Supporting Information

Oxidative modification of the potential G-quadruplex sequence in the *PCNA* gene promoter can turn on transcription

Samuel C.J. Redstone, Aaron M. Fleming, and Cynthia J. Burrows*

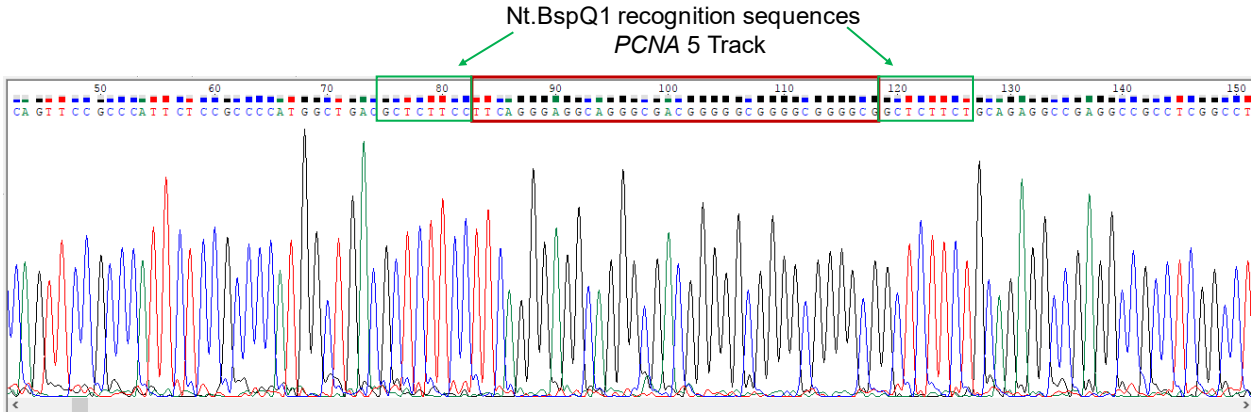
Department of Chemistry, University of Utah, Salt Lake City, UT 84112-0850, USA

*To whom correspondence should be addressed

E-mails: burrows@chem.utah.edu

Table of Contents

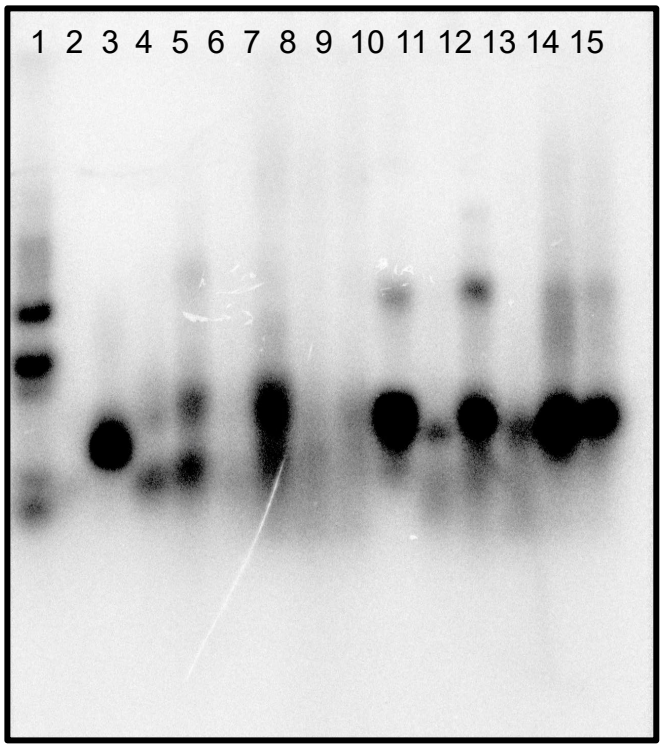
Figure S1. Example of a Sanger sequencing chromatogram



Sanger sequencing chromatogram showing successful incorporation of the *PCNA* 5-track and flanking Nt.BspQ1 recognition sequences in the target plasmid. The sequence inserted was 5`-pTT CAG GGA GGC A**G**G GC**G** ACG GGG GC**G** GGG CGG GGC GG CTC TTC T and the red Gs are the positions studied with OG and p = phosphate.

Figure S2. Native PAGE analysis of lesion-containing *PCNA* 4- and 5-track G4s

Lane	ODN
1	Ladder
2	G5 WT
3	G4 WT
4	G5 F Loop 2
5	G5 F Loop 1
6	G5 F Core
7	G5 OG Loop 2
8	G5 OG Loop 1
9	G5 OG Core
10	G4 F Loop 2
11	G4 F Loop 1
12	G4 F Core
13	G4 OG Loop 2
14	G4 OG Loop 1
15	G4 OG Core

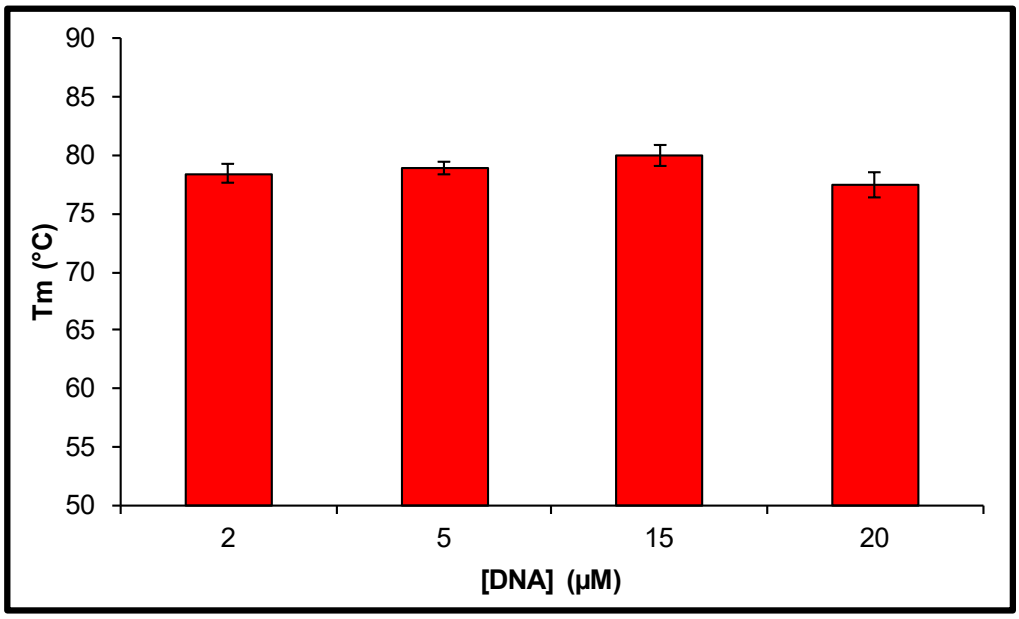


Non-denaturing PAGE analysis of lesion-containing *PCNA* 4- and 5-track sequences. *We would like to reiterate that the five-track *PCNA* sequence in all of its modified forms was extremely challenging to work with, and this is why a number of the lanes failed to give satisfactory bands. Previous studies have determined that folded G4s migrate faster than the single strands (e.g., Shrestha, P.; et al., *Nucleic Acids Res.* **2014**, *42*, 7236-7246).

Strand	**Sequence**
PCNA G5	5`-CAGGGAGGCA**G**GGC**G**ACGGGGGC**G**GGGCGGGGCG-3`
PCNA G4	5'-CA**G**GGC**G**ACGGGGGC**G**GGGCGGGGCG- 3'

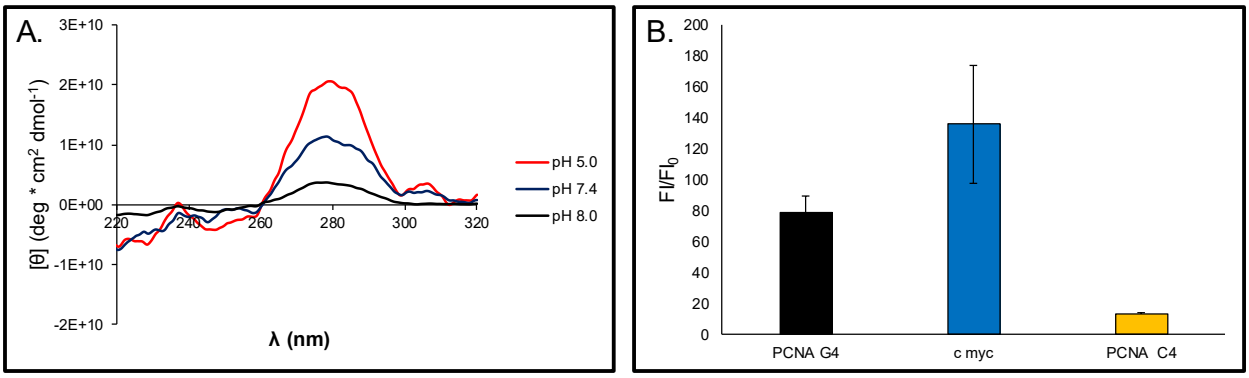
The Gs labeled red represent the sites that OG or F was studied individually.

Figure S3. T_m data supporting the *PCNA* G4 adopts an intramolecular fold



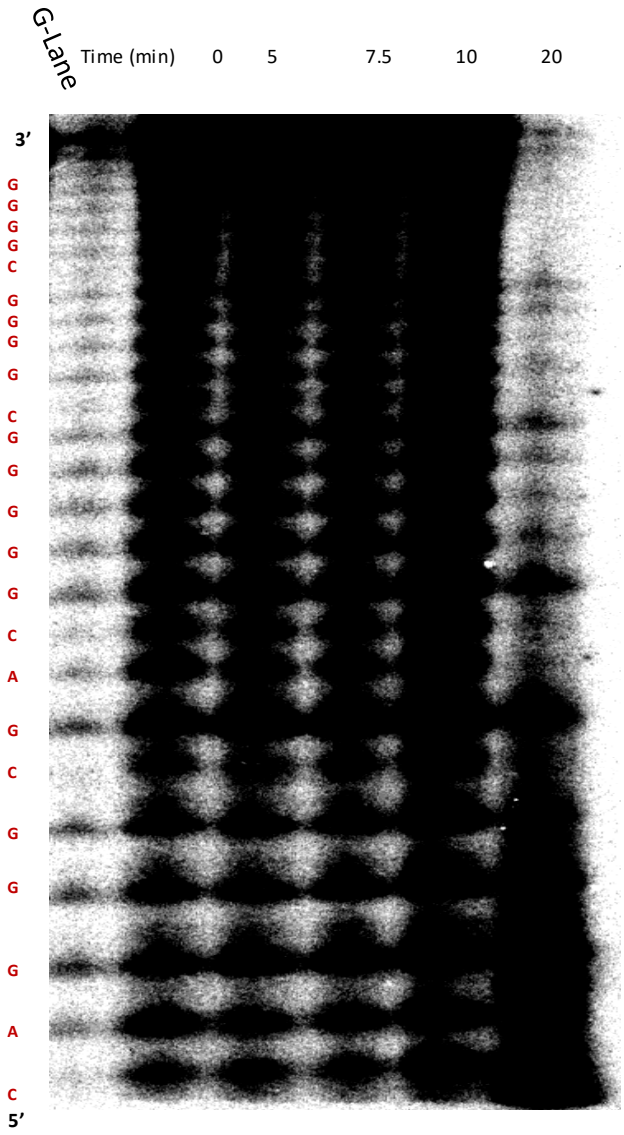
The T_m values measured when the ODN concentration was increased to support intramolecular G4 folding. The lack of a T_m value change across the concentrations studied support this claim. The *PCNA* G4 sequence studied was 5'-CAG GGC GAC GGG GGC GGG GCG GGG CG- 3'

Figure S4. Thioflavin-T data of *PCNA* 4-track and CD data for *PCNA* i-motif



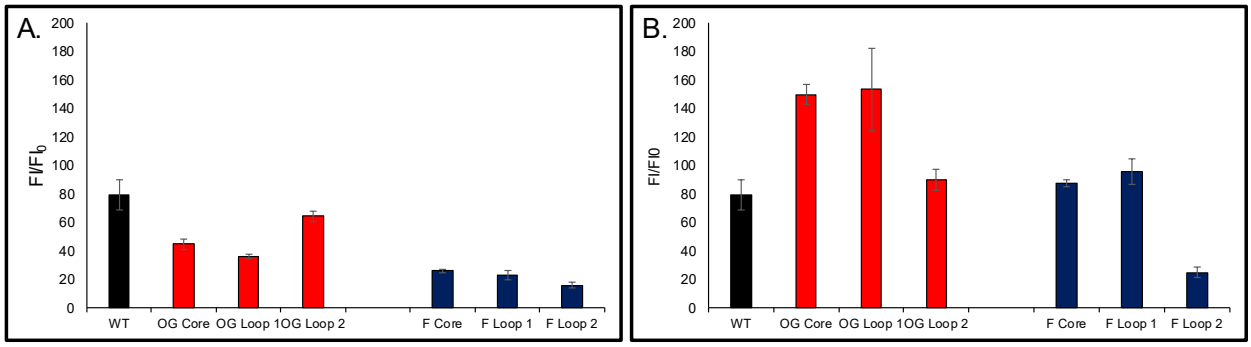
(**A**) CD spectra on the C-rich complementary strand showing the *PCNA* i-motif to be single stranded at the pH conditions of the study on the basis of the lambda max for this sequence to be at 278 nm at pH 5.0, 7.4, and 8.0. (**B**) Data showing the thioflavin-T fluorescence enhancement values for the *PCNA* G4 and C-rich complementary strand compared to the *c-MYC* positive control. The low fluorescence enhancements found for the C-rich strand supports the single-strand state for this sequence and not an i-motif fold. The *PCNA* G4 sequence studied was 5'-CAG GGC GAC GGG GGC GGG GCG GGG CG- 3', the *PCNA* C4 sequence was 5`-CG CCC CGC CCC GCC CCC GTC GCC CTG-3`, and the *c-MYC* sequence was 5`-GA GGG T GGGG A GGG T GGGG AA-3`.

Figure S5. Overexposed oxidation gel to visualize the Maxam-Gilbert G lane



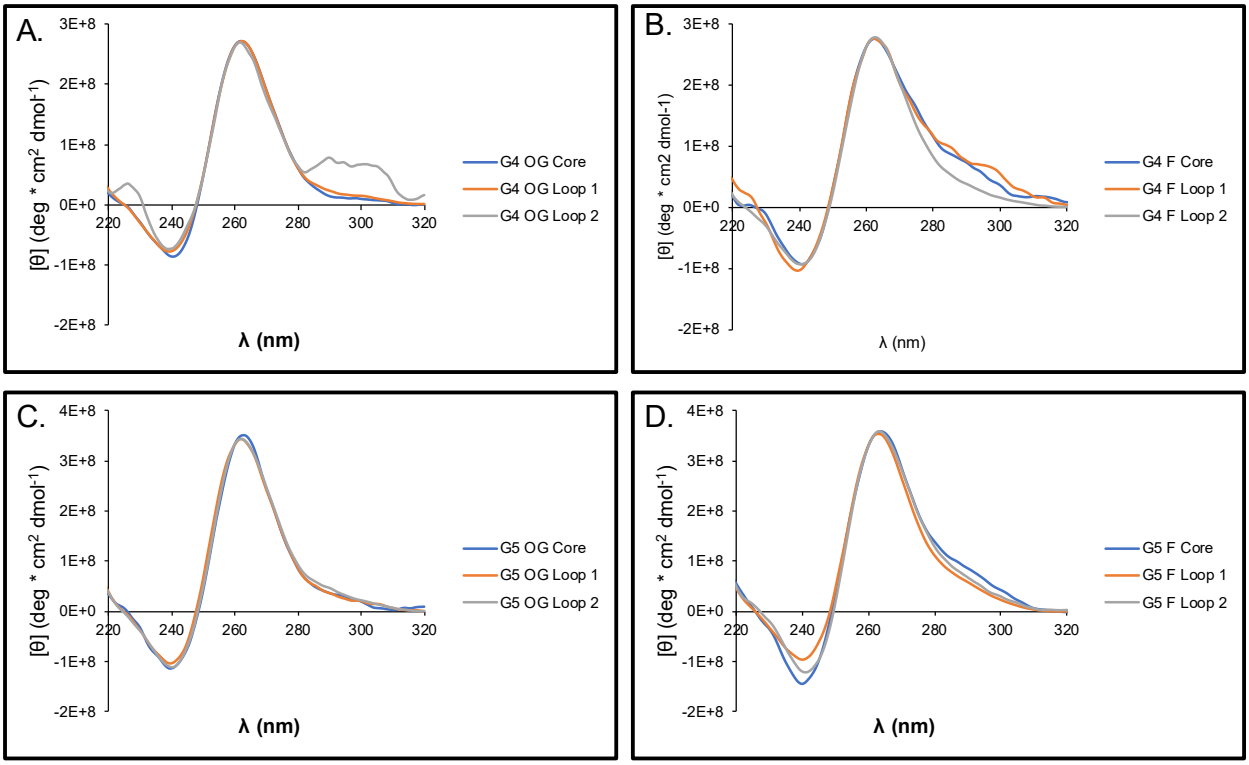
This is the overexposed PAGE shown in Figure 5 to illustrate the Maxam-Gilbert G lane (left lane) utilized to identify the G nucleotides oxidized.

Figure S6. Thioflavin-T experimentation of lesion-containing *PCNA* 4- and 5-track G4s



Thioflavin-T experimentation of lesion-containing *PCNA* 4- and 5-track G4s. (**A**) The *PCNA* 4-track data and (**B**) the *PCNA* 5-track data. The *PCNA* G4 sequence studied was 5'-CA**G** GGC **G**AC GGG GGC **G**GG GCG GGG CG- 3' and the G5 sequence studied was 5`-CAGGGAGGCA**G**GGC**G**ACGGGGGC**G**GGGCGGGGCG-3`; the Gs that are red represent the sites in which OG or F was studied.

Figure S7. CD spectra of the lesion-containing *PCNA* 4- and 5-track G4s



CD spectra of the lesion-containing *PCNA* 4- and 5-track G4s. (**A**) The *PCNA* 4-track with OG inserted in the locations of interest, (**B**) the *PCNA* 4-track with the abasic site analog F inserted in the locations of interest, (**C**) the *PCNA* 5-track with OG inserted in the locations of interest, and (**D**) the *PCNA* G5 with F inserted in locations of interest. The *PCNA* G4 sequence studied was 5'-CA**G** GGC **G**AC GGG GGC **G**GG GCG GGG CG- 3' and the G5 sequence studied was 5`-CAG GGA GGC A**G**G GC**G** ACG GGG GC**G** GGG CGG GGC G-3`; the Gs that are red represent the sites in which OG or F was studied.