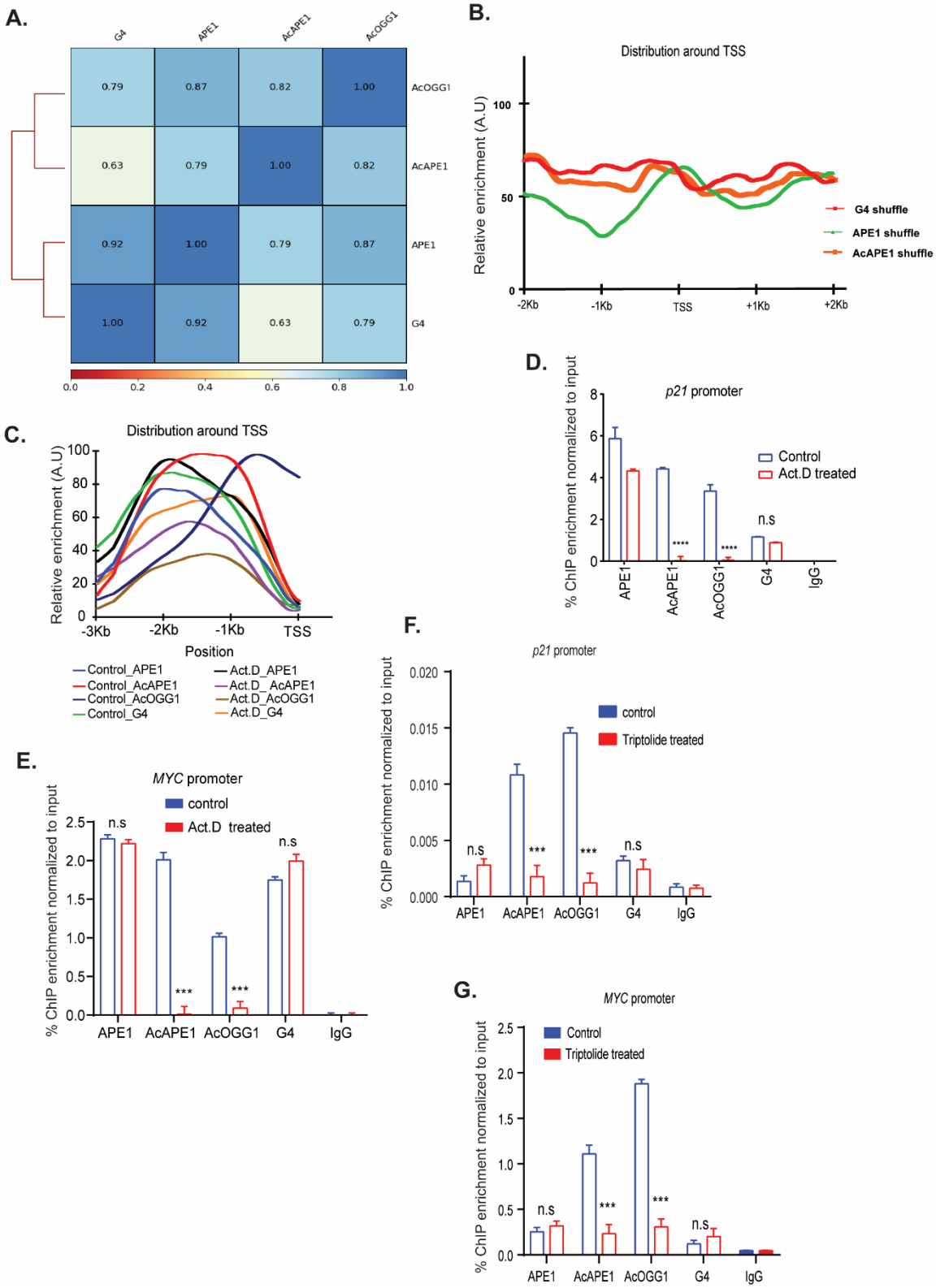
xylene cyanol and heating at 95 °C for 5 min. The intact substrate and cleaved product were separated by electrophoresis in an 8 M urea denaturing 20% polyacrylamide gel. The gels were visualized in the Odyssey Li-COR machine.



Supplementary Figure 1: Genome-wide mapping of endogenous AP site damage by AP-Seq and binding of APE1 and AcAPE1. A) Left, experimental schematic for mapping AP sites (AP-seq) in the genome. AP-sites in DNA were tagged using a biotin-conjugated aldehyde reactive probe (ARP), which specifically recognizes AP-sites (16). Biotin-tagged AP-sites were pulled down with streptavidin beads. The enriched DNA was processed for sequencing and mapped to the reference human genome (hg19). Right,

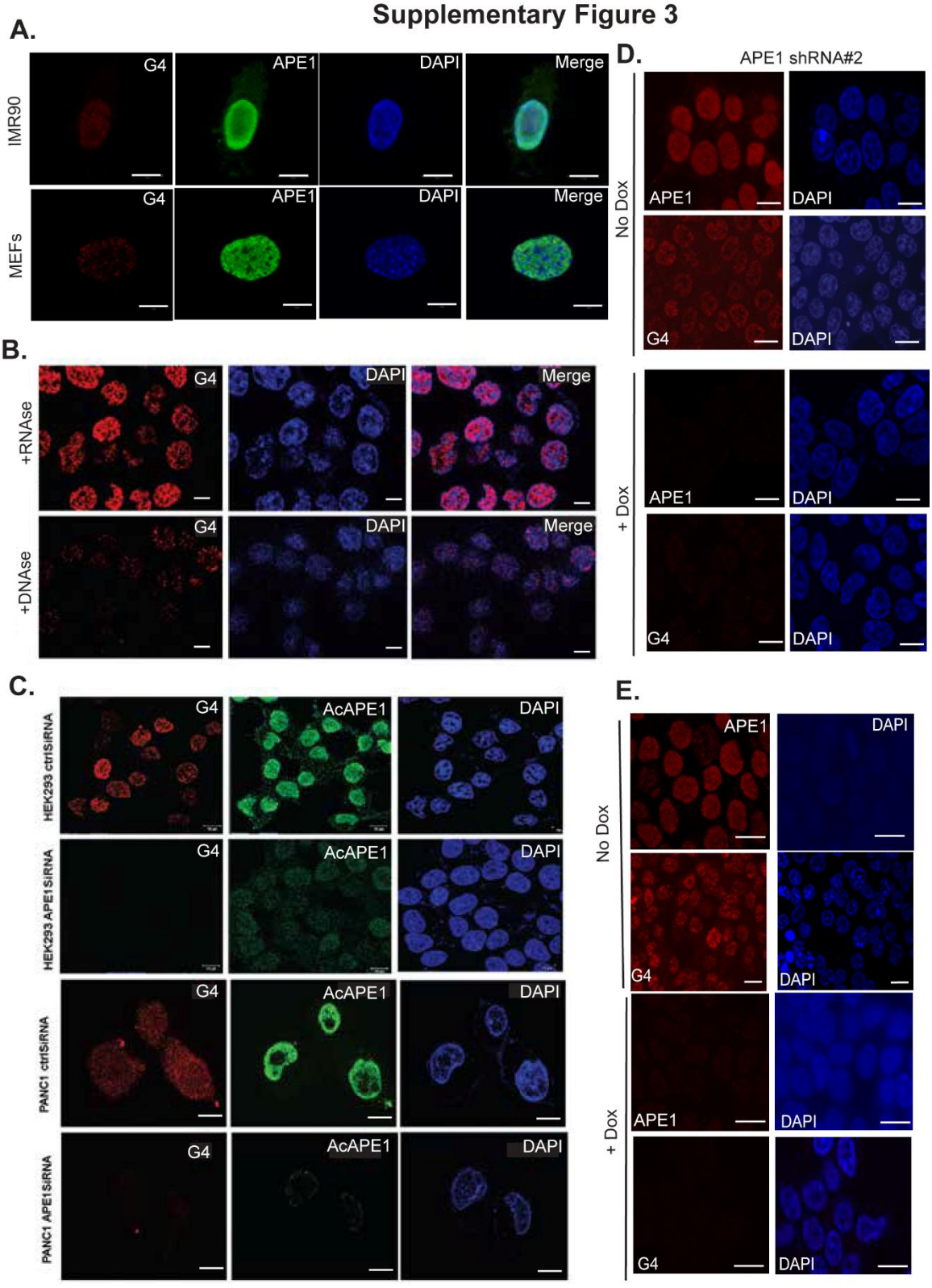
schematic for mapping genome-wide binding or enrichment of APE1 and AcAPE1 (Repair-seq) by ChIP-seq. **B)** Representative region of chromosome 8 showing AP site and AcAPE1 profiles compared to IgG in control HCT116 and isogenic HCT116^{APE1shRNA} (AcAPE1-APE1KD) cells. The shaded box highlights the disappearance of AcAPE1 peaks in APE1 KD cells. **C)** The heatmap displays the Spearman's correlation between APE1 and AcAPE1 ChIP-seq and AP-seq (ARP). **D)** Analysis of distribution of endogenous AP site damage and AcAPE1 relative to promoter, gene body, and intergenic regions in HCT116 cells. **E & F)** Distribution of APE1 and AcAPE1 binding in 100 Kb bins in individual chromosomes. The color scale represents the $\log_2$ fold change of normalized APE1 and AcAPE1 enrichment over background (Relative enrichment).

Supplementary Figure 2



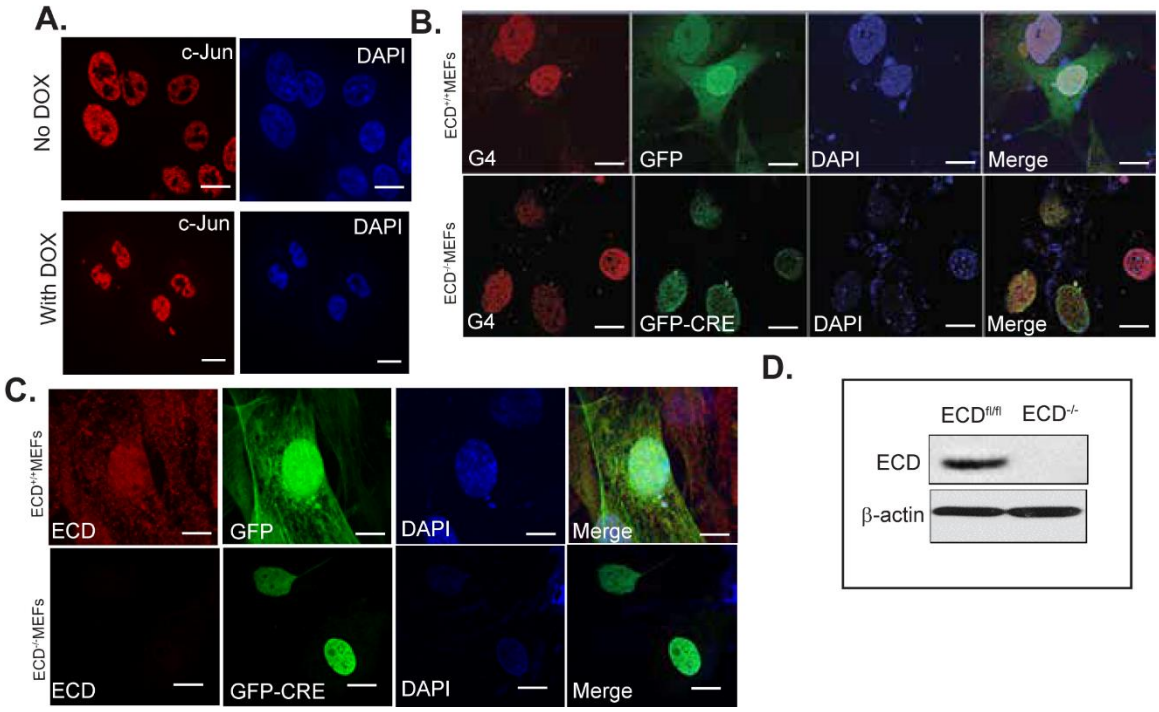
Supplementary Figure 2: Inhibition of transcription does not affect G4 occurrence and APE1 binding. A) The heatmap displays the Spearman correlation between AcOGG1, APE1, AcAPE1, and G4. **B)** Distribution of randomly shuffled G4, APE1, and AcAPE1 peaks +/-2000 bp over the TSS. **C)** Metaprofiles of the relative enrichment of APE1, AcOGG1, AcAPE1, and G4 with respect to +/- 2000 bp over the TSSs of ~ 28000 protein-coding genes before and after inhibition of transcription elongation with Actinomycin D (Act. D, 10 µg/ mL for 2 hours) in A549 cells. **D, E, F, and G)** Binding of APE1, AcAPE1, and AcOGG1 and occurrence of G4 were validated by RT-ChIP-PCR with respect to the *P21* and *MYC* gene promoters before and after Actinomycin D or Triptolide (500 nM for 4 hours) treatment. The *p*-values were calculated using an unpaired student's t-test (****p<0.0001, ***p<0.001, **p<0.01, ^{ns}p>0.05). Error bars denote +SD**.**

Supplementary Figure 3



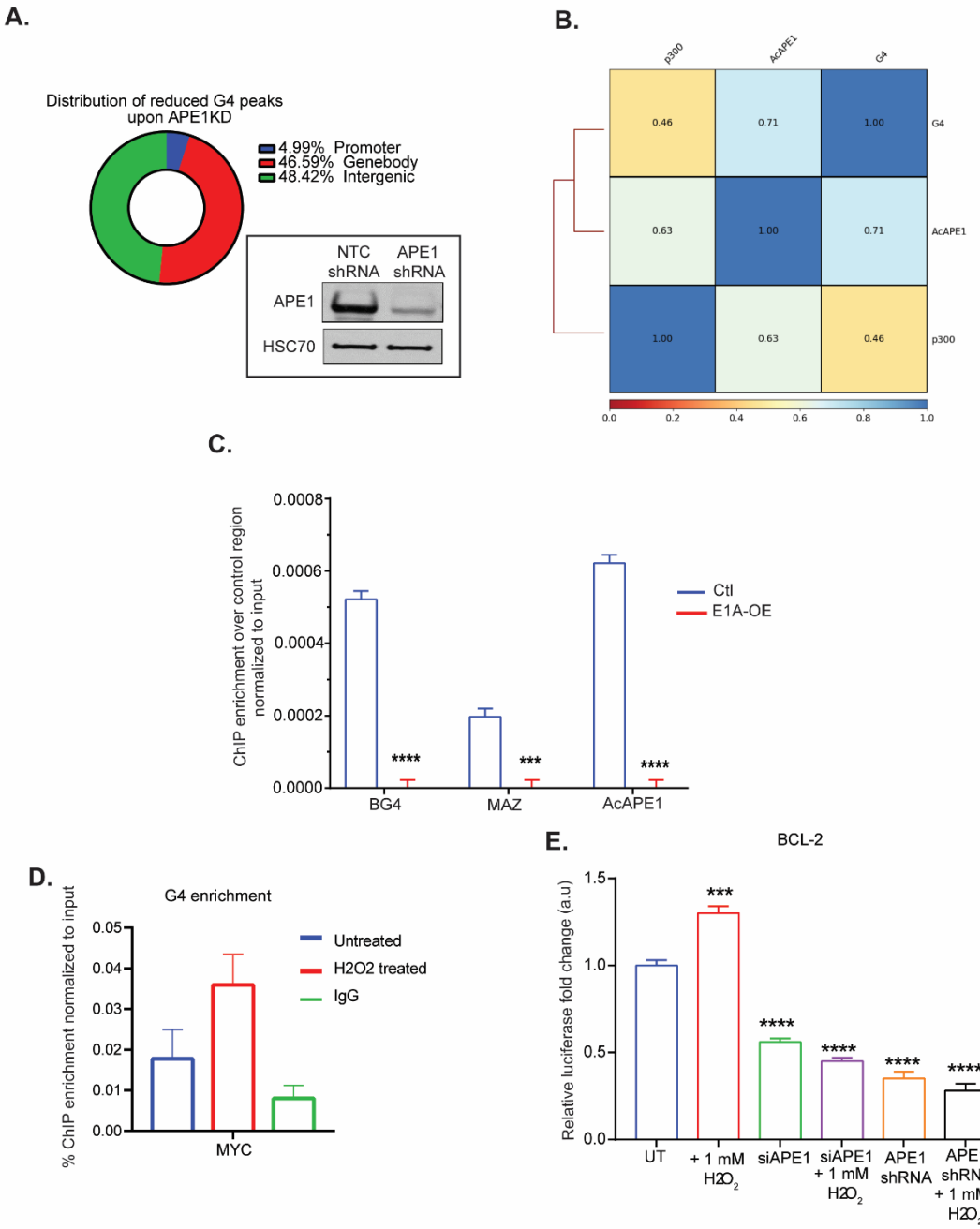
Supplementary Figure 3: Base excision repair protein APE1 regulates G4 formation and stabilization in cells. A) 3D SIM images of IMR-90 and MEF cells immunostained with α-1H6 (a G4 structure-specific antibody) and α-APE1 Abs and counterstained with DAPI. **B)** Immunofluorescence (IF) images of A549 cells treated with RNase A or DNase, immunostained with α-1H6 and counterstained with DAPI. **C)** IF images of control siRNA and APE1siRNA transfected HEK-293 and PANC1 cells immunostained with α-1H6 and α-AcAPE1 and counterstained with DAPI. **D &E)** Confocal microscopy image of HEK-293T and HCT116 cells expressing a Doxycycline inducible APE1shRNA immunostained with α-1H6 and α-AcAPE1 after Doxycycline (2 μg/mL) treatment for 2 days.

Supplementary Figure 4



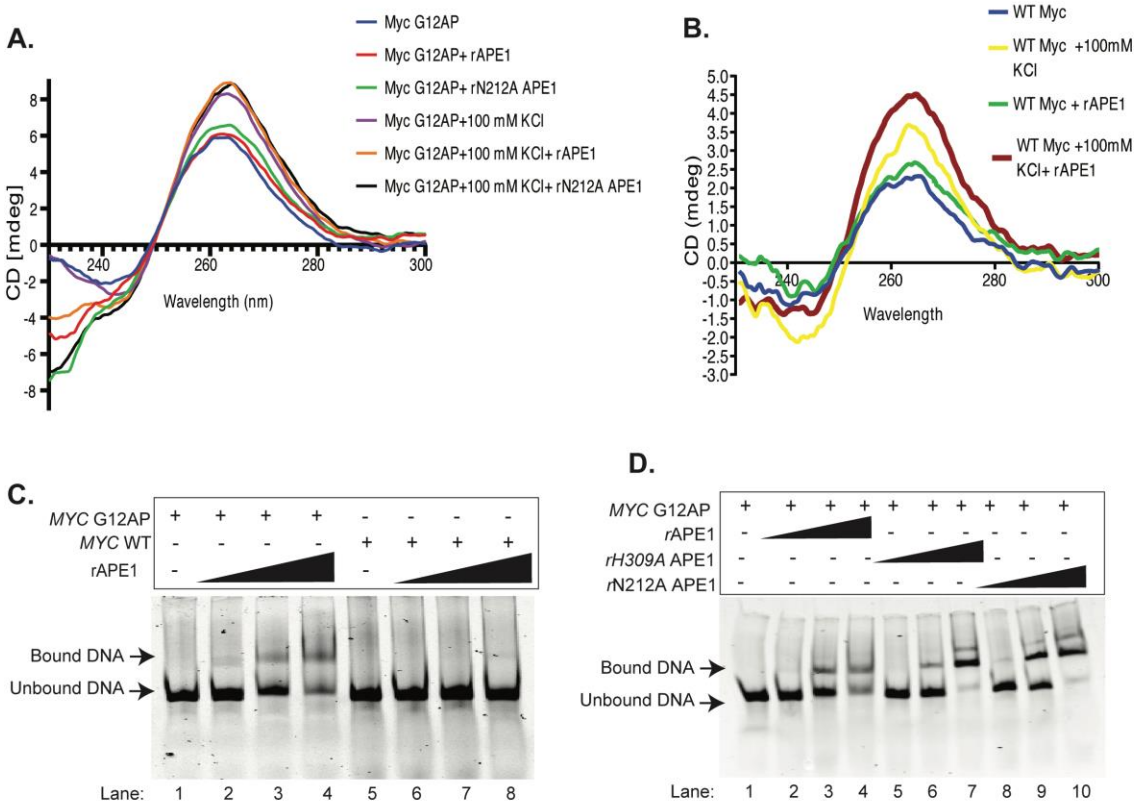
Supplementary Figure 4: Ecdysoneless (*ECD*) gene deletion does not affect G4 formation in cells. A) HCT116^{APE1shRNA} cells with or without Doxycycline (2 μg/mL) treatment were immunostained with α-c-Jun. **B & C)** *ECD* $^{fl/fl}$ MEFs expressing adeno-control GFP or adeno-Cre GFP (green) (17) were stained with α-1H6 (red) or α-ECD (red) Abs. **D)** The level of ECD in these cell extracts was examined by Western blot analysis with α-ECD and α-β-actin (loading control).

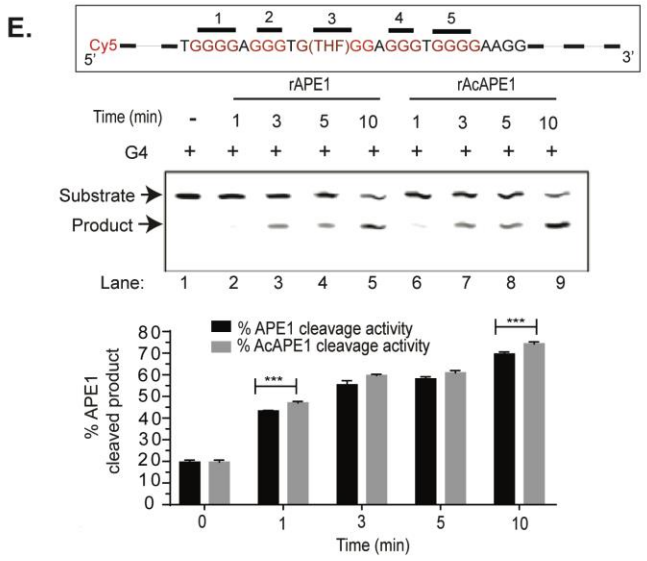
Supplementary Fig. 5



Supplementary Figure 5: **Oxidative damage and base excision repair protein APE1 regulates G4 mediated gene expression. A)**. A549 cells expressing non-targeting control shRNA (NTCshRNA) and APE1-specific shRNA (APE1shRNA) under a Doxycycline-inducible promoter were treated with Doxycycline (2 μg/mL) for 3 days and ChIP was performed with α-BG4.Distribution of reduced G4 enriched (log_2 fold change ≤ 0.5, adjusted p value≤ 0.05) peaks in the promoter, gene body, and intergenic regions in the genome upon APE1 knockdown. APE1 level in these cell extracts was examined by Western blot analysis with α-APE1 and α-HSC70 (loading control) **B)** Spearman's correlation between p300, G4, and AcAPE1 ChIP-seq. **C)** E1A was overexpressed (E1A-OE) in HCT116 cells for 48 hours, and ChIP-qPCR was performed to examine the enrichment of G4, MAZ, and AcAPE1 on the *KRAS* G4 promoter. **D)** Primer directed ChIP-qPCR performed with respect to the *MYC* gene promoter region to examine enrichment of G4 in A549 cells before or after H_2O_2 treatment. **E)** *BCL-2* G4 promoter-luciferase reporter vector and pRL-TK-renila luciferase were co-transfected in control, transient APE1siRNA (siAPE1), and constitutively expressing APE1shRNA HCT116 cells. After 48 hours, cells were treated with 1 mM H_2O_2 for 1 hour and promoter luciferase activity was measured.

Supplementary Fig. 6





Supplementary Figure 6: **Effect of APE1 binding on *MYC*-G4 without an AP site. A)** CD spectra of *MYC* G12AP at 20 °C in the presence of 100 mM KCl alone or in combination with 1 µg of recombinant APE1 or N212A APE1. **B)** CD spectra of *MYC* WT at 20 °C in the absence or presence of 100 mM KCl, 1 µg of recombinant APE1, or both. **C)** Electrophoretic mobility shift assay using increasing amounts of recombinant APE1 (6.4, 64, and 640 ng) was performed to analyze binding to Cy5-labeled WT and AP site containing *MYC* G4 oligonucleotides. **D)** Electrophoretic mobility shift assay using increasing amounts (6.4, 64, and 640 ng) of recombinant WT or H309A or N212A APE1 was performed to analyze binding to Cy5-labeled WT and AP site containing *MYC* G4 oligonucleotides. **E)** Comparison of time-dependent cleavage activity of recombinant APE1 and AcAPE1 on a Cy5-labeled 75-mer c-*MYC* quadruplex DNA (substrate, S). Formation of a 37-mer cleaved product (P) was shown at different time points. Quantitation of the average cleaved product from three independent experiments are shown.

Dataset S1 (separate file). List of overlapping G4 and AcAPE1 enriched sequences that have a PQS score ≥ 20.

Dataset S2 (separate file). List of overlapping between differentially expressed (D.E.) genes and genes with overlapping enrichment of AcOGG1, AcAPE1, and G4.

SI References

1. A. Bhattacharyya *et al.*, Acetylation of apurinic/apyrimidinic endonuclease-1 regulates Helicobacter pylori-mediated gastric epithelial cell apoptosis. *Gastroenterology* **136**, 2258-2269 (2009).
2. A. Henderson *et al.*, Detection of G-quadruplex DNA in mammalian cells. *Nucleic acids research* **42**, 860-869 (2014).
3. R. Chattopadhyay *et al.*, Regulatory role of human AP-endonuclease (APE1/Ref-1) in YB-1-mediated activation of the multidrug resistance gene MDR1. *Molecular and cellular biology* **28**, 7066-7080 (2008).
4. K. K. Bhakat, S. K. Mokkapati, I. Boldogh, T. K. Hazra, S. Mitra, Acetylation of human 8-oxoguanine-DNA glycosylase by p300 and its role in 8-oxoguanine repair in vivo. *Molecular and cellular biology* **26**, 1654-1665 (2006).
5. A. A. Olou *et al.*, Mammalian ECD Protein Is a Novel Negative Regulator of the PERK Arm of the Unfolded Protein Response. *Molecular and cellular biology* **37** (2017).
6. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357 (2012).
7. Y. Zhang *et al.*, Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**, R137 (2008).
8. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics (Oxford, England)* **31**, 2382-2383 (2015).
9. E. D. Stavrovskaya *et al.*, StereoGene: rapid estimation of genome-wide correlation of continuous or interval feature data. *Bioinformatics (Oxford, England)* **33**, 3158-3165 (2017).
10. F. Ramirez *et al.*, deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* **44**, W160-165 (2016).
11. G. Marsico *et al.*, Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic acids research* **47**, 3862-3874 (2019).
12. M. Allhoff, K. Sere, F. P. J, M. Zenke, G. C. I, Differential peak calling of ChIP-seq signals with replicates with THOR. *Nucleic acids research* **44**, e153 (2016).
13. K. K. Bhakat, T. Izumi, S. H. Yang, T. K. Hazra, S. Mitra, Role of acetylated human AP-endonuclease (APE1/Ref-1) in regulation of the parathyroid hormone gene. *The EMBO journal* **22**, 6299-6309 (2003).

14. D. Dutta *et al.*, Cell penetrating thiazole peptides inhibit c-MYC expression via site-specific targeting of c-MYC G-quadruplex. *Nucleic acids research* **46**, 5355-5365 (2018).
15. A. K. Mantha *et al.*, Unusual role of a cysteine residue in substrate binding and activity of human AP-endonuclease 1. *Journal of molecular biology* **379**, 28-37 (2008).
16. H. Atamna, I. Cheung, B. N. Ames, A method for detecting abasic sites in living cells: age-dependent changes in base excision repair. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 686-691 (2000).
17. J. H. Kim *et al.*, Role of mammalian Ecdysoneless in cell cycle regulation. *The Journal of biological chemistry* **284**, 26402-26410 (2009).