Recognition of G-Quadruplexes by the FANCJ AKKQ Peptide

Submitted by

Laura Campbell

To

The Honors College

Oakland University

In partial fulfillment of the

Requirement to graduate from

The Honors College

Mentor: Dr. Colin G Wu

Department of Chemistry

Oakland University

December 3, 2021

Abstract

G-quadruplexes (G4s) are stable structures formed by Hoogsteen hydrogen bonding between four stretches of three or more consecutive guanines. If left accumulated, G4s can disrupt DNA replication, repair, and RNA transcription. Additionally, guanines are susceptible to damage from reactive oxygen species that cause guanine to thymine transversion mutations. FANCJ is a DNA helicase that participates in the repair of interstrand DNA crosslinks in human cells. FANCJ has been shown to support DNA replication through guanine rich sequences by unfolding G4 structures. Previous research has shown that FANCJ possesses an "AKKQ" amino acid motif that target G4s and is thought to target 8-oxoguanine modified G4 sequences (8oxoG4s). We hypothesize that the molecular recognition of FANCJ to 8oxoG4s is correlated to the DNA damage position, sequence composition, and stability of the substrate. To test this, we have measured the binding affinity of the FANCJ AKKQ peptide to G4 DNA that can adopt parallel, antiparallel, or hybrid spatial configurations, and we have systematically replaced the guanine bases with 8oxoGs. Circular dichroism spectroscopy (CD) was used to define the conformation of the G4 DNA, and CD data were collected as a function of solution conditions to determine the melting temperature. Fluorescence spectroscopy was used to measure the binding affinities of FANCJ AKKQ to the library of G4 DNA. These experimental outcomes will provide a detailed view of how 8oxoGs affect the physical properties of G4s and their molecular recognition by FANCJ.

Introduction

G-Quadruplexes

G-Quadruplexes (G4s) are stable structures found in guanine-rich DNA sequences. In single-stranded DNA, guanines can interact with each other through an interaction called Hoogsteen hydrogen bonding to form this secondary DNA structure. These structures can form from sequences containing four stretches of three or more consecutive guanines and are further stabilized through the presence of monovalent cations, such as potassium and sodium ions. As shown in figure 1A, the general sequence for G4s has been identified as $(GGGN_{1-7})_4$, with N representing any DNA base [1]. This sequence has led to an estimate of over 400,000 G4 motifs being identified in the human genome [2]. Although G4s are involved in gene regulation, in accumulation, G4s can be toxic to cells as they stall DNA repair, replication, and RNA transcription [3].

G4s can adopt several different spatial conformations depending on solution conditions and intervening loop sequence [4]. As shown in figure 1B, the main three conformations have been classified as parallel, antiparallel, and hybrid. These different spatial conformations have distinct optical properties and give different intensity signals when examined under CD spectroscopy [5, 6]. As shown in figure 1C, parallel conformations give a maximum signal around 262 nm and a minimum around 240 nm. Antiparallel conformations give two peak signals around 295 nm and 262 nm and a minimum around 260 nm. Finally, hybrid conformations give a peak at 295 nm, a shoulder around 260 nm, and a minimum at 240 nm.

CD spectroscopy can also be used to determine the melting temperature of G4 structures [7]. Figure 2A shows that as temperature increases, the intensity of spectral peaks and troughs

decreases. This decreases as a function of temperature can then be analyzed to determine the melting temperature, shown in figure 2B. The melting temperature of G4 structures is directly correlated to their thermal stability.

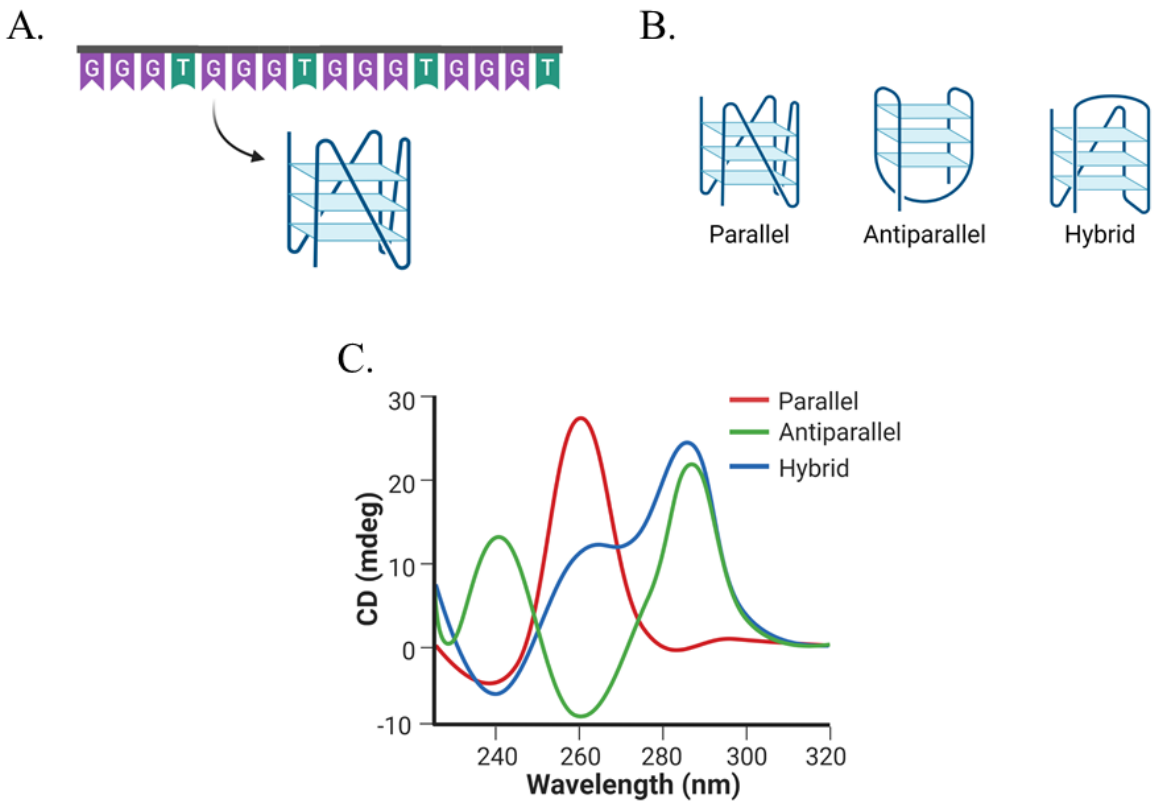

Figure 1. **(A).** Graphic representation of a general sequence for a G4. **(B).** Graphic representation of three different intra-strand G4 conformations. **(C).** Graphic representation of CD spectra for parallel, antiparallel, and hybrid G4 conformations.

Guanines are also susceptible to oxidative damage when exposed to reactive oxygen species. A guanine that has been exposed to oxidative damage is commonly referred to as an 8-oxoguanine (8oxoG). If unrepaired, these 8oxoGs can cause guanine to thymine transversion mutations during DNA replication [8]. While 8oxoGs are still able to form G4 structures, if the damage causes a significant decrease in stability, it may leave the DNA vulnerable to cellular nucleases and lead to untimely shortening of DNA [9, 10].

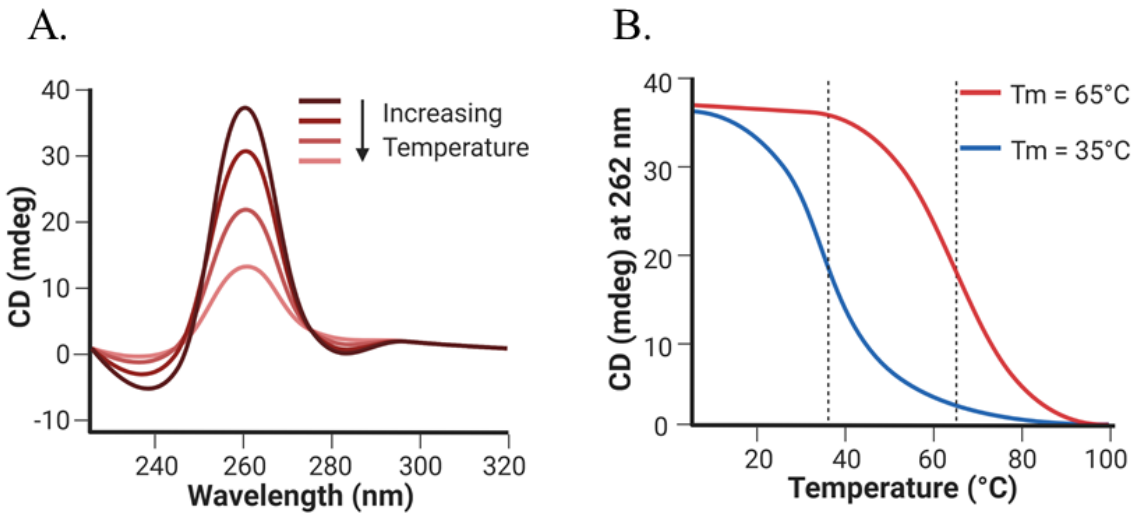

Figure 2. **(A.)** Graphic representation of the CD spectra of a parallel G4 conformation as temperature increases. **(B.)** Graphic representation of a thermal melt graph of two G4 structures.

This work will examine the thermal stability of G4 sequences with varying intervening sequence lengths as well as those same sequences with oxidative damage at the first, third or fifth guanine. It is expected that oxidative damage will decrease the thermal stability compared to the undamaged sequence. It will also be examined how intervening sequence length affects G4 conformation and thermal stability. These sequences will be examined in the presence of potassium ions and in the presence of sodium ions.

FANCJ

FANCJ is an iron-sulfur containing human DNA repair helicase that participates in the repair of DNA damage through the DNA cross-link repair pathway [11]. Mutations to the gene can result in the development of Fanconi anemia and an increased risk of cancer [11-13]. Previous studies have shown FANCJ coordinates the repair of DNA intermediates that arise from the nucleotide excision repair pathway [14-16]. Specifically, FANCJ has been shown to bind to G4

DNA and supports DNA replication through G4 rich regions by unwinding G4 structures [14, 17]. The current model for FANCJ supported replication is illustrated in figure 3B. When DNA polymerases encounter a folded G4 structure, the protein is blocked from adding the next base pair. FANCJ recognizes the G4 structure, binds and unfolds it, and recruits REV1, a translesion polymerase, to continue with replication [18].

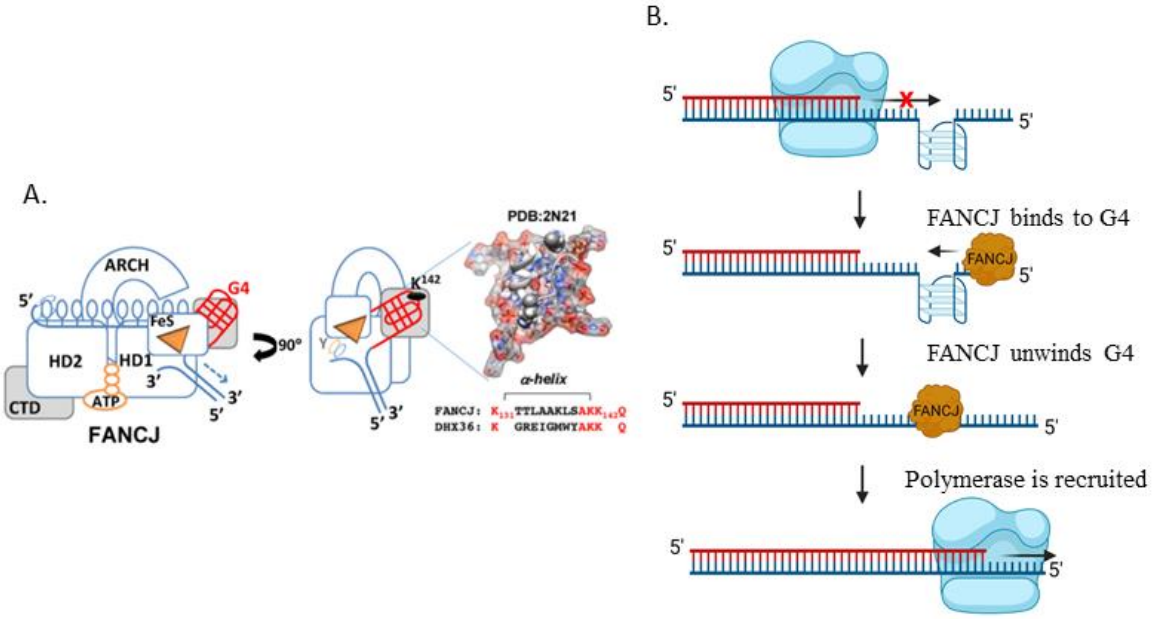

Figure 3. **(A.)** Molecular model of FANCJ. FANCJ possesses a G4 recognition site capable of binding to G4 structures. FANCJ also possesses an AKKQ motif similar in structure and sequence to the AKKQ binding motif found in G4 binding protein DHX36 [18]. **(B).** Model for FANCJ supported DNA replication through G4 rich sequences. DNA polymerase stalls when it encounters a G4 structure. FANCJ binds to the G4 structure and unwinds it. Another polymerase is recruited to continue with replication.

An unrelated RNA helicase, RHAU, has been identified as able to bind to G4s and 8oxoG4s through an AKKQ targeting motif in which two lysine residues anchor into the DNA backbone [19, 20]. FANCJ has also been identified as possessing a similar AKKQ motif at K141/K142 [16].

This motif is known to be responsible for the binding of FANCJ to G4 structures however it is unknown if this motif can also identify 8oxoG4s. This motif is illustrated in figure 3A.

This work will examine the interactions of FANCJ AKKQ with G4s and 8oxoG4s of varying intervening sequence lengths and damage position. These interactions will then be related back to G4 conformation and thermal stability to further characterize the recognition of G4 and 8oxoG4 sequences by FANCJ AKKQ.

Methods

Buffers

All buffers were prepared with reagent-grade chemicals and water purified with a Smart2Pure 6 UV/UF system. Solutions were further filtered through a 0.22-micron PES filter.

Peptide and DNA oligos

FANCJ AKKQ peptide, 129-PEKTTLAAKLSAKKQASIW-147, was synthesized by Genscript, and the powder dissolved into 20 mM HEPES pH 7.5, 150 mM KCl, 5 mM TCEP, and 5% (v/v) glycerol. Protein concentration was determined using a NanoDrop One UV-Vis microvolume spectrophotometer and molar extinction coefficient. Peptide samples were aliquoted, frozen, and stored at -20 °C. Oligodeoxyribonucleotides were purchased from Integrated DNA Technologies and concentrations were determined similarly to the above stated. DNA substrates were stored at 4 °C. A list of sequences used is described in the table below:

Table of DNA sequences

Name	DNA Sequence 5' to 3'
GGGT	GGGTGGGTGGGTGGGT
GGGT 8oxo1	T/i8oxodG/GGTGGGTGGGTGGGT
GGGT 8oxo3	GG/i8oxodG/TGGGTGGGTGGGT
GGGT 8oxo5	GGGTG/i8oxodG/GTGGGTGGGT
GGGTT	GGGTTGGGTTGGGTTGGGTT
GGGTT 8oxo1	T/i8oxodG/GGTTGGGTTGGGTTGGGTT
GGGTT 8oxo3	GG/i8oxodG/TTGGGTTGGGTTGGGTT
GGGTT 8oxo5	GGGTTG/i8oxodG/GTTGGGTTGGGTT
TTAGGG	TTAGGGTTAGGGTTAGGGTTAGGG
TTAGGG 8oxo1	TTA/i8oxodG/GGTTAGGGTTAGGGTTAGGG
TTAGGG 8oxo3	TTAGG/i8oxodG/TTAGGGTTAGGGTTAGGG
TTAGGG 8oxo5	TTAGGGTTAG/i8oxodG/GTTAGGGTTAGGG

Circular Dichroism Spectroscopy (CD)

CD measurements were performed on a JASCO J-810 spectropolarimeter equipped with a PTC-423S peltier system and Koolance Liquid Cooling System. G4 DNA samples were dialyzed into 20 mM Boric Acid, 150 mM KCl or NaCl. CD spectra scans were taken from 220 nm to 320 nm at 25 °C with DNA samples of 0.3 mg/mL. Five traces for each DNA sample were taken and averaged. A reference scan of buffer alone was subtracted from the averaged scan. Temperature melt scans were taken from 15 °C to 95 °C at 262 nm (for parallel conformations) or 295 nm (for antiparallel and hybrid conformation). Samples were heated to 95 °C and allowed to cool overnight prior to the first melt. Three trials for each DNA sample were taken and averaged.

Melting temperature was determined using SigmaPlot as an analysis program and the model described by Greenfield for a two-state transition of a monomer from a folded to an unfolded state [21].

Fluorescence Spectroscopy

Fluorescence spectroscopy measurements were performed on a Cary Eclipse Fluorescence Spectrophotometer equipped with a PCB 1500 water peltier system. Spectra were taken from 300 nm to 450 nm (excitation at 280 nm) at 25 °C. A scan of buffer alone, 20 mM HEPES pH 7.5, 150 mM KCl, 5 mM TCEP, and 5% (v/v) glycerol, was used as a baseline. FANCJ AKKQ peptide (5 μM) was titrated with additions of a DNA substrate. The mixture was allowed to equilibrate for 3 minutes before taking each spectra scan.

Observed fluorescence quenching (ΔF) was calculated using the following equation, with F_0 representing total fluorescence from the FANCJ peptide alone and F_i representing the total fluorescence after the *i*th addition of DNA:

$$\Delta F = \frac{F_0 - F_i}{F_0}$$

Binding isotherms were calculated by plotting ΔF versus total DNA concentration. The resulting curve was fit to a 1:1 interaction model using Scientist 2.0 software to calculate an equilibrium K. The equilibrium K was calculated using the following formulas, where *A* was the amplitude of fluorescence quenching and *K* was the equilibrium association constant. D_f and D_t were the free and total concentration of DNA, while P_f and P_t described the free and total peptide concentration. Fluorescence titrations were performed in triplicate. The reported K values were determined from the average and 95% confidence interval of the three independent data-sets.

$$\Delta F_{obs} = A\left(\frac{KD_f}{1+KD_f}\right)$$
$$D_t = D_f(1+KP_f)$$
$$P_t = P_f(1+KD_f)$$

Results

G-Quadruplex Conformation and Thermal Stability

CD spectroscopy was used to determine G4 conformation and thermal stability. Three undamaged DNA sequences were chosen based on intervening sequence length and relevant importance. The sequence TTAGGG was chosen because it is the human telomeric DNA sequence and has significant biological relevance. The sequence GGGT was chosen because the single thymidine base produces a compact G4 structure. Additionally, GGGT is a HIV-1 integrase aptamer and is of interest to study [22]. Finally, the sequence GGGTT is an artificial sequence and was designed to have an intervening sequence of intermediate length. For each sequence, damaged positions were chosen where an 8-oxoguanine modified base was substituted for the first (8oxo1), third (8oxo3), or fifth (8oxo5) guanine. These positions were chosen because previous studies showed that, for the human telomeric sequence, 8oxo5 produced the most decrease in thermal stability, 8oxo1 produced the least decrease to thermal stability and 8oxo3 produced an intermediate level of thermal stability [10, 23]. Additionally, these damaged positions place an 8-oxoguanine in each of the three tetrads formed in a G4 structure. These positions are visualized in Figure 4.

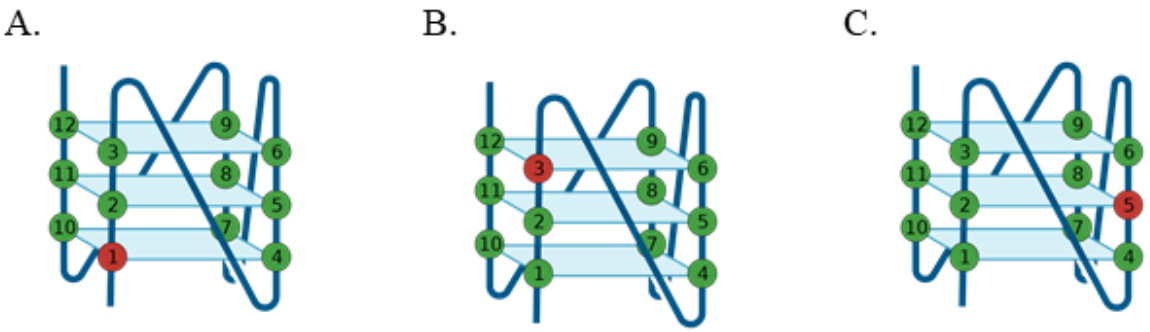

Figure 4. Graphic representation of a parallel G4 conformation with an 8-oxoguanine modified base substituted at the **(A.)** first guanine (8oxo1), **(B).** third guanine (8oxo3), and **(C).** fifth guanine (8oxo5).

Figure 5 shows the CD spectra and thermal melts of all DNA sequences of interest in the presence of KCl. It was found that all GGGT sequences in KCl produced a parallel conformation with a peak around 262 nm (Figure 5A left). Damaged GGGT sequences did not show any structural changes based on the overlaying CD spectra traces. Thermal melts show that GGGT has a melting temperature of 88.4 ± 8.4 °C (Figure 5A right and Table 1). 8oxo1 produced the highest decrease in thermal stability of ΔT_m = -19.0 °C while 8oxo3 produced the lowest change in thermal stability of ΔT_m = -8.1 °C.

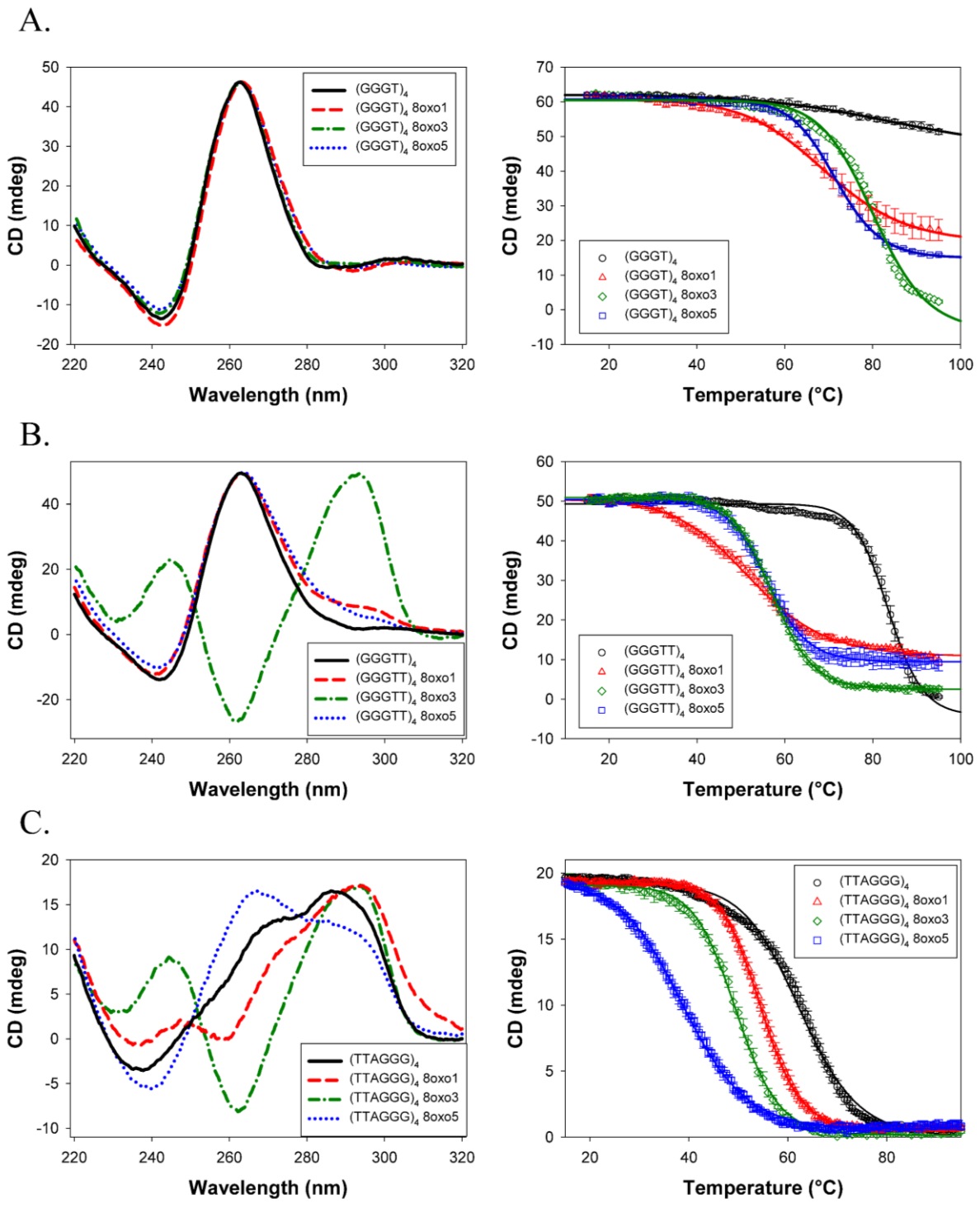

Figure 5. All experiments performed in the presence of 150 mM KCl. **(A.) (**left) CD spectra of $(GGGT)_4$, 8oxo1, 8oxo3, and 8oxo5 taken from 320 to 220 nm. (right) Temperature melt of same sequences. Melt taken at 262 nm from 15 to 90 °C. **(B.)** (left) CD spectra of $(GGGTT)_4$, 8oxo1, 8oxo3, and 8oxo5 taken from 320 to 220 nm. (right) Temperature melt of same sequences. Melt taken at 262 nm for all sequences except 8oxo3, which was taken at 295 nm. **(C.)** (left) CD spectra of $(TTAGGG)_4$, 8oxo1, 8oxo3, and 8oxo5 taken from 320 to 220 nm. (right) Temperature melt of the same sequences. Melt taken at 295 nm.

Figure 5B (left) shows that GGGTT also forms a parallel conformation in the presence of KCl. 8oxo1 and 8oxo5 modified sequences also produce a parallel conformation with slight structural changes based on the differences in traces in the 300 to 280 nm range. The 8oxo3 modified sequence changes GGGTT to an antiparallel sequence. Damage at the third position likely causes a significant amount of structural instability to cause the sequence to adopt a different conformation in which it is more stable. Thermal melts show that GGGTT has a melting temperature of 83.5 ± 0.7 °C (Figure 5B right and Table 1). Similar to GGGT in KCl, 8oxo1 also produced the highest decrease in thermal stability of ΔT_m = -31.5 °C while 8oxo3 produced the lowest change in thermal stability of ΔT_m = -26.3 °C.

Figure 5C (left) shows that TTAGGG forms a hybrid conformation in the presence of KCl. 8oxo1 and 8oxo5 modified sequences also form a hybrid conformation with significant structural changes. This is shown by 8oxo1 by the 295 nm peak shift to the right and double trough at 260 nm and 240 nm. 8oxo5 experiences depression in its peak at 295 nm and an increase in its shoulder at 260 nm and trough at 240 nm. Similar to GGGTT, the 8oxo3 modified sequence experienced a change to the antiparallel conformation. Thermal melts show that TTAGGG has a melting temperature of 62.5 ± 0.5 °C (Figure 5C right and Table 1). Concurrent with previous studies, 8oxo5 also produced the highest decrease in thermal stability of ΔT_m = -23.6 °C while 8oxo1 produced the lowest change in thermal stability of ΔT_m = -8.0 °C [10].

Table 1. Summary of G4 conformation and meting temperature in 150 mM KCl

Name	Conformation	Melting Temperature (°C)	Δ Melting Temperature (°C)
GGGT	Parallel	88.4 ± 8.4	—
GGGT 8oxo1	Parallel	69.4 ± 0.6	-19.0
GGGT 8oxo3	Parallel	80.3 ± 0.4	-8.1
GGGT 8oxo5	Parallel	71.6 ± 0.4	-16.8
GGGTT	Parallel	83.5 ± 0.7	—
GGGTT 8oxo1	Parallel	52.0 ± 1.4	-31.5
GGGTT 8oxo3	Antiparallel	57.2 ± 0.1	-26.3
GGGTT 8oxo5	Parallel	55.9 ± 0.3	-27.6
TeloG4	Hybrid	62.5 ± 0.5	—
TeloG4 8oxo1	Hybrid-esk	54.5 ± 0.5	-8.0
TeloG4 8oxo3	Antiparallel	49.5 ± 0.3	-13.0
TeloG4 8oxo5	Hybrid-esk	38.9 ± 0.9	-23.6

Figure 6 shows the CD spectra and thermal melts of all the sequences of interest in the presence of NaCl. Figure 6A (left) shows that all GGGT sequences also form parallel conformations in NaCl. All oxidatively damaged sequences show a slight decrease in structural stability as shown by the differences in traces in the 300 to 280 nm region. Thermal melts show that GGGT has a melting temperature of 70.3 ± 0.7 °C (Figure 6A right and Table 2). Similar to GGGT in KCl, 8oxo1 also produced the highest decrease in thermal stability of ΔT_m = -27.8 °C while 8oxo3 produced the lowest change in thermal stability of ΔT_m = -13.9 °C.

Figure 6B (left) shows that GGGTT forms a parallel conformation in the presence of NaCl. All 8oxoG modified sequences had a change in conformation. 8oxo1 and 8oxo5 produced an antiparallel conformation while 8oxo5 produced a hybrid conformation. Thermal melts show that GGGTT has a melting temperature of 44.8 ± 1.2 °C (Figure 6B right and Table 2). 8oxo1 and 8oxo3 did not show a significant change in thermal stability, with 8oxo1 even having an increase in thermal stability of ΔT_m = 2.4 °C. This is likely because 8oxo1 and 8oxo3 experienced a change

in conformation. 8oxo5 produced the largest change in thermal stability of ΔT_m = -17.3 °C. The thermal melt for GGGTT 8oxo5 was taken over -5 to 75 °C because it did not appear fully folded at 15 °C.

Table 2. Summary of G4 conformation and meting temperature in 150 mM NaCl

Name	**Conformation**	**Melting Temperature (°C)**	**Δ Melting Temperature (°C)**
GGGT	Parallel	70.3 ± 0.7	—
GGGT 8oxo1	Parallel	42.5 ± 0.5	-27.8
GGGT 8oxo3	Parallel	56.4 ± 0.4	-13.9
GGGT 8oxo5	Parallel	44.0 ± 0.3	-26.3
GGGTT	Parallel	44.8 ± 1.2	—
GGGTT 8oxo1	Antiparallel	47.2 ± 0.3	2.4
GGGTT 8oxo3	Antiparallel	44.6 ± 0.2	-0.2
GGGTT 8oxo5	Hybrid-esk	27.5 ± 0.8	-17.3
TeloG4	Antiparallel	52.1 ± 0.3	—
TeloG4 8oxo1	Antiparallel	47.8 ± 0.6	-4.3
TeloG4 8oxo3	Antiparallel	42.2 ± 0.7	-9.9
TeloG4 8oxo5	Antiparallel	33.3 ± 0.4	-18.8

Finally, Figure 6C (left) shows that TTAGGG forms an antiparallel conformation in NaCl. 8oxoG modified sequences also folded into an antiparallel conformation with slight structural instabilities. 8oxo1 shows a slight shift in the trough to the right at 260 nm, 8oxo3 shows a slight increase in intensity in its peak at 240 nm and trough at 260 nm, and 8oxo5 shows a decrease in intensity of its trough at 260 nm and a shift of its peaks towards the center. Thermal melts show that TTAGGG has a melting temperature of 52.1 ± 0.3 °C (Figure 6C right and Table 2). Concurrent with previous studies and KCl data shown above, 8oxo5 produced the highest decrease in thermal stability of ΔT_m = -18.8 °C while 8oxo1 produced the lowest change in thermal stability of ΔT_m = -4.3 °C.

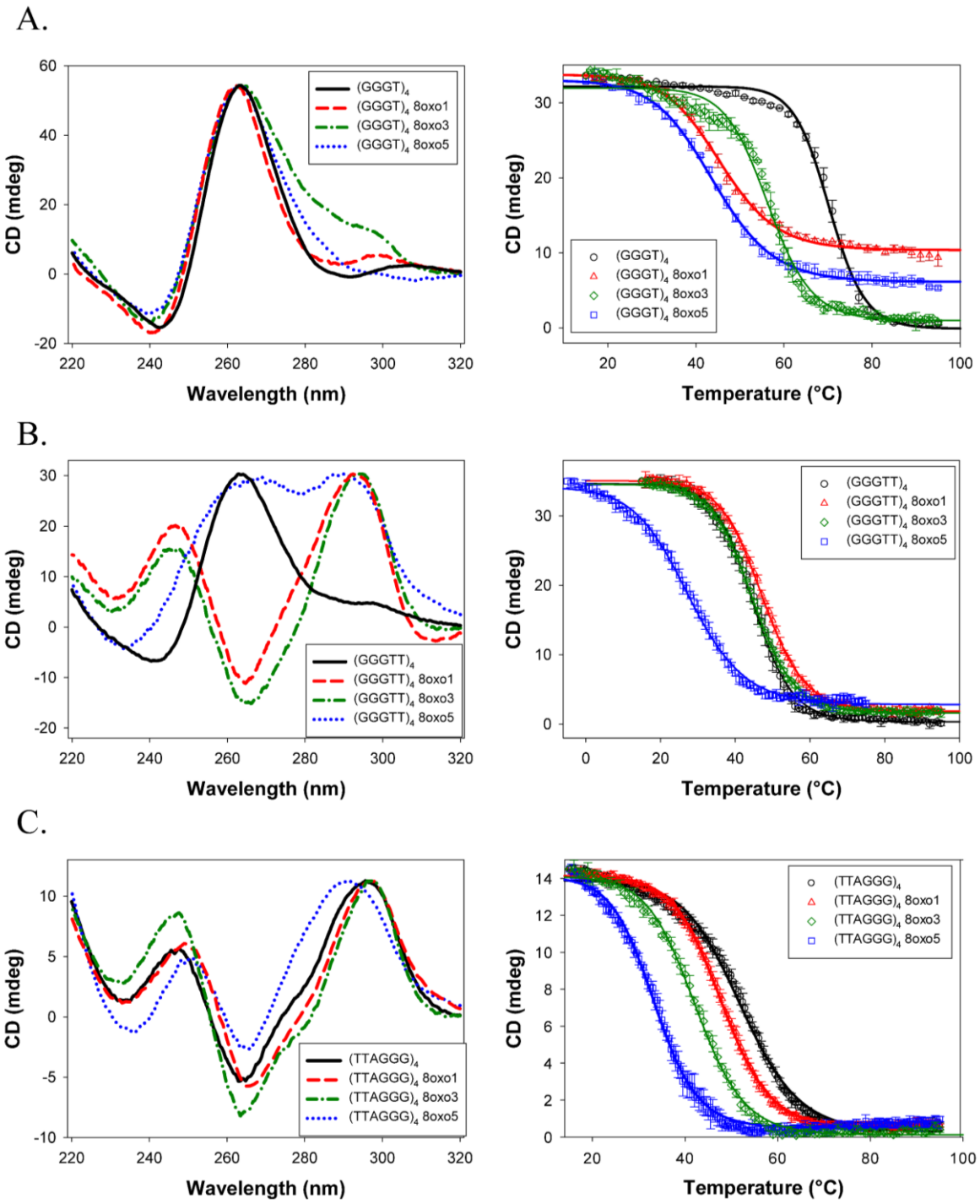

Figure 6. All experiments performed in the presence of 150 mM NaCl. **(A.) (**left) CD spectra of $(GGGT)_4$, 8oxo1, 8oxo3, and 8oxo5 taken from 320 to 220 nm. (right) Temperature melt of same sequences. Melt taken at 262 nm from 15 to 90 °C. **(B.)** (left) CD spectra of $(GGGTT)_4$, 8oxo1, 8oxo3, and 8oxo5 taken from 320 to 220 nm. (right) Temperature melt of same sequences. Melt taken at 262 nm for all sequences except 8oxo3, which was taken at 295 nm. **(C.)** (left) CD spectra of $(TTAGGG)_4$, 8oxo1, 8oxo3, and 8oxo5 taken from 320 to 220 nm. (right) Temperature melt of the same sequences. Melt taken at 295 nm.

FANCJ AKKQ Binding

Fluorescence spectroscopy titrations were used to determine interaction strength between AKKQ and G4 sequences. As shown in figure 7A, the AKKQ sequence examined was modified to contain a tryptophan residue, W, at the C-terminal end of the sequence. Due to its aromatic properties, tryptophan produces a fluorescence peak from 300 nm to 450 nm. When bound the peptide binds to DNA, this signal is quenched, and the strength of binding can be determined from the decrease in fluorescence intensity plotted against increasing DNA concentration. The resulting curve was fit to a 1:1 equilibrium binding interaction model using a nonlinear least squared analysis to generate a K value. This K value represents an equilibrium K value and a larger K value can be correlated to a stronger interaction strength between peptide and DNA. This is demonstrated in figure 7B-C.

Figure 8A shows the binding interactions between FANCJ AKKQ and GGGT. There is a clear shift in the binding curve from the native sequence to the damaged sequences, indicating an increase in binding affinity. These binding affinities are listed in Table 3. AKKQ had the strongest interaction strength with GGGT 8oxo1 with a K value of $7.7 \pm 0.7 \times 10^5\ M^{-1}$. This was followed by 8oxo5 with a K value of $4.4 \pm 0.6 \times 10^5\ M^{-1}$, then 8oxo3 with a K value of $3.2 \pm 0.1 \times 10^5\ M^{-1}$ and had the weakest binding affinity to the undamaged sequence with a K value of $2.7 \pm 0.4 \times 10^5\ M^{-1}$.

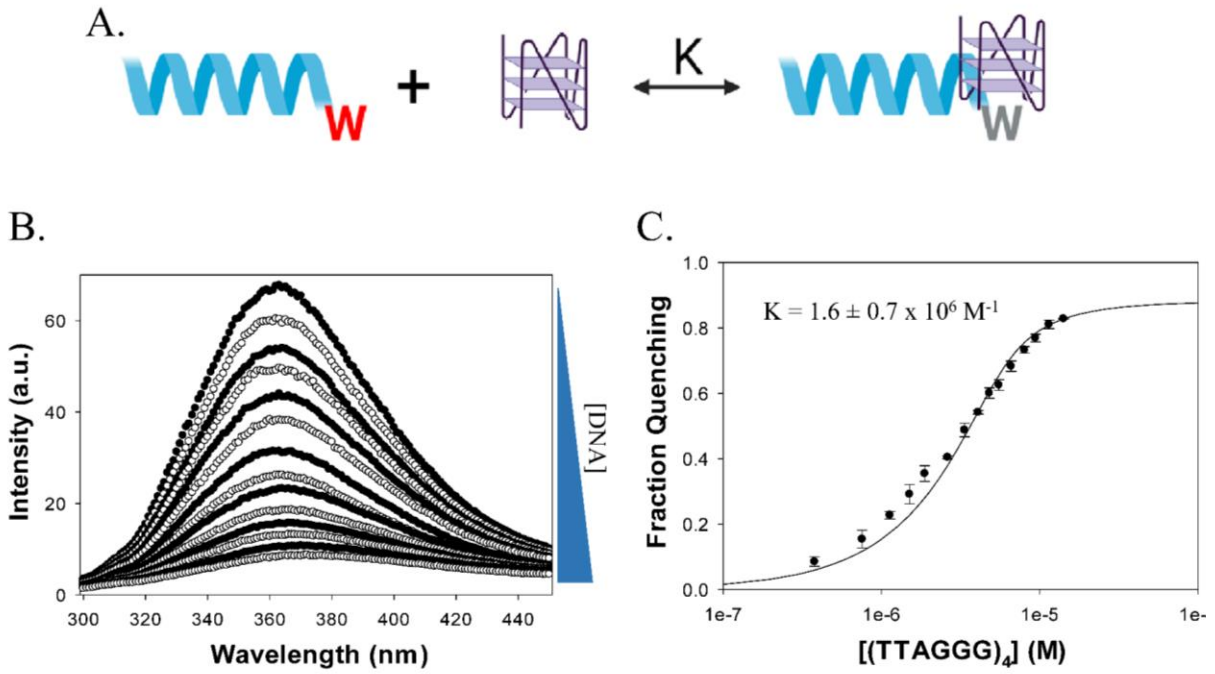

Figure 7. **(A.)** Molecular model of fluorescence spectroscopy titrations. AKKQ is equipped with a tryptophan residue that fluoresces. When bound to DNA the residue is quenched and fluorescence intensity decreases. **(B.)** Titration of AKKQ with $(GGGT)_4$. As concentration of DNA increases, intensity peak decreases. **(C.)** Analysis of fluorescence titration. The decrease in integrated peak is plotted as a function of DNA concentration. The resulting isotherm is fit to a 1:1 binding interaction model to generate an equilibrium K value.

Figure 8B shows the binding interactions between FANCJ AKKQ and GGGTT. There is a slight shift in the binding curve from FANCJ AKKQ binding to the native sequence and respective 8oxoG4 sequences. However, there is no change in interaction between FANCJ AKKQ and different damage positions. These binding affinities are also summarized in Table 3. The interaction of FANCJ AKKQ with the undamaged sequence had a K value of $6.0 \pm 0.9 \times 10^5\ M^{-1}$. Each respective 8oxoG4 sequence had similar K values of approximately $8\text{-}9 \times 10^5\ M^{-1}$.

Figure 8C shows the binding interactions between FANCJ AKKQ and TTAGGG. From the graph, there is no shift in the curve between the undamaged and each damaged sequence,

indicating that FANCJ AKKQ binds with relatively the same affinity to each sequence. The K values listed in Table 3 further support this with each K value being within error of 1.3 x 10^6 M^{-1}.

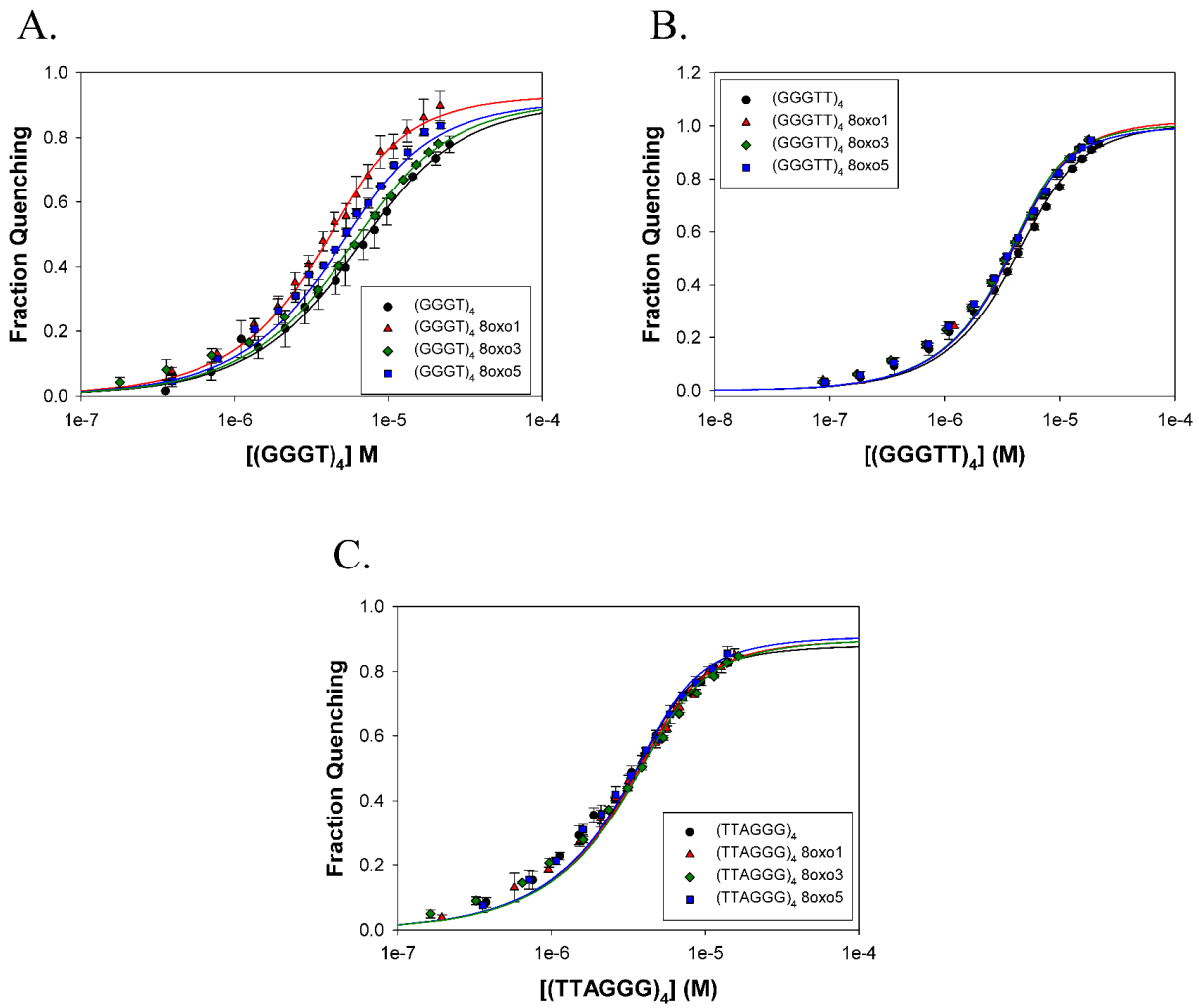

Figure 8. Summary of FANCJ AKKQ-G4 binding by fluorescence spectroscopy. Trials performed in 20 mM HEPES pH 7.5, 150 mM KCl, 5 mM TCEP, and 5% (v/v) glycerol **(A.)** FANCJ AKKQ equilibrium binding to GGGT and respective 8oxoG4 sequences. **(B.)** FANCJ AKKQ binding to GGGTT and respective 8oxoG4 sequences. **(C.)** FANCJ AKKQ binding to TTAGGG and respective 8oxoG4 sequences.

Table 3. Summary of FANCJ AKKQ-G4 binding by fluorescence spectroscopy

Name	K (M^{-1})	R^2
GGGT	$2.7 \pm 0.4 \times 10^5$	0.997
GGGT 8oxo1	$7.7 \pm 0.7 \times 10^5$	0.998
GGGT 8oxo3	$3.2 \pm 0.1 \times 10^5$	0.997
GGGT 8oxo5	$4.4 \pm 0.6 \times 10^5$	0.997
GGGTT	$6.0 \pm 0.9 \times 10^5$	0.997
GGGTT 8oxo1	$8.2 \pm 1.0 \times 10^5$	0.997
GGGTT 8oxo3	$9.6 \pm 1.0 \times 10^5$	0.996
GGGTT 8oxo5	$9.3 \pm 0.8 \times 10^5$	0.997
TTAGGG	$1.6 \pm 0.7 \times 10^6$	0.996
TTAGGG 8oxo1	$1.3 \pm 0.4 \times 10^6$	0.997
TTAGGG 8oxo3	$1.2 \pm 0.1 \times 10^6$	0.996
TTAGGG 8oxo5	$1.4 \pm 0.5 \times 10^6$	0.997

Discussion

G-Quadruplexes

The CD data for the human telomeric sequences is concurrent with those of previous studies in that TTAGGG and respective 8oxoG4 sequences form hybrid and antiparallel conformations in potassium ions and antiparallel conformations in NaCl. Additionally, 8oxo5 modified sequences produced the most thermal instability while 8oxo1 produced the least in both KCl and NaCl solutions [10]. The pattern of decrease in thermal stability was also seen by the GGGTT in sodium ions. Additionally, while this sequence formed a parallel conformation, its 8oxoG4 analogs all formed antiparallel or hybrid conformations.

A different stability pattern was seen for the other substrates. GGGT sequences in KCl and NaCl and GGGTT sequences in KCl formed predominately parallel conformations, with only GGGTT 8oxo3 forming an antiparallel conformation. These sequences in their respective ions all

followed the same decrease in thermal stability as well, with the 8oxo1 modified sequence producing the most thermal instability and 8oxo3 producing the least.

Based on these results, it is likely that the effect of an 8-oxoguanine on the thermal stability of the G4 structure depends on sequence position and G4 conformation. G4 sequences that formed the compact parallel conformations were most destabilized by damage to the first guanine, which resides in the bottom tetrad of the G4 structure, and were least destabilized by damage to the third guanine, which resides in the top tetrad. On the other hand, G4 sequences that formed the more flexible antiparallel or hybrid conformation were most destabilized by damage to the fifth guanine, which resides in the middle tetrad, and were least destabilized by damage to the first guanine.

Of monovalent cations, potassium and sodium ions are of the most physiological relevance and are most commonly studied in relation to G4 conformation and stability. While GGGT and its respective 8oxoG4 derivatives formed parallel conformations in both KCl and NaCl solutions, GGGTT and TeloG4 sequences mostly did not fold into the same conformations. Additionally, thermal stability for all sequences was lower in NaCl than in KCl. This shows that sodium ions are less effective at stabilizing the G4 structure. In the presence of only sodium ions, less stable G4 sequences are unable to adopt the rigid conformations that are formed in potassium ions and settle into less stable conformations. These results are concurrent with previous studies, and it is likely that while both ions are capable of stabilizing G4 structures, potassium ions are of optimal size and ionic strength to do so [24].

Another interesting observation is that in KCl there was a correlation between intervening sequence length and thermal stability of the native sequences. GGGT had the highest melting temperature while TTAGGG had the lowest melting temperature. This indicates that as the length

of the intervening sequence increases, thermal stability of the G4 structure decreases. This observation is in agreement with that of other data [25, 26]

FANCJ AKKQ and G-Quadruplex Binding

The results suggest that FANCJ AKKQ can target G4s and 8oxoG4s, however, its molecular recognition of these substrates depends on sequence composition and damage position. Since the fluorescence titrations were performed in the presence of KCl, binding affinities can only be compared to CD data that was also performed in KCl.

For TTAGGG, which forms the less rigid hybrid or antiparallel conformation, AKKQ bound with relatively the same affinity to damaged and undamaged sequences. In this case, damage to the G4 structure or thermal stability due to oxidative damage does not seem to affect AKKQ binding. Since hybrid and antiparallel structures better expose the DNA loop, this indicates that AKKQ preferentially targets the intervening sequence.

For GGGTT, which forms predominately parallel conformations, AKKQ bound stronger to the 8oxoG4 sequences than its native sequence. However, there was no significant change in affinity between different damage positions. Assuming FANCJ AKKQ is targeting the intervening sequence, this increase in binding can be explained by destabilization of the G4 structure. Although 8oxo1 and 8oxo5 still formed parallel conformations, the traces from the CD spectra (figure 5b left) show destabilization in the 280 to 300 nm range. This destabilization could have resulted in the intervening sequence being more exposed than in an undamaged parallel conformation, thereby allowing FANCJ AKKQ to target it better.

For GGGT, which forms a compact parallel conformation, AKKQ bound stronger to the damaged sequences than its native sequence. Additionally, the increase in binding affinity between

AKKQ and 8oxoG4 damage positions correlates to the decrease in thermal stability with 8oxoG4 damage position. 8oxo1, which showed the greatest decrease in thermal stability, had the strongest interaction with AKKQ. 8oxo3, which showed the least decrease in thermal stability, had the weakest interaction with AKKQ of the damaged substrates. Since there was no change in structural stability, as indicated by the CD spectra in figure 5a (left), this suggests that AKKQ is targeting the 8-oxoguanine position.

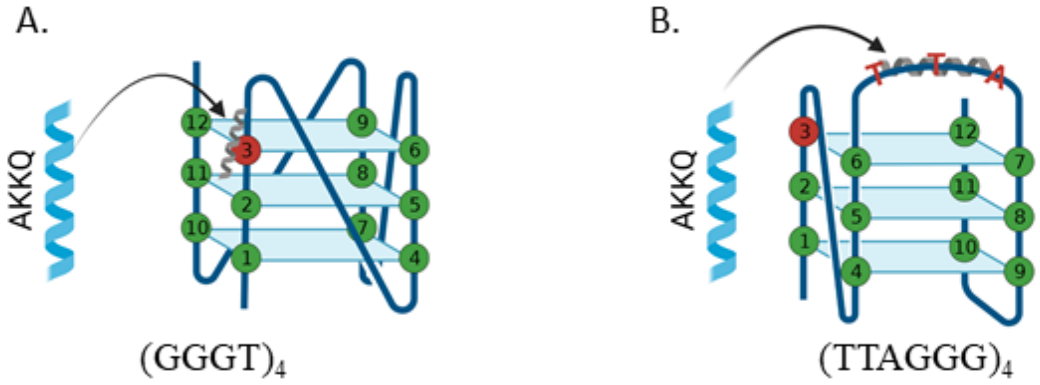

Figure 9. Proposed binding interaction model between FANCJ AKKQ and 8oxoG4 sequences. **(A.)** For G4 sequences that form compact structures with short intervening sequences, such as GGGT, AKKQ preferentially targets the 8-oxoguanine. **(B.)** For G4 sequences that expose the intervening sequence, such as TTAGGG, AKKQ preferentially targets the intervening sequence and ignores the 8-oxoguanine.

Taken together, this suggests that FANCJ AKKQ targets the intervening sequences. Additionally, between the undamaged sequences, binding interaction increased as DNA loop sequence increased. This further supports that FANCJ AKKQ binds preferentially to the intervening sequence. A proposed interaction model for AKKQ and 8oxoG4 sequences is illustrated in figure 9. For compact G4 sequences, such as GGGT, AKKQ targets the 8-oxoguanine

site preferentially and promotes the repair of such sequences. For G4s that expose the intervening sequences, such as with antiparallel or hybrid conformation, AKKQ binds preferentially to this sequence and ignores the 8-oxoguanine.

Since FANCJ is a DNA repair helicase, it was expected that FANCJ AKKQ would bind preferentially to 8-oxoguanines to promote their repair. However, this was only the case for GGGT and the results suggest that FANCJ AKKQ actually targets the intervening sequence. Past work has shown that FANCJ interacts with the translesion polymerase REV1 for the successful replication of G4 rich DNA strands. Studies have shown that REV1 is also able to unwind G4 structures [27]. It is possible that FANCJ and REV1 are responsible for the repair of different types of G4 DNA. While FANCJ binds stronger to G4s with exposed DNA loops, REV1 may be responsible for the unwinding of more compact G4 structures. Future work with this project should aim to characterize REV1's interactions with G4s and 8oxoG4s. Additionally, for the scope of this project G4 damage positions were only examined at the first, third, and fifth guanine. For the future of this project, it will be interesting to see if the trends discovered here continue if 8-oxoguanines are systematically substituted for each G4 position.

References

1. Huppert JL, and Balasubramanian S, Prevalence of quadruplexes in the human genome. *Nucleic Acids Res*, 2005. 33(9): p. 2908–2916.

2. Lombardi PE, Holmes A, Verga D, Teulade-Fichou M.P, Nicolas A. and Londoño-Vallejo A, Thermodynamically stable and genetically unstable G quadruplexes are depleted in genomes across species. *Nucleic Acids Res*, 2019. 47(12): p. 6098–6113.

3. Rhodes D and Lipps H, G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res*, 2015. 43(18): p. 8627-8637.

4. Hazel P, Huppert J, Balasubramanian S and, Neidle S, Loop-Length-Dependent Folding of G-Quadruplexes. *J. Am. Chem. Soc.* 2004. 126(50): p. 16405–16415.

5. Burge S, Parkinson GN, Hazel P, Todd AK and Neidle S, Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*, 2006. 34(19): p. 5402-15.

6. Del Villar-Guerra R, Trent JO and Chaires JB, G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew Chem Int Ed Engl*, 2018. 57(24): p. 7171-7175.

7. Rachwal PA, Fox KR, Quadruplex melting. *Methods.* 2007. 43(4): p. 291-301.

8. Neeley WL and Essigmann JM, Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chem. Res. Toxicol*, 2006. 19: p. 491–505.

9. Bielskute S, Plavec J and Podbevsek P, Impact of Oxidative Lesions on the Human Telomeric G-Quadruplex. *J Am Chem Soc*, 2019. 141(6): p. 2594-2603.

10. Vorlickova M, Tomasko M, Sagi A.J, Bednarova K and Sagi J, 8-oxoguanine in a quadruplex of the human telomere DNA sequence. *FEBS J*, 2012. 279(1): p. 29-39.

11. Brosh R.M Jr. and Cantor SB, Molecular and cellular functions of the FANCJ DNA helicase defective in cancer and in Fanconi anemia. *Front Genet*, 2014. 5: p. 372.

12. Palovcak A, Liu W, Yuan F and Zhang Y, Maintenance of genome stability by Fanconi anemia proteins. *Cell Biosci,* 2017. 7: p. 8.

13. Datta A and Brosh RM Jr, Holding All the Cards-How Fanconi Anemia Proteins Deal with Replication Stress and Preserve Genomic Stability. *Genes (Basel),* 2019. 10(2).

14. Wu Y, Shin-ya K, and Brosh RM, Jr., FANCJ helicase defective in Fanconia anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol Cell Biol,* 2008. 28(12): p. 4116-28.

15. Cantor SB, and Xie J, Assessing the link between BACH1/FANCJ and MLH1 in DNA crosslink repair. *Environ Mol Mutagen,* 2010. 51(6): p. 500-7.

16. Sarkies P, Murat P, Phillips LG, Patel KL, Balasubramanian S and Sale JE, FANCJ coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. *Nucleic Acids Res,* 2012. 40(4): p. 1485-98.

17. Castillo Bosch P, Segura-Bayona S, Koole W, van Heteren J.T., Dewar J.M., Tijsterman M. and Knipscheer P., FANCJ promotes DNA synthesis through G-quadruplex structures. *EMBO J,* 2014. 33(21): p. 2521-33.

18. Wu CG, Maria S, G-quadruplex recognition and remodeling by the FANCJ helicase, *Nucleic Acids Research*, 2016. 44(18): p. 8742–8753.

19. Lattmann S, Giri B, Vaughn JP, Akman SA and Nagamine Y, Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU. *Nucleic Acids Res,* 2010. 38(18): p. 6219-33.

20. Heddi B, Cheong VV, Martadinata H, and Phan AT, Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex complex. *Proc Natl Acad Sci USA,* 2015. 112(31): p. 9608-13.

21. Greenfield N, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc*, 2006. 1: p. 2527–2535.

22. Jing N, Rando RF, Pommier Y, Hogan ME, Ion Selective Folding of Loop Domains in a Potent Anti-HIV Oligonucleotide. *Biochemistry*, 1997, 36(41): p. 12498-12505.

23. Petraccone L, Spink C, Trent JO, Garbett NC, Mekmaysy CS, Giancola C and Chaires JB, Structure and stability of higher-order human telomeric quadruplexes. *J Am Chem Soc*, 2011. 133(51): p. 20951-61.

24. Hardin CC, Watson T, Corregan M, and Bailey C, Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG$_3$GCG). *Biochemistry*, 1992 31: p. 833–841.

25. Guédin A, De Cian A, Gros J, Lacroix L, and Mergny JL, Sequence effects in single-base loops for quadruplexes. *Biochimie.* 2008. 90(5): p. 686-696.

26. Guédin A, Gros J, Alberti P, and Mergny J, How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Research.* 2010 38: p. 7858-7868.

27. Eddy S, Ketkar A, Zafar MK, Maddukuri L, Choi JY, and Eoff RL, Human Rev1 polymerase disrupts G-quadruplex DNA. *Nucleic Acids Research*, 2014 42(5): p. 3272–3285.