MECHANISMS OF G-QUADRUPLEX RECOGNITION AND UNFOLDING BY THE FANCJ HELICASE AND REV1 POLYMERASE

by

KAITLIN ANTHA LOWRAN

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN BIOMEDICAL SCIENCES: HEALTH AND ENVIRONMENTAL CHEMISTRY

2024

Oakland University
Rochester, Michigan

Doctoral Advisory Committee:

Colin G. Wu, Ph.D., Chair
Adam Avery, Ph.D.
Luis G. Villa-Diaz, Ph.D.

To my parents, brother, and partner

ACKNOWLEDGMENTS

I would like to extend my heartfelt gratitude to several individuals who have supported and encouraged me throughout the journey of completing this dissertation.

First and foremost, I want to express my deepest appreciation to my mentor, Dr. Colin Wu. Your guidance, wisdom, and unwavering support have been indispensable throughout every stage of this research. Your insights and constructive feedback have not only refined this dissertation but have also shaped my growth as a scholar. I am truly fortunate to have had the opportunity to learn from you.

To my beloved family, Mom, Dad, Nate, and Ashley, your love, patience, and encouragement have sustained me through the challenges and triumphs of this academic pursuit. Your unwavering belief in my abilities has been a constant source of strength, and I am deeply grateful for your sacrifices and unwavering support.

I am also immensely grateful to my fellow lab members Laura and Emma for being a pillar of strength and a source of inspiration during this journey. Your words of encouragement, understanding, and occasional distractions have kept me motivated when the journey seemed daunting.

Last but certainly not least, I want to thank Zak. Your unwavering love, understanding, and patience have been my anchor throughout this process. Your belief in me never wavered, even during the most challenging times, and for that, I am profoundly grateful.

To all those mentioned above and to countless others who have supported me in ways seen and unseen, I extend my sincerest thanks. Your contributions have not gone unnoticed, and I am indebted to each of you for helping me realize this achievement.

Kaitlin Antha Lowran

ABSTRACT

MECHANISMS OF G-QUADRUPLEX RECOGNITION AND UNFOLDING BY THE FANCJ HELICASE AND REV1 POLYMERASE

by

KAITLIN ANTHA LOWRAN

Adviser: Colin G. Wu, Ph.D.

G-quadruplexes (G4s) are secondary structures formed by guanine-rich nucleic acid sequences. Accumulated G4s in cells can disrupt DNA replication, RNA transcription, and other cellular processes. G4 DNA can adopt parallel, antiparallel, or hybrid conformations depending on solution conditions and base sequence. Moreover, reactive oxygen species can readily convert guanine to 8-oxoguanine (8oxoG) to form 8oxoG4s. However, the precise mechanisms by which repair proteins recognize and remove such lesions remain unclear. The FANCJ DNA helicase possesses an AKKQ motif that functions as a G4-targeting site that binds to G4s and 8oxoG4s. Additionally, FANCJ interacts with the REV1 translesion synthesis polymerase to facilitate replication across G4-forming sites. Although REV1 can also target and disrupt G4s, it is unknown whether the activities of FANCJ and REV1 are redundant or if they act on specific G4 substrates. We hypothesized that the presence of 8oxoG would impact the folding and stability of parallel, antiparallel, and hybrid G4s in specific positions within the G4. Furthermore, we anticipated that FANCJ and REV1 would have distinct 8oxoG4 binding preferences that can be correlated with the physical properties of the 8oxoG4s. We

predicted that the collaborative action of FANCJ and REV1 would repair tightly-folded 8oxoG4 structures, whereas REV1 alone may suffice for disrupting the loosely folded 8oxoG4s. Our results revealed that introducing 8oxo1 significantly reduced the thermal stability of both parallel and antiparallel G4 structures. While 8oxo5 destabilized hybrid structures. FANCJ and REV1 both preferred parallel G4 and 8oxoG4s due to their greater stability and tighter folding compared to hybrid or antiparallel conformations. Furthermore, FANCJ had the lowest affinity for antiparallel 8oxo3 G4s, while REV1 had a lower affinity for hybrid 8oxo1 G4s. These findings indicate that the presence of 8oxoG modification in G4 structures may hinder the binding of these proteins. The study revealed that both FANCJ and REV1 have distinct substrate-specific preferences for unfolding G4s. FANCJ demonstrated lower unfolding rates when 8oxo5 was present within the G4 stack, indicating that the position of the lesion plays a critical role in the unfolding efficiency. REV1, on the other hand, showed a lower unfolding rate for more stable G4 structures like the GGGT G4, but faster rates for the oxidized constructs. This research provides significant insights into the specific roles and mechanisms of FANCJ and REV1 in unfolding 8oxoG-damaged G4s, contributing to our understanding of DNA repair mechanisms.

TABLE OF CONTENTS

TABLE OF CONTENTS—Continued

TABLE OF CONTENTS—Continued

TABLE OF CONTENTS—Continued

TABLE OF CONTENTS—Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES—Continued

LIST OF ABBREVIATIONS

CTD	C-terminal domain
DTT	Dithiothreitol
G4	G-quadruplex
MW	Molecular weight
WB	Western Blot
FPLC	Fast Protein Liquid Chromatography
kDa	Kilodalton
SDS-PAGE	Sodium dodecyl-polyacrylamide gel electrophoresis
G4BP	G-quadruplex binding protein
HR	Homologous Recombination
TLS	Translesion Synthesis
TB	Terrific Broth
LB	Laruia-Bertani Broth
IPTG	isopropyl-β-D-1-thiogalactopryanoside
SF2	Super-family 2
PMSF	Phenylmethylsulfonyl Fluoride
FA	Fanconi Anemia
HF	High Five
ATP	Adenosine Triphosphate
GST	Glutathione S-tranferase

CHAPTER ONE

INTRODUCTION

1.0 G-quadruplexes (G4s)

1.0.1 History

Traditionally, DNA adopts a double helix structure, first described by Watson and Crick in 1953 [1]. While this structure is most commonly observed, DNA can also adopt non-canonical structures, such as G-quadruplexes (or G4s). Sequences rich in guanine nucleotides that can form G4s within both DNA and RNA strands. Early observations by German chemist Ivar Bang revealed that guanine molecules have the ability to self-associate, leading to the formation of a gel from highly concentrated guanylic acid monomers [2]. Fifty years later Gellert and colleagues, using X-ray diffraction on guanylic acid, unveiled that the aggregates observed by Bang resulted from the assembly of tetrameric elements, forming large helical structures [3]. Subsequent studies elucidated that four guanines could arrange into a planar tetrad. The structure initially observed by Bang and later confirmed by Gellert is now referred to as a G-quartet, capable of forming in both RNA and DNA under physiological conditions.

1.0.2 Topology and Structure

The standard nucleotide formula for G4s is: $G_xN_{1-7}\ G_xN_{1-7}\ G_xN_{1-7}\ G_xN_{1-7}$ where X (>2) represents the number of guanines in the sequence, and N is the number of nucleotides in the loop sequence between the tetrads [4, 5]. The loop length significantly affects the stability of G4 structures; shorter loops tend to result in more stable

configurations [6, 7]. Most stable G4s have loops of up to 7 nucleotides, although G4s with loops up to 21 nucleotides have also been identified [8].

In G4s, four guanines can arrange in a co-planar structure known as G-tetrads. Each of the four guanines interacts with an adjacent guanine via Hoogsteen hydrogen bonding [3, 9, 10] . Specifically, N1 and N2 of one guanine interact with the O6 and N7 of the other guanine [11](Figure 1.1). However, hydrogen bonding alone isn't sufficient for stability; G4 structures are further stabilized by the overlap of π orbitals of stacked guanines and the presence of a monovalent cation centrally coordinated to the O6 of the guanines [12]. The stabilization order of monovalent cations is K+ (positioned between G-tetrads) > Na+ (within the tetrad plane) > Li+ (unknown) [11, 13]. Stable G4 structures consist of at least three tetrads, though two-tetrad G4s have also been observed. The stability of G4s can be assessed through various experimental techniques such as NMR, UV, and CD melting experiments.

In addition to the factors affecting G4 stability, such as loop length and the presence of monovalent cations, oxidative damage can also play a significant role in the stability of G-quadruplexes [14]. Guanine has the lowest redox potential and reactive oxygen species can induce oxidation of guanine to form 8-oxoguanine. The 8oxoG can pair with both adenine and cytosine causing a G to T transverse mutation [15]. Additionally, the introduction of 8-oxoguanine into guanine-rich sequences can disrupt the Hoogsteen hydrogen bonding and π-π stacking interactions that stabilize G4 structures [16]. The incoporation of 8oxoG into G4s may cause a change in the structure or stability that could affect cellular processes such as replication, transcription, and

telomere maintenance [17, 18].

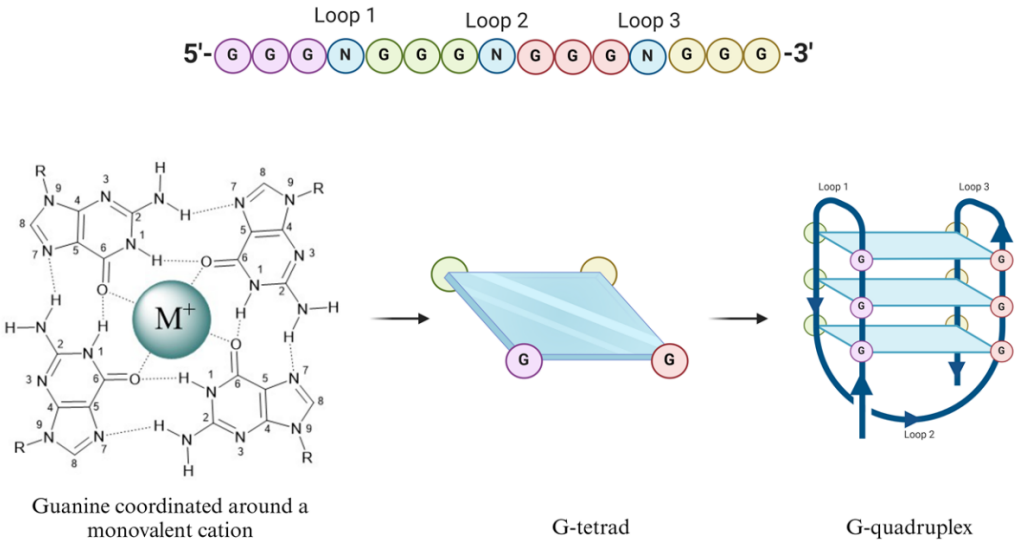

Figure1.1: Formation of G-quadruplex

The topology of G-quadruplexes can vary based on the number and orientation of the strands involved. G4 structures can be formed by a single DNA or RNA strand folding into a quadruplex (intramolecular) or by multiple strands coming together (intermolecular). The topology is further classified based on strand directionality: parallel, where all strands run in the same direction; antiparallel, where strands run in opposite directions; or a hybrid of both [19]. Parallel G-quadruplexes typically feature all strands oriented in the same direction, leading to a compact and stable structure with a central channel where cations, such as potassium or sodium, reside to stabilize the G-quartets [20]. Antiparallel G-quadruplexes, on the other hand, exhibit alternating strand orientations, which results in a more extended and flexible topology [21]. This flexibility can influence the biological roles of G-quadruplexes, affecting DNA replication and transcriptional processes [22]. Hybrid G-quadruplexes combine features of both parallel

and antiparallel topologies and are often found in complex genomic regions. These hybrids can form under physiological conditions and play crucial roles in regulating gene expression and maintaining genome stability [23]. Recent studies using high-resolution techniques like X-ray crystallography and NMR spectroscopy have provided detailed insights into the molecular structures of G-quadruplexes, enhancing our understanding of their biological functions and the impact of their topology on stability and drug interactions [24]. A visual representation of the three G4 topologies and their corresponding loops is provided in Figure 1.2.

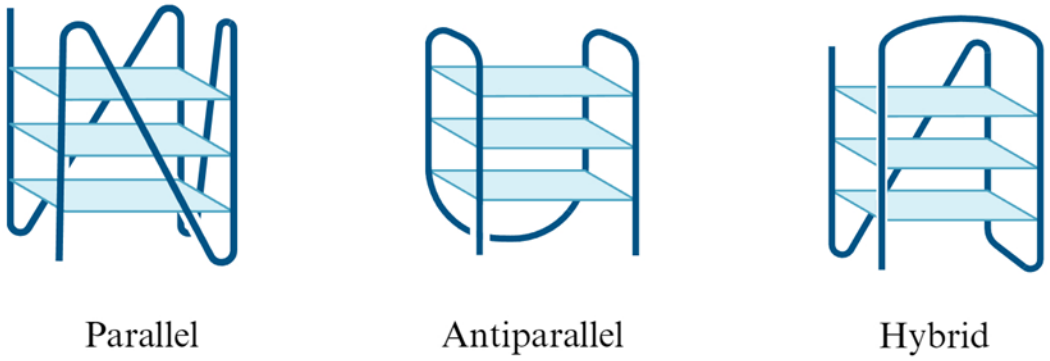

Figure 1.2: Graphical representation of the three G4 conformations

1.0.3 Biological Relevance

1.0.3.1 G4s in the genome

The detection and study of G4s within the human genome have unveiled their significant presence and potential roles in various genomic processes. Computational methods have predicted over 370,000 G4 sequence motifs, particularly enriched in regions associated with genome regulation, such as promoters, telomeres, and untranslated regions [5, 25]. Telomeres are DNA-protein complexes at the ends of

eukaryotic chromosomes. Telomeres, specifically, contain tandem repeats of guanine-rich sequences, with the conserved sequence in vertebrates being TTAGGG[26-28].

The first *in vivo* detection method of G4s utilized G4 structure-specific antibodies, such as BG4, a monoclonal single-chain antibody. BG4 has been widely employed to investigate the presence of G4 structures in the genome due to its high affinity and specificity for G4s [23, 29-31]. Immunostaining with BG4 in human cells illustrated the presence of G4s in all cell phases, with the majority of foci in the S phase [23]. Additionally, small molecules like 3,6-bis(1-Methyl-4-vinylpyridinium) carbazole diiodide (BMVC) have been developed as fluorescent probes to detect G4 structures within cells [32-35]. While initially designed for detecting G4 structures in human telomeres, BMVC has also been found to bind to G4 structures in the MYC promoter [35]. These findings were supported by colocalization studies with the BG4 antibody [34].

1.0.3.2 Epigenetic Control

Epigenetic markers are modifications to DNA and histone proteins that regulate or silence genes without altering the primary DNA sequence. These markers include DNA methylation, histone modification, and nucleosome positioning, which collectively modulates gene expression. Epigenetic features can persist through cell division or be induced be environmental stimuli. Alterations in these markers have been linked to cancer and neurodevelopmental disorders.

Recently, G4s have garnered attention for their role in epigenetic programming and chromatin remodeling. The formation of G4s in the genome is associated with chromatin accessibility, with open chromatin facilitating G4 formation. CpGs refer to

cytosines methylated at the fifth carbon of the pyrimidine ring, often followed by guanine residues [36]. Genomic regions with a high frequency of CpGs are known as CpG islands (CGIs). The relationship between CGIs and G4s was investigated using ChIP-seq, revealing significant overlap between the two [37]. Interestingly, G4-CGI overlaps were found to be markedly hypomethylated in open chromatin sites [37]. Additionally, G4-CGI sites were found to overlap with or be proximal to DNA methyltransferase 1 (DMNT1) binding sites. DMNT1 methylates Cs at CpG dinucleotide and shows a preference for binding to G4s *in vitro* [37]. These findings suggest the recruitment and inhibition of DMNT1 at hypomethylated G4.

1.0.4 G4-binding proteins

The presence of G4s at genome regulatory sites, coupled with the observation that G4 formation in cells is contingent upon cell type and state, suggest that G4s are under the regulation of cellular proteins. Numerous proteins have been identified to interact with G4s, leading to the development of databases such as the G4-interacting proteins database (G4IPDB). To identify these proteins, many studies have relied on affinity proteomics experiments utilizing G4 oligomers as baits to isolate G4 binders from cellular lysates [38-40]. Others have employed computational analysis of genomic protein binding sites to assess the presence of G4 motifs within these sites [41-43]. G4 binding proteins (G4BPs) can interact with either DNA, RNA, or possibly both.

Several G4BPs are known to bind to telomeres and unfold them to maintain their length and integrity. For instance, the protein POT1 (Protection of Telomeres 1) binds to the 3’ overhang of telomeric repeats and regulates its G4-unfolding partner TPP1 [44]. Another protein involved in unfolding telomeric G4 DNA is replication protein A (RPA),

although it cannot unfold G4s in the presence of POT-TPP1 complexes[45]. Additionally, it's noteworthy that human telomerase reverse transcriptase (hTERT) acts as a G4 resolvase to facilitate proper telomere replication.

In addition to the previously mentioned G4-binding proteins, several others play significant roles in interacting with G4 structures. BRCA1, renowned for its involvement in DNA repair pathways, has been found to bind to G4 structures, suggesting its potential contribution to genome stability and maintenance [46]. Another notable protein, hnRNP A1, part of the heterogeneous nuclear ribonucleoprotein family, is able to unfold G4s to allow telomerase binding for telomere lengthening [47]. FMRP, associated with Fragile X syndrome, has been shown to bind to RNA G4 structures, suggesting a role in mRNA translation regulation [48]. The inclusion of these G4-binding proteins highlights the complexity of G4-mediated processes and their regulation in various cellular contexts.

Beyond telomeres, G4BPs play crucial roles in replication and transcription. They are recruited to G4 sites to unwind them, facilitating efficient DNA replication and maintaining genome integrity. However, while G4s may serve to prevent the uncoupling of leading and lagging-strand polymerases, they can also act as barriers hindering further replication. Helicases, such as FANCJ, can unfold G4 structures to allow replication to proceed uninterrupted. Further information on helicases will be provided in the subsequent section.

1.1 FANCJ Helicase

1.1.1 Helicases

Helicases, essential motor proteins, harness the chemical energy of ATP to generate mechanical energy, driving the unwinding of DNA molecules and the

remodeling of RNA molecules [49-51]. Their pivotal roles span across numerous biological processes, including replication, recombination, transcription, repair, and other nucleic acid metabolism pathways [52, 53]. Initially discovered in 1976 by Abdel Monem et al., with the identification of *Escherichia coli* Tral, hundreds of helicases have since been found in various organisms [54].

Helicases typically possess conserved domains, such as the arginine finger ATP-associated residues, facilitating their ATP-dependent functions [55, 56]. Classifications of helicases into six superfamilies (SF1-6) are based on conserved sequences SF3-6 comprise ring-shaped hexameric helicases, while SF1 and SF2 encompass non-ring enzymes [57]. Further differentiation can be made based on their ATPase types. For instance, SF4 and SF5 helicases are categorized as RecA-like ATPases, where as SF3 and SF6 helicases belong to the AAA+ ATPase group [55, 58]. Notably, the SF2 type helicases represents the largest subfamily and will be the primary focus in the subsequent section.

1.1.1.1 SuperFamily 2

Helicases belonging to SF2, characterized by their "helicase signature motifs," encompass two main domains and seven conserved primary sequence motifs within their amino acid sequence [49]. The central motor core, composed of RecA-like folds termed helicase domain 1 (HD1) and helicase domain 2 (HD2), facilitates the conservation of ATP energy into directional motion crucial for their various biological functions [58-60]. These helicases primarily interact with the phosphate-sugar backbone of nucleic acids through conserved sequence motifs, with additional base contacts established by residues located within the helicase core or in distinct domains.

SF2 helicases exhibit remarkable versatility, participating in diverse cellular processes such as DNA replication, repair, recombination, RNA transcription, and RNA processing. Notably, they demonstrate substrate specificity, proficiently recognizing and unwinding a wide array of nucleic acid structures. SF2 helicases can unwind duplex DNA, RNA duplexes, RNA-DNA hybrids, and complex RNA secondary structures like hairpins and G-quadruplexes [61-64]. This substrate specificity is vital for maintaining genome stability, regulating gene expression, and ensuring efficient RNA metabolism within the cell.

Furthermore, SF2 helicases play pivotal roles in orchestrating the assembly and function of large protein complexes involved in nucleic acid transactions. They contribute to the recruitment and coordination of other proteins and enzymes, facilitating critical biological processes such as DNA replication fork progression, RNA polymerase movement along DNA, and the resolution of intricate DNA and RNA secondary structures. Through their multifaceted functions, SF2 helicases are integral components of cellular machinery, essential for genome maintenance and gene expression regulation.

SF2 encompasses a diverse array of helicase proteins, each with unique functions and substrate specificities. One notable SF2 helicase is the human RECQ4, it is a highly conserved 3'-5' helicase that functions include initiation of DNA replication, multiple DNA damage repair pathways, and maintenance of telomeric and mitochondrial DNA. RECQ4 mutations of this protein have been associated with genetic disorders such as Rothmund-Thomson, Werner, Bloom, and Baller-Gerold syndromes [65, 66]. Another prominent SF2 helicase is the DEAD-box RNA helicase, DDX3X, involved in various aspects of RNA metabolism, including mRNA splicing, translation initiation, and RNA

transport [67]. DDX3X has been implicated in cancer progression and viral infections [68]. Additionally, SF2 helicases such as WRN and BLM are known for their roles in DNA repair and genome stability maintenance, with mutations in these genes linked to Werner syndrome and Bloom syndrome, respectively [65, 66, 69] . The SF2 helicase family also includes members like FANCJ, involved in DNA repair through homologous recombination, and XPD, a component of the transcription factor IIH complex essential for transcription initiation [70, 71]. These diverse SF2 helicase proteins highlight the breadth of functions and biological pathways regulated by this superfamily, underscoring their importance in cellular physiology and disease pathology. With an understanding of SF2 helicases and their versatile functions, we now delve into the specific characteristics and roles of one notable member, the FANCJ helicase.

1.1.2 FANCJ helicase: a crucial player in DNA repair.

FANCJ helicase, a 1249 amino acid protein, belongs to the SF2 superfamily and holds significant importance in DNA repair mechanisms. Its discovery was initially spurred by its interaction with the tumor suppressor BRCA1 [72]. The structure of FANCJ comprises of seven motifs in its N-terminal region, displaying strong homology with the DEAH family of helicases, notably including Xerderma Pigmentosum complementation group D (XPD) [70, 73]. Like XPD, FANCJ features a conserved iron-sulfur (Fe-S) cluster, a crucial structural component implicated in DNA unwinding, and an arch domain[70, 74, 75] [76]. Moreover, the C-terminal region of FANCJ encompasses amino acids forming the BRCA1-interacting region (BRCT) [77].

Figure 1.3: Outline of FANCJ helicase

Biochemical studies have elucidated the dual DNA-dependent ATPase and helicase activities of FANCJ [78]. It effectively unwinds DNA duplexes in a 5' to 3' direction but necessitates backbone continuity in both strands as it progresses along the DNA [78, 79]. Interestingly, FANCJ can sense both DNA strands during unwinding, and this requirement for backbone continuity can be surmounted by augmenting the amount of single-strand DNA at the 5'end, facilitating FANCJ loading onto the DNA [79].

Regulatory factors, such as Replication Protein A (RPA), stimulate FANCJ helicase activity, promoting unwinding of longer duplexes [80]. However, *in vitro* studies demonstrate that the addition of BRCA1 has no effect on FANCJ helicase activity [78, 80]. FANCJ exhibits proficiency in recognizing complex DNA structures present at replication forks or intermediate structures formed during HR.

While FANCJ demonstrates a preference for binding to and unwinding forked duplex substrates common to DNA replication forks [81], it also efficiently unwinds model structures mimicking D-loops formed during HR [81]. However, FANCJ is incapable of resolving Holliday junction structures formed during HR [81]. Additionally, it can disrupt protein-DNA complexes akin to those formed during DNA replication and

repair [82], implying its ability to recognize and unwind DNA structures commonly encountered during these processes.

The role of FANCJ in HR-mediated DNA double strand break (DSB) repair is underscored by its genetic interaction with BRCA1. Phosphorylated during the S-phase of the cell cycle, FANCJ remains bound to BRCA1 after the generation of DNA DSBs [83]. Immunofluorescence microscopy reveals discrete nuclear foci of FANCJ, co-localizing with BRCA1 at DNA damage sites in human cells. Notably, the brightness of FANCJ foci is diminished in BRCA1 null cells, indicating BRCA1 has a role in the localization of FANCJ into nuclear foci [72]. Conversely, BRCA1 foci are diminished and delayed in FANCJ-deficient cells [83].

Beyond its role in DSB repair, FANCJ is also crucial for addressing other complex DNA damages, such as interstrand crosslinks (ICLs), which are particularly deleterious forms of DNA damage that impede the separation of DNA strands, essential for vital processes such as transcription and replication [84, 85]. These crosslinks can be induced by various chemotherapeutic agents and environmental mutagens, and the role of FANCJ through the Fanconi anemia (FA) pathway is critical in mediating their repair. The FANCJ helicase activity helps unwind DNA at crosslink sites, facilitating the recruitment and proper function of other essential repair proteins, and the helicase has been shown to unfold DNA-protein crosslinks [86].

1.1.3 FANCJ and G-Quadruplexes

Besides its ability to catalytically unwind duplex DNA structures and branched D-loop structures, which are intermediates in homologous recombination, FANCJ also acts as a resolvase for G-quadruplex (G4) structures [87]. In humans, FANCJ plays a crucial

role in suppressing DNA damage and genomic instability caused by G4s. While several DNA helicases can unwind intermolecular G4s, FANCJ is distinctive in its ability to resolve the entropically favored intramolecular G4s [87-91]. In *C. elegans*, deficiency in dog-1 (the ortholog of FANCJ) causes deletions of G4 DNA and leads to stalling of DNA replication [92, 93]. A study using *Xenopus* egg extract, found that FANCJ was required for unwind of stable G4s, and depletion of FANCJ resulted in stalling at G4 structures [90].

The mechanism for FANCJ recognizing and unwinding G4 structures involves an intricate interaction with the DNA backbone and the G4 motif. An 'AKKQ' motif found in DHX36 can recognize G4 sites, this motif has also been identified in FANCJ [91, 94, 95]. The two lysine residues in DHX36 form a loop that tightly anchors to the G4s, facilitating robust binding and unwinding. Studies have shown that FANCJ can induce a partial unfolding of G4 structures, which is essential for the subsequent recruitment of repair proteins [91]. This partial unfolding allows for the exposure of otherwise sequestered DNA bases, making them accessible for repair and replication processes [96].

Recent research further reveals that FANCJ, in coordination with replication protein A (RPA), resolves G4 structures more efficiently [80]. This synergy not only facilitates the unwinding of G4s but also promotes the binding of RPA [97]. Moreover, FANCJ has an important role in unfolding G-quadruplexes with its ability to collaborate with REV1, a translesion polymerase, at G4-forming sites. According to a recent model, FANCJ repeatedly unfolds G4s at stalled replication forks, enabling repair polymerases like REV1 to effectively bind [91, 94]. The enhanced understanding of the involvement

of FANCJ in G-quadruplex resolution provides profound insights into its broader role in cellular mechanisms, highlighting potential therapeutic targets, particularly for diseases characterized by genomic instability, such as cancer.

1.1.4 Role in Fanconi anemia

Fanconi anemia (FA) is a rare genetic disorder, first described by Guido Fanconi in 1927, characterized by defects in DNA repair mechanisms [98]. This disorder typically manifests with diverse physical abnormalities and an increased risk of hematological issues. FA is diagnosed in approximately 10 individuals per million, with an average age of diagnosis around 7 years [99]. The disorder is notable for major abnormalities in two-thirds of patients, including skeletal defects, reduced fertility, and hematological issues, with bone marrow failure typically appearing by the age of 10. The average lifespan of patients is approximately 20 years, primarily due to bone marrow failure. Additionally, patients with FA are at an elevated risk for conditions like acute myeloid leukemia.

To date, defects in 23 genes have been identified in FA patients, contributing to a spectrum of repair deficits at DNA interstrand crosslinks (ICLs), stalled replication forks, and double-strand breaks (DSBs) [100]. Despite these advancements, the specific molecular functions of many FA proteins within the repair pathway remain largely unexplored. However, the biochemical activities of FANCJ have provided significant insights into the mechanisms of FA-mediated DNA repair.

FANCJ plays a critical role within the FA repair pathway by addressing specific types of DNA damage such as ICLs and G4s. Upon detecting DNA damage, FANCJ is rapidly recruited to the damage site through its interactions with other FA pathway proteins, including FANCD2 and BRCA1. Once localized, FANCJ unwinds DNA

surrounding the damage, thus facilitating the access of repair factors and enzymes necessary for effective repair. Additionally, FANCJ has the ability to resolve secondary DNA structures like G4s is vital, as it prevents the replication and transcription machinery from stalling, thereby enhancing the repair efficiency. Through its coordinated action with other FA proteins and its interaction with BRCA1, FANCJ significantly contributes to the repair of ICLs and the stabilization of replication forks, which are crucial for maintaining genomic stability and preventing the progression of cancer in FA patients.

1.1.5 FANCJ Summary

As we have seen, FANCJ plays a critical role in maintaining genomic stability by resolving G-quadruplex structures and facilitating DNA repair in Fanconi anemia. One of the key interactions in this process involves REV1, a translesion polymerase that works closely with FANCJ at G4-forming sites. This collaboration is not only pivotal for the unwinding and stabilization of complex DNA structures but also underscores the integrated network of repair mechanisms that safeguard cellular integrity. Building on the importance of this interaction, we will now delve deeper into the role of REV1. This next section will explore its unique biochemical properties, its contributions to translesion synthesis, and its broader implications in DNA repair pathways, highlighting how it supports the cellular response to genomic stress alongside FANCJ.

1.2 REV1 polymerase

1.2.1 Translesion Synthesis polymerases

Translesion synthesis (TLS) polymerases, specialized DNA polymerases, are crucial for enabling DNA replication to continue past sites of DNA damage, despite their

generally lower fidelity compared to high-fidelity replicative polymerases [101]. These polymerases are adept at bypassing various DNA lesions, such as thymine dimers, apurinic/apyrimidinic (AP) sites, and bulky adducts [102]. In the human genome, there are several TLS polymerases including polymerase eta (Pol η), polymerase iota (Pol ι), polymerase kappa (Pol κ), and polymerase zeta (Pol ζ), each capable of bypassing different types of damage [103, 104] .

Pol η, in particular, shows a unique flexibility in its active site, enabling it to bypass G4 structures more effectively than replicative polymerases [105]. Additionally, REV1 stands out for its ability to incorporate cytosine opposite various lesions and has been observed to aid in the unfolding of G4 structures [106-109]. Despite the vital role of TLS polymerases in preventing replication fork collapse and ensuring the completion of DNA replication, especially in guanine-rich regions, their error-prone nature can introduce mutations during replication past G4 structures.

1.2.2 Discovery and Structure of REV1 Polymerase

REV1 polymerase is a specialized DNA polymerase enzyme involved in DNA repair and replication processes. The discovery of REV1 came from early studies in yeast, particularly *Saccharomyces cerevisiae* in the 1970s [110]. REV1 was initially identified in a genetic screen for mutants with increased sensitivity to DNA-damaging agents, and its role in TLS became apparent due to its involvement in bypassing DNA damage. Jacobs et. al. lab found that REV1 played an essential role in the generation of G to C transversions downstream of the Ung2 (uracil-DNA glycosylase 2) pathway [111]. REV1 is unique among DNA polymerases due to its ability to incorporate deoxycytosines opposite of damaged nucleotides, such as ^{6}O-megG, a guanine with a

large adduct at the C8 or N2 position, or abasic sites [112]. This property enables the bypass of lesions that would otherwise stall replicative polymerases.

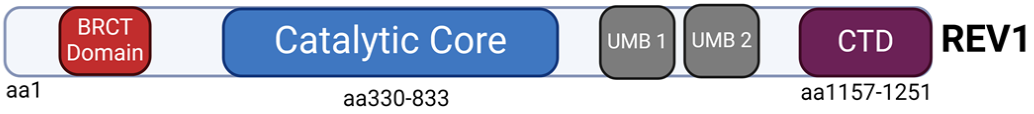

Figure 1.4: Outline of REV1 polymerase

Following the initial discoveries in yeast, REV1 was found to be conserved across eukaryotes, including humans, where it retains its function as a TLS polymerase. The REV1 gene is located on the short arm of chromosome 2 in humans, specifically at the 2q11.2 locus. The human REV1 gene spans approximately 100 kilobases and has 23 exons encoding REV1 protein consisting of 1,251 amino acids. The structural features of REV1 include the catalytic domain responsible for DNA polymerase activity (aa 330-833), the unique C-terminal deoxycytidyl transferase (dCTPase) domain, and regions that mediate interactions with other proteins involved in DNA damage response, such as the ubiquitin-binding UBM domain. At the N-terminus, REV1 contains a BRCA1 C-terminus (BRCT) domain, which is implicated in protein-protein interactions and is essential for mediating the recruitment of REV1 to DNA damage sites [113, 114].

1.2.3 Functions and Substrates of REV1 polymerase

REV1 polymerase, a key component of the translesion synthesis (TLS) pathway, is instrumental in allowing DNA replication to bypass lesions on the DNA template that typically stall high-fidelity polymerases [114, 115]. REV1 helps to continue synthesis past many different replication block and in response to nucleotide depletion [116].

REV1 has the ability to insert deoxycytidines opposite a variety of DNA lesions, including apurinic/apyrimidinic (AP) sites, thymine glycols, 8-oxoguanine, and intrastrand crosslinks, is crucial for maintaining DNA integrity under stress conditions [117, 118].

In addition to its well-established role in bypassing conventional DNA lesions, recent research has highlighted REV1 being capable of interacting with G4s [106, 119]. An important role of REV1 is the recruitment of multiple TLS, by the C-terminal domain (CTD), to sites of replication stress. Deletion of the REV1 CTD reduced the efficiency of bypassing G4s, and the complete loss of the protein's ability to histone cycle. A study revealed that REV1 can disrupt G4 DNA structures through the catalytic core, thereby aiding in the continuation of DNA replication across regions that would otherwise be prone to stalling and genomic instability [120]. This same study found mutations of the insert-2 motif in REV1, 54 amino acids that serve as a flap over the hydrophobic pocket of REV1, produced changes in binding to G4s [120].

The ability of REV1 to bind to and unfold G4 structures enhances our understanding of its broader role in genomic maintenance. This function is particularly significant in regions of the genome where G4 structures are prevalent and likely to form, such as telomeres and promoter regions of oncogenes. By resolving these structures, REV1 not only prevents potential replication stress and genomic instability but also presents itself as a target for therapeutic strategies, especially in cancer, where genomic stability is often compromised.

1.3 Dissertation Hypothesis

This research hypothesizes that the oxidative modification of guanine to 8-oxoguanine (8oxoG) significantly impacts the folding and stability of G-quadruplexes in distinct genomic hotspots. Moreover, we anticipate that the helicase FANCJ and the translesion synthesis polymerase REV1 exhibit unique binding preferences and interactions with 8oxoG-modified G4 structures. We predict that FANCJ and REV1 together are required to repair tightly folded 8oxoG4 structures, whereas REV1 alone may be sufficient to disrupt loosely folded 8oxoG4s. This hypothesis underscores the potential therapeutic implications of targeting these protein interactions in diseases characterized by oxidative stress and genomic instability, such as cancer.

CHAPTER TWO

MATERIALS AND METHODS

2.0 Introduction

Protein purification is a vital process for this study of FANCJ and REV1, aimed at obtaining the proteins in a highly pure form for subsequent functional assays. This chapter focuses on expressing and isolating FANCJ and REV1 using affinity chromatography. The expression of these recombinant proteins in a suitable host system, followed by their purification using affinity tags, ensures high specificity and yield. The methods section is organized into four subsections: protein expression, cell lysis and extraction, affinity chromatography, and protein analysis. This structured approach provides a comprehensive and clear presentation of the techniques used to purify FANCJ and REV1, facilitating detailed biochemical and structural studies.

2.1 Buffers and Reagents

All buffers and reagents were prepared with reagent-grade chemicals and type I water from a Smart2Pure 6 UV system (ThermoFisher). Solutions were then sterilized through a 0.22 µm PES filter. A primary buffer used in experiments is G4 reaction (G4 Rxn) buffer (20 mM HEPES pH 7.5, 150 mM KCl, 5 mM TCEP, 5% glycerol), unless stated otherwise.

2.2 Peptides and Oligos Preparations

List of oligos and peptides used can be found in corresponding chapters. Oligos were purchased from Integrated DNA Technologies (IDT). DNA oligos were prepared in G4 Rxn buffer, there concentration was then determined using a NanoDrop One UV-Vis

microvolume spectrophotometer (ThermoFisher) with their respective molar extinction coefficients (ε_{260}), and then stored at 4°C. Before using, G4 DNA was heated to 95°C for ten minutes and then slowly cooled to room temperature to ensure G4-formation.

Peptides were synthesized by Genscript. The lyophilized powered were resuspended in G4 Rxn buffer, filtered with 0.22 μm PES filter, and concentrations were then determined spectrophotometrically with the molar extinction coefficients (ε_{280}). Peptides were aliquoted, frozen in liquid nitrogen, and then stored at -20°C.

2.3 FANCJ Expression and Purification

There were two FANCJ constructs used, FANCJ-Flag and GST-FANCJ-Flag. The gene encoding for the human FANCJ protein was inserted into pET vector or pvL1392 with dual tags of GST at the N-terminus and Flag at the C-terminus was synthesized by VectorBuilder.

2.3.1 Expression of the pET-GST-FANCJ-Flag construct

GST-FANCJ-Flag was expressed in NiCo21 (DE3) cell line from NEB. Transformed cells were grown in Terrific Broth (TB) in the presence of 50 μg/L ampicillin at 37°C, shaking at 250 rpm until the OD_{600} reached 0.4-0.6. Cells were induced with 1mM of isopropyl-β-D-1-thiogalactopryanoside (IPTG) and grown at 16°C for ~18-22 hours. The cells were then harvested by centrifugation at 4,000 rpm at 4°C for 1 hour. The cells were then resuspended with cell resuspension (CR) buffer (20mM Tris HCl pH 7.5, 10% sucrose), centrifuged at 4,000 rpm at 4°C for 20 mins, then stored at -80°C until purification.

2.3.2 Virus propagation, amplification, and expression of the pvL1392-FANCJ-Flag and pvL1392-GST-FANCJ-Flag

Seed 0.9-1 x 10^6 of Sf9 cell line (*Spodoptera frugiperda*) in a well of a six-well plate. Incubate the plate undisturbed at 28°C for 20 mins. During this time, prepare the co-transfection mixtures using the BestBac transfection kit (Expression Systems). Pipette 100 µL of the transfection medium (from kit) into two sterile 1.5 mL centrifuge tubes (Label one tube A and the other tube B). To test tube A, add 5 µL of BestBac DNA and 2 µg of the FANCJ-Flag or GST-FANCJ-Flag DNA. Into tube B, add 6 µL of Expres2TR transfection agent, and incubate both tubes at room temperature for 5 mins. After 5 mins, combine tubes A and B, gently mixing then incubate at room temperature for 20 minutes. Add 800 µL of transfection medium to the mixture to have a total volume of 1 mL. Aspirate the medium from the cells and then add the transfection mixture and leave at 28°C for 4 hours for the FANCJ-Flag construct and 5 hours for the GST-FANCJ-Flag, and then replace with fresh culture medium. Incubate for 4-5 days and then collect the supernatant by centrifugation at 3,000 rpm for 5 mins, and then filter with 0.22 µm syringe filter. This is the P0 baculovirus stock to be amplified. To amplify, 1 mL of P0 virus was added to 25 mL of suspended Sf9 cells at 2 x 10^6 cell/mL and 95% viable, and then collected when the viability is 20% or lower, this was P1 virus. To generate P2 virus, add 5 mL of P1 virus to 100 mL of Sf9 cells that are 95% viable and have a confluency of 2 x 10^6 cell/mL, then collect the cells when the viability reached lower than 20%.

To express FANCJ, the insect cell line is switched High Five (HF) (BTI-TN-5B1-4) that are derived from *Trichoplusia ni*, these cells allowed for an increase in protein

production compared to Sf9 cells. High Five cells required Express Five SFM media (ThermoFisher) supplemented with 18 mM of L-glutamine. 10 mL of P2 virus was added to 200 mL of suspended HF cells at 2 x 10^6 cell/mL at 95% viability, when the cells reached 85% viability they were harvested by centrifugation at 2000 rpm, 30 mins, at 4°C.

2.3.3 Purification of GST-FANCJ-Flag

GST-FANCJ-Flag was purified from both the bacterial construct in Nico21 cells and from the insect cell construct in HighFive cells. Beyond the lysis step, the purification method used was the same and will be described below.

To lyse NiCo 21 cells, they were resuspended in GST lysis (GL) buffer (75mM Tris HCl pH 7.5, 300 mM NaCl, 5% glycerol, 1% NP40) with the addition of 1 mM phenylmethylsulfonyl fluoride (PMSF). The suspension was then homogenized using a high-pressure homogenizer (EmulsiFlex-5, Alvestin) two passes at 15,000 psi. The lysed cells were centrifuged at 19,000 rpm for 1 hour at 4°C.

To lyse the HF cells, the cells were resuspended in Lysis Buffer (10 mM Tris HCl pH 7.5, 130 mM NaCl, 10 mM NaPi, 5% glycerol, 1% Triton X-100) with the addition of two 50 mL Pierce Protease Inhibitor tablets per 50mL of lysate and 1 mM of PMSF. The lysate is left to rotate for 2 hours at 4°C. The lysed cells were clarified by centrifugation at (18,500 rpm at 4°C) for 30 mins.

The supernatant first filtered using 0.45 μm PES membrane syringe filter and then loaded on to a 5 mL GSTrap HP (GE Healthcare) on the AKTA Start FPLC (GE Healthcare), which was equilibrated with the 30 mL of GL buffer, and the protein was eluted with GST elution (GE) buffer (75 mM Tris HCl pH 7.5, 150 mM NaCl, 5%

glycerol, 50 mM Glutathione) over a 30 mL linear gradient. Fractions containing FANCJ were then incubated with FLAG antibody resin (Millipore Sigma) for 2 hours at 4°C. The resin was washed three times with flag equilibrium (FEq) buffer (25 mM Tris HCl pH 7.5, 300 mM NaCl, 0.1% NP40, 5% Glycerol). The resin was washed a second time with Flag Wash (FW) buffer (25 mM Tris HCl pH 7.5, 30 mM NaCl, 0.1% NP40, 5% Glycerol). FANCJ was then eluted with 100 µg/mL 3x FLAG peptide (Med Chem Express) in flag elution (FEl) buffer (25 mM Tris HCl, 30 mM NaCl, 15% Glycerol). The purity of FANCJ was assessed via SDS-PAGE with Coomassie staining and protein concentration was determined via nanodrop using the extinction coefficient ε_{280} = 181,375 $M^{-1}cm^{-1}$. Samples containing purified protein were dialyzed into G4 Rxn buffer with 20% glycerol, aliquots were frozen in liquid nitrogen and stored at -80°C.

2.4 Expression and purification of REV1 Core and REV1 CTD

2.4.1 Expression of the pET-6xHis-hREV1-CTD (aa 1157-1251) construct

hREV1-CTD was expressed in NiCo21 (DE3) cell line from NEB. Transformed cells were grown in LB in the presence of 50 ug/L kanamycin at 37°C, shaking at 250 rpm until the OD_{600} reached 0.4-0.6. Cells were induced with 1mM IPTG and grown at 16°C for ~18-22 hours. The cells were then harvested by centrifugation at 4,000 rpm at 4°C for 1 hour. The cells were then resuspended with CR buffer, centrifuged at 4,000 rpm at 4°C for 20 mins, then stored at -80°C until purification.

For purification, cell pellet was lysed with lysis binding buffer (20 mM NaPi pH 7.5, 300 mM NaCl, 30 mM Imidazole, 1 mM DTT, 5 % glycerol +1 % NP40 buffer) and 1mM PMSF. The lysate was then sonicated on ice at 45% amplitude using a Fisher 505 dismembrator (Fisher Scientific). After sonicating for 30 mins total (15 sec on, 45 sec

off), the lysate was clarified by centrifugation (18,000 rpm at 4°C) for 1 hour. The supernatant was filtered using a 0.45 μm PES membrane syringe filter.

The supernatant was then applied to a Ni-NTA column (Macherey-Nagel) for the affinity chromatography. The filtered supernatant was mixed with Nickel beads for 15 minutes and then collected as flow through (FT). The first wash used 75 mL of A1 buffer (20 mM NaPi pH 7.5, 300 mM NaCl, 30 mM Imidazole, 1 mM DTT, 5 % glycerol). The second wash used 75 mL of A2 buffer (20 mM NaPi pH 7.5, 150 mM NaCl, 30 mM Imidazole, 1 mM DTT, 5 % glycerol). To elute, 25 mL of the B1 buffer (20 mM NaPi pH 7.5, 150 mM NaCl, 1 M Imidazole, 1 mM DTT, 5 % glycerol) was added and then mixed for 30 mins to detach REV1 CTD from the column. REV1 CTD was then dialyzed into storage buffer G4 Rxn buffer with 20% glycerol.

2.4.2 Expression and Purification of the pBG101-6xHis-GST-hREV1-Core (aa 330-833) and E466A/Y470A construct

The hREV1 core and E466A/Y470A variant construct were provided by Dr. Robert Eoff from the University of Arkansas. The hREV1-Core was expressed in BL21 (DE3) cell line from NEB. Transformed cells were grown in Lysogeny Broth (LB) in the presence of 50 ug/L kanamycin at 37°C, shaking at 250 rpm until the OD_{600} reached 0.4-0.6. Cells were induced with 1 mM IPTG and grown at 18°C for ~18-22 hours. The cells were then harvested by centrifugation at 4,000 rpm at 4°C for 1 hour. The cells were then resuspended with CR buffer, centrifuged at 4,000 rpm at 4°C for 20 mins, then stored at -80°C until purification. The expression of the E466A/Y470A was the same as the procedure described above.

For the purification, the scheme described below was used for both the wild-type hREV1 core and the hREV1 core E466A/Y470A mutant.

The frozen cell stock was resuspended with lysis buffer A1 and with the addition of 1 mM PMSF. The lysate was homogenized at 15,000 psi and passed through twice. Lysate was then centrifuged at 19,000 rpm for 1 hour at 4°C. Supernatant was then applied to a 20 mL IMAC column (GE Healthcare) that was charged with Ni^{2+} and equilibrated with A1 buffer using the FPLC. The column was then washed with 200 mL of A2 buffer. hREV1 core was eluted with B1 buffer over a 200 mL linear gradient. Fractions containing hREV1 core were identified by SDS-PAGE, combined, and then loaded onto a 5 mL GSTrap. The purification scheme is repeated as with the GST-FANCJ-Flag construct. The fractions were then combined, and purity was confirmed by SDS-PAGE. Protein concentration was determined via nanodrop using the extinction coefficient ε_{280} = 77,240 $M^{-1}cm^{-1}$. Samples containing purified protein were dialyzed into G4 Rxn buffer with 20% glycerol, aliquots were frozen in liquid nitrogen and stored at -80°C.

2.5 Western Blots

All blots described were done with the Trans-Blot Turbo System, the RTA transfer kit, and protocol from Bio-Rad Laboratories. The 1L 1x stock transfer buffer was prepared by combining 200 mL of 5x transfer buffer, 200 mL ethanol, and 600 mL of nanopure water.

The PVDF membrane was immersed in 100% ethanol and then transferred into 30 mL of 1x transfer buffer. Two transfer stacks were hydrated with 50 mL of 1x transfer buffer for 5 min. The blotting sandwich was built from layering, in order of anode to

cathode: transfer stack, PVDF membrane, SDS-PAGE, transfer stack. The sandwich was placed into the cassette and run at 25V, 30 min, at room temperature. Post transfer, the PVDF membrane was rocked for an hour in 15 mL of Blocker Casein in TBS (ThermoFisher). Next, the membrane was washed with 15 mL TBST (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween) for 10 mins repeated two additional times. Then, 5 μL of the HRP-conjugated antibody was applied with 10 mL of TBST and the membrane was then rocked for 1 hour. The membrane was washed as stated above.

Detection of bands was achieved by pipetting 4 mL of the Clarity Western ECL substrate (2 mL of the peroxide reagent, 2 mL of the lumino/enhancer reagent) (Bio-Rad) and allowed it to develop for 5 mins. The antibodies used in this work include the Anti-BACH/BRIP1-HRP (Cusabio), Anti-GST-HRP (Cytiva), Anti-Flag-HRP antibody (R&D Systems). The blots were visualized with Azure Biosystems Azure 600 and analyzed used BioRad Image Lab 6.0 to analyze the blots.

2.6 Results

2.6.1 FANCJ-Flag

The first FANCJ construct expressed and purified was pvL1392-FANCJ-Flag, obtained from Dr. Robert Brosh at the NIH. However, this construct presented challenges in purification and maintaining a sufficient concentration for experiments. Initially, the protein was expressed and purified in Sf9 cells, followed by attempts in ExpiSf9 cells, and finally in HF cells. One factor contributing to the difficulties was that the pvL1392-FANCJ-Flag construct was not codon-optimized for insect cells. To address this, the construct was sent to VectorBuilder for optimization to enhance protein expression. After codon optimization, the new FANCJ-Flag construct was expressed in HF cells and

purified using Flag affinity chromatography. The elution sample was analyzed on a 4-20% SDS-PAGE gel for Coomassie staining and Western blot analysis, as shown in Figure 2.1. The purity of the protein post purification was ~95%. However, the reproducibility of pure protein was difficult and so the construct was switch to the GST-FANCJ-Flag.

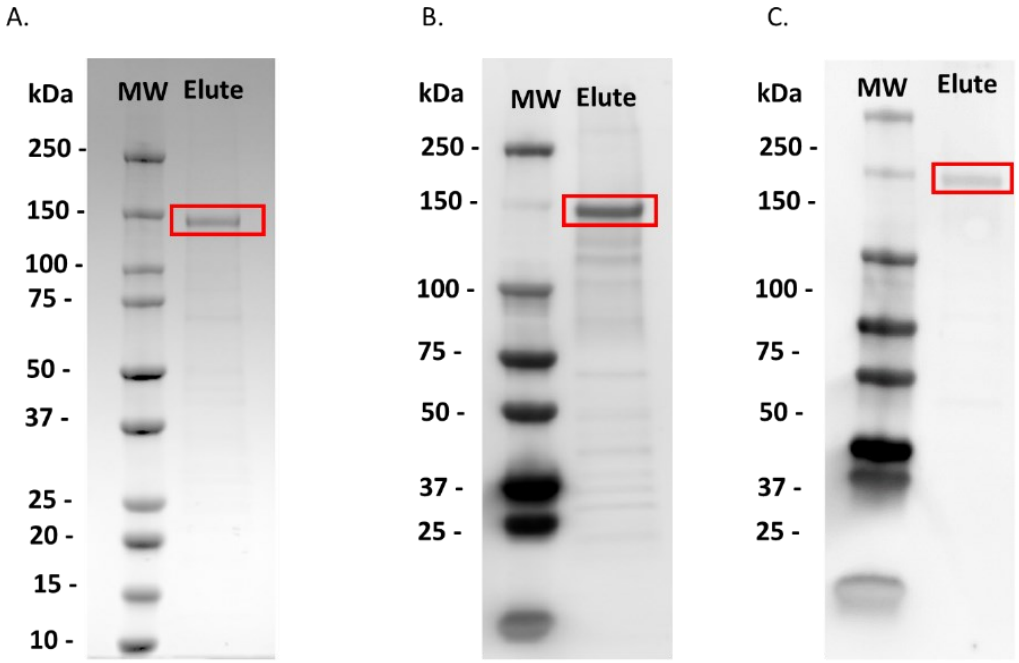

Figure 2.1: SDS-PAGE and Western blots of expressed and purified FANCJ-Flag: A. The gel is a 4-20% SDS - PAGE showing expressed and purified FANCJ-Flag from High Five cells via Flag affinity chromatography. The first lane contains molecular weight marker and elution (Elute). B. Western blot using Flag antibodies. C. Western blot using FANCJ antibodies to ensure the protein is FANCJ.

2.6.2 GST-FANCJ-Flag

GST-FANCJ-Flag construct is in both a pvL1392 vector for insect cells and in pET vector for *E. coli*. Both constructs were codon optimized for the corresponding cell type.

GST-FANCJ-Flag was expressed in NiCo21 cells. After expression, the construct purified first using GST affinity chromatography and then Flag. Figure 2.2 the SDS-PAGE. The purity level was assessed based on the SDS-PAGE, and the purity was >80%. The western blots indicate the band of interest is FANCJ and that there are several truncations of the protein that are not prominent in the SDS-PAGE.

GST-FANCJ-Flag was expressed in HF cells. After expression the cells were purified only using GST affinity chromatography. Figure 2.3 shows the SDS-Page and Western blost for post purification. The expression level is lower for this construct; however, this construct had the best purity of >90% based off on the analysis of the SDS-PAGE. This construct will be used for the DNA unfolding assays in the chapter 5.

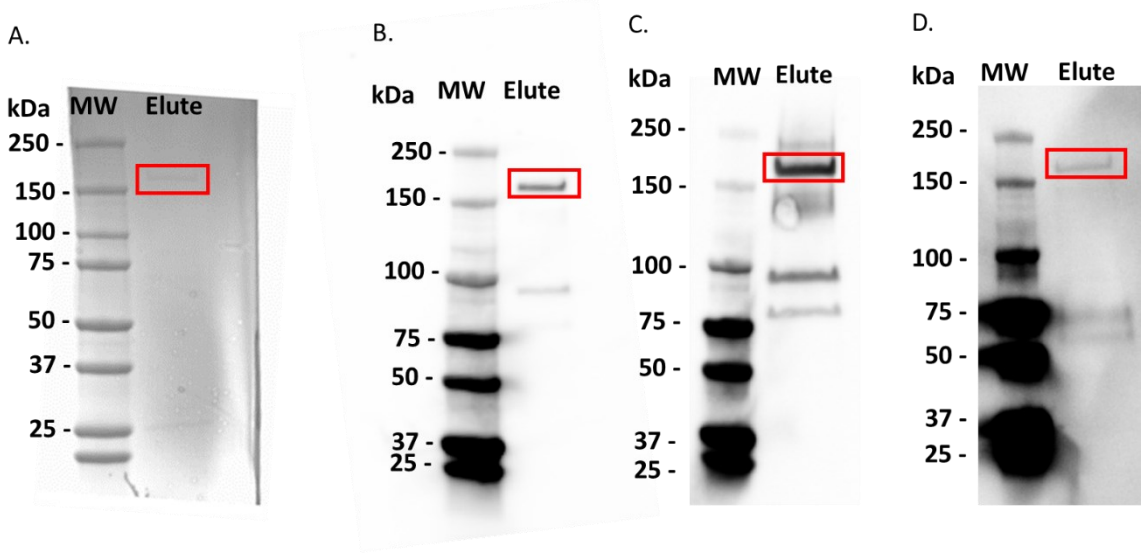

Figure 2.2: SDS-PAGE and Western blots of expressed and purified GST-FANCJ-Flag from High Five Cells: A. The gel is a 4-20% SDS - PAGE showing expressed and purified FANCJ-Flag from High Five cells via GST affinity chromatography. The first lane contains molecular weight marker and elution (Elute). B. Western blot using GST antibodies. C. Western blot using Flag antibodies D. Western blot using FANCJ antibodies to ensure the protein is FANCJ.

2.6.3 6x His-REV1 CTD, 6x His-GST-REV1 Core , and REV1 Core E466A/Y470A

6x His-GST-REV1 Core and the E466A Y470A variant were expressed and purified from BL21 (DE3) cells. Both constructs were purified using GST affinity chromatography. The elution of the constructs was run on a 12% SDS-PAGE and have 90% purity. REV1 CTD was expressed and purified from NiCo21 cells. The protein was purified using Ni affinity chromatography. The protein is over 90% pure (Figure 2.4).

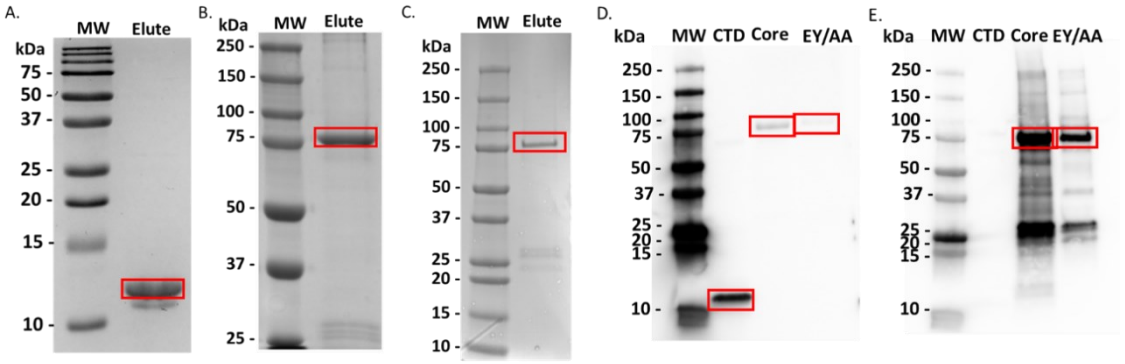

Figure 2.3: SDS-PAGE and Western blots of expressed and purified REV1 Core and REV1 CTD: A. The gel is a 4-20% SDS - PAGE showing expressed and purified REV1 CTD using Nickle affinity chromatography. B. 4-20% SDS-PAGE showing purified REV1 Core. C. 4-20% SDS-PAGE showing purified REV1 Core E466A Y470A D. Western blot using His antibodies. E. Western blot using GST antibodies.

2.7 Conclusion

In this chapter, the optimization of protein expression and purification methods for FANCJ and REV1 were described. The primary goal was to maximize the yield and purity of these proteins for further study. Various forms of affinity chromatography were used to achieve the desired level of purity, including different techniques to optimize the conditions for each protein. Despite facing challenges such as managing solubility and minimizing nonspecific binding, strategies were found through adjustments in expression systems and purification protocols. The purity of FANCJ and REV1 was evaluated using

SDS-PAGE, and Western blotting confirmed the identity of the target proteins. For the following studies FANCJ-Flag will be used for the binding studies and the GST-FANCJ-Flag will be used for the unfolding studies. Overall, the optimization process resulted in efficient and high-yield purification, setting a solid foundation for subsequent research on these proteins.

CHAPTER THREE

CONFORMATION AND STABILITY OF 8-OXO MODIFIED G-QUADRUPLEXES

3.0 Introduction

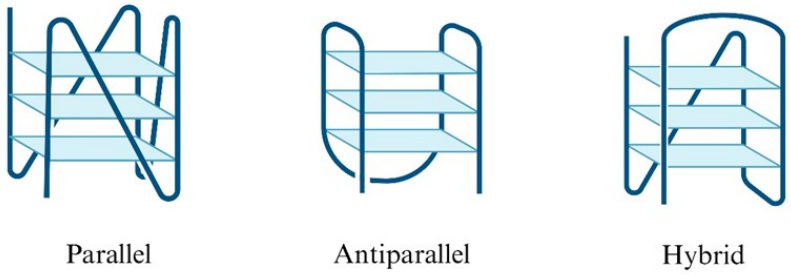

Figure 3.1: G-quadruplex conformations

G-quadruplexes (G4s) can adopt a multitude of spatial conformations that are governed by different parameters, such as the presence of monovalent cations or the sequence of the intervening loops [121-124]. The most typical conformations, as illustrated in Figure 3.1, are classified as parallel, antiparallel, and hybrid. Each conformation exhibits distinct optical properties, which can be discerned using circular dichroism (CD) spectroscopy [125, 126]. Specifically, parallel conformations exhibit a maximum absorption signal around 262 nm and a minimum around 240 nm. In contrast, antiparallel conformations display two peak signals at approximately 295 and 240nm, with a trough at 260 nm. The hybrid conformations are characterized by a peak at 295 nm, a shoulder around 260 nm, and a minimum at 240 nm as seen in Figure 3.2B. Additionally, it is important to note that G-quadruplexes can exist as a mixture of these

conformations, leading to composite spectroscopic profiles that reflect contributions from each structural type. This complexity is significant in both research and application, as it affects the interpretation of CD spectroscopic data and the understanding of G-quadruplex functions.

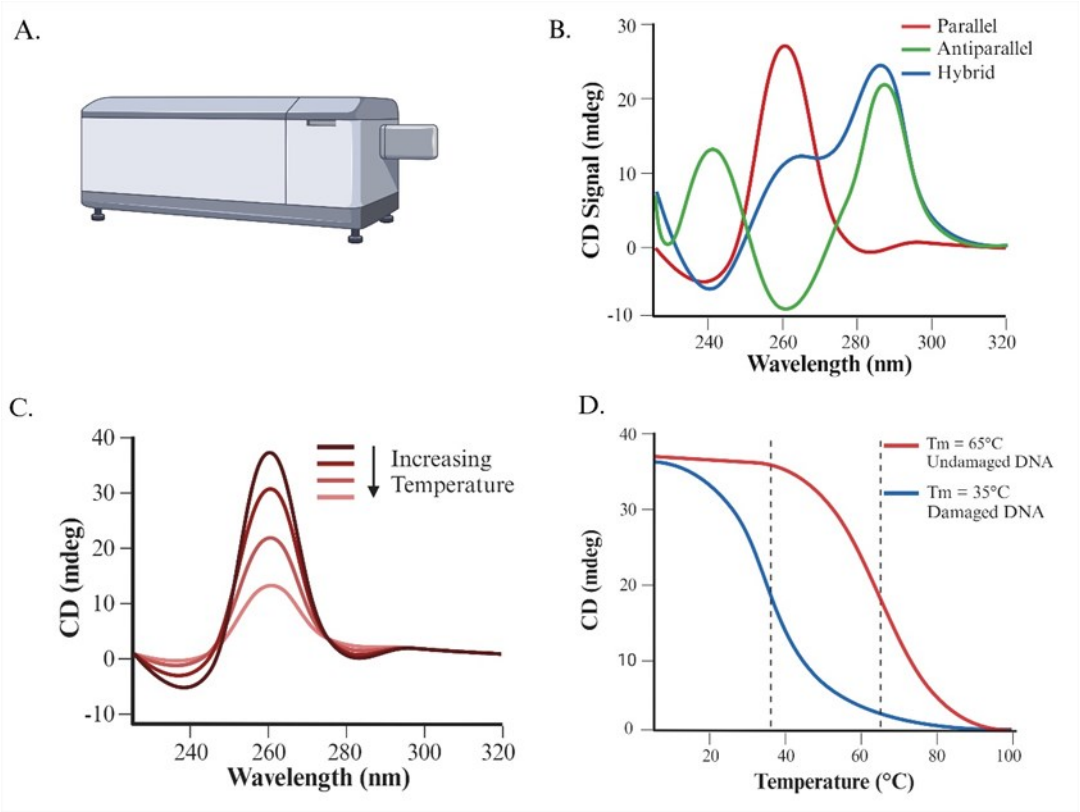

Figure 3.2: Graphic representation of CD experiments A. Example of a CD spectrometer B. Example of CD spectra for parallel, antiparallel, and hybrid G4 structures. C. Example spectra as G4s are unfolded and the temperature increases. D. Graphical representation of a thermal melt for damaged and undamaged G4 DNA.

Circular dichroism (CD) spectroscopy is a technique used to study the chiral properties of molecules and their secondary structures [127]. CD spectroscopy measures the differential absorption of left-handed and right-handed circularly polarized light, which can provide insights into the overall fold and conformation of biomolecules. It is particularly valuable for assessing the stability of G-quadruplexes (G4s) by determining

their melting temperature (Tm) [21, 128-130]. As demonstrated in Figure 3.2 C, the intensity of the spectral peak associated with G-quadruplex structure diminishes as the G4 unfolds with the increasing of temperature. This decrease in intensity, plotted as a function of temperature, allows for the analysis of the melting temperatures, as depicted in Figure 3.2D. The melting temperature of G4 structures is directly correlated with their thermal stability, indicating that a higher Tm signifies greater stability.

Guanine bases are particularly vulnerable to oxidative damage from reactive oxygen species (ROS) due to their lowest oxidation potential among the DNA bases, making them a common target in the genome [131]. 8-oxoguanine (8oxoG) is a prevalent oxidative lesion resulting from this vulnerability, generated through various endogenous and exogenous sources. The accumulation of 8oxoG within DNA is a significant concern because it mispairs with adenine during DNA replication, leading to G:C to T:A transversion mutations [132]. These mutations can disrupt gene function and expression, potentially resulting in altered protein synthesis and functionality. Over time, the persistence of such mutations within regulatory genes may lead to oncogenesis or contribute to the progression of age-related diseases.

Moreover, 8oxoG can also incorporate into G4s, influencing their stability and structural conformation. Studies have shown that 8oxoG may decrease the thermal stability of G-quadruplexes, which could alter their biological functions [17, 133]. This destabilization may render the DNA more susceptible to nuclease activity, potentially leading to DNA breaks and genomic instability [14]. Additionally, the presence of 8oxoG in G-quadruplexes can affect the binding of proteins and small molecules, which are critical for the biological roles of these structures [14, 131]. The integration of 8oxoG

into G4s highlights the intricate interplay between oxidative stress and genomic regulation, underscoring the importance of understanding these processes to develop targeted therapeutic strategies.

This study will explore the thermal stability of G4 sequences with varying intervening sequence lengths and to examine the impact of 8oxoG at the first, third, or fifth guanine position within these sequences. It is anticipated that oxidative damage will compromise the thermal stability compared to undamaged sequences. Additionally, the study will investigate how the loop length influences the conformation and stability of G4 structures in the presence of both potassium and sodium ions. This comprehensive approach will help elucidate the intricate relationship between sequence variation, ionic conditions, and oxidative stress, providing deeper insights into the structural dynamics of and stability of G4s.

3.1 Materials

3.1.1 DNA Oligos

The G4 DNA oligos were synthesized by IDT (Integrated DNA Technologies), and the specific sequences are listed in Table 3.1. All G4 DNA samples were dialyzed into 20 mM Boric Acid, 150 mM KCl or NaCl, at pH 7.5. After dialysis, the samples were heated to 95°C for 10 minutes and then allowed to cool overnight prior to the temperature melting experiments. DNA concentrations were assessed using NanoDrop One (Thermo Fisher Scientific), utilizing the 260 nm molar extinction coefficients provided by IDT.

Name	Sequence 5' – 3'
GGGT	GGGTGGGTGGGTGGGT
GGGT 8oxo1	T/i8oxodG/GGTGGGTGGGTGGGT
GGGT 8oxo3	GG/i8oxodG/TGGGTGGGTGGGT
GGGT 8oxo5	GGGTG/i8oxodG/GTGGGTGGGT
GGGTT	GGGTTGGGTTGGGTTGGGTT
GGGTT 8oxo1	/i8oxodG/GGTTGGGTTGGGTTGGGTT
GGGTT 8oxo3	GG/i8oxodG/TTGGGTTGGGTTGGGTT
GGGTT 8oxo5	GGGTTG/i8oxodG/GTTGGGTTGGGTT
TeloG4	TTAGGGTTAGGGTTAGGGTTAGGG
TeloG4 8oxo1	TTA/i8oxodG/GGTTAGGGTTAGGGTTAGGG
TeloG4 8oxo3	TTAGG/i8oxodG/TTAGGGTTAGGGTTAGGG
TeloG4 8oxo5	TTAGGGTTAG/i8oxodG/GTTAGGGTTAGGG

Table 3.1: List of G-quadruplex sequences

3.2 Method

3.2.1 CD spectroscopy

CD measurements were conducted using a JASCO J-815 spectropolarimeter, equipped with a PTC-423s Peltier system and a Koolance Liquid Cooling System. CD spectra were scanned from 220 nm to 320 nm at 25°C using a 1.0-mm path length quartz cuvette with a reaction volume of 200 uL. The concentration of the DNA samples was 0.3 mg/mL. An initial reference scan of the buffer was performed to be subtracted from the DNA spectra scans. Five traces were collected for each DNA sample and subsequently averaged.

For temperature melts, scans were conducted from 15°C to 95°C at 1°C/min. Maximum absorbance was used for the melting experiments, at 262 nm for parallel structures, or at 295 nm for antiparallel and hybrid structures. These melts were repeated

in triplicate for each DNA construct (n=3), the error is based on the standard deviation of the three trials. The temperature melts were analyzed using the following model described by Greenfield [134]:

$$m=tm+273.15 \quad (1)$$

$$t=v+273.15 \quad (2)$$

$$k=\exp((h/(1.987*t))*((t/m)-1)) \quad (3)$$

$$y=k/(1+k) \quad (4)$$

$$f=((u-l)*y)+l \quad (5)$$

Where h is the enthalpy, u is the mean residue ellipticity of 100% folded DNA, l is mead residue ellipticity of unfolded DNA, and tm is the T_M in °C. These equations account for a two-state transition of the G4 from a folded to an unfolded state and was programmed into SigmaPlot.

3.3 Results

CD spectroscopy was utilized to assess the conformation and thermal stability of G4 structures. Three undamaged DNA sequences were selected based on the length of intervening sequences and their biological significance. The sequence $(TTAGGG)_4$ was chosen because of its role as the human telomeric DNA sequence, which has been studied extensively and has significant biological relevance. The sequence $(GGGT)_4$ was selected due to the compact G4 structure formed by the single thymidine loop and is also an aptamer of HIV-1 integrase. Lastly, the artificial sequence GGGTT was designed with an intervening sequence of intermediate length to explore variations in stability.

For each sequence, specific positions were modified with an 8-oxoguanine to study the impact on thermal stability. These modifications were made at the first (8oxo1),

third (8oxo3), and fifth (8oxo5) guanine positions. Based on findings from previous studies on the human telomeric sequence, 8oxo5 significantly decreased thermal stability, 8oxo1 was the least destabilizing, and 8oxo3 resulted in intermediate destabilization. Additionally, these modifications were strategically placed to introduce an 8-oxoguanine within one of the three tetrads of the G4 structure, further exploring the structural implications.

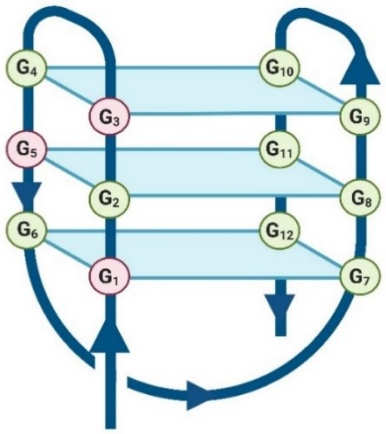

Figure 3.3: 8-oxoguanine positions chosen within the G4. 8-oxoguanine was placed at the first, third, or fifth position within the G4 and are highlighted in red.

It was found that GGGT sequences in KCl produced parallel structures with a peak at 262 nm (Figure 3.4 A). The conformations of the damaged GGGT sequences did not change based on the overlaying CD spectra traces. Thermal melts show that GGGT has a melting temperature of 88.4 ± 8.4˚C. 8oxo1 produced the highest decrease in thermal stability with the ΔT_m= -19.0 ± 8.4˚C, while 8oxo3 produced the lowest change in thermal stability of ΔT_m= -8.1 ± 8.4˚C.

The GGGTT sequence also forms parallel conformation in the presence of KCl. The 8oxo1 and 8oxo5 modified sequences also produced a parallel conformation with slight structural changes based on the differences in traces in the 300 to 280 nm range.

However, the 8oxo3 modified GGGTT sequence changed to an antiparallel sequence. The thermal melts show that GGGTT has a melting temperature of 83.5 ± 0.7˚C. Similar to GGGT in KCl, 8oxo1 produced the highest decrease in thermal stability with a ΔT_m= -31.5 ± 1.6˚C, while 8oxo3 produced the lowest with the ΔT_m= -26.3 ± 0.7˚C.

The TTAGGG sequence forms hybrid structures in the presence of KCl, and the 8oxo1 and 8oxo5 remain in the hybrid conformation. Similarly to GGGTT, the addition of the 8oxo3 damage caused a conformational change to an antiparallel structure. The melting temperature for TTAGGG is 62.5 ± 0.5˚C, and 8oxo5 had the highest decrease in thermal stability of ΔT_m= -23.6 ± 0.8˚C.

Substrate	Conformation	Tm (°C)	ΔTm (°C)
GGGT	Parallel	88.4 ± 8.4	———
GGGT 8oxo1	Parallel	69.4 ± 0.6	-19.0 ± 8.4
GGGT 8oxo3	Parallel	80.3 ± 0.4	-8.1 ± 8.4
GGGT 8oxo5	Parallel	71.6 ± 0.4	-16.8 ± 8.4
GGGTT	Parallel	83.5 ± 0.7	———
GGGTT 8oxo1	Parallel	52.0 ± 1.4	-31.5 ± 1.6
GGGTT 8oxo3	Antiparallel	57.2 ± 0.1	-26.3 ± 0.7
GGGTT 8oxo5	Parallel	55.9 ± 0.3	-27.6 ± 0.8
TeloG4	Hybrid	62.5 ± 0.5	———
TeloG4 8oxo1	Hybrid	54.5 ± 0.5	-8.0 ± 0.7
TeloG4 8oxo3	Antiparallel	49.5 ± 0.3	-13.0 ± 0.6
TeloG4 8oxo5	Hybrid	38.9 ± 0.6	-23.6 ± 0.8

Table 3.2: Summary of G4 conformations and melting temperatures in KCl

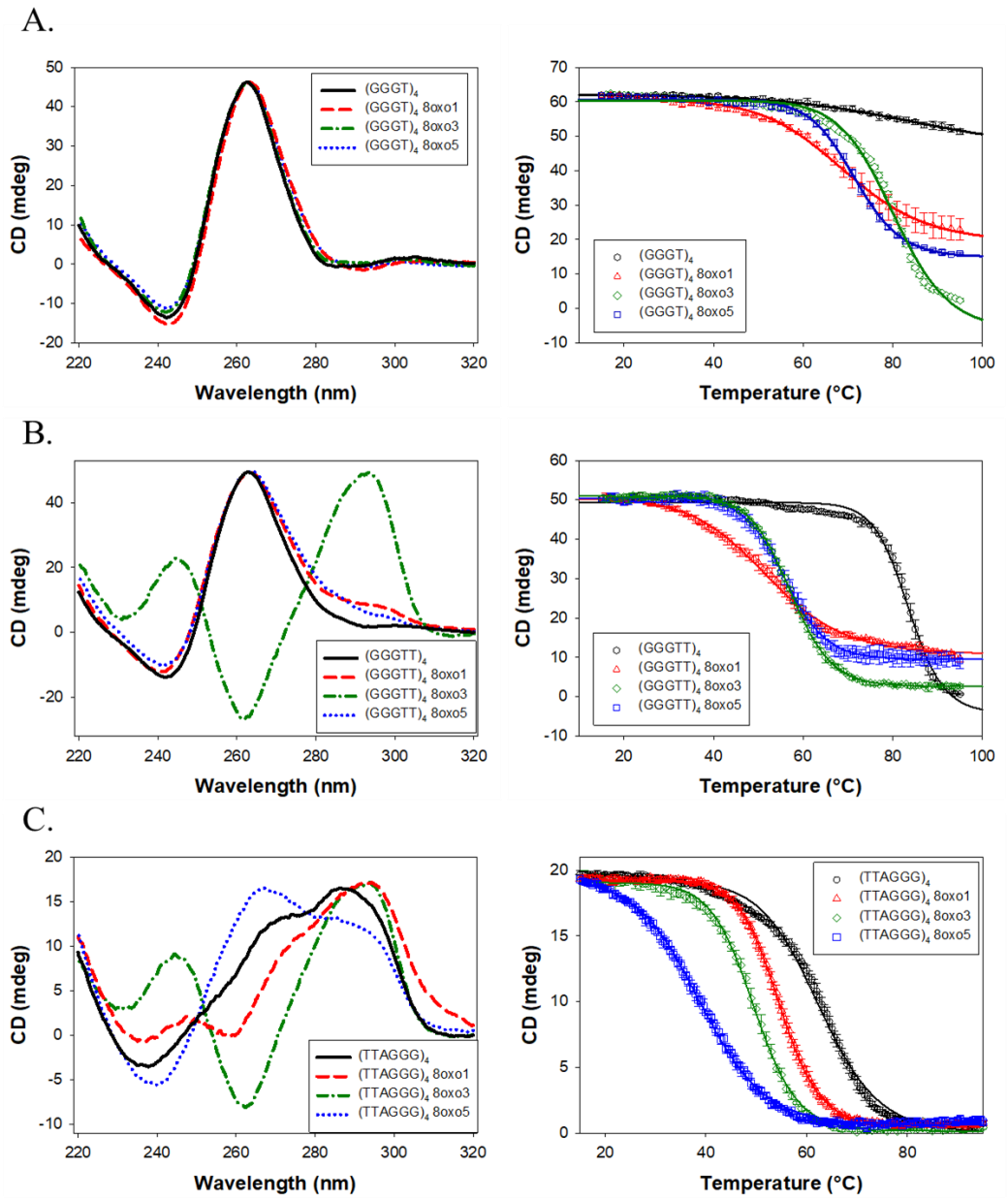

Figure 3.4 KCl CD spectra and melts. A. CD spectra and melts of GGGT G4 DNA. B. GGGTT C.TTAGGG

Figure 3.5 shows the CD spectra and thermal melts in the presence of Na^+ ion. It was found that GGGT sequences in NaCl produced a parallel structures with a peak at 262 nm. The conformations of the damaged GGGT sequences did not change based on

the overlaying CD spectra traces. Thermal melts show that GGGT has a melting temperature of 70.3 ± 0.7˚C. 8oxo1 produced the highest decrease in thermal stability with the ΔT_m= -27.8 ± 0.9˚C, while 8oxo3 produced the lowest change in thermal stability of ΔT_m= -13.9 ± 0.8 ˚C.

Substrate	Conformation	Tm (°C)	ΔTm (°C)
GGGT	Parallel	70.3 ± 0.7	———
GGGT 8oxo1	Parallel	42.5 ± 0.5	-27.8 ± 0.9
GGGT 8oxo3	Parallel	56.4 ± 0.4	-13.9 ± 0.8
GGGT 8oxo5	Parallel	44.0 ± 0.3	-26.3 ± 0.8
GGGTT	Parallel	44.8 ± 1.2	———
GGGTT 8oxo1	Antiparallel	47.2 ± 0.3	2.4 ± 1.2
GGGTT 8oxo3	Antiparallel	44.6 ± 0.2	-0.2 ± 1.2
GGGTT 8oxo5	Hybrid-esque	27.5 ± 0.8	-17.3 ± 1.4
TeloG4	Antiparallel	52.1 ± 0.3	———
TeloG4 8oxo1	Antiparallel	47.8 ± 0.6	-4.3 ± 0.7
TeloG4 8oxo3	Antiparallel	42.2 ± 0.7	-9.9 ± 0.8
TeloG4 8oxo5	Antiparallel	33.3 ± 0.4	-18.8 ± 0.5

Table 3.3: Summary of G4 conformations and melting temperature in NaCl

The GGGTT formed parallel conformation in NaCl. However, 8oxo1 and 8oxo3 sequences formed antiparallel conformation, while 8oxo5 formed hybrid-esk conformation. The thermal melts show that GGGTT has a melting temperature of 44.8 ± 1.2˚C. 8oxo5 produced the highest decrease in thermal stability with a ΔT_m= -17.3 ± 1.4 ˚C, while 8oxo3 produced the lowest with the ΔT_m= -0.2 ± 1.2 ˚C.

The TTAGGG and all damaged constructs formed antiparallel structures in NaCl. The melting temperature for TTAGGG is 52.1 ± 0.3˚C, and 8oxo5 had the highest decrease in thermal stability of ΔT_m= -18.8 ± 0.5˚C, while 8oxo1 produced the lowest with the ΔT_m= -4.3 ± 0.7 ˚C.

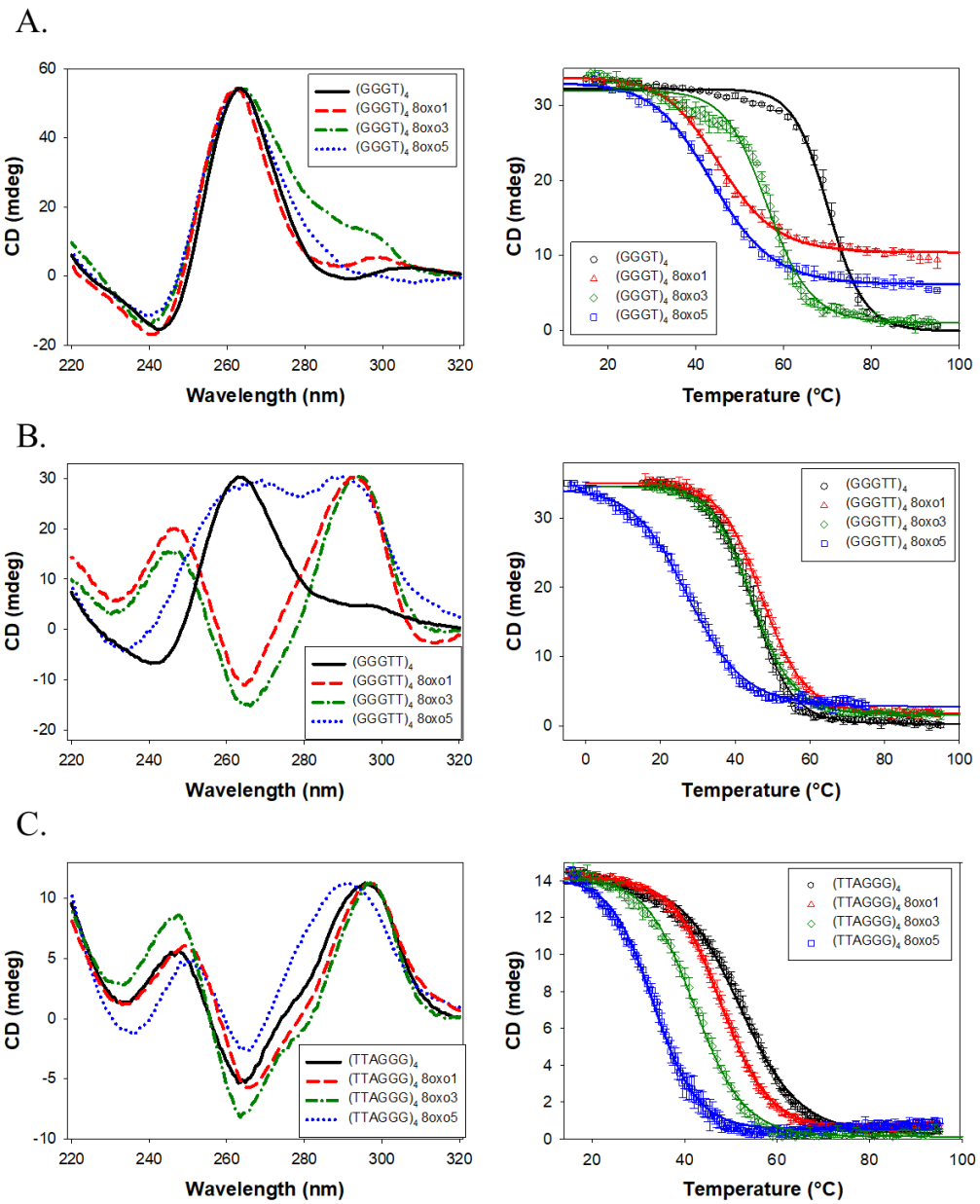

Figure 3.5: KCl CD spectra and melts. A. CD spectra and melts of GGGT G4 DNA. B. GGGTT C.TTAGGG

3.4 Discussion

Previous CD studies for the human telomeric sequences align with the current CD results, both agreeing that TTAGGG and its 8oxoG4 variants form hybrid and antiparallel conformations in potassium ions, and antiparallel conformations in NaCl [14, 135, 136]. Moreover, the 8oxo5-modified sequences exhibit the greatest thermal instability, while 8oxo1-modified sequences show the least instability in both KCl and NaCl solutions. A similar pattern of decreased thermal stability is observed with the GGGTT sequences, which the 8oxoG4 analogs adopt antiparallel or hybrid conformations in NaCl.

Different stability patterns emerged for the other substrates. In both KCl and NaCl, GGGT sequences, along with GGGTT sequences in KCl, primarily form parallel conformations. However, GGGTT 8oxo3 adopts an antiparallel conformation. Damage at the third position may cause a change in the guanine to go from a *syn* to an *anti*-conformation or vice versa in order to preserve the integrity of the structure [14]. These sequences show a consistent decrease in thermal stability, with 8oxo1-modified sequences inducing the most instability, and 8oxo3 the least.

These findings suggest that the impact of 8-oxoguanine on the thermal stability of G-quadruplex structures depends on the sequence position and G4 conformation. Sequences forming compact parallel conformations are most destabilized by damage to the third guanine, located at the top of the tetrad. Conversely, sequences with more flexible antiparallel or hybrid conformations are most destabilized by damage to the fifth guanine, located in the middle tetrad, and least by damage to the first guanine.

Among monovalent cations, potassium and sodium ions are the most physiologically relevant and commonly studied concerning G4 conformations and

stability. While GGGT and its respective 8oxoG4 derivatives form parallel conformations in both KCl and NaCl, GGGTT and TeloG4 sequences generally do not adopt the same conformations. Moreover, all sequences exhibit lower thermal stability in NaCl than in KCl, suggesting that sodium ions are less effective at stabilizing G4 structures. In the presence of only sodium ions, less stable G4 sequences fail to adopt the rigid conformations seen in potassium ions and settle into less stable conformations. These results align with prior studies, indicating that although both ions can stabilize G4 structures, potassium ions are optimally sized and have the appropriate ionic strength for this purpose.

An additional observation in KCl is the correlation between the intervening sequence length and the thermal stability of native sequences. GGGT exhibits the highest melting temperature, while TTAGGG has the lowest. This suggests that increasing the length of the intervening sequence reduces the thermal stability of the G4 structure, a finding that is consistent with other studies.

For future studies, incorporating more 8-oxoguanines within the sequences to determine how it affects the stability in all the positions. Additionally, looking into how 8-oxoguanines affect the stability of other known G4 sequences, such as MyC G4 DNA.

CHAPTER FOUR

RECOGNITION OF G-QUADRUPLEXES BY FANCJ HELICASE AND REV1 POLYMERASE

4.0 Introduction

FANCJ is a potent resolvase of G4 DNA. Missense mutations in the Fe-S cluster of FANCJ domain that are linked to FA or associated with cancer result in the inactivation of its helicase activity on G4 or duplex DNA substrates [137, 138]. There is much interest in G4 DNA, given its unique structure and potential to interfere with DNA synthesis, transcription or recombination [139]. FANCJ suppresses DNA damage and genomic instability caused by G4 in human cells [73, 137]. While FANCJ effectively mitigates DNA damage and instability caused by G4 structures, its similarity to other helicases, such as RHAU (or DHX36), highlights the diverse mechanisms through which cells recognize and interact with G4 and related RNA structures. The RHAU helicase recognizes G4 RNA through a special motif at the protein N-terminus [140]. In solution, this site adopts an α-helix that is followed by an AKKQ loop [95]. It is thought that the helix stacks on top of a guanine tetrad while lysine residues within the AKKQ are anchored into the RNA backbone. FANCJ possesses an AKKQ motif that is flanked by a comparable G4-targeting peptide [91]. However, it is unknown if this motif or if the full-length protein can also identify 8oxoG4s.

In a current model, FANCJ repeatedly unfolds and allows the G4 to refold at stalled replication fork to keep the strand clear for a repair polymerase [91] (Figure 4.1). The REV1 translesion DNA synthesis polymerase has been suggested for this role

because it prefers to incorporate cytosine across from any templated base [111]. REV1 also has a higher affinity for G4-containing DNA and disrupts its structure upon loading [106].

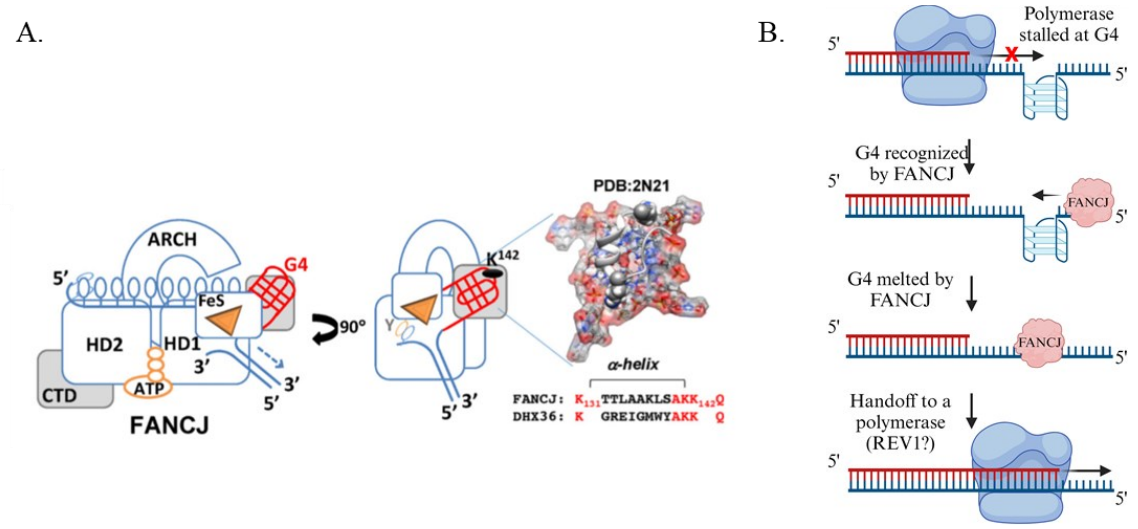

Figure 4.1: Graphical representation of FANCJ model A. Molecular model of FANCJ. FANCJ possesses an AKKQ motif that recognizes and binds to G4 structures. B. Model for DNA synthesis through G4-forming regions by FANCJ. DNA polymerase stalls at a G4 site, FANCJ then binds and unfolds the structure. Then another polymerase, which we believe is REV1, is recruited to continue with replication.

Repair polymerases are commonly recruited to DNA damage via interactions with the PCNA sliding clamp [141]. Their contacts are mediated through a PCNA Interacting Peptide (or PIP) motif [142]. PIP-boxes have loosely defined consensus residues of a glutamine in position 1, a hydrophobic residue in position 4 (L, I, or M), aromatic residues in position 7 and 8 (F or Y). There have been proteins that still bound PCNA but lacked positions 1, 7, or 8 [143]. Diverging PIP-like sequences can bind to other repair proteins including REV1 [113, 141, 144]. Because FANCJ has a potential PIP-motif based on its primary amino acid sequence (SWSSFNSLGQYFTGKIP), we hypothesized that this site can function either as a canonical PIP that recruits REV1 through

interactions with PCNA, or as an REV1 Interacting Region (RIR) that brings the polymerase directly to G4. To test this, we used a biolayer interferometry assay (BLI) to monitor the molecular interactions between the FANCJ PIP, REV1, and PCNA.

This work will examine the interactions of FANCJ and REV1 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 and REV1.

4.1 Materials

4.1.1 Buffers

All buffers and reagents were prepared with reagent-grade chemicals and double distilled water that was further purified with a Smart2Pure 6 UV/UF system (ThermoFisher, Waltham, MA, USA). Solutions were sterilized through a 0.22 micron PES filter. Experiments were performed at 25°C in Buffer H (20 mM HEPES pH 7.5, 150 mM KCl, 5 mM TCEP, and 5% (v/v) glycerol) unless otherwise specified.

4.1.2 FANCJ Peptides

Name	Sequence
bioFANCJ PIP	/Biosg/SWSSFNSLGQYFTGKIP
bioFANCJ PIP AA	/Biosg/SWSSFNSLGQAATGKIP
FANCJ AKKQ	SPEKTTLAAKLSAKKQASIW

Table 4.1: Peptide sequences

FANCJ peptides were synthesized by Genscript (Piscataway, NJ). The lyophilized powders were dissolved in Buffer H, and the concentrations were determined using a

Nanodrop One UV-Vis microvolume spectrometer (ThermoFisher) and the molecular extinction coefficients (E280). Peptide were aliquoted, frozen, and stored at -20°C.

4.1.3 G4 DNA Oligos

Name	Sequence 5' – 3'
bioGGGT	/5Biosg/GGGTGGGTGGGTGGGT
bioGGGT 8oxo1	/5Biosg/T/i8oxodG/GGTGGGTGGGTGGGT
bioGGGT 8oxo3	/5Biosg/GG/i8oxodG/TGGGTGGGTGGGT
bioGGGT 8oxo5	/5Biosg/GGGTG/i8oxodG/GTGGGTGGGT
bioGGGTT	/5Biosg/GGGTTGGGTTGGGTTGGGTT
bioGGGTT 8oxo1	/5Biosg//i8oxodG/GGTTGGGTTGGGTTGGGTT
bioGGGTT 8oxo3	/5Biosg/GG/i8oxodG/TTGGGTTGGGTTGGGTT
bioGGGTT 8oxo5	/5Biosg/GGGTTG/i8oxodG/GTTGGGTTGGGTT
bioTeloG4	/5Biosg/TTAGGGTTAGGGTTAGGGTTAGGG
bioTeloG4 8oxo1	/5Biosg/TTA/i8oxodG/GGTTAGGGTTAGGGTTAGGG
bioTeloG4 8oxo3	/5Biosg/TTAGG/i8oxodG/TTAGGGTTAGGGTTAGGG
bioTeloG4 8oxo5	/5Biosg/TTAGGGTTAG/i8oxodG/GTTAGGGTTAGGG

Table 4.2: DNA oligos used for biolayer interferometry

The G4 DNA oligos for BLI were synthesized by IDT (Integrated DNA Technologies), and the specific sequences are listed in Table 4.1. The samples were heated to 95°C for 10 minutes and then allowed to cool overnight prior to the temperature melting experiments. DNA concentrations were assessed using NanoDrop One (Thermo Fisher Scientific), utilizing the 260 nm molar extinction coefficients provided by IDT. G4 DNA used for fluorescence spectroscopy can be found in Table 3.1 in Chapter 3.

4.2 Methods

4.2.1 Fluorescence Spectroscopy

Equilibrium binding experiments were conducted using a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies) with the samples maintained at 25 ◦C using a PCB 1500 water Peltier system (Agilent). Buffer H was used to establish a baseline. Then, 5 μM of FANCJ peptide was placed into a microcuvette. Fluorescence spectra were collected from 300 nm to 450 nm, with λ_{ex} at 280 nm and slit widths of 5 nm.

The DNA substrate was then titrated into the cuvette. Each mixture was allowed to equilibrate for 3 minutes before data collection commenced. Fluorescence quenching was quantified using the following equation:

$$\Delta Fobs = (F0 - Fi)/ F0 \quad (1)$$

F0 represents the initial fluorescence signal of the peptide alone, and Fi is the fluorescence intensity after each addition of DNA.

Binding isotherms were created by plotting fluorescence quenching observed (ΔFobs) against the total DNA concentration, adjusted for sample dilution. These curves were analyzed using a 1:1 interaction model through Scientist 2.0 software (Micromath). The equilibrium K was calculated using the following formulas, where A is the amplitude of fluorescence quenching and K was the equilibrium association constant. D_f and D_t are the free and total concentration of DNA, while P_f and P_t describe the free and total peptide concentrations. All fluorescence titrations were performed in triplicate. The reported values are based on the average and the 95% confidence interval from the three independent datasets.

4.2.2 Biolayer Interferometry (BLI)

Biolayer Interferometry (BLI) is an advanced, label-free analytical technique that measures interactions between biomolecules in real time by detecting changes in light interference. This occurs when a sample binds to the biosensor, altering the optical thickness of the biosensor layer and consequently the light interference pattern. This shift is continuously measured, producing a dynamic graph that illustrates the kinetics of the molecular interactions (Figure 4.1).

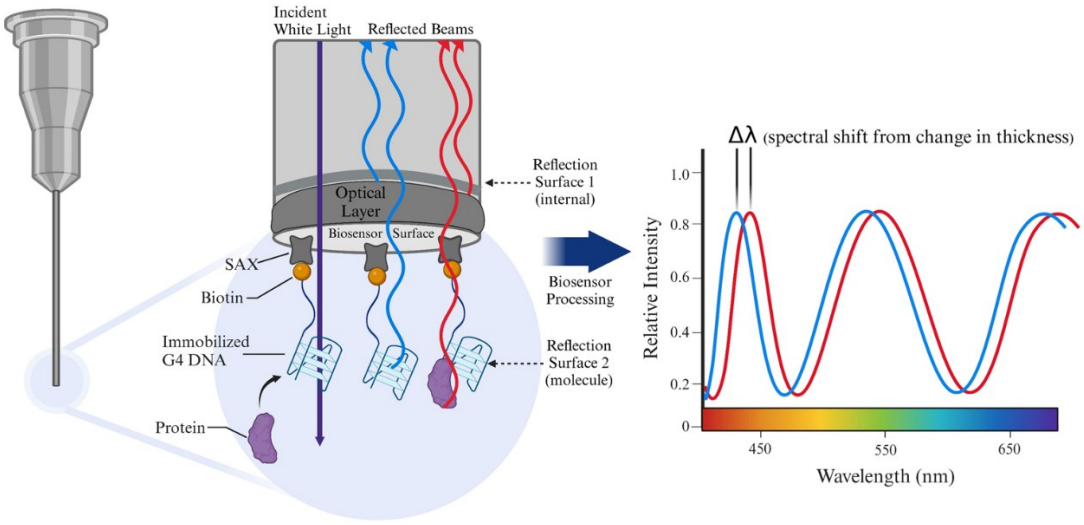

Figure 4.2: Overview of a bio-layer interferometry setup

BLI assays were carried out using a BLItz instrument (Sartorius). Streptavidin-coated biosensors, obtained from Sartorius, were hydrated in Buffer H for 10 min, following which the baseline optical interference was established. Biosensors were then loaded with 150 nM of either biotinylated G4 DNA or FANCJ (PIP or GST-labeled) over a period of 2 minutes and subsequently washed with Buffer H. The binding reactions commenced once the biosensor were immersed in analyte solutions at specified concentrations. After reaching equilibrium, the biosensors were placed back in the buffer

to observe the dissociation of the complexes. Figure 4.2 Illustrates a representative time course, demonstrating the typical progression of the biomolecular interactions observed during the assay.

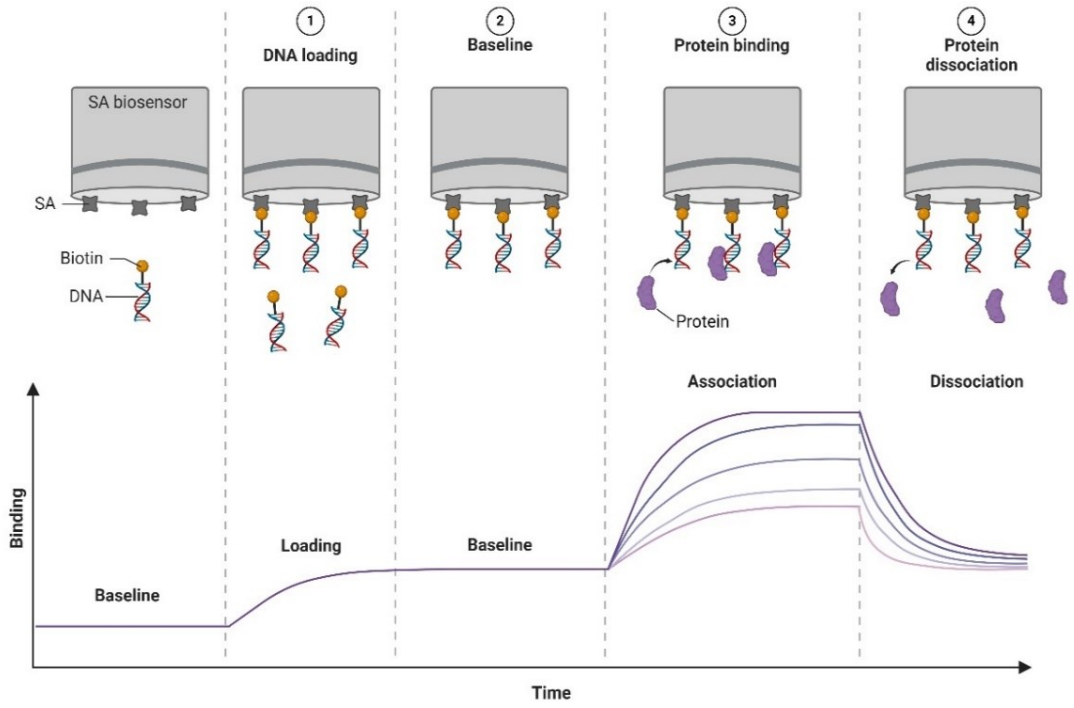

Figure 4.3: Overview of BLI data collection

The experiments were conducted in triplicates to ensure reliability. BLI binding isotherms were then created by plotting the plateau values of the association phase against the total analyte concentration. The plateau values were fitted to a 1:1 interaction model.

$$R = (R\_0/k\_obs\,)(1 - e^{-k_{obs}(t-t_0)}),\ k_{obs} = k_{on}[Protein] + k_{off} \quad (1)$$

$$R = R_0 e^{-k_{off}(t-t_0)} \quad (2)$$

The association phases are fit to Eq. (1), in which the binding response signal (R) is described by an exponential increase function with time constant k_{obs}. The dependence of k_{obs} on FANCJ or REV1 concentration is used to determine the association rate constant

k_{on}. The dissociation phases are analyzed by an exponential decay function, Eq. (2). Since the binding G4 complex is a unimolecular process, the dissociation rate constant k_{off} is independent of protein concentration. The ratio of the two rate constants is used to calculate the affinity for the G4 interaction ($KD = k_{off}/k_{on}$). The parameters reported are derived from the average and the 95% confidence interval of the three datasets.

4.3 Results

4.3.1 FANCJ AKKQ peptide binding to G4 DNA

To test whether the FANCJ AKKQ motif alone can target G4 structures, we used fluorescence spectroscopy to monitor the binding interactions between an FANCJ peptide and a G4 formed by the human telomeric sequence $(TTAGGG)_4$. A tryptophan residue was substituted for tyrosine at the C-terminal end of SPEKTTLAAKLSAKKQASIW (FANCJ aa 128–147) to serve as a reporter (Figure 4.3 A). The free peptide produced a fluorescence peak from 300 to 450 nm; equilibrium titrations of $(TTAGGG)_4$ quenched this signal with increasing DNA concentration (Figure 4.3 B). The extent of fluorescence quenching was used to generate a binding isotherm (Figure 4.3 C). Nonlinear least squares analysis of this curve to a single-site interaction model resulted in an affinity value of $K = 1.6 \pm 0.7 \times 10^6\ M^{-1}$.

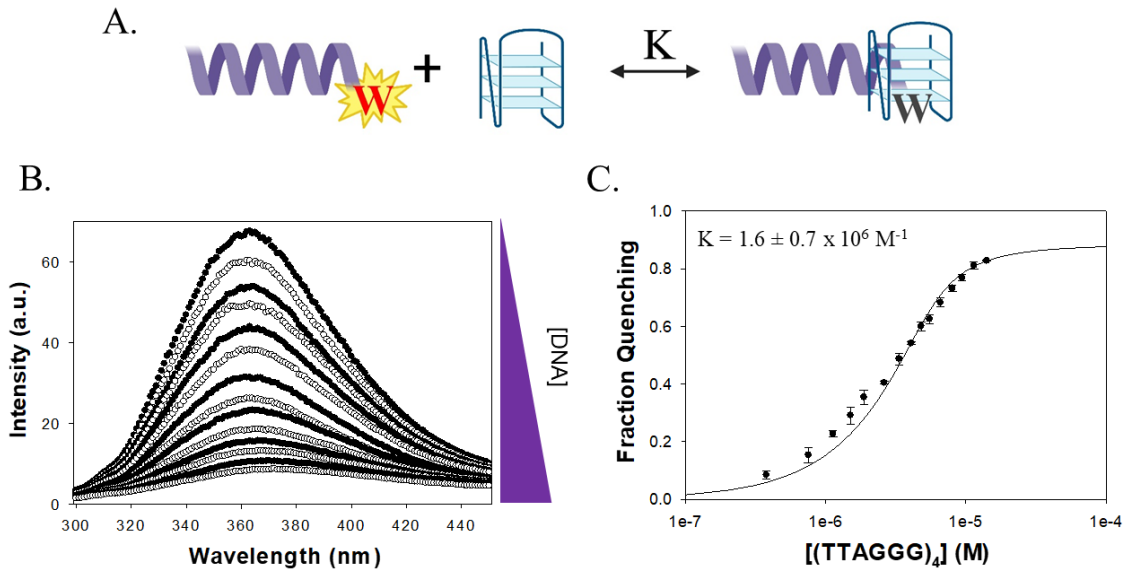

Figure 4.4: Overview of fluorescence spectroscopy study A. A tryptophan residue was incorporated into the AKKQ peptide as a fluorescence probe. B. The peptide produced a large fluorescence peak alone in Buffer H. This signal was quenched upon titration of the human telomeric G4 $(TTAGGG)_4$. C. The extent of fluorescence quenching was plotted against total DNA concentration to generate a binding isotherm. Nonlinear least squares analysis of this curve to a 1:1 binding model.

To further understand how specific modifications affect FANCJ AKKQ-G4 interactions, we examined the incorporation of 8-oxoguanine.We used 8oxoG4 substrates with a damaged base at the first (8oxo1), third (8oxo3), or fifth (8oxo5) guanine position for fluorescence titrations. The FANCJ AKKQ peptide bound to all 8oxoG4 substrates with similar affinities as the native substrate; $K_{8oxo1} = 1.3 \pm 0.4 \times 10^6\ M^{-1}$, $K_{8oxo3} = 1.2 \pm 0.1 \times 10^6\ M^{-1}$, and $K_{8oxo5} = 1.4 \pm 0.5 \times 10^6\ M^{-1}$ (Figure 4.4 A). The equilibrium binding and CD results suggested that the FANCJ AKKQ targeted the TTA loops of $(TTAGGG)_4$, since 8-oxoguanine altered the hybrid G4 configuration but had a negligible effect on peptide binding (Figure 4.4 A).

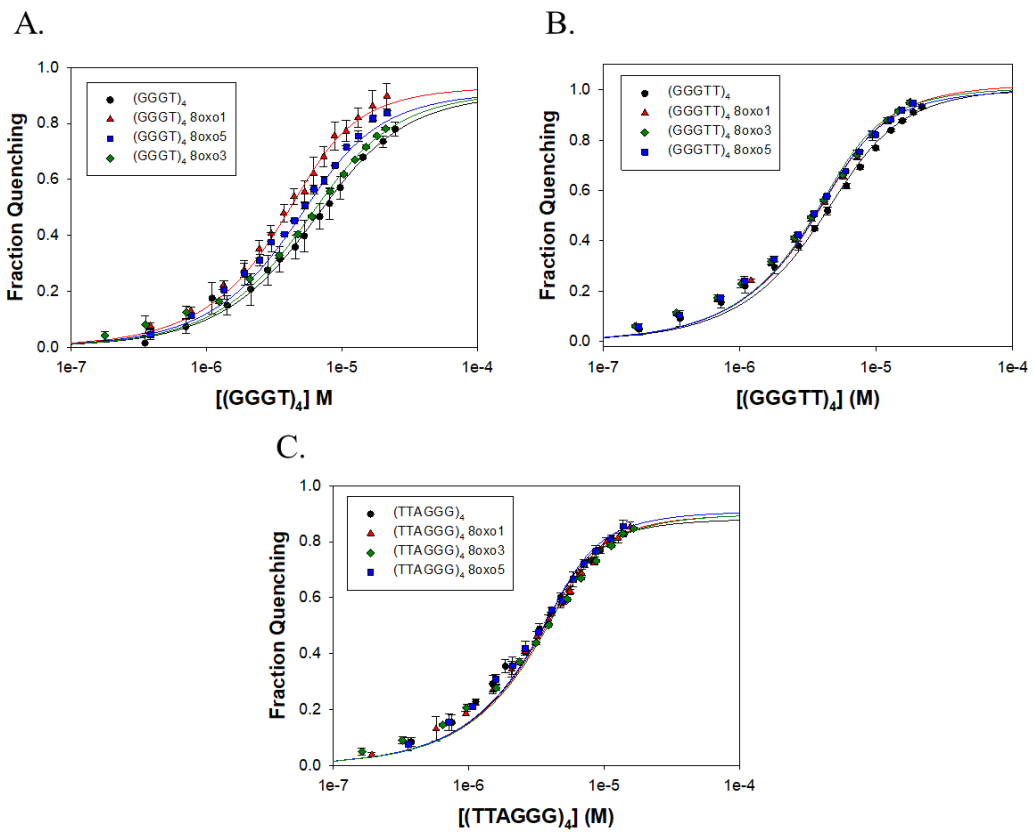

Figure 4.5: Fluorescence binding isotherms for FANCJ AKKQ A. FANCJ AKKQ equalibrium binding to $(GGGT)_4$ and respective 8oxoG4 constructs B. FANCJ AKKQ binding to $(GGGTT)_4$ and the 8oxoG4s C. FANCJ AKKQ binding to $(TTAGGG)_4$ and the 8oxoG4 sequences.

We then tested the binding interactions between FANCJ AKKQ and different sequences. Figure 4.4 A illustrates the binding interactions between FANCJ AKKQ and $(GGGT)_4$. A clear shift in the binding curve from the native sequence to the damaged sequences indicated an increase in binding affinity. These binding affinities are listed in Table 4.3. The strongest interaction was observed with GGGT 8oxo1 with $K_{8oxo1} = 7.7 \pm 0.7 \times 10^5\ M^{-1}$. This was followed by interactions with 8oxo5 ($K = 4.4 \pm 0.6 \times 10^5\ M^{-1}$) and 8oxo3 ($K = 3.2 \pm 0.1 \times 10^5\ M^{-1}$). The weakest binding affinity was to the undamaged sequence, with $K = 2.7 \pm 0.4 \times 10^5\ M^{-1}$.

Substrate	K (M^{-1})	A	R^2
$(GGGT)_4$	2.7 ± 0.4 x 10^5	0.93 ± 0.04	0.997
$(GGGT)_4$ 8oxo1	7.7 ± 0.7 x 10^5	0.93 ± 0.01	0.998
$(GGGT)_4$ 8oxo3	3.2 ± 0.1 x 10^5	1.0 ± 0.00	0.997
$(GGGT)_4$ 8oxo5	4.4 ± 0.6 x 10^5	0.91 ± 0.03	0.997
$(GGGTT)_4$	6.0 ± 0.9 x 10^5	1.0 ± 0.02	0.997
$(GGGTT)_4$ 8oxo1	8.2 ± 1.0 x 10^5	1.0 ± 0.01	0.997
$(GGGTT)_4$ 8oxo3	9.6 ± 1.0 x 10^5	1.0 ± 0.01	0.996
$(GGGTT)_4$ 8oxo5	9.3 ± 0.8 x 10^5	1.0 ± 0.01	0.997
$(TTAGGG)_4$	1.6 ± 0.7 x 10^6	0.88 ± 0.05	0.996
$(TTAGGG)_4$ 8oxo1	1.3 ± 0.4 x 10^6	0.90 ± 0.04	0.997
$(TTAGGG)_4$ 8oxo3	1.2 ± 0.1 x 10^6	0. ± 0.0	0.996
$(TTAGGG)_4$ 8oxo5	1.4 ± 0.5 x 10^6	0.91 ± 0.05	0.997

Table 4.3: Summary of fluorescence spectroscopy AKKQ binding study in 150mM KCl

Similarly, Figure 4.4 C shows the binding interactions between FANCJ AKKQ and $(GGGTT)_4$. There is a small shift in the binding isotherm from AKKQ binding to GGGTT and the respective 8oxoG4 sequences. However, there is no change in the interaction between FANCJ AKKQ and the different damage positions. The interaction of FANCJ AKKQ with the undamaged sequence had a K value of 6.0 ± 0.9 x 10^5 M^{-1}. Each respective 8oxoG4 sequence had similar K values of approximately 8-9 x 10^5 M^{-1}. Table 4.3 summarizes K values for each construct.

4.3.2 FANCJ binding to G4 DNA

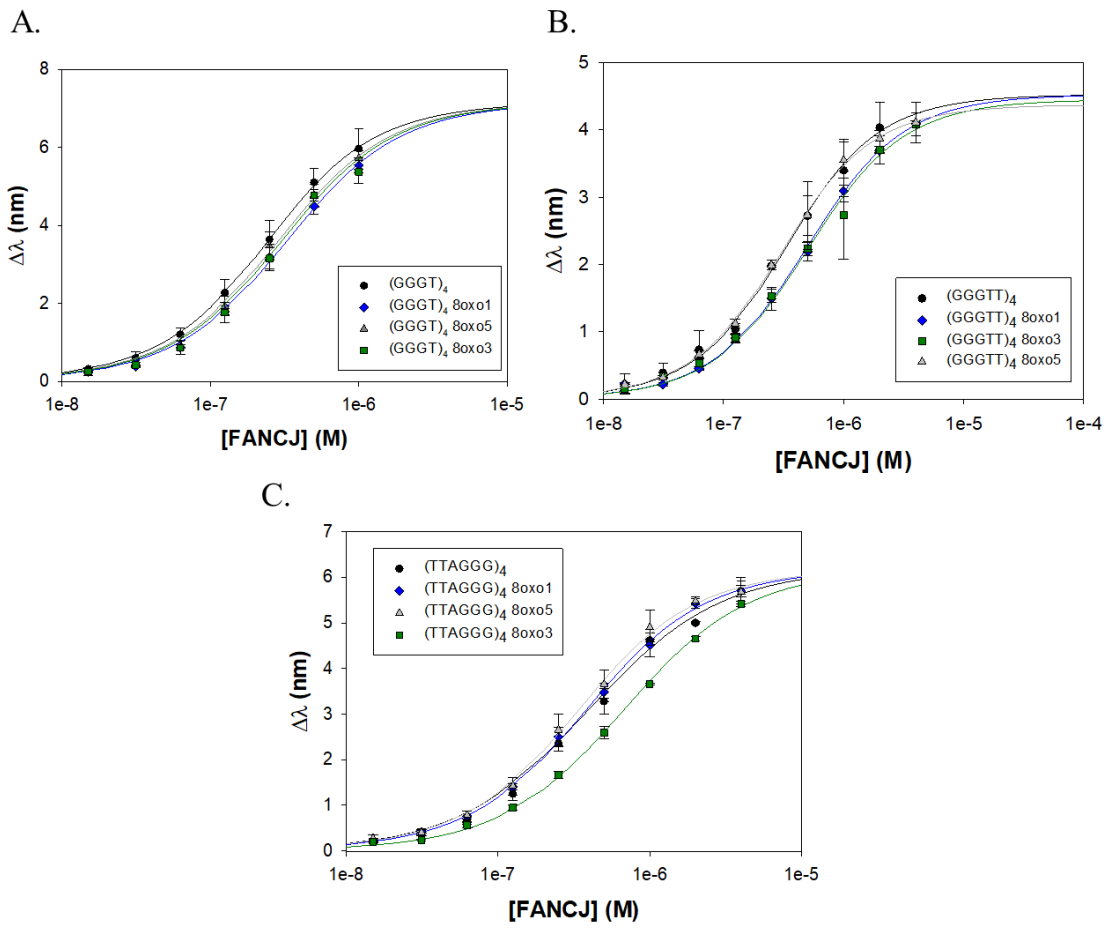

Figure 4.6: BLI binding isotherms with full-length FANCJ. A. FANCJ equilibrium binding to $(GGGT)_4$ and the 8oxoG4 constructs B. FANCJ binding to $(GGGTT)_4$ and the respective 8oxoG4s C. FANCJ binding to (TTAGGG)4 and respective 8oxoG4s.

To test whether full-length FANCJ can target 8oxoG4 structures, we used biolayer interferometry (BLI) to monitor the binding interactions between FANCJ and the human telomeric sequence $(TTAGGG)_4$. The plateau $\Delta\lambda$ values at the end of the binding phase were plotted against FANCJ concentrations to generate an isotherm. Nonlinear least squares analysis of this curve to a 1:1 interaction model resulted in an equilibrium constant of $K = 2.4 \pm 0.3 \times 10^6$. Further analysis revealed that FANCJ bound

to G4 with 8oxo1 and 8oxo5 substrates, yielding $K_{8oxo1} = 3.3 \pm 0.7 \times 10^6\ M^{-1}$ and $K_{8oxo5} = 3.8 \pm 0.8 \times 10^6\ M^{-1}$ (Figure 4.5 C). However, there was reduced binding for the 8oxo3 modified $(TTAGGG)_4$, with $K_{8oxo3} = 1.7 \pm 0.2 \times 10^6\ M^{-1}$.

Substrate	K (M^{-1})	A	R^2
$(GGGT)_4$	6.1 ± 1.3 x 10^6	6.6 ± 0.4	0.999
$(GGGT)_4$ 8oxo1	4.1 ± 0.9 x 10^6	7.2 ± 0.5	0.998
$(GGGT)_4$ 8oxo3	4.5 ± 0.5 x 10^6	6.5 ± 0.8	0.996
$(GGGT)_4$ 8oxo5	4.9 ± 0.8 x 10^6	5.9 ± 0.2	0.998
$(GGGTT)_4$	3.9 ± 0.5 x 10^6	4.3 ± 0.6	0.997
$(GGGTT)_4$ 8oxo1	2.4 ± 0.5 x 10^6	4.5 ± 0.1	0.999
$(GGGTT)_4$ 8oxo3	2.3 ± 0.2 x 10^6	4.5 ± 0.1	0.992
$(GGGTT)_4$ 8oxo5	4.4 ± 1.2 x 10^6	4.4 ± 0.3	0.997
$(TTAGGG)_4$	2.4 ± 0.3 x 10^6	5.1 ± 0.2	0.998
$(TTAGGG)_4$ 8oxo1	3.3 ± 0.7 x 10^6	5.5 ± 0.1	0.999
$(TTAGGG)_4$ 8oxo3	1.7 ± 0.2 x 10^6	4.5 ± 0.1	0.999
$(TTAGGG)_4$ 8oxo5	3.8 ± 0.8 x 10^6	6.2 ± 0.2	0.999

Table 4.4: Summary of BLI binding study with full-length FANCJ

Next, we tested FANCJ with the same sequences as AKKQ. As shown in Figure 4.6 A, the binding interactions between FANCJ and $(GGGT)_4$ were analyzed. There was a small shift in binding when FANCJ interacted with GGGT and the respective 8oxoG4 sequences. However, there was no significant change in interaction between the different 8oxoG4s. FANCJ exhibited the highest affinity for the native strand, with K = 6.0 ± 0.3 x $10^6\ M^{-1}$. In contrast, it showed lower affinities to the 8oxoG-damaged structures. Notably,

there was no difference in binding between 8oxo1 ($K = 4.8 \pm 0.5 \times 10^6\ M^{-1}$), 8oxo3 ($K = 4.5 \pm 1.2 \times 10^6\ M^{-1}$), or 8oxo5 ($K = 4.8 \pm 0.7 \times 10^6\ M^{-1}$).

Additionally, Figure 4.6 B illustrates the binding interactions between FANCJ and $(GGGTT)_4$. There was minimal difference in binding interactions between the native and the damaged sequences. The interaction of FANCJ with the undamaged sequence had a K value of $3.9 \pm 0.5 \times 10^6\ M^{-1}$. Each respective 8oxoG4 sequence had similar K values of approximately 2-4 $\times 10^6\ M^{-1}$. The binding affinities are listed in Table 4.4.

4.3.3 REV1 Core binding to G4 DNA

Next, we tested to see if the REV1 Core domain can bind to 8oxoG4 structures, we used biolayer interferometry (BLI) to monitor the binding interactions between REV1 core and the $(TTAGGG)_4$ sequence. The plateau $\Delta\lambda$ values at the end of the binding phase were plotted against REV1 Core concentrations to generate an isotherm. Nonlinear least squares analysis of this curve to a 1:1 interaction model resulted in an equilibrium constant of $K = 4.2 \pm 0.9 \times 10^6\ M^{-1}$. Further analysis revealed that REV1 bound to G4 with 8oxo3 and 8oxo5 substrates, yielding $K_{8oxo3} = 3.9 \pm 0.1 \times 10^6\ M^{-1}$ and $K_{8oxo5} = 3.5 \pm 0.4 \times 10^6\ M^{-1}$ (Figure 4.7 C). However, there was reduced binding for the 8oxo1 modified $(TTAGGG)_4$, with $K_{8oxo1} = 1.7 \pm 0.4 \times 10^6\ M^{-1}$.

Next, we tested REV1 Core with the same sequences as FANCJ. As shown in Figure 4.7 A, the binding interactions between REV1 and $(GGGT)_4$ were analyzed. There was no shift in binding when REV1 Core interacted with GGGT and the respective 8oxoG4 sequences. REV1 exhibited the highest affinity for the construct overall, with $K = 3.8 \pm 0.4 \times 10^7\ M^{-1}$. Notably, there was no difference in binding between 8oxo1 ($K = 2.4 \pm 0.5 \times 10^7\ M^{-1}$), 8oxo3 ($K = 1.8 \pm 0.1 \times 10^7\ M^{-1}$), or 8oxo5 ($K = 2.4 \pm 0.3 \times 10^7\ M^{-1}$).

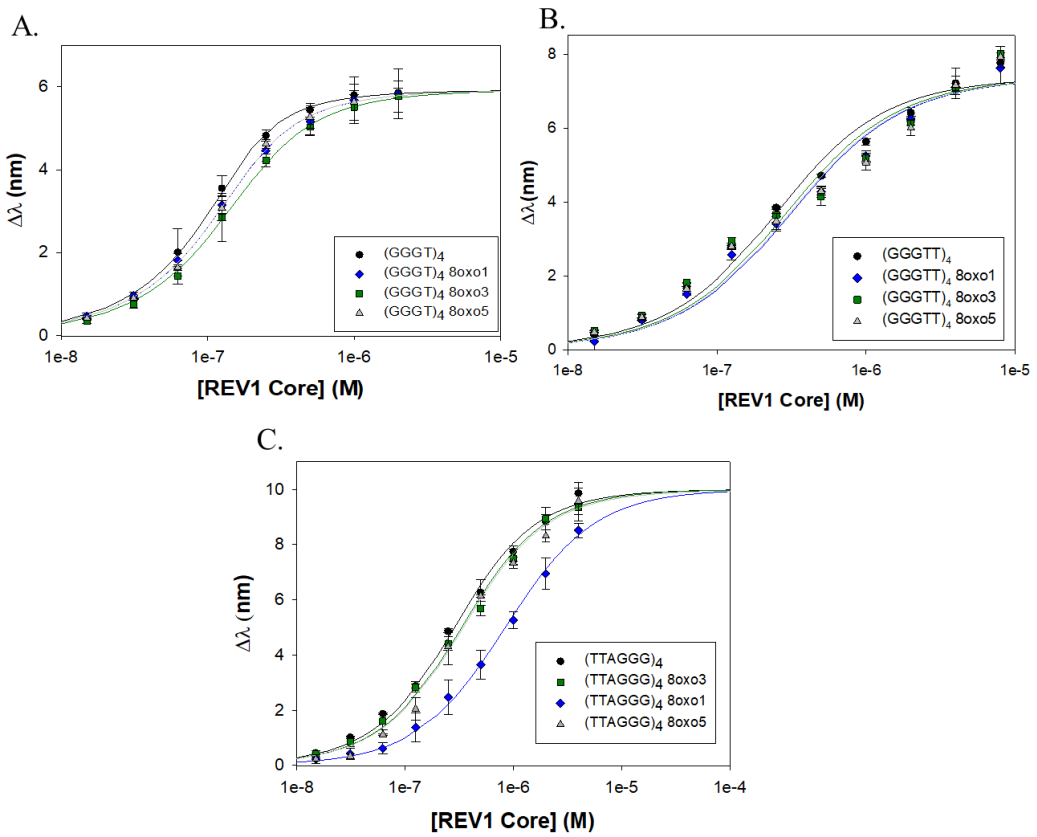

Figure 4.7: BLI binding isotherms with REV1 Core (aa 330-833). A. REV1 equalibrium binding to $(GGGT)_4$ and the 8oxoG4 constructs B. REV1 binding to $(GGGTT)_4$ and the respective 8oxoG4s C. REV1 binding to $(TTAGGG)_4$ and respective 8oxoG4s.

Additionally, Figure 4.7 B illustrates the binding interactions between REV1 Core and $(GGGTT)_4$. There was minimal difference in binding interactions between the native and the damaged sequences. The interaction of REV1 with the undamaged sequence had a K value of 5.7 ± 0. x 10^6 M^{-1}. Further analysis revealed that REV1 bound to G4 with 8oxo1 and 8oxo5 substrates, yielding $K_{8oxo1} = 4.3 \pm 1.0 \times 10^6$ M^{-1} and $K_{8oxo5} = 4.3 \pm 0.2 \times 10^6$ M^{-1}. However, there was reduced binding for the 8oxo1 modified $(TTAGGG)_4$, with $K_{8oxo1} = 7.0 \pm 0.4 \times 10^6$ M^{-1}.

Substrate	K (M^{-1})	A	R^2
$(GGGT)_4$	3.8 ± 0.4 x 10^7	5.4 ± 0.1	0.998
$(GGGT)_4$ 8oxo1	2.4 ± 0.5 x 10^7	5.9 ± 0.6	0.999
$(GGGT)_4$ 8oxo3	1.8 ± 0.1 x 10^7	5.0 ± 0.3	0.998
$(GGGT)_4$ 8oxo5	2.4 ± 0.3 x 10^7	4.3 ± 0.6	0.999
$(GGGTT)_4$	5.7 ± 0.2 x 10^6	7.3 ± 0.2	0.994
$(GGGTT)_4$ 8oxo1	4.3 ± 1.0 x 10^6	7.4 ± 0.3	0.993
$(GGGTT)_4$ 8oxo3	7.0 ± 0.4 x 10^6	6.1 ± 0.2	0.986
$(GGGTT)_4$ 8oxo5	4.3 ± 0.2 x 10^6	6.9 ± 0.3	0.987
$(TTAGGG)_4$	4.2 ± 0.9 x 10^6	7.9 ± 0.3	0.998
$(TTAGGG)_4$ 8oxo1	1.7 ± 0.4 x 10^6	10.1 ± 1.7	0.997
$(TTAGGG)_4$ 8oxo3	3.9 ± 0.1 x 10^6	7.0 ± 0.4	0.996
$(TTAGGG)_4$ 8oxo5	3.5 ± 0.4 x 10^6	10.0 ± 0.4	0.996

Table 4.5: Summary of BLI binding study with REV1 core

4.3.4 REV1 Core E466A/Y470A binding to G4s

To test how the E466A/Y470A (EA/YA) double mutant within the hydrophobic pocket of REV1 effects G4 binding, we measure the K values for the native G4 DNA substrates. EA/YA had the strongest affinity for GGGT (K = 7.2 ± 2.7 x 10^7 M^{-1}) and to GGGTT (K = 4.9 ± 2.1 x 10^7 M^{-1}). EA/YA bound the weakest to TTAGGG construct with K = 4.4 ± 2.4 x 10^6 M^{-1}. The overall binding of the mutants was higher than wild-type REV1 core. Table 4.5 Summary of BLI binding study with REV1 core

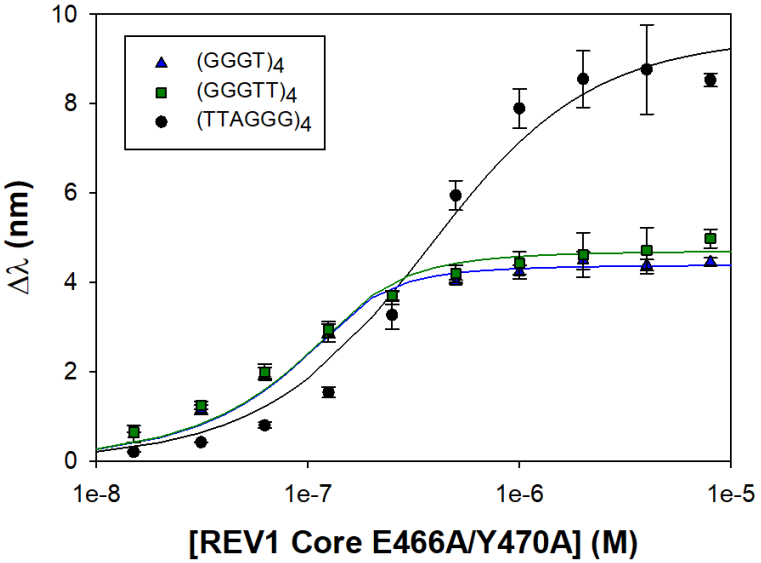

Figure 4.8: BLI isotherms of REV1 Core E466A/Y470A to the native sequences

Substrate	K (M^{-1})	A	R^2
bio(GGGT)$_4$	7.2 ± 2.7 x 10^7	4.4 ± 0.1	0.996
bio(GGGTT)$_4$	4.9 ± 2.1 x 10^7	4.7 ± 0.1	0.993
bio(TTAGGG)$_4$	4.4 ± 2.4 x 10^6	9.6 ± 1.2	0.996

Table 4.6: Summary of BLI binding equilibriums with REV1 Core EA/YA

4.3.5 FANCJ PIP recruits REV1 and PCNA

It was suggested that FANCJ possessed a PCNA-interaction peptide (PIP) within amino acids 1001—1017 (SWSSFNSLGQYFTGKIP), based on the underlined consensus residues. PIP-containing proteins bind directly to PCNA to coordinate DNA replication and repair. However, it is now evident that non-canonical PIP-boxes can serve as hotspots to recruit other proteins, including the REV1 translesion synthesis polymerase. To test if the PIP-like motif in FANCJ functions as a PIP or an REV1-

interacting region (RIR), we compared how this peptide bound to PCNA versus REV1 by BLI. Biotinylated FANCJ PIP was immobilized on a streptavidin-coated biosensor and experiments were performed as described above. As seen in the truncated time-courses, REV1 CTD produced a larger binding values compared to PCNA (Figure 4.9). Our BLI results indicated that more REV1 CTD molecules bound to the FANCJ PIP than PCNA. This difference is further magnified as REV1 CTD has a molecular mass of ~12 kDa, while a human PCNA trimer is about seven times its size, at ~86 kDa [145, 146].

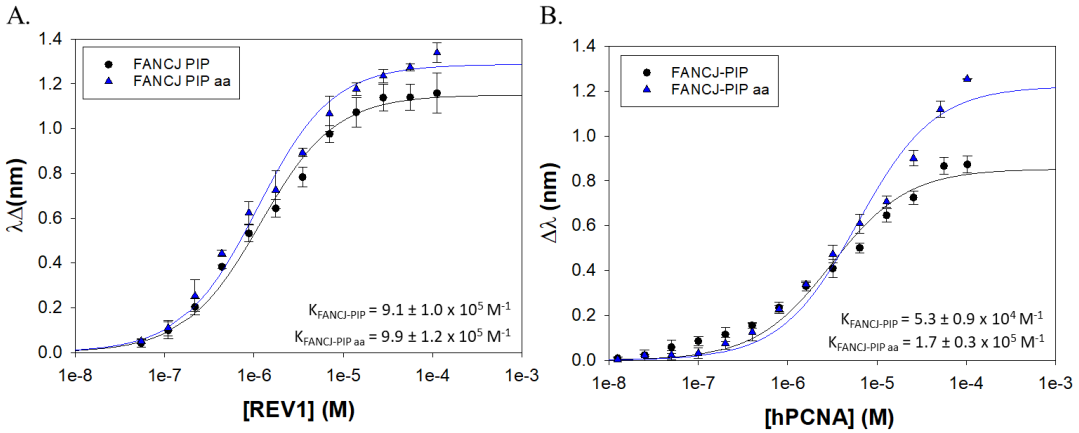

Figure 4.9: FANCJ PIP and PIP AA binding isotherms A. REV1 CTD binding to FANCJ PIP and PIP AA B. PCNA binding to FANCJ PIP and PIP AA.

Next, we tested the effects of a FANCJ PIP AA peptide on its interactions with REV1 CTD and PCNA. The consecutive consensus aromatic residues (YF) within the potential PIP-motif were replaced with alanine amino acids (AA). The isotherms showed that REV1 CTD bound to FANCJ PIP AA more versus PCNA (Figure 4.9). BLI experiments with PIP and PIP AA were repeated as a function of REV1 CTD and PCNA concentration to measure their corresponding equilibrium constant values. The PIP AA mutation resulted in a small increase in affinity for REV1 CTD but the binding of PCNA was weakened twofold.

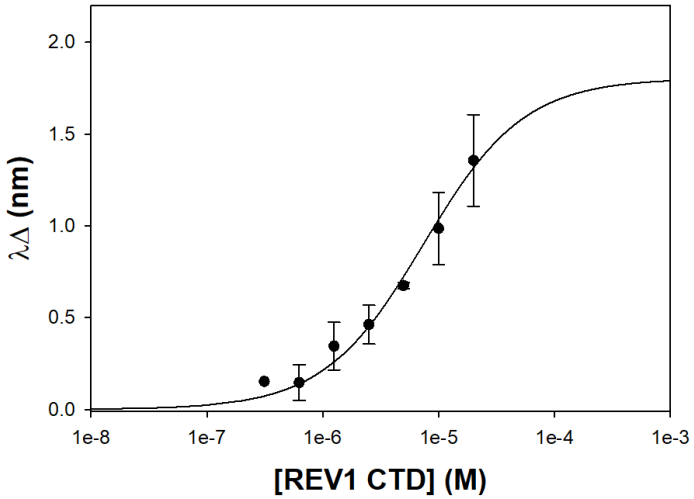

Figure 4.10: Binding isotherm of REV1 CTD binding to full-length FANCJ

Finally, we tested if the REV1 CTD could bind to the full-length FANCJ. To study the interaction, GST-FANCJ-Flag was bound to anti-GST BLI sensor tips. REV1 CTD bound to FANCJ with K = 1.9 ± 0.9 x 10^5 M^{-1}. Our results suggest that the PIP sequence in FANCJ has the biochemical properties of a canonical PIP-box that binds to PCNA, but it functions better as an RIR-motif that interacts with the REV1 polymerase.

4.4 Discussion

4.4.1 FANCJ AKKQ peptide binds to G4s

The fluorescent spectroscopy results indicate that FANCJ AKKQ peptide alone can target G4s and 8oxoG4s, although 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 obtained under the same conditions.

For the human telomeric G4 $(TTAGGG)_4$, which formed the less rigid hybrid conformation, the AKKQ peptide bound with similar affinity for the native and the 8oxoG4 substrates. Neither the damage to the G4 nor change in the thermal stability, because of the oxidative damage, does not seem to affect AKKQ binding. In this case, it is reasoned that FANCJ AKKQ bound specifically to the TTA loop of the $(TTAGGG)_4$ sequence, which would explain why the peptide was insensitive to the presence of 8oxoG in the rest of the quadruplex.

For the $(GGGTT)_4$ constructs, which formed predominately parallel conformations, AKKQ bound more strongly 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. 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 the $(GGGT)_4$ substrates, which formed compact parallel conformations, AKKQ also bound stronger to the 8oxoG damaged sequences than to the native sequence. Unlike $(GGGTT)_4$, there is an increase in binding affinity between AKKQ and 8oxoG4 damage positions can be correlated 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, this suggests that AKKQ is targeting the 8-oxoguanine position.

Taken together, this suggests that FANCJ AKKQ targets the intervening sequences. By decreasing the loop length that connected the guanine tetrads, the physical contact available to FANCJ were scarce, this is indicated by the binding interaction increased as the 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 4.11. 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.

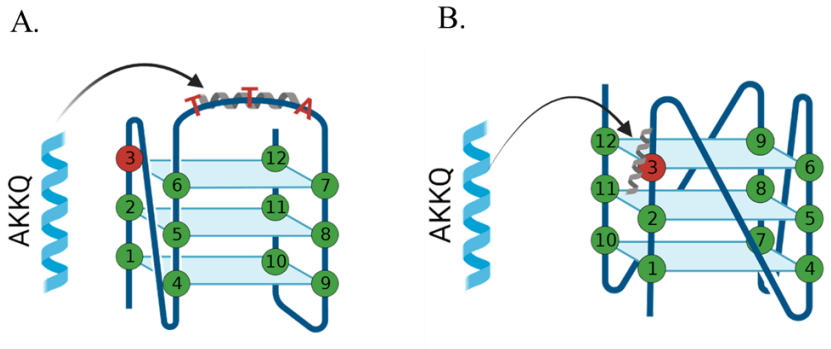

Figure 4.11: Proposed binding model between FANCJ AKKQ peptide and G4 sequences. A. G4 sequences that have an intervening sequence, such as the TTA of G4telo, AKKQ will preferentially target the intervening sequence and not bind to the 8oxoG. B. For G4 sequences that tightly folded and stable structures with short loops, such as GGGT, AKKQ will bind the 8oxoG.

4.4.2 FANCJ binds to 8oxoG4s

The aim of this study was to see if the full-length FANCJ binds to G4 structures similarly to the AKKQ peptide. The BLI results indicate that full-length FANCJ can

indeed bind to 8oxoG4s. Since the BLI study was performed in the presence of KCl, binding affinities can only be compared to CD data obtained under the same conditions.

For the human telomeric G4 $(TTAGGG)_4$, which formed a less rigid hybrid conformation, FANCJ bound with similar affinity to the native, 8oxo1, and 8oxo5 substrates. FANCJ exhibited the weakest binding to 8oxo3 compared to the native and other damaged sequences. Interestingly, 8oxo3 changed conformation from hybrid to antiparallel, potentially causing steric hindrance for protein binding due to the 8oxo3 damage located at the top of the stack.

For the $(GGGTT)_4$ constructs, which predominately formed parallel conformations, FANCJ bound more tightly to the native and 8oxo5 damaged sequence, and slightly weaker to the 8oxo1 and 8oxo3 sequences. The 8oxo3 construct also formed an antiparallel structure. Although this difference is not as great with the telomeric G4, FANCJ still bound weaker than the other damages. In this case, it is reasoned that FANCJ bound specifically to the TT loop of the $(TTAGGG)_4$ sequence, which would explain why the peptide was insensitive to the presence of 8oxoG in the rest of the quadruplex. Finally, for the $(GGGT)_4$ substrates, which formed compact parallel conformations, FANCJ bound the tightest over all the other G4 constructs to the native substrate. FANCJ had slightly weaker but equal affinities for the damaged G4s.

In summary, the AKKQ peptide alone binds specifically to the intervening loop of the G4 structure, with its binding influenced by loop length and modifications such as 8oxoG. In contrast, the full-length protein binds to the G4 structure in a manner that is less dependent on loop length but highly sensitive to the conformation and the position of the 8oxoG within the tetrad. This difference may be attributed to the full-length protein

having multiple binding domains and the potential conformational constraints of the AKKQ peptide within the full-length protein. These findings highlight the complexity of protein-DNA interactions and the importance of studying both isolated peptides and full-length proteins to fully understand their binding mechanisms.

4.4.3 REV1 Core binds to 8oxoG4s

This study aimed to determine if REV1 core could also bind and recognize 8oxoG-damaged G4 structures. Previous work by Ketkar et al. demonstrated that the REV1 core exhibits preferential binding to parallel structures, followed by hybrid structures, with the least affinity for antiparallel structures [120]. Therefore, we hypothesized that damage causing a change in conformation to antiparallel would result in decreased binding.

For the $(TTAGGG)_4$ constructs, REV1 bound with equal affinity to the native, 8oxo3, and 8oxo5 substrates. Interestingly, REV1 did not exhibit the weakest binding to the 8oxo3 G4, which had an antiparallel conformation, but rather to the 8oxo1 G4 DNA, which was a hybrid structure. For the $(GGGTT)_4$ constructs, REV1 exhibited stronger binding to the 8oxo3 G4, and slightly weaker to the native, 8oxo1, and 8oxo5 sequences. The 8oxo3 construct also formed an antiparallel structure, which is interesting because the previous study found that REV1 bound the weakest to antiparallel structures. The addition of the 8oxoG seems to influence the binding regardless of the G4 structure. Lastly, the $(GGGT)_4$ DNA, REV1 bound tighter to these substrates than to $(TTAGGG)_4$ or $(GGGT)_4$. REV1 exhibited the strongest binding to the native strand and slightly weaker to the damaged G4s. The difference in the binding preference of REV1 may be due to the specific sequence of the G4s and their conformation, rather than conformation

alone. Both GGGT and GGGTT constructs are parallel structures, yet REV1 bound most tightly to the GGGT construct. Additionally, within the GGGTT construct, REV1 bound the tightest to the 8oxo1 antiparallel G4. This suggests that conformation alone does not dictate binding preference.

For the REV1-Core E466A/Y470A mutant, we only tested the native sequences. Previous work by Ketkar et al. in 2021 found that E466A mutation decreased binding to G4s by ~6 fold, while the Y470A mutation decreased affinity for G4 DNA about ~4 fold [120]. Combing the two mutations was expected to decrease the binding affinity or result in no binding at all. A later study by Ketkar et al. in 2023, found that the double mutant diminished the preferential binding to G4 substrates. The study by Ketkar et al in 2023 did not use the GGGT or GGGTT sequence, but they did use a sequence similar to our TTAGGG, but it was not exact and longer in sequence. Comparison of the wild-type REV1 core binding to TTAGGG, Ketkar had a binding of 410 nM, while for our study we have 240 nM, which is about a 2 fold difference. In our study, the EA/YA bound to TTAGGG with an affinity for 230 nM, while in the Ketkar study the mutant bound at 700 nM, which is about a 3 fold difference. The difference in binding between the two studies is most likely due to the difference in the DNA sequence used for the TTAGGG study.

4.4.4 FANCJ PIP binds to REV1 CTD

Classical PIP-boxes bind to PCNA, a strategy used to recruit repair polymerases to damaged DNA sites [142]. Because PCNA forms a trimeric ring, it can carry up to three different repair enzymes through PIP interactions, functioning as a molecular tool belt [147]. This enables the complex to switch polymerases to bypass damage-specific

lesions. Some PIP-like sequences can bind directly to other repair proteins without the need for a PCNA hub.

FANCJ possesses a potential PIP-motif based on its amino acid sequence (SWSSFNSLGQYFTGKIP) [91, 144]. Our BLI analysis showed that the FANCJ PIP bound to both REV1 CTD and PCNA. However, interactions with REV1 CTD showed stronger binding affinity than to PCNA, suggesting that the formation of a PIP-CTD complex was more favorable. Additionally, REV1 CTD bound to the full-length FANCJ within the same magnitude as REV1 binding as it did to the PIP peptide. Interestingly, the PIP AA mutant reduced the binding constant for PCNA by half but resulted in a small increase in affinity for REV1 CTD. However, corresponding double alanine substitutions within the PIP motifs of Polη and Polι have been shown to eliminate PCNA binding entirely [148]. Since FANCJ PIP AA still retained binding interactions with PCNA, it likely does not function as a canonical PIP-box. Similar mutations in RIRs have either reduced or prevented binding of REV1 [109, 141].

Taken together, the FANCJ PIP exhibits biochemical properties of both PIPs and RIRs, but the consecutive aromatic amino acids within the broadly defined PIP consensus sequence are not essential for protein binding. It is possible that protein-protein interactions are mediated by more than the two consecutive consensus aromatic residues (YF) and possibly interacts with an up-stream phenylalanine, similar to APIM motifs [149]. Having a PIP-box that can bind to both PCNA and REV1 may indicate that FANCJ forms a ternary complex with both proteins to process G4s.

CHAPTER FIVE

UNFOLDING OF G-QUADRPLEXES BY FANCJ AND REV1

5.0 Introduction

G-quadruplexes (G4s) pose significant challenges to essential cellular processes such as replication, transcription, and translation due to the greater forces required to unfold them compared to the capabilities of polymerases. Notably, the helicase FANCJ has been shown to unwind both intramolecular and intermolecular G-quadruplexes *in vitro*, highlighting its potential role in maintaining genomic stability by resolving these structures. Because FANCJ is a helicase, it requires ATP for the unfolding activity. It was previously found that a single-stranded DNA overhang is required before the G4 for FANCJ to dock and proceed to unfold [91]. Similarly, REV1 can mechanically disrupt and prevent the refolding of G4 structures *in vitro,* and can disrupt the G4 without the addition of CTP [106, 120]. While both FANCJ and REV1 are known to play roles in unfolding G4s, their efficiency in handling oxidative damaged G4 structures remains unclear.

An intriguing characteristic of G4s is their ability to selectively bind hemin (iron (III)-protopor-phyrin IX) through π-π stacking interactions, forming peroxidase-mimicking G4 DNAzymes (or RNAzymes) [150]. These DNAzymes/RNAzymes exhibit catalytic activity similar to natural peroxidases, which are critical in aerobic metabolism [151]. Compared to natural peroxidases, G4 DNAzymes/RNAzymes offer advantages such as small size, ease of synthesis, straightforward manipulation. These features make them a powerful catalytic toolkit in biosensing, biomaterials, and biomolecular devices

[152-154]. In the context of studying G4 unfolding, hemin-bound G4s serve as a useful model to understand how FANCJ and REV1 interact with and resolve these structures.

3,6-bis(1-Methyl-4-vinylpyridinium) carbazole diiodide (BMVC) is a G4-specific ligand and the first fluorescent probe to detect G4 structures in human cells . BMVC adjusts its pyridine rings to adopt a contracted conformation to perfectly match the three bases of a G-tetrad, resulting in optimized π-π stacking with the external G-tetrad.

This dissertation chapter investigates the unfolding of 8-oxoguanine damaged G-quadruplexes and examines the ability of FANCJ helicase and REV1 polymerase to interact with and unfold these structures. Both FANCJ and REV1 are known to play roles in unfolding G4s, yet their efficiency in handling oxidatively damaged G4 structures remains unclear. Specifically, this research seeks to determine if FANCJ and REV1 can still unfold 8-oxoG containing G-quadruplexes and whether there is a difference in the rate of unfolding between the proteins.

5.1 Materials

5.1.1 Hemin Assay Materials

The 2,2′-azinobis (3-ethylbenzothiazoline-6-sulfonic acid ammonium salt (ABTS) and hemin were purchased from MedChem Express. Hemin was dissolved in DMSO, diluted to 20 mM, and stored in a dark place at 4 °C. ABTS was dissolved in Buffer C (25 mM HEPES-NaOH pH 7.4, 100 mM KCl, 2 mM MgCl2, 0.05% Triton X-100, 1% DMSO), diluted to 25 mM, and stored at 4°C.

5.1.2 BMVC Assay Materials

3,6-Bis (1-methyl-4vinylpyridinium) carbazole diiodide (BMVC) was purchased from MedChem Express. BMVC was dissolved in DMSO, having a concentration of 10 μM, and stored in a dark place at 4 °C.

5.1.3 Oligonucleotides

Name	Sequence 5' – 3'
T15 GGGT Cy3	TTTTTTTTTTTTTTGGGTGGGTGGGTGGGT/3Cy3/
T15 GGGT 8oxo1 Cy3	TTTTTTTTTTTTTTT/i8oxodG/GGTGGGTGGGTGGGT/3Cy3/
T15 GGGT 8oxo3 Cy3	TTTTTTTTTTTTTTGG/i8oxodG/TGGGTGGGTGGGT/3Cy3/
T15 GGGT 8oxo5 Cy3	TTTTTTTTTTTTTTGGGTG/i8oxodG/GTGGGTGGGT/3Cy3/
T15 TeloG4 Cy3	TTTTTTTTTTTTTTTTAGGGTTAGGGTTAGGGTTAGGG/3Cy3/
T15 TeloG4 8oxo1 Cy3	TTTTTTTTTTTTTTTTA/i8oxodG/GGTTAGGGTTAGGGTTAGGG/3Cy3/
T15 TeloG4 8oxo3 Cy3	TTTTTTTTTTTTTTTTAGG/i8oxodG/TTAGGGTTAGGGTTAGGG/3Cy3/
T15 TeloG4 8oxo5 Cy3	TTTTTTTTTTTTTTTTAGGGTTAG/i8oxodG/GTTAGGGTTAGGG/3Cy3/
GGGT Trap	CCCACCCACCCA
T4Gelo Trap	CCCTAACCCTAA

Table 5.1: List of G-quadruplex sequences used for unfolding assays

The G4 DNA oligos were synthesized by IDT (Integrated DNA Technologies), and the specific sequences are listed in Table 5.1. G4 DNA was designed with a T15 ssDNA overhang for protein docking. All G4 DNA samples were dialyzed into Buffer C. After dialysis, the samples were heated to 95°C for 10 minutes and then allowed to cool

overnight prior to the unfolding experiments. DNA concentrations were assessed using NanoDrop One (Thermo Fisher Scientific), utilizing the 260 nm molar extinction coefficients provided by IDT.

5.2 Methods

5.2.1 G-quadruplex/Hemin colorimetric unfolding assay

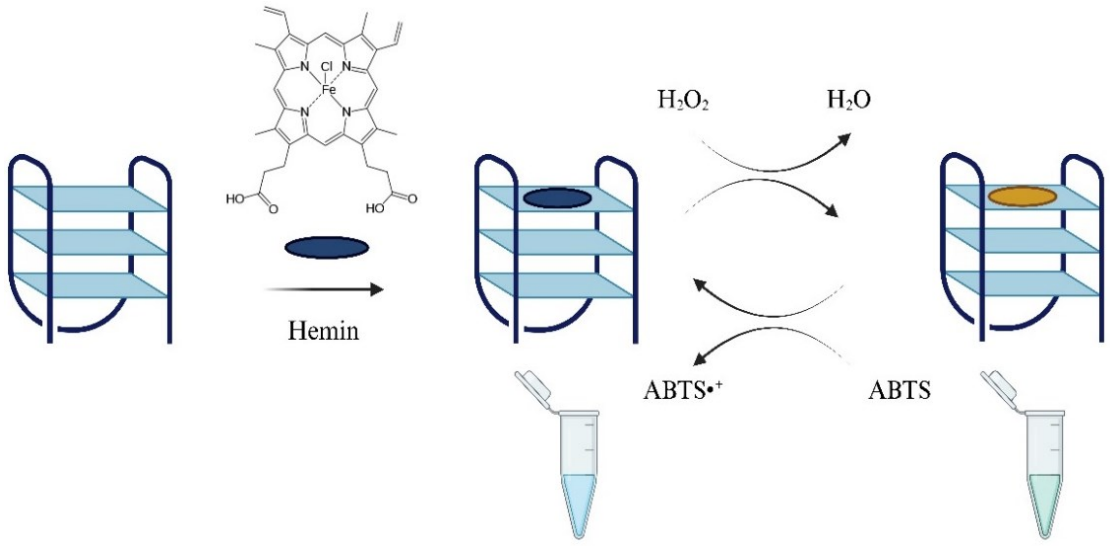

Figure 5.1: Graphic representation of the G-quadruplex-hemin DNAzyme system.

The unfolding reactions were carried out in a 96-well UV plate from Greiner Bio-One, with a final volume of 100 µL per well. Buffer C was used for these experiments. For the G4s, 250 nM of DNA and 250 nM of Hemin were added to each well. After incubation at 25 °C for 30 min, varying concentrations of REV1 Core or FANCJ was added and incubated for an additional 5 mins. Next, 10 mM ATP and 265 mM of trap DNA were added and incubated for 10 mins. Finally, 2 mM ABTS and 2 mM H_2O_2 were added to trigger the colorimetric reaction. The absorption spectra was collected from 400-

500 nm and were recorded by a BioTek Synergy H1 Microplate Reader. Each assay was repeated in triplicates (n=3).

5.2.2 BMVC/G-Quadruplex fluorescence unfolding assay

Unfolding reaction was completed in a 96-well clear bottom plates from ThermoFisher Scientific, with a final volume of 100 µL for each well. For the G4s, 250 nM of DNA and 0.1 µM of BMVC was added to each well. After an incubation at 25 °C for 5 min, REV1 Core or FANCJ was added at varying concentrations at incubated for 5 mins. Then, 10 mM ATP and 265 mM of trap DNA, and let incubate for 10 mins. The fluorescence signal was recorded 575 nm by a BioTek Synergy H1 Microplate Reader. Each assay was repeated in triplicates (n=3).

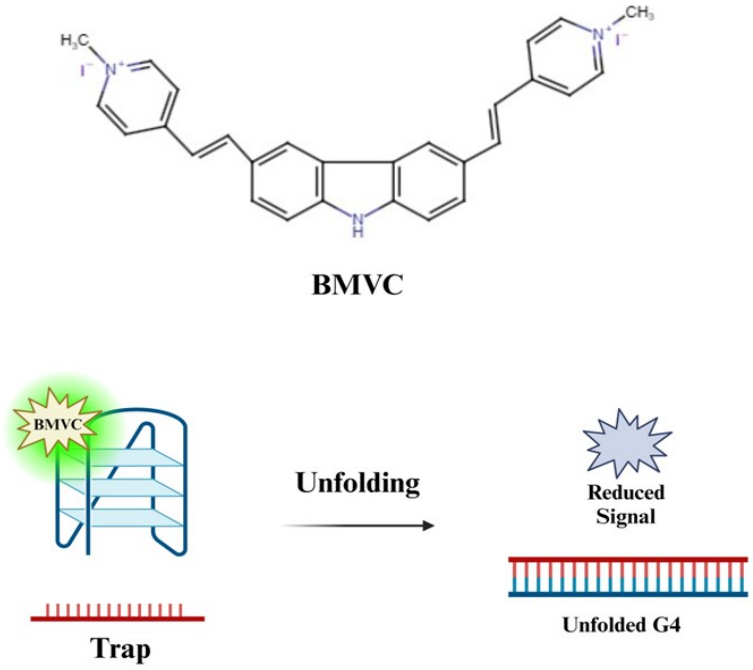

Figure 5.2: Graphic representation of the G-quadruplex-BMVC unfolding system

5.3 Results

5.3.1 FANCJ Unfolding

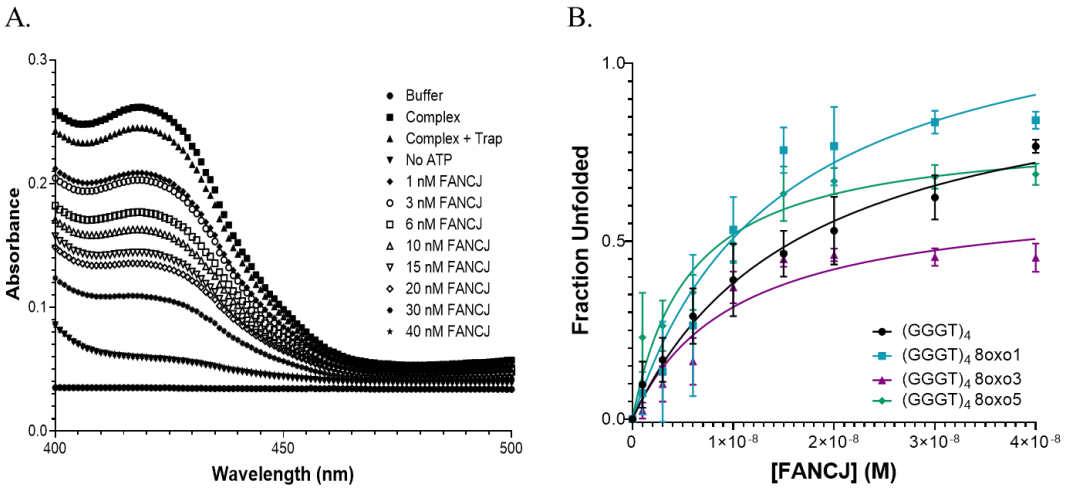

Figure 5.3: FANCJ unfolding study of the GGGT constructs A. FANCJ was added in increasing concentrations before triggering the colorimetric reaction and measured from 400 nm to 500 nm. B. Maximum absorbance at 420 nm for each construct was used to compare the unfolding activity.

Using the G-quadruplex-hemin DNAzyme system, we examined the effect of FANCJ unfolding on the G-quadruplexes formed by T15$(GGGT)_4$. When FANCJ was added to the reaction mixture, the amount of ABTS·+ decreased significantly as the concentration of FANCJ increased (Figure 5.3 A). This indicates that FANCJ bound to the G-quadruplex-hemin complex and altered its conformation and activity. The data for FANCJ unfolding of GGGT and the damaged constructs reveals significant variations in the rates of unfolding across different substrates (Figure 5.3 B). Specifically, FANCJ exhibits the highest rate of unfolding for the 8oxo3 substrate, as evidenced by the lowest K_m = 1.0 x 10^{-8} M. In contrast, the unfolding rate for 8oxo5 is the lowest (K_m = 5.6 x 10^{-8} M), albeit still within a similar order of magnitude, indicating a relatively less favorable but still significant interaction.

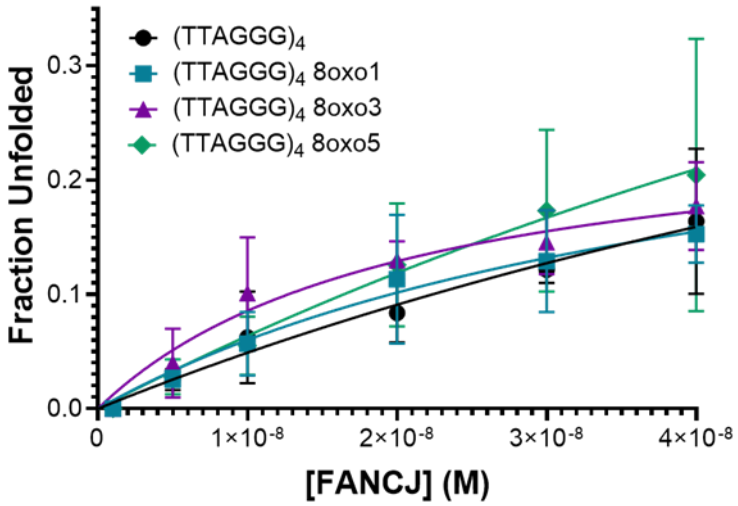

Figure 5.4: FANCJ unfolding study of the TTAGGG constructs

Using the G-quadruplex-BMVC system, we examined the effect of FANCJ unfolding on the G-quadruplexes formed by T15(TTAGGG)$_4$. When FANCJ was added to the reaction mixture, the fluorescence signal from the BMVC decreased as the concentration of FANCJ increased. This indicates that FANCJ bound to the G-quadruplex-BMVC complex and altered its conformation and activity. The data for FANCJ unfolding of T15-(TTAGGG)$_4$-Cy3 and the 8oxoG damaged constructs had relatively similar rates of unfolding across different substrates (Figure 5.4). With the exception of 8oxo5, FANCJ exhibits a lower unfolding rate of K_m =1.3 x 10^{-7} M.

Furthermore, for the unfolding studies, it was noticed that without the presence of ATP, FANCJ was still able to unfold G4s with FANCJ at max concentration. So, to test if it is concentration dependent, we replicated GGGT/hemin unfolding assay without ATP, and repeated the experiment in triplicates. FANCJ was still able to unfold the structure with a K_m = 3.1 x 10^{-9} M, which is ~5 fold faster than with ATP.

Substrate	K_m (M)	R^2
$(GGGT)_4$	$1.7 \pm 0.3 \times 10^{-8}$	0.99
$(GGGT)_4$ 8oxo1	$1.4 \pm 0.5 \times 10^{-8}$	0.96
$(GGGT)_4$ 8oxo3	$1.0 \pm 0.4 \times 10^{-8}$	0.93
$(GGGT)_4$ 8oxo5	$5.6 \pm 1.4 \times 10^{-8}$	0.96
$(TTAGGG)_4$	$1.2 \pm 0.9 \times 10^{-8}$	0.98
$(TTAGGG)_4$ 8oxo1	$4.4 \pm 1.6 \times 10^{-8}$	0.98
$(TTAGGG)_4$ 8oxo3	$2.1 \pm 0.7 \times 10^{-8}$	0.97
$(TTAGGG)_4$ 8oxo5	$1.3 \pm 0.7 \times 10^{-7}$	0.99

Table 5.2: Unfolding rates for FANCJ

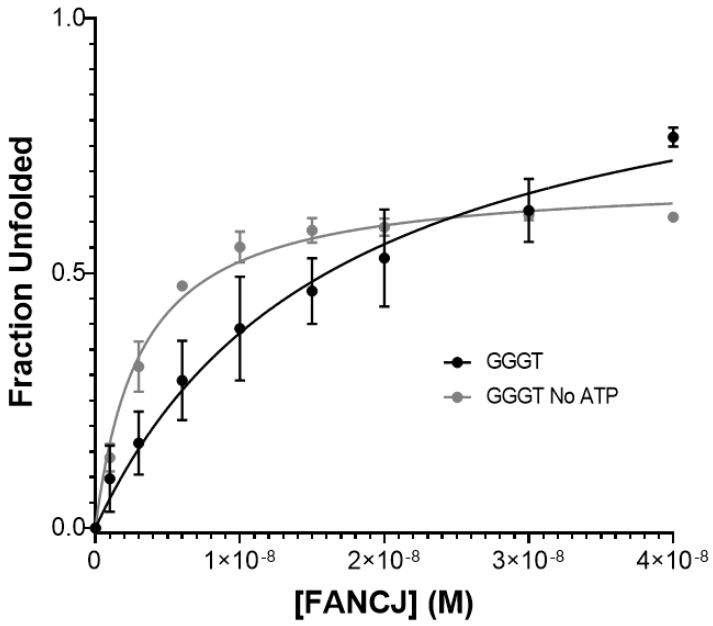

Figure 5.5: G4 DNA unfolding study with and without 10 mM ATP for FANCJ. With ATP $K_m = 1.7 \pm 0.3 \times 10^{-8}$ M, while no ATP $K = 3.1 \pm 0.4 \times 10^{-9}$ M

5.3.2 REV1 Unfolding

The unfolding dynamics of the T15-$(GGGT)_4$-Cy3 construct by REV1 Core were investigated using a hemin/G-quadruplex colorimetric assay. The results revealed significant differences in the unfolding rates among the various GGGT substrates. REV1 core had the lowest unfolding rate for the GGGT native sequence with a K_m value of 1.4 x 10^{-5} M. While the damage sequences had unfolding rates of 2-4 x 10^{-7} M. This suggest that the presence of the 8oxoG did not impede unfolding but rather aided in unfolding.

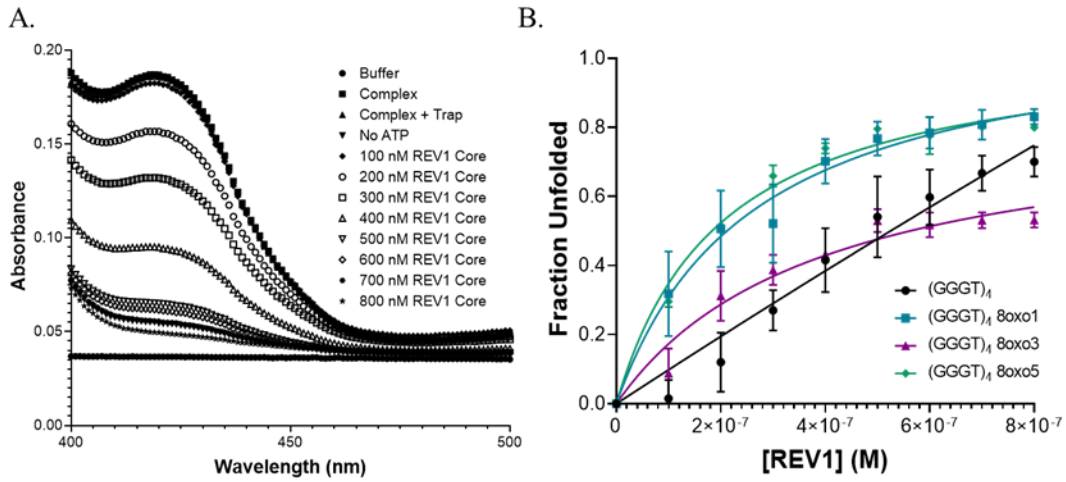

Figure 5.6: REV1 Core unfolding study of the GGGT constructs A. REV1 Core was added in increasing concentrations before triggering the colorimetric reaction and measured from 400 nm to 500 nm. B. Maximum absorbance at 420 nm for each construct was used to compare the unfolding activity.

The unfolding dynamics of the T15-$(TTAGGG)_4$-Cy3 construct by REV1 were analyzed using the BMVC/G-quadruplex fluorescence assay. The fluorescence signal, measured at 575 nm, provided insights into the interaction between REV1 and the G-quadruplex structures. The results demonstrated that REV1 effectively unfolded the TTAGGG G-quadruplex constructs, with varying efficiencies depending on the presence of oxidized guanine lesions. The fluorescence data showed a significant decrease in the

BMVC signal as the concentration of REV1 increased, indicating successful binding and unfolding of the G-quadruplex by REV1.

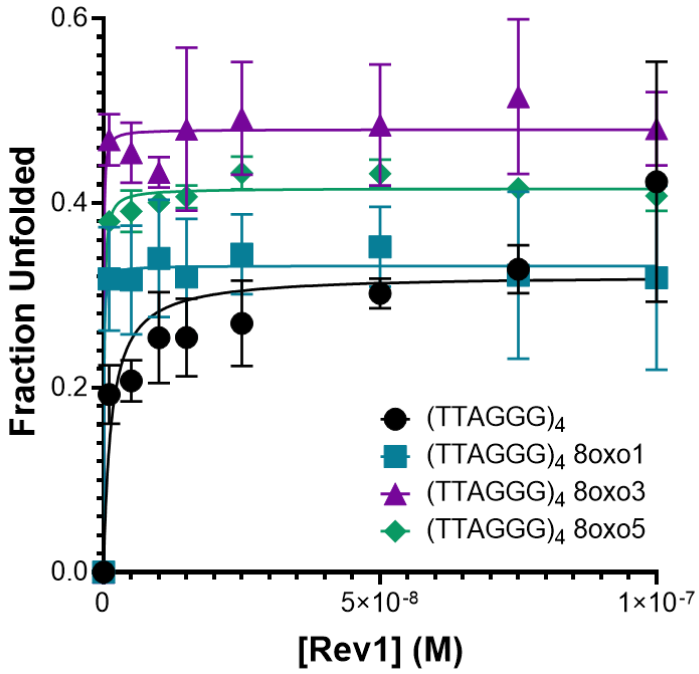

Figure 5.7: REV1 Core unfolding study of TTAGGG constructs. Fluorescence at 580 nm for each construct was used to compare the unfolding activity.

The results of the BMVC unfolding study demonstrated that REV1 effectively unfolded the T15 $(TTAGGG)_4$ G-quadruplex substrates with relatively quick unfolding rates overall. However, there were notable differences in unfolding among the different damaged constructs. The native sequence had the lowest unfolding rate with a Km value of 1.4×10^{-9}, while the oxidized constructs were quicker to unfold. Amongst the damaged constructs, 8oxo5 had the lowest unfolding rate with a Km value of 1.1×10^{-10}, suggesting a less favorable interaction compared to the other oxidized forms. In contrast, the unfolding rates for 8oxo1 and 8oxo3 were the fastest, with Km 4-5 $\times 10^{-11}$, indicating that REV1 could effectively recognize and process these oxidized lesions.

Substrate	K_m (M)	R^2
$(GGGT)_4$	$1.4 \pm 0.6 \times 10^{-5}$	0.97
$(GGGT)_4$ 8oxo1	$2.6 \pm 0.5 \times 10^{-7}$	0.99
$(GGGT)_4$ 8oxo3	$3.9 \pm 1.4 \times 10^{-7}$	0.96
$(GGGT)_4$ 8oxo5	$2.0 \pm 0.1 \times 10^{-7}$	0.99
$(TTAGGG)_4$	$1.4 \pm 0.9 \times 10^{-9}$	0.89
$(TTAGGG)_4$ 8oxo1	$5.0 \pm 0.5 \times 10^{-11}$	0.99
$(TTAGGG)_4$ 8oxo3	$4.1 \pm 0.6 \times 10^{-11}$	0.97
$(TTAGGG)_4$ 8oxo5	$1.1 \pm 0.4 \times 10^{-10}$	0.99

Table 5.3 Unfolding rates for REV1 Core

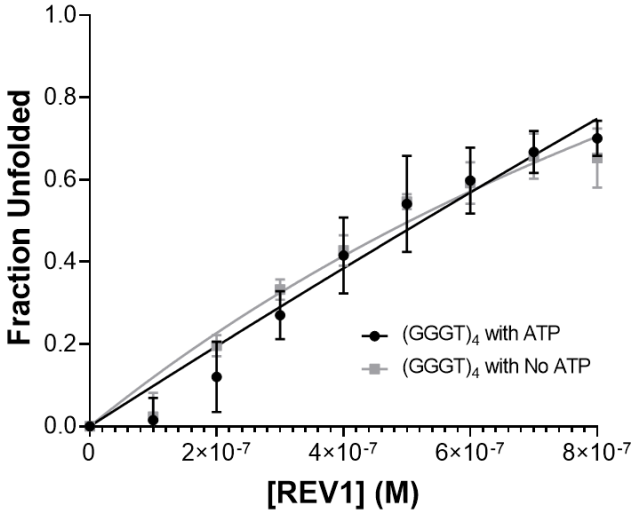

Figure 5.8: G4 DNA unfolding study with and without 10 mM ATP for REV1. With ATP $K_m = 1.4 \pm 0.3 \times 10^{-5}$ M, while no ATP $K = 1.9 \pm 0.4 \times 10^{-6}$ M

It was previously found that the REV1 Core did not require CTP to disrupt G4s. To test the impact of ATP, we repeated the T15-$(GGGT)_4$ unfolding with and without ATP. In the presence of ATP, the dissociation constant for GGGT was $K_m = 1.4 \times 10^{-5}$ M,

while in the absence of ATP K = 1.88 x 10^{-6} M. REV1 Core was able to disrupt the G4 without ATP present.

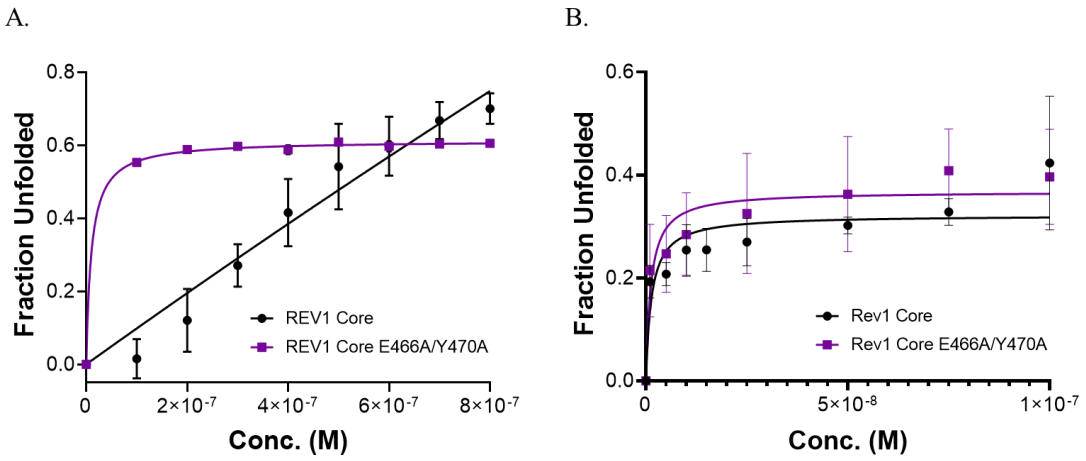

Figure 5.9: REV1Core E466A/Y470A unfolding study A. Unfolding of T15-(GGGT) by REV1 Core (K_m = 1.4 x 10^{-5} M) and the EA/YA mutant (K_m = 1.3 ± 0.2 x 10^{-8} M). B. Unfolding of T15-$(TTAGGG)_4$ by REV1 Core (K_m = 1.4 x 10^{-9} M) and the EA/YA mutant (K_m = 1.2 ± 0.5 x 10^{-9} M).

Next, we investigated if the REV1 Core E466A/Y470A retained unfolding capabilities (Figure 5.8). To compare to the binding study, only the native sequence was used. The EA/YA mutants unfolded the T15-(GGGT)-Cy3 construct with a rate of K_m = 10.3 x 10^{-9} M, which is a much faster unfolding rate compared to the wild-type. For the T15-$(TTAGGG)_4$-Cy3, the EA/YA mutant unfolded the G4 with the same rate as wild-type REV1 Core (K_m = ~ 1 x 10^{-9} M).

5.4 Discussion

This study investigated the unfolding dynamics of G-quadruplexes (G4s) using the FANCJ helicase and REV1 polymerase, with a particular focus on their interactions with 8-oxoguanine (8oxoG) damaged G4s. The results reveal significant substrate-

specific preferences in the unfolding rates of both proteins, underscoring their specialized roles in DNA repair and damage tolerance mechanisms.

For FANCJ, the lowest rate of unfolding was observed with the 8oxo5 substrate (K = 5.62E-09 M), suggesting a weaker interaction with this substrate. Comparing the binding data of FANCJ (discussed in Chapter 4) to the 8oxoG GGGT substrates and the unfolding data, it was expected that because FANCJ exhibited the weakest binding to all 8oxoG constructs, there would be an overall lower unfolding rate compared to the native sequence. However, FANCJ unfolded the native, 8oxo1, and 8oxo3 G4s at equal rates, indicating that the presence of 8oxoG itself does not affect unfolding process; rather, its position within the stack is critical.

In the unfolding of T15-$(TTAGGG)_4$-Cy3 by FANCJ, the lowest unfolding rates were observed for the native and the 8oxo5, both with K_m values of 1 x 10^{-7} M. The binding data for these constructs show minimal difference in binding between the native, 8oxo1, and 8oxo5. Similar to the GGGT constructs, the binding does not necessarily impact the unfolding rate. Interestingly, the unfolding rate of the 8oxo5 substrate for both TTAGGG and GGGT was lower than the other constructs. For the TTAGGG constructs, 8oxo5 was the most destabilizing, whereas for the GGGT constructs, 8oxo5 was the second most destabilizing. In this case of FANCJ unfolding, the heavily destabilizing 8oxo5 had the lowest unfolding rates.

For the REV1 core, the lowest unfolding rate was observed with the native GGGT substrate (K_m = 1.4 x 10^{-5} M), while the unfolding rates for the 8oxoG substrates remained consistent (K_m = 2-4 x 10^{-7} M). The GGGT G4 is the most stable structure of the G4s used in this study. The slower unfolding rate may be due to the high stability of

the structure. Comparatively, for the binding data (Chapter 4), it was observed that the REV1 core had an order of magnitude higher binding for the GGGT/8oxoG constructs compared to the other G4 DNA. Notably, REV1 had the highest affinity for the native GGGT substrate. The increase in binding may not be solely due to its preference for parallel structures, but for G4s that are the most stable and difficult to unfold.

For the TTAGGG constructs, REV1 core unfolded these constructs faster than the GGGT overall. The lowest unfolding rate was for the native sequence (K_m = 1.4 x 10^{-9}), while the 8oxo1 and 8oxo3 had the fastest unfolding rates of K_m = 4-5 x 10^{-11} M. The TTAGGG constructs are less stable than GGGT constructs, which may contribute to the faster rates of unfolding. As with the GGGT, the native TTAGGG construct had a lower unfolding rate compared to the 8oxoG substrates. Looking at the binding of native TTAGGG (K = 4.2 x 10^6 M^{-1}), it is relatively equal in binding to 8oxo3 (K = 3.9 x 10^6 M^{-1}). While the model does not track as closely, the undamaged sequence that REV1 had the most affinity is more challenging for REV1 to unfold.

Next we investigated if the REV1 Core E466A/Y470A mutant was able to disrupt GGGT and TTAGGG. For the GGGT construct, EA/YA protein was able to unfold the construct faster than the wild-type, for the TTAGGG unfolded at an equal rate. The expectation of this mutation was that these residues interact with the planar face of G-quartets and could influence binding and unfolding. Previous work by Ketkar et al., study the mutations individually and found that the activity of Y470A was still comparable to the wild-type enzyme, but the E466A mutant had reduced activity of about 50%. While it was unexpected that the double mutant could unfold GGGT faster, the fact that it retained the ability to unfold was not surprising.

The ability for REV1 and FANCJ to disrupt G4s without the need for ATP may be related to the features of G4s. Insight as to how this may be possible could be found with the single-molecule study that looked at the mechanical features of G4 folding/unfolding. This study found that the primary energetic barrier to G4 unfolding involved disrupting only one or a few tetrad associated guanines. In the case of REV1, by sequestering just one guanine base from the rest of the G4 it could be enough to destabilize, but may need additional enzymes to fully unfold. Previous study by Eddy et al., found that REV1 core could still disrupt G4s without the need for CTP. They hypothesized that REV1 is able to sequester guanines within the protein which can cause the disruption. As for FANCJ, the initial loading of the protein, if near the start of the G4, may be bulky enough to disrupt the initial guanines to allow for unfolding of the G4. Now, the addition of ATP may allow these enzymes to properly unfold through the G4 instead of just the initial disruption, which may account for the difference in unfolding rates.

In conclusion, this study has provided significant insights into the dynamics of G-quadruplex (G4) unfolding mediated by FANCJ helicase and REV1 polymerase, particularly in the context of 8-oxoguanine (8oxoG) damaged G4s. The findings underscore the substrate-specific preferences and roles of these proteins in DNA repair and damage tolerance mechanisms. FANCJ showed varying unfolding rates depending on the position of 8oxoG within the G4 stack, highlighting the critical influence of the 8oxoG position rather than its mere presence. Similarly, REV1 displayed differential unfolding rates, with a notable lower unfolding rate for more stable G4 structures such as the GGGT G4. The unexpected faster unfolding of GGGT by the REV1 Core

E466A/Y470A mutant suggests complex interactions at the molecular level that warrant further investigation. The ability of REV1 and FANCJ to disrupt G4s without ATP suggests a potential mechanism involving the initial disruption of guanines, which may be enhanced by ATP for complete unfolding. Future studies should delve deeper into the molecular interactions and structural dynamics of other G4 DNA.

CHAPTER SIX
CONCLUSION AND FUTURE DIRECTION

6.1 Conclusion

This study aimed to determine the impact of 8oxoG on the folding and stability of parallel, antiparallel, and hybrid G4 structures, depending on its specific positions within the G4. Additionally, we sought to establish that FANCJ and REV1 exhibit distinct binding preferences for 8oxoG4, correlating with the physical properties of these structures. We predicted that the combined action of FANCJ and REV1 would be necessary to repair tightly folded 8oxoG4 structures, while REV1 alone might be sufficient to disrupt more loosely folded 8oxoG4s.

For the G4 stability and conformation study, the presence of 8oxoG in the first position (8oxo1) was the most destabilizing for the parallel sequences, while the fifth position was the most destabilizing for hybrid sequences. These findings suggested that the impact of 8-oxoguanine on the thermal stability of G-quadruplex structures depends on the sequence position and G4 conformation.

This study determined the ability of FANCJ and REV1 to recognize and bind to 8oxoG4s. FANCJ and REV1 exhibited the same preferential binding to the parallel GGGT sequence. However, FANCJ had the least preference for antiparallel 8oxo3 G4s, while REV1 had the least binding for hybrid 8oxo1 G4s. This suggests that while the proteins can act independently on parallel G4s, they may need to recruit one another in the presence of 8oxoG. This is further highlighted by the unfolding studies. FANCJ had the lowest unfolding rate for 8oxo5 G4 DNA, while REV1 had no issues with those

constructs. Conversely, REV1 had difficulty unfolding native G4 DNA, and FANCJ had no issues. These unfolding characteristics indicate a complementary action between the two proteins. When one protein falters in unfolding, the other can be recruited to unfold the G4.

6.2 Future Direction

For future directions of this work, the interaction between FANCJ and REV1 needs to be further tested. A method to study their interactions in the cell would be to use co-immunoprecipitation (co-IP). This method is a technique that is used to identify physiological protein-protein interactions by using protein-specific antibodies to capture the proteins that are bound to a specific target. Using this method would give conformation that the two proteins interact in the cell. Additionally, determining if FANCJ and REV1 interact at G4 sites within the cell would also need to be established. Constructs for FANCJ and REV1 have been developed with different fluorescent tags. To determine interaction, FANCJ-BFP, REV1-m-Cherry, and BMVC could be used for co-localization studies.

An additional future study would be to include the GGGTT sequence in the unfolding studies to determine if the pattern seen here continues with the other sequence. Also, the study should incorporate other G4 sequences such as the *MYC* G4, regulates the MYC gene, which is implicated in many cancers. Using other G4 DNA may give rise to sequence specific binding for FANCJ and REV1 that is not seen in this study. The future unfolding studies should also include FANCJ and REV1 being incorporated into the same unfolding assay to determine if there is a synergistic effect that facilitates the rescue of G4 unfolding.

APPENDIX

IBC APPROVAL LETTER

cayuse Hazard Safety

Reference: IBC Protocol # IBC-00001119

Protocol Title: Mutational analysis of DNA repair proteins

Principal Investigator: Wu, Colin

Approval Period: 02/20/2024 - 02/20/2027

Dear Professor Wu, Colin:
The Institutional Biosafety Committee (IBC), responsible for the review of research involving materials of human origin, recombinant DNA, and biohazardous, infectious or toxic materials, has reviewed the original proposal, noted the revisions provided by you upon committee's request (if applicable), and approved the project referenced above. This is your formal approval notice; please save it for your records.

The proposed activities are approved for a period of three years, starting 02/20/2024 and ending 02/20/2027, and assigned the approval # IBC-00001119. Please be aware that protocols expire at 12:01 AM on the expiration date. The highest BSL level approved for the protocol can be found on the Overview page of the protocol form.

Any unanticipated problems, changes in, or halting of the activities proposed in the project, as described, must be promptly reported to the IBC or the biosafety officer.

In addition, all faculty, staff, and students involved in the project must be trained and qualified to conduct the project in a responsible manner. Please note, in this regard, that you, the Principal Investigator, are fully responsible at all times for limiting or restricting access to laboratories and for assessing, in each circumstance, who may enter or work in a laboratory when biohazardous materials are in use.

Prior to initiating any work on this approved project, please contact me by email (jdhallau@oakland.edu) to schedule a startup Biosafety Audit. This audit will verify that all biosafety controls described in your application have been implemented.

Please let me know if you have any questions. Thank you.

Sincerely,

Janell Hallauer
Biosafety Officer
Laboratory Safety and Compliance

REFERENCES

1. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-738.

2. Bang, I., *Untersuchungen über die Guanylsäure.* Biochem. Z, 1910. **26**: p. 293-311.

3. Gellert, M., M.N. Lipsett, and D.R. Davies, *HELIX FORMATION BY GUANYLIC ACID.* Proceedings of the National Academy of Sciences, 1962. **48**(12): p. 2013-2018.

4. Todd, A.K., M. Johnston, and S. Neidle, *Highly prevalent putative quadruplex sequence motifs in human DNA.* Nucleic acids research, 2005. **33**(9): p. 2901-2907.

5. Huppert, J.L. and S. Balasubramanian, *Prevalence of quadruplexes in the human genome.* Nucleic acids research, 2005. **33**(9): p. 2908-2916.

6. Mukundan, V.T. and A.T. Phan, *Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences.* Journal of the American Chemical Society, 2013. **135**(13): p. 5017-5028.

7. Sengar, A., et al., *Structure of a (3+ 1) hybrid G-quadruplex in the PARP1 promoter.* Nucleic acids research, 2019. **47**(3): p. 1564-1572.

8. Maizels, N. and L.T. Gray, *The G4 genome.* PLoS genetics, 2013. **9**(4): p. e1003468.

9. Arnott, S., R. Chandrasekaran, and C.M. Marttila, *Structures for polyinosinic acid and polyguanylic acid.* Biochemical Journal, 1974. **141**(2): p. 537-543.

10. Zimmerman, S.B., G.H. Cohen, and D.R. Davies, *X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid.* Journal of Molecular Biology, 1975. **92**(2): p. 181-192.

11. Sen, D. and W. Gilbert, *Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis.* Nature, 1988. **334**(6180): p. 364-366.

12. Neidle, S. and S. Balasubramanian, *Quadruplex nucleic acids*. Vol. 7. 2006: Royal Society of Chemistry.

13. Sundquist, W.I. and A. Klug, *Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops.* Nature, 1989. **342**(6251): p. 825-829.

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

15. Ohno, M., et al., *8-oxoguanine causes spontaneous de novo germline mutations in mice.* Scientific Reports, 2014. **4**(1): p. 4689.

16. Pavlova, A.V., et al., *Impact of G-Quadruplexes on the Regulation of Genome Integrity, DNA Damage and Repair.* Biomolecules, 2021. **11**(9): p. 1284.

17. Vorlickova, M., et al., *8-oxoguanine in a quadruplex of the human telomere DNA sequence.* FEBS J, 2012. **279**(1): p. 29-39.

18. Lech, C.J., et al., *Effects of site-specific guanine C8-modifications on an intramolecular DNA G-quadruplex.* Biophys J, 2011. **101**(8): p. 1987-98.

19. Lane, A.N., et al., *Stability and kinetics of G-quadruplex structures.* Nucleic Acids Res, 2008. **36**(17): p. 5482-515.

20. Phan, A.T., V. Kuryavyi, and D.J. Patel, *DNA architecture: from G to Z.* Curr Opin Struct Biol, 2006. **16**(3): p. 288-98.

21. Olsen, C.M., W.H. Gmeiner, and L.A. Marky, *Unfolding of G-quadruplexes: energetic, and ion and water contributions of G-quartet stacking.* J Phys Chem B, 2006. **110**(13): p. 6962-9.

22. Balasubramanian, S., L.H. Hurley, and S. Neidle, *Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?* Nat Rev Drug Discov, 2011. **10**(4): p. 261-75.

23. Biffi, G., et al., *Quantitative visualization of DNA G-quadruplex structures in human cells.* Nat Chem, 2013. **5**(3): p. 182-6.

24. Day, H.A., P. Pavlou, and Z.A. Waller, *i-Motif DNA: structure, stability and targeting with ligands.* Bioorg Med Chem, 2014. **22**(16): p. 4407-18.

25. Huppert, J.L. and S. Balasubramanian, *G-quadruplexes in promoters throughout the human genome.* Nucleic Acids Res, 2007. **35**(2): p. 406-13.

26. Wang, Y. and D.J. Patel, *Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex.* Structure, 1993. **1**(4): p. 263-82.

27. Parkinson, G.N., M.P. Lee, and S. Neidle, *Crystal structure of parallel quadruplexes from human telomeric DNA.* Nature, 2002. **417**(6891): p. 876-80.

28. Blackburn, E.H., *Telomeres and their synthesis.* Science, 1990. **249**(4968): p. 489-90.

29. Tateishi-Karimata, H., K. Kawauchi, and N. Sugimoto, *Destabilization of DNA G-Quadruplexes by Chemical Environment Changes during Tumor Progression Facilitates Transcription.* J Am Chem Soc, 2018. **140**(2): p. 642-651.

30. Wang, M., et al., *Odorant Sensory Input Modulates DNA Secondary Structure Formation and Heterogeneous Ribonucleoprotein Recruitment on the Tyrosine Hydroxylase and Glutamic Acid Decarboxylase 1 Promoters in the Olfactory Bulb.* J Neurosci, 2017. **37**(18): p. 4778-4789.

31. Das, K., M. Srivastava, and S.C. Raghavan, *GNG Motifs Can Replace a GGG Stretch during G-Quadruplex Formation in a Context Dependent Manner.* PLoS One, 2016. **11**(7): p. e0158794.

32. Chang, T.C. and C.C. Chang, *Detection of G-quadruplexes in cells and investigation of G-quadruplex structure of d(T2AG3)4 in K+ solution by a carbazole derivative: BMVC.* Methods Mol Biol, 2010. **608**: p. 183-206.

33. Tseng, T.Y., et al., *Binding of Small Molecules to G-quadruplex DNA in Cells Revealed by Fluorescence Lifetime Imaging Microscopy of o-BMVC Foci.* Molecules, 2018. **24**(1).

34. Tseng, T.Y., et al., *The G-quadruplex fluorescent probe 3,6-bis(1-methyl-2-vinyl-pyridinium) carbazole diiodide as a biosensor for human cancers.* Sci Rep, 2018. **8**(1): p. 16082.

35. Liu, W., et al., *Structures of 1:1 and 2:1 complexes of BMVC and MYC promoter G-quadruplex reveal a mechanism of ligand conformation adjustment for G4-recognition.* Nucleic Acids Res, 2019. **47**(22): p. 11931-11942.

36. Christman, J.K., *5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy.* Oncogene, 2002. **21**(35): p. 5483-95.

37. Mao, S.Q., et al., *DNA G-quadruplex structures mold the DNA methylome.* Nat Struct Mol Biol, 2018. **25**(10): p. 951-957.

38. Byrd, A.K., et al., *Evidence That G-quadruplex DNA Accumulates in the Cytoplasm and Participates in Stress Granule Assembly in Response to Oxidative Stress.* J Biol Chem, 2016. **291**(34): p. 18041-57.

39. Cogoi, S., et al., *Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins.* Nucleic Acids Res, 2008. **36**(11): p. 3765-80.

40. Pagano, B., et al., *Identification of novel interactors of human telomeric G-quadruplex DNA.* Chem Commun (Camb), 2015. **51**(14): p. 2964-7.

41. Gray, L.T., et al., *G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD.* Nat Chem Biol, 2014. **10**(4): p. 313-8.

42. Brázda, V., et al., *The Amino Acid Composition of Quadruplex Binding Proteins Reveals a Shared Motif and Predicts New Potential Quadruplex Interactors.* Molecules, 2018. **23**(9).

43. Huang, Z.L., et al., *Identification of G-Quadruplex-Binding Protein from the Exploration of RGG Motif/G-Quadruplex Interactions.* J Am Chem Soc, 2018. **140**(51): p. 17945-17955.

44. Huang, W., et al., *Yin Yang 1 contains G-quadruplex structures in its promoter and 5′-UTR and its expression is modulated by G4 resolvase 1.* Nucleic Acids Res, 2012. **40**(3): p. 1033-49.

45. Ray, S., et al., *G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding.* Proc Natl Acad Sci U S A, 2014. **111**(8): p. 2990-5.

46. Brazda, V., et al., *Strong preference of BRCA1 protein to topologically constrained non-B DNA structures.* BMC Mol Biol, 2016. **17**(1): p. 14.

47. Hudson, J.S., et al., *Recognition and binding of human telomeric G-quadruplex DNA by unfolding protein 1.* Biochemistry, 2014. **53**(20): p. 3347-56.

48. Goering, R., et al., *FMRP promotes RNA localization to neuronal projections through interactions between its RGG domain and G-quadruplex RNA sequences.* eLife, 2020. **9**: p. e52621.

49. Pyle, A.M., *Translocation and unwinding mechanisms of RNA and DNA helicases.* Annu Rev Biophys, 2008. **37**: p. 317-36.

50. Patel, S.S. and I. Donmez, *Mechanisms of Helicases.* J. Biol. Chem., 2006. **281**(27): p. 18265-18268.

51. Lohman, T.M. and K.P. Bjornson, *Mechanisms of helicase-catalyzed DNA unwinding.* Annu Rev Biochem, 1996. **65**: p. 169-214.

52. Gauss, P., et al., *DNA helicase requirements for DNA replication during bateriophage T4 infection.* Journal of Bacteriology, 1994. **176**: p. 1667-1672.

53. Biswas, E.E., P.-H. Chen, and S.B. Biswas, *DNA helicase associated with DNA polymerase α: isolation by a modified immunoaffinity chromatography.* Biochemistry, 1993. **32**: p. 13393-13398.

54. Abdel-Monem, M. and H. Hoffmann-Berling, *Enzymatic unwinding of DNA. 1. Purification and characterization of a DNA-dependent ATPase from Escherichia coli.* Eur J Biochem, 1976. **65**(2): p. 431-40.

55. Singleton, M.R., M.S. Dillingham, and D.B. Wigley, *Structure and Mechanism of Helicases and Nucleic Acid Translocases.* Annu Rev Biochem, 2007. **76**: p. 23-50.

56. Buttner, K., S. Nehring, and K.P. Hopfner, *Structural basis for DNA duplex separation by a superfamily-2 helicase.* Nat Struct Mol Biol, 2007. **14**(7): p. 647-52.

57. O'Donnell, M.E. and H. Li, *The ring-shaped hexameric helicases that function at DNA replication forks.* Nature Structural & Molecular Biology, 2018. **25**(2): p. 122-130.

58. Gao, Y. and W. Yang, *Different mechanisms for translocation by monomeric and hexameric helicases.* Curr Opin Struct Biol, 2020. **61**: p. 25-32.

59. von Hippel, P.H. and E. Delagoutte, *A General Model for Nucleic Acid Helicases and their "Coupling" within Macromolecular Machines.* Cell, 2001. **104**: p. 177-190.

60. Fairman-Williams, M.E., U.P. Guenther, and E. Jankowsky, *SF1 and SF2 helicases: family matters.* Curr Opin Struct Biol, 2010. **20**(3): p. 313-24.

61. Seidel, R., et al., *Motor step size and ATP coupling efficiency of the dsDNA translocase EcoR124I.* Embo J, 2008.

62. Myong, S., et al., *Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase.* Science, 2007. **317**(5837): p. 513-6.

63. Zhang, X.D., et al., *Escherichia coli RecQ is a rapid, efficient, and monomeric helicase.* J. Biol. Chem., 2006. **281**(18): p. 12655-63.

64. Fairman, M.E., et al., *Protein displacement by DExH/D "RNA helicases" without duplex unwinding.* Science, 2004. **304**(5671): p. 730-4.

65. Cheok, C.F., et al., *Roles of the Bloom's syndrome helicase in the maintenance of genome stability.* Biochem. Soc. Trans., 2005. **33**(Pt 6): p. 1456-9.

66. Budhathoki, J.B., et al., *RecQ-core of BLM unfolds telomeric G-quadruplex in the absence of ATP.* Nucleic Acids Res, 2014. **42**(18): p. 11528-45.

67. Mo, J., et al., *DDX3X: structure, physiologic functions and cancer.* Molecular Cancer, 2021. **20**(1): p. 38.

68. Khadivjam, B., et al., *RNA helicase DDX3X modulates herpes simplex virus 1 nuclear egress.* Communications Biology, 2023. **6**(1): p. 134.

69. Brosh, R.M., Jr., et al., *Functional and physical interaction between WRN helicase and human replication protein A.* J Biol Chem, 1999. **274**(26): p. 18341-50.

70. Fan, L., et al., *XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations.* Cell, 2008. **133**(5): p. 789-800.

71. Wu, Y., A.N. Suhasini, and R.M. Brosh, Jr., *Welcome the family of FANCJ-like helicases to the block of genome stability maintenance proteins.* Cell Mol Life Sci, 2009. **66**(7): p. 1209-22.

72. Cantor, S.B., et al., *BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function.* Cell, 2001. **105**(1): p. 149-60.

73. London, T.B., et al., *FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts.* J Biol Chem, 2008. **283**(52): p. 36132-9.

74. Liu, H., et al., *Structure of the DNA repair helicase XPD.* Cell, 2008. **133**(5): p. 801-12.

75. Wolski, S.C., et al., *Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD.* PLoS Biol, 2008. **6**(6): p. e149.

76. Rudolf, J., et al., *The DNA repair helicases XPD and FancJ have essential iron-sulfur domains.* Mol Cell, 2006. **23**(6): p. 801-8.

77. Yu, X., et al., *The BRCT Domain Is a Phospho-Protein Binding Domain.* Science, 2003. **302**(5645): p. 639-642.

78. Cantor, S., et al., *The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations.* Proceedings of the National Academy of Sciences, 2004. **101**(8): p. 2357-2362.

79. Gupta, R., et al., *Inhibition of BACH1 (FANCJ) helicase by backbone discontinuity is overcome by increased motor ATPase or length of loading strand.* Nucleic Acids Res, 2006. **34**(22): p. 6673-83.

80. Gupta, R., et al., *FANCJ (BACH1) helicase forms DNA damage inducible foci with replication protein A and interacts physically and functionally with the single-stranded DNA-binding protein.* Blood, 2007. **110**(7): p. 2390-8.

81. Gupta, R., et al., *Analysis of the DNA substrate specificity of the human BACH1 helicase associated with breast cancer.* J Biol Chem, 2005. **280**(27): p. 25450-60.

82. Sommers, J.A., et al., *FANCJ uses its motor ATPase to destabilize protein-DNA complexes, unwind triplexes, and inhibit RAD51 strand exchange.* J Biol Chem, 2009. **284**(12): p. 7505-17.

83. Peng, M., et al., *The FANCJ/MutLalpha interaction is required for correction of the cross-link response in FA-J cells.* EMBO J, 2007. **26**(13): p. 3238-49.

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

85. Awate, S., et al., *FANCJ compensates for RAP80 deficiency and suppresses genomic instability induced by interstrand cross-links.* Nucleic Acids Res, 2020. **48**(16): p. 9161-9180.

86. Yaneva, D., et al., *The FANCJ helicase unfolds DNA-protein crosslinks to promote their repair.* Molecular Cell, 2023. **83**(1): p. 43-56.e10.

87. Wu, Y., K. Shin-ya, and R.M. Brosh, 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.

88. Sarkies, P., et al., *FANCJ coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA.* Nucleic Acids Res, 2012. **40**(4): p. 1485-98.

89. Bharti, S.K., et al., *Specialization among iron-sulfur cluster helicases to resolve G-quadruplex DNA structures that threaten genomic stability.* J Biol Chem, 2013. **288**(39): p. 28217-29.

90. Castillo Bosch, P., et al., *FANCJ promotes DNA synthesis through G-quadruplex structures.* EMBO J, 2014. **33**(21): p. 2521-33.

91. Wu, C.G. and M. Spies, *G-quadruplex recognition and remodeling by the FANCJ helicase.* Nucleic Acids Res, 2016. **44**(18): p. 8742-8753.

92. Kruisselbrink, E., et al., *Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCJ-defective C. elegans.* Curr Biol, 2008. **18**(12): p. 900-5.

93. Cheung, I., et al., *Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA.* Nat Genet, 2002. **31**(4): p. 405-9.

94. Lowran, K., et al., *Assembly of a G-Quadruplex Repair Complex by the FANCJ DNA Helicase and the REV1 Polymerase.* Genes (Basel), 2019. **11**(1).

95. Heddi, B., et al., *Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex complex.* Proc Natl Acad Sci U S A, 2015. **112**(31): p. 9608-13.

96. Mailand, N., *A DNA helicase remodeling proteins: How DNA-protein crosslink repair unfolds via FANCJ.* Molecular Cell, 2023. **83**(1): p. 3-5.

97. Peng, M., et al., *Opposing Roles of FANCJ and HLTF Protect Forks and Restrain Replication during Stress.* Cell Rep, 2018. **24**(12): p. 3251-3261.

98. Wiedemann, H.R., *Guido Fanconi (1892-1979) in memoriam.* Eur J Pediatr, 1979. **132**(3): p. 131-2.

99. Bagby, G.C., *Multifunctional Fanconi proteins, inflammation and the Fanconi phenotype.* EBioMedicine, 2016. **8**: p. 10-11.

100. Niraj, J., A. Färkkilä, and A.D. D'Andrea, *The Fanconi Anemia Pathway in Cancer.* Annu Rev Cancer Biol, 2019. **3**: p. 457-478.

101. Hoeijmakers, J.H., *Genome maintenance mechanisms for preventing cancer.* Nature, 2001. **411**(6835): p. 366-74.

102. Friedberg, E.C., A.R. Lehmann, and R.P. Fuchs, *Trading places: how do DNA polymerases switch during translesion DNA synthesis?* Mol Cell, 2005. **18**(5): p. 499-505.

103. Ogi, T. and A.R. Lehmann, *The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair.* Nat Cell Biol, 2006. **8**(6): p. 640-2.

104. Ogi, T., et al., *Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells.* Mol Cell, 2010. **37**(5): p. 714-27.

105. Bétous, R., et al., *Role of TLS DNA polymerases eta and kappa in processing naturally occurring structured DNA in human cells.* Mol Carcinog, 2009. **48**(4): p. 369-78.

106. Eddy, S., et al., *Human Rev1 polymerase disrupts G-quadruplex DNA.* Nucleic Acids Res, 2014. **42**(5): p. 3272-85.

107. Masuda, Y. and K. Kamiya, *Biochemical properties of the human REV1 protein.* FEBS Lett, 2002. **520**(1-3): p. 88-92.

108. Nelson, J.R., C.W. Lawrence, and D.C. Hinkle, *Deoxycytidyl transferase activity of yeast REV1 protein.* Nature, 1996. **382**(6593): p. 729-31.

109. Ohashi, E., et al., *Identification of a novel REV1-interacting motif necessary for DNA polymerase kappa function.* Genes Cells, 2009. **14**(2): p. 101-11.

110. Prakash, L., *Effect of Genes Controlling Radiation Sensitivity on Chemically Induced Mutations in SACCHAROMYCES CEREVISIAE.* Genetics, 1976. **83**(2): p. 285-301.

111. Krijger, P.H., et al., *Rev1 is essential in generating G to C transversions downstream of the Ung2 pathway but not the Msh2+Ung2 hybrid pathway.* European Journal of Immunology, 2013. **43**(10): p. 2765-2770.

112. Nair, D.T., et al., *DNA Synthesis across an Abasic Lesion by Yeast Rev1 DNA Polymerase.* Journal of Molecular Biology, 2011. **406**(1): p. 18-28.

113. Boehm, E.M. and M.T. Washington, *R.I.P. to the PIP: PCNA-binding motif no longer considered specific: PIP motifs and other related sequences are not distinct entities and can bind multiple proteins involved in genome maintenance.* Bioessays, 2016.

114. Guo, C., et al., *REV1 protein interacts with PCNA: significance of the REV1 BRCT domain in vitro and in vivo.* Mol Cell, 2006. **23**(2): p. 265-71.

115. Pustovalova, Y., et al., *Interaction between the Rev1 C-Terminal Domain and the PolD3 Subunit of Polzeta Suggests a Mechanism of Polymerase Exchange upon Rev1/Polzeta-Dependent Translesion Synthesis.* Biochemistry, 2016. **55**(13): p. 2043-53.

116. Nayak, S., et al., *Inhibition of the translesion synthesis polymerase REV1 exploits replication gaps as a cancer vulnerability.* Sci Adv, 2020. **6**(24): p. eaaz7808.

117. Johnson, R.E., et al., *Identification of APN2, the Saccharomyces cerevisiae homolog of the major human AP endonuclease HAP1, and its role in the repair of abasic sites.* Genes Dev, 1998. **12**(19): p. 3137-43.

118. Johnson, R.E., et al., *Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions.* Nature, 2000. **406**(6799): p. 1015-9.

119. Schiavone, D., et al., *Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells.* EMBO J, 2014. **33**(21): p. 2507-20.

120. Ketkar, A., et al., *Human Rev1 relies on insert-2 to promote selective binding and accurate replication of stabilized G-quadruplex motifs.* Nucleic Acids Res, 2021. **49**(4): p. 2065-2084.

121. Bhattacharyya, D., G. Mirihana Arachchilage, and S. Basu, *Metal Cations in G-Quadruplex Folding and Stability.* Frontiers in Chemistry, 2016. **4**.

122. Hazel, P., et al., *Loop-length-dependent folding of G-quadruplexes.* J Am Chem Soc, 2004. **126**(50): p. 16405-15.

123. Smargiasso, N., et al., *G-Quadruplex DNA Assemblies: Loop Length, Cation Identity, and Multimer Formation.* Journal of the American Chemical Society, 2008. **130**(31): p. 10208-10216.

124. Guédin, A., et al., *How long is too long? Effects of loop size on G-quadruplex stability.* Nucleic Acids Res, 2010. **38**(21): p. 7858-68.

125. Burge, S., et al., *Quadruplex DNA: sequence, topology and structure.* Nucleic Acids Res, 2006. **34**(19): p. 5402-15.

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

127. Del Villar-Guerra, R., R.D. Gray, and J.B. Chaires, *Characterization of Quadruplex DNA Structure by Circular Dichroism.* Curr Protoc Nucleic Acid Chem, 2017. **68**: p. 17 8 1-17 8 16.

128. Olsen, C.M., H.-T. Lee, and L.A. Marky, *Unfolding Thermodynamics of Intramolecular G-Quadruplexes: Base Sequence Contributions of the Loops.* The Journal of Physical Chemistry B, 2009. **113**(9): p. 2587-2595.

129. Olsen, C.M. and L.A. Marky, *Monitoring the Temperature Unfolding of G-Quadruplexes by UV and Circular Dichroism Spectroscopies and Calorimetry Techniques*, in *G-Quadruplex DNA: Methods and Protocols*, P. Baumann, Editor. 2010, Humana Press: Totowa, NJ. p. 147-158.

130. Rachwal, P.A. and K.R. Fox, *Quadruplex melting.* Methods, 2007. **43**(4): p. 291-301.

131. Fortini, P., et al., *8-Oxoguanine DNA damage: at the crossroad of alternative repair pathways.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2003. **531**(1): p. 127-139.

132. Bruner, S.D., D.P.G. Norman, and G.L. Verdine, *Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA.* Nature, 2000. **403**(6772): p. 859-866.

133. Cheong, V.V., et al., *Xanthine and 8-oxoguanine in G-quadruplexes: formation of a G.G.X.O tetrad.* Nucleic Acids Res, 2015. **43**(21): p. 10506-14.

134. Greenfield, N.J., *Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions.* Nat Protoc, 2006. **1**(6): p. 2527-35.

135. Ambrus, A., et al., *Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution.* Nucleic Acids Res, 2006. **34**(9): p. 2723-35.

136. Balagurumoorthy, P. and S.K. Brahmachari, *Structure and stability of human telomeric sequence.* J Biol Chem, 1994. **269**(34): p. 21858-69.

137. Wu, Y., et al., *Fanconi anemia Group J mutation abolishes its DNA repair function by uncoupling DNA translocation from helicase activity or disruption of protein-DNA complexes.* Blood, 2010.

138. Odermatt, D.C., et al., *Cancer-associated mutations in the iron-sulfur domain of FANCJ affect G-quadruplex metabolism.* PLoS Genet, 2020. **16**(6): p. e1008740.

139. Varshney, D., et al., *The regulation and functions of DNA and RNA G-quadruplexes.* Nature Reviews Molecular Cell Biology, 2020. **21**(8): p. 459-474.

140. Lattmann, S., et al., *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.

141. Boehm, E.M., et al., *The Proliferating Cell Nuclear Antigen (PCNA)-interacting Protein (PIP) Motif of DNA Polymerase eta Mediates Its Interaction with the C-terminal Domain of Rev1.* J Biol Chem, 2016. **291**(16): p. 8735-44.

142. Masuda, Y., et al., *Different types of interaction between PCNA and PIP boxes contribute to distinct cellular functions of Y-family DNA polymerases.* Nucleic Acids Res, 2015. **43**(16): p. 7898-910.

143. Armstrong, A.A., F. Mohideen, and C.D. Lima, *Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2.* Nature, 2012. **483**(7387): p. 59-63.

144. Prestel, A., et al., *The PCNA interaction motifs revisited: thinking outside the PIP-box.* Cell Mol Life Sci, 2019.

145. De March, M., et al., *Structural basis of human PCNA sliding on DNA.* Nat Commun, 2017. **8**: p. 13935.

146. Pozhidaeva, A., et al., *NMR structure and dynamics of the C-terminal domain from human Rev1 and its complex with Rev1 interacting region of DNA polymerase eta.* Biochemistry, 2012. **51**(27): p. 5506-20.

147. Boehm, E.M., M. Spies, and M.T. Washington, *PCNA tool belts and polymerase bridges form during translesion synthesis.* Nucleic Acids Res, 2016. **44**(17): p. 8250-60.

148. Hishiki, A., et al., *Structural basis for novel interactions between human translesion synthesis polymerases and proliferating cell nuclear antigen.* J Biol Chem, 2009. **284**(16): p. 10552-60.

149. Gilljam, K.M., et al., *Identification of a novel, widespread, and functionally important PCNA-binding motif.* J Cell Biol, 2009. **186**(5): p. 645-54.

150. Poon, L.C., et al., *Guanine-rich RNAs and DNAs that bind heme robustly catalyze oxygen transfer reactions.* J Am Chem Soc, 2011. **133**(6): p. 1877-84.

151. Alizadeh, N., A. Salimi, and R. Hallaj, *Hemin/G-Quadruplex Horseradish Peroxidase-Mimicking DNAzyme: Principle and Biosensing Application.* Adv Biochem Eng Biotechnol, 2020. **170**: p. 85-106.

152. Wang, F., C.H. Lu, and I. Willner, *From cascaded catalytic nucleic acids to enzyme-DNA nanostructures: controlling reactivity, sensing, logic operations, and assembly of complex structures.* Chem Rev, 2014. **114**(5): p. 2881-941.

153. Willner, I., et al., *DNAzymes for sensing, nanobiotechnology and logic gate applications.* Chem Soc Rev, 2008. **37**(6): p. 1153-65.

154. Mao, X., et al., *Activity modulation and allosteric control of a scaffolded DNAzyme using a dynamic DNA nanostructure.* Chem Sci, 2016. **7**(2): p. 1200-1204.