Mechanisms of DNA Translocases in the Repair of Damaged Replication Forks

By

Akosua Agyeman Badu-Nkansah

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

August, 2016

Nashville, Tennessee

Approved:

David K. Cortez, Ph.D.

Brandt F. Eichman, Ph.D.

Nicholas J. Reiter, Ph.D.

Scott W. Hiebert, Ph.D.

Katherine L. Friedman, Ph.D.

To my parents and my brother, Kwabena

ACKNOWLEDGEMENTS

This work would not be possible without the financial support provided by the Biochemical and Chemical Training for Cancer Research grant T32 CA9582-25 and the NCI Supplement to Promote Diversity in Health-Related Research R01CA102729.

I want to thank my mentor, Dr. David Cortez, for his guidance. I want to also thank the members of my committee, Dr. Christine Eischen, Dr. Brandt Eichman, Dr. Katherine Friedman, Dr. Scott Hiebert and Dr. Nicholas Reiter, for the feedback that they provided during committee meetings.

To current and former Cortez lab members, Carol Bansbach Robbins, Jamie Couch, Remy Bétous, Clint Carroll, Kamakoti Bhat, and Cory Holland thank you for helping me with *in vitro* DNA binding, ATPase and fork remodeling enzymatic assays that were used throughout my studies on SMARCAL1 and ZRANB3 and generating complementation assays for studying ZRANB3 function in cells. Aaron Mason, a former post doctoral trainee in Dr. Brandt Eichman's lab, helped me tremendously in purifying the substrate recognition domain construct.

Bianca Sirbu, thank you for your encouragement and positivity throughout the years. Thank you Gina Kavanaugh for always being someone I could talk to about science and life. The Budo Taijutsu lessons were really fun! Lisa Poole, your pastries and Korean dumplings lifted my spirits and thank you for reminding me that birthdays should always be celebrated! You, Kamakoti Bhat and Thomas Bass are fantastic scientists and I am sure that all of you will be tremendously successful in your future career endeavors. To Jessica Luzwick, we entered the Cortez Lab together and we are graduating together. Thank you for being a great lab member and friend.

Kareem Mohni, Huzefa Dungrawala, and Sarah Wessel served as excellent resources for help with troubleshooting experiments and discussing scientific literature. You guys are amazing scientists and thank you for your help.

Gloria Glick and Nancy Zhoa are the reason for the why the Cortez lab is so efficient. Thank you for your help over these years.

Lastly, thank you to my mom, my dad, my brother, Dr. Samuel Adunyah and Mrs. Patrice Adunyah and all of my Nashville friends. Thank you for all the fun times, support and up lifting words.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

9-1-1	RAD9 - RAD1- HUS1
APIM	ALKB homolog 2 PCNA interaction motif
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATPase	ATP hydrolysis activity
ATR	Ataxia telangiectasia mutated and Rad3-related
ATRIP	ATR-interacting protein
BRCT	Breast cancer-1/BRCA1 C-terminal
BLM	Bloom Syndrome protein
CDC6	Cell division cycle 6
CDC7	Cell division cycle 7
CDK	Cyclin dependent kinase
CDT1	Chromatin licensing and DNA replication factor 1
CHK2	Checkpoint kinase 2
CPT	Camptothecin
CRISPR	Clustered regularly interspaced short palindromic repeats
CtIP	CtBP (C-terminal binding protein) interacting protein
DBD	DNA binding domain
DBF4	DumbBell former; part of the DDK (DBF4-CDC7) complex
DDK	DBF4-dependent kinase
DDT	DNA damage tolerance
DMEM	Dulbecco’s modified eagle medium
DNA-PK	DNA-dependent protein kinase
DPB11	DNA polymerases B (II)
DSB	Double stranded break

E2	Ubiquitin conjugating complex
E3	Ubiquitin ligase complex
EMSA	Electrophoretic mobility shift assay
ERFS	Early replicating fragile sites
Exo	Exonuclease domain
FBS	Fetal bovine serum
FEN1	Flap structure specific endonuclease 1
γH2AX	H2AX phosphoserine 139
G1	Gap 1 phase (Cell cycle phase)
GFP	Green fluorescent protein
GINS	Go, Ichi, Ni, and San complex
gRNA	guide RNA
H_2O_2	Hydrogen Peroxide
HARP	HepA-related protein
HEK293T	Human embryonic kidney cell line
HIRAN	HIP116, RAD5P N-terminal domain
HLTF	Helicase-like transcription factor
HNH	Histidine-asparagine-histidine endonuclease domain
HP1	HARP1 domain
HP2	HARP2 domain
HRDC	Helicase and RNase D-like C-terminal domain
HU	Hydroxyurea
ICL	Interstrand crosslink
ISWI	Imitation SWI
KIX	Kinase inducible domain interacting domain
LIG1	DNA Ligase 1
MCM	Mini-chromosome maintenance
MonoUB	Monoubiquitination
MMS	Methyl-methanesulphonate

MMS2	Methyl-methanesulphonate sensitivity 2
Mock	No protein control
MOT1	Modifier of transcription 1
MRD	Mismatch recognition domain
MRN	MRE11-RAD50-NBS1
MSH6	MUTS homolog 6
MT1	L760A/D761A/I762A
MT2	W790A/S791A/S792A
MUS81	MUS81 structure-specific endonuclease subunit
MUTS	Mutator S
NZF	NPL4 zinc finger motif
OB	Oligonucleotide/Oligosaccharide binding domain
ORC	Origin of replication complex
PARP	Poly ADP ribose polymerase
PBS	Phosphate buffer saline
PCNA	Proliferating cell nuclear antigen
PEI	Polyethylenimine
PHD	PHD-finger domain (binds histones)
PIP box	PCNA interacting protein box motif
POL α-Primase	Polymerase alpha- primase complex
POL δ	Polymerase delta
POL ε	Polymerase epsilon
POL ζ	Polymerase zeta (B-family)
POL η	Polymerase nu (Y-family)
POL ι	Polymerase iota (Y-family)
POL κ	Polymerase kappa (Y-family)
PolyUb	Polyubiquitination
Pre-RC	Pre-replication complex
RAD5	Radiation sensitive 5

RAD17	Checkpoint clamp loader
RBD	RPA binding domain
RECA	Recombinase A
RECG	ATP-dependent DNA helicase RECG
RECQ	ATP-dependent DNA helicase RECQ
RECQ1/RecQL1	RECQ-like helicase 1
RECQ4/RecQL4	RECQ-like helicase 4
RECQ5/RECQL5	RECQ-like helicase 5
RECQ5β/BRCv	BRC repeat variant
REV1	REV1, DNA directed polymerase
RFC	Replication factor c
RING	Really interesting new gene, zinc binding domain
ROS	Reactive oxygen species
RPA	Replication protein A
RPE	Retinal pigment epithelium cell line
RQC	RECQ C-terminal domain
RTS	Rothmund Thomson syndrome
SA	Strand annealing domain
SAM	S-adenosylmethionine
SF2	Super family 2
SHPRH	SF2 histone linker PHD RING helicase (HLTF, RAD5 homolog)
SLD	Synthetically lethal with DPB11
SLD2	SLD2 like domain
SLX4	SLX4 structure-specific endonuclease subunit
SMARCAL1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a-like 1
SNF2	Sucrose non-fermenting 2 family of proteins
SNF2P	Sucrose non-fermenting-2 protein

SRD	Substrate recognition domain
SRI	SET2 RPB1-interacting domain
TBE	Tris/Borate/EDTA buffer
TOPBP1	Topoisomerase II binding protein 1
TOPO I	Topoisomerase I
TOPO II	Topoisomerase II
U2OS	Human osteosarcoma epithelial cells
UBC13	E2 ubiquitin-conjugating protein UBC13
UV	Ultraviolet radiation
UVSW	UV repair and recombination protein
WRN	Werner Syndrome protein
WT	Wild type
XPB	Xeroderma Pigmentosum, complementation group B; ERCC3
ZRANB3/AH2	Zinc finger RAN binding protein 2- type containing 3/ Annealing helicase 2

CHAPTER I

INTRODUCTION

Genome replication is a challenging and highly complex process. In addition to unraveling over 6 billion base pairs from tightly packaged nucleosome complexes and faithfully replicating, in a timely fashion, over 12 billion individual nucleotides, the DNA replication machinery must also navigate roadblocks such as damaged DNA, secondary DNA structured complexes and DNA breaks. Many serious diseases are directly attributed to problems involving DNA replication, such as cancer. Mechanism based research studying proteins functioning in such a fundamental process will (1) establish a clear understanding of how DNA replication is conducted in the context of replication impediments and (2) generate effective therapies to treat diseases (1).

The DNA within every cell experiences thousands of assaults from endogenous and exogenous factors that cause DNA damage (2). Reactive oxygen species or substances (ROS) such as hydrogen peroxide (H_2O_2) generated from metabolizing endogenous macromolecules or from exogenous chemical exposure are forms of DNA damaging factors that cause single and double stranded DNA breaks that lead to replication fork stalling (3). Additionally, direct interference of DNA synthesis due to deoxyribonucleotide pool reduction or polymerase inhibition also hinders replication fork progression. A stalled replication fork triggers a signal that activates the replication stress and DNA damage response pathway that functions to fix damaged DNA and resolve stalled replication forks.

This dissertation is a study of the mechanisms of proteins that function in replication stress and DNA damage tolerance pathways triggered by replication fork stalling. First I will give an overview of DNA replication with a focus on the major

replisome protein players. Then I will discuss mechanisms that cope with damaged DNA and replication stress while actively synthesizing nascent DNA and a class of enzymes that act to stabilize and restart stalled replication forks. This dissertation focuses on the SNF2 class of DNA translocases, specifically the identification of a DNA binding domain that governs the enzymatic activities of one of the SNF2 protein members, ZRANB3. I will then conclude by discussing a model describing ZRANB3 function in the replication stress and DNA damage tolerance pathways, methods for testing this model and new projects resulting from my work.

DNA Replication

Initiation

DNA replication is initiated at sites of the genome called origins. With the exception of *Saccharomyces cerevisiae* (*S.cerevisiae*), whose origins are defined by a consensus DNA sequence, the location of origin sites in higher eukaryotes is poorly defined. A seven subunit protein complex ORC-CDC6 (origin recognition complex-cell division cycle 6) binds at origin sites during late mitosis and early G1 phases of the cell cycle, thereby marking locations in the genome where DNA replication may commence (1). The origin becomes “licensed” or poised to fire upon loading of the MCM2-7 (mini-chromosome maintenance) helicase complex to double stranded DNA by CDT1 (chromatin licensing and DNA replication factor 1). Helicase loading triggers CDC6 mediated ATP hydrolysis, which releases both CDC6 and CDT1 from DNA and allows the ORC-MCM2-7 complex to form the pre-replicative complex (Pre-RC). An increase in the activity of the S-phase specific kinases DDK (DBF4-CDC7) and CDK (cyclin dependent kinase) recruits factors that function in helicase activation and origin firing. Specifically, biochemical reconstitution of origin firing using *S. cerevisiae* proteins

exquisitely shows that DDK phosphorylation of the MCM2-7 complex promotes SLD3/7 and CDC45 binding to phosphorylated MCM2-7; DDK/CDK activity also promotes the recruitment of SLD2 (Synthetically lethal with DPB11-1), DPB11 (DNA Polymerase B (II)), GINS (Go, Ichi, Ni and San complex), POL ε (polymerase epsilon) and MCM10 (minichromosome maintenance 10 replication initiation factor) (4). Helicase activation and local unwinding of the DNA duplex recruits the singled strand binding protein RPA (replication protein A), TOPO II (topoisomerase II) and the POL α-Primase complex (polymerase alpha-primase) which collectively promote origin firing and replication of the leading template strand (1,4). In a cell, additional components including POL δ (polymerase delta), RFC (replication factor C) and PCNA (proliferating cell nuclear antigen) are required for lagging strand synthesis, protein recruitment and enhancing polymerase processivity, which is explained in further detail in the “elongation” section.

Elongation

DNA synthesis is catalyzed by DNA polymerases and this process is supported by several factors that collectively comprise the core replication machinery. During the initiation step, the POL α-primase complex synthesizes RNA-DNA primers on the lagging and leading template strands that are extended by the lagging strand polymerase, POL δ (polymerase delta), and the leading strand polymerase POL ε (polymerase epsilon). Prior to POL δ and POL ε loading onto their respective template strands, PCNA is loaded onto DNA by RFC. PCNA binding to DNA polymerases δ and ε ensures that the polymerases are constantly attached to the template strands while synthesizing nascent DNA, and serves as a protein scaffold to localize factors to DNA synthesis sites such as those involved in DNA damage repair and maintaining replication fidelity (5). Replication of the lagging strand template requires the endonuclease FEN1 (Flap structure-specific endonuclease 1) to process Okazaki fragments by cleaving flaps

that are produced by POL δ replicating through and displacing the RNA-DNA primer synthesized by POL α on the preceding Okazaki fragment. “Matured” Okazaki fragments, lacking RNA-DNA primer sequences and adjacent leading strand DNA fragments synthesized by POL ε are joined by LIG1 (DNA ligase 1). Single stranded DNA generated on the lagging strand is bound and protected by RPA; supercoiled DNA and DNA catenations generated upstream of the helicase and replication fork are relieved by TOPO I and II (topoisomerase).

The DNA Damage Response

The replication machinery may stall once it encounters DNA damage such as DNA breaks, DNA base adducts and secondary DNA structures. Depending on the type of damage, a signal is produced and triggers the activation of a highly conserved DNA damage response signaling or S-phase checkpoint pathway that is transduced by three key kinases, ATR, ATM and DNA-PK. These kinases orchestrate various cellular responses involved in maintaining genome integrity such as repairing damaged DNA, programmed cell cycle arrest, and cell death. ATM (Ataxia Telangiectasia Mutated) responds to double stranded DNA breaks (DSB) and becomes activated by interacting with the MRN (MRE11-RAD50-NSB1) complex bound at broken DNA ends. ATM phosphorylates targets primarily involved in homologous recombination mediated repair of DSBs (6). DNA-PK (DNA-dependent protein kinase) activity is stimulated by KU70/KU80 heterodimers bound at DSB ends and promotes non-homologous end joining repair of DSBs (7).

ATR (Ataxia-Telangiectasia mutated and RAD3-related) functions in a more global response to DNA damage and replication stress than ATM and DNA-PK. ATR activation is triggered by the concerted function of multiple proteins that localize to

stalled replication forks. Upon fork stalling, large stretches of single stranded DNA are generated upstream of the fork due to uncoupled DNA synthesis and helicase activities. RPA localizes to and binds single stranded DNA generated at stalled replication forks. RPA coated single stranded DNA recruits ATR and its obligate binding partner ATRIP (ATR-interacting protein). The RAD9-RAD1-HUS1 (9-1-1) heterotrimeric ring complex is loaded by RAD17 (RAD17 checkpoint clamp loader) onto 5' primer ends at stalled replication forks. TOPBP1 (topoisomerase II-binding protein 1), a protein containing BRCT (breast cancer-1/BRCA1 C-terminal) phospho-protein binding domains I and II, binds phosphorylated RAD9 of the 9-1-1 complex. TOPBPI stimulates ATR activation by binding ATRIP through its ATR activation domain. ATR activation induces cell cycle arrest, prevents fork collapse, promotes fork restart and slows origin firing (8).

The Replication Stress Response

Replication forks that stall upon confronting an obstacle impeding its progression induce replication stress. For a cell to proliferate it must replicate its genome to completion and transfer a copy to the daughter cell; replication stress compromises this process. Therefore, the cell induces a replication stress response to stabilize and restart stalled forks. Apart from some replication stressors that generate stalled forks such as ICLs (interstrand DNA crosslinks), the replication stress response is mostly mediated by ATR since replication fork stalling typically generates RPA coated singled stranded DNA (9). The cell uses both ATR dependent and independent mechanisms to protect, stabilize and restart stalled replication forks.

Causes of replication stress

Forms of replication stress that are inherent with DNA synthesis and processes involved in repairing damaged DNA include low deoxyribonucleotide levels, single strand

nicks, gaps and stretches of single stranded DNA (9). As mentioned previously, DNA is a highly reactive macromolecule that is chemically manipulated by byproducts of endogenous metabolic processes and from exogenous sources (2,3). DNA damage from endogenous sources occur at a higher frequency than from exogenous sources (3). Nucleotide oxidation and spontaneous base depurination by ROS, protein-DNA crosslinks and DNA adducts formed by metabolic aldehydes and DNA methylation induced by S-adenosylmethionine (SAM) are examples of common DNA modifications resulting from endogenous sources that cause replication stress (2,3).

RNA transcription and DNA synthesis operate simultaneously. Therefore sites where the RNA and DNA machinery collide are sources of replication stress. Early replicating fragile sites (ERFS) located at highly transcribed early replicating regions, are prone to DSB formation, potentially due to such collisions (9,10). Also, regions of the genome containing long genes (>800 kilobases) requiring more than one cell cycle to be completely transcribed experience collisions between replication and transcription complexes. These collisions create common fragile sites that are prone to breakage due to DNA:RNA hybrid (R-loop) formation (11,12).

Secondary DNA structures such as two-stranded hairpins, guanine rich G4 quadraplexes, left-hand Z form DNA, cruciform structures and three-stranded triplex DNA are direct obstacles to progressing replication forks and can cause genome instability (13). Although some DNA structures function to mark transcription start sites, failure to resolve these structures once they are encountered by a replication fork can cause replication stress and create deleterious DSBs (13).

Mechanisms promoting stability and restart of stalled replication forks

Multiple mechanisms have evolved to resolve replication stress (Figure 1.1-Figure 1.2). Ultimately, the main objective is to prevent prolonged fork stalling and promote replication fork recovery and unperturbed DNA synthesis until the entire genome is replicated. To achieve this, the replication machinery of an adjacent converging replication fork can synthesize DNA up to the location of the stall thereby ensuring that regions adjacent to the stalled fork are replicated. The replication stress response can also fire adjacent dormant origins as a means to replicate regions surrounding a stalled fork (14). Another mechanism that is used to restart stalled forks involves re-priming past the DNA block as a way to continue replicating DNA (15).

The DNA damage tolerance (DDT) pathway includes mechanisms that allow the replication machinery to tolerate and replicate through damaged DNA and replication stressors encountered by progressing forks or preserve fork stability until the damage or replication stressor is resolved. In eukaryotes, Y family (REV1, POL κ, POL η, POL ι) and B family (POL ζ) translesion polymerases function to synthesize across and past DNA lesions (16). DNA damage response mediated PCNA monoubiquitination (monoUB) on residue K164 recruits translesion polymerases to stalled replication forks (5,16). Unlike the canonical replicative polymerases δ and ε, translesion polymerases are specialized to accommodate modified DNA bases on template DNA within their active site, such as bulky base adducts and pyrimidine dimers, and incorporate bases to extend the nascent DNA strand to replicate past the damage (16). Translesion polymerases are more mutagenic and less processive than the replicative polymerases. Therefore, once the lesion is bypassed and the growing nascent strand is extended a

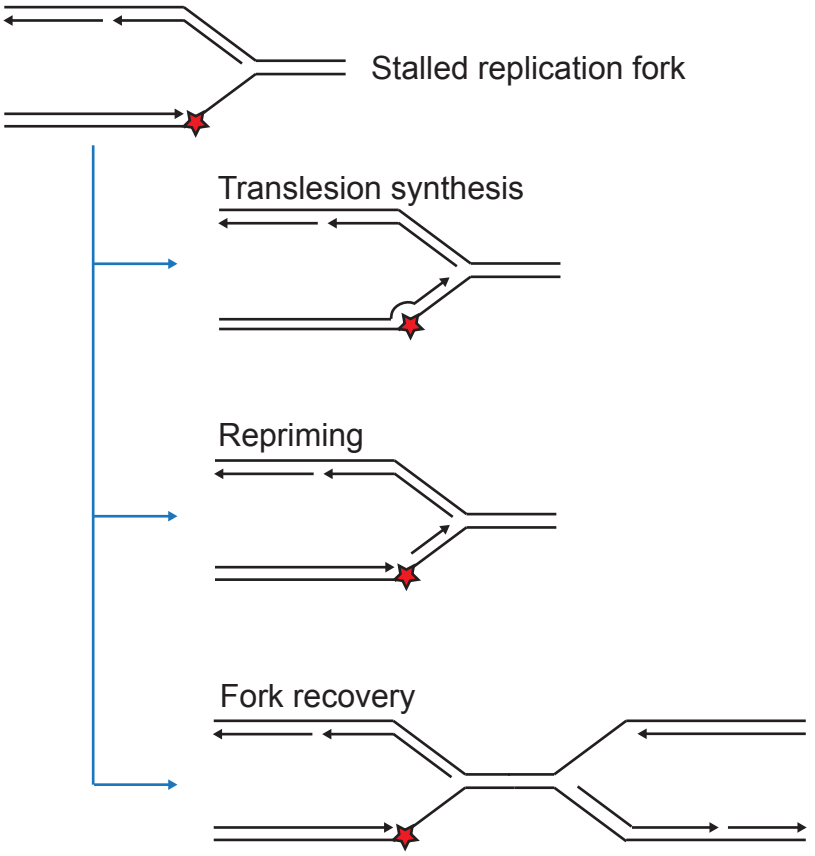


Figure 1.1 Mechanisms of restarting stalled replication forks. Depiction of translesion synthesis, reprimming and fork recovery processes to restart stalled forks.

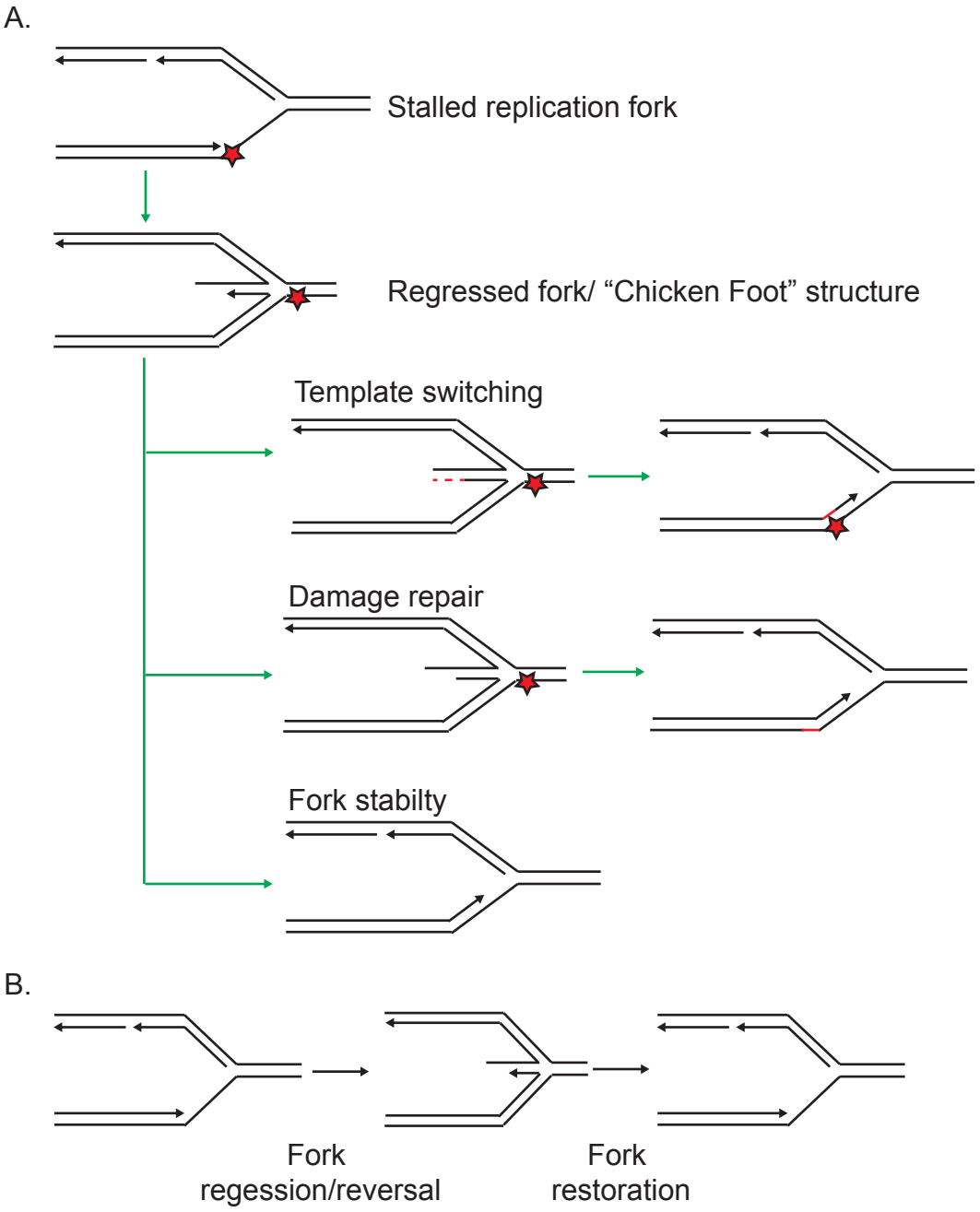


Figure 1.2 Fork remodeling mechanisms to restart stalled replication forks. (A) Mechanisms to restart stalled replicaton forks involving fork regression/reversal include template switching, DNA repair and fork stability. (B) Depiction of fork regression/reversal and fork restoration reactions catalyzed by fork remodelers *in vitro.*

couple nucleotides past the damage, the replicative polymerases take over to continue replicating the template strand (16).

The replication stress response can be artificially induced and studied by treating cells with replication stressor agents. For example, Hydroxyurea (HU) inhibits ribonucleotide reductase, which converts ribonucleotides to deoxyribonucleotides, and in turn lowers deoxyribonucleotide pools and stalls replication forks. Importantly, treating cells with HU activates ATR; therefore making HU a powerful reagent to study ATR pathway function and the replication stress response. Replication forks in cells treated for prolonged periods of time with HU (16-24 hours) accumulate DSBs due to nucleolytic processing, and undergo fork collapse (17). Collapsed replication forks are generally defined as forks that are unable to restart DNA replication. Mus81 mediated cleavage of persistently stalled forks creates intermediates to promote recombination mediated DNA synthesis to replicate and repair DNA in regions containing collapsed forks (18). However, replication forks in cells treated for shorter periods of time with HU (2-4 hours) are able to restart upon HU removal and thereby recover from the treatment (17,19). This result suggests that there are mechanisms in place allowing stalled replication forks to tolerate replication stress and restart DNA synthesis once the stall is resolved. One mechanism by which stalled replication forks undergo fork restart involves a process called fork remodeling.

Fork Remodeling

Fork remodeling is a DNA damage tolerance mechanism by which complementary parental and nascent DNA at a stalled replication fork are re-annealed or regressed into a “chicken foot” structure; once the cause of the stall is resolved, the “chicken foot” structure is then remodeled into a functional replication fork (Figure 1.2)

(19,20). A regressed or “chicken foot” structure functions as a common intermediate to multiple modes of fork restart (19,20) (Figure 1.2 A-B). Evidence for fork remodeling and the proteins involved in this process are well characterized in *E.coli* (21). *E.coli* cells contain only one origin of replication, making *E.coli* highly dependent on mechanisms that restart stalled forks and recover collapsed replication forks for its survival.

In yeast, regressed forks are hypothesized to form as a consequence of a dysfunctional S-phase checkpoint. RAD53 (CHK2 in mammalian cells) deficient budding yeast strains treated with HU were shown to accumulate regressed forks. Upon HU removal, regressed forks persisted and failed to restart (22,23). Similarly, CDS2 (CHK2 in mammalian cells) deficient fission yeast strains accumulate regressed forks that collapse when treated with HU and MMS (methyl methanesulphonate-methylates DNA bases) (24). Alternatively, regressed forks were induced in checkpoint proficient budding yeast treated with the topoisomerase II poison CPT (camptothecin) (25). Regressed forks in this context are thought to prevent the formation of double stranded DNA breaks. Therefore, in yeast, it is unclear whether fork regression is 1) a mechanism that functions to stabilize and resolve stalled forks, as in *E.coli*, or 2) as a negative consequence of a compromised checkpoint to replication stress.

Fork reversal in higher eukaryotes promotes fork stability and replication restart (20). U2OS cells (Human osteosarcoma epithelial cells) and *Xenopus laevis* egg extracts treated with 25nM CPT had increased levels of reversed forks compared to untreated U2OS cells and *Xenopus laevis* extracts (25). U2OS cells treated with 25nM CPT also had slower replication rates and undetectable DSBs. PARP (Poly-ADP-ribose polymerase 1) binds DNA nicks that form in response to topoisomerase inhibition and recruits DNA repair proteins to break sites; inhibiting PARP in U2OS cells and mouse embryonic fibroblasts significantly decreased reversed fork formation and increased

DSBs in cells treated with 25nM CPT. In this context, fork reversal is necessary to prevent DSB formation and promote fork stability (25). Furthermore, regressed forks were observed in U2OS cells and the untransformed mammalian cell line RPE (retinal pigment epithelium) in response to various replication fork stalling agents including HU, MMC (mitomycin C – interstrand crosslinking agent), aphidicolin (DNA polymerase inhibitor), H_2O_2, MMS and UV (ultraviolet radiation-dimerizes adjacent pyrimidine DNA bases) (26). These data suggest that fork reversal is a global response to many forms of DNA damage and replication stressors in mammalian cells.

Fork Remodelers

There is ample biochemical evidence on enzymes that can catalyze fork remodeling to replication fork type DNA structures *in vitro* and genetic evidence showing that these enzymes function to maintain genome stability. These enzymes are members of the SF2 (superfamily 2) family of chromatin remodelers, DNA translocases and helicases (27). Members of the SF2 superfamily of proteins share a common dual lobed ATPase domain that hydrolyzes ATP molecules to carryout most enzymatic activities (27). The ATPase domain is strongly conserved with the ATPase domain present in the *E.coli* recombinase RECA (27). For my studies, I primarily focused on the SNF2 subfamily of proteins within the SF2 superfamily class of proteins present in mammalian cells. When relevant, I make functional parallels to the well-characterized RECQ family of enzymes, and prokaryotic functional homologs RECG and UVSW (Figure 1.3).

RECQ Helicases

The RECQ-like family of helicases in mammalian cells constitutes a group of five proteins (WRN, BLM, RECQ1, RECQ4, RECQ5) that display ubiquitous functionality

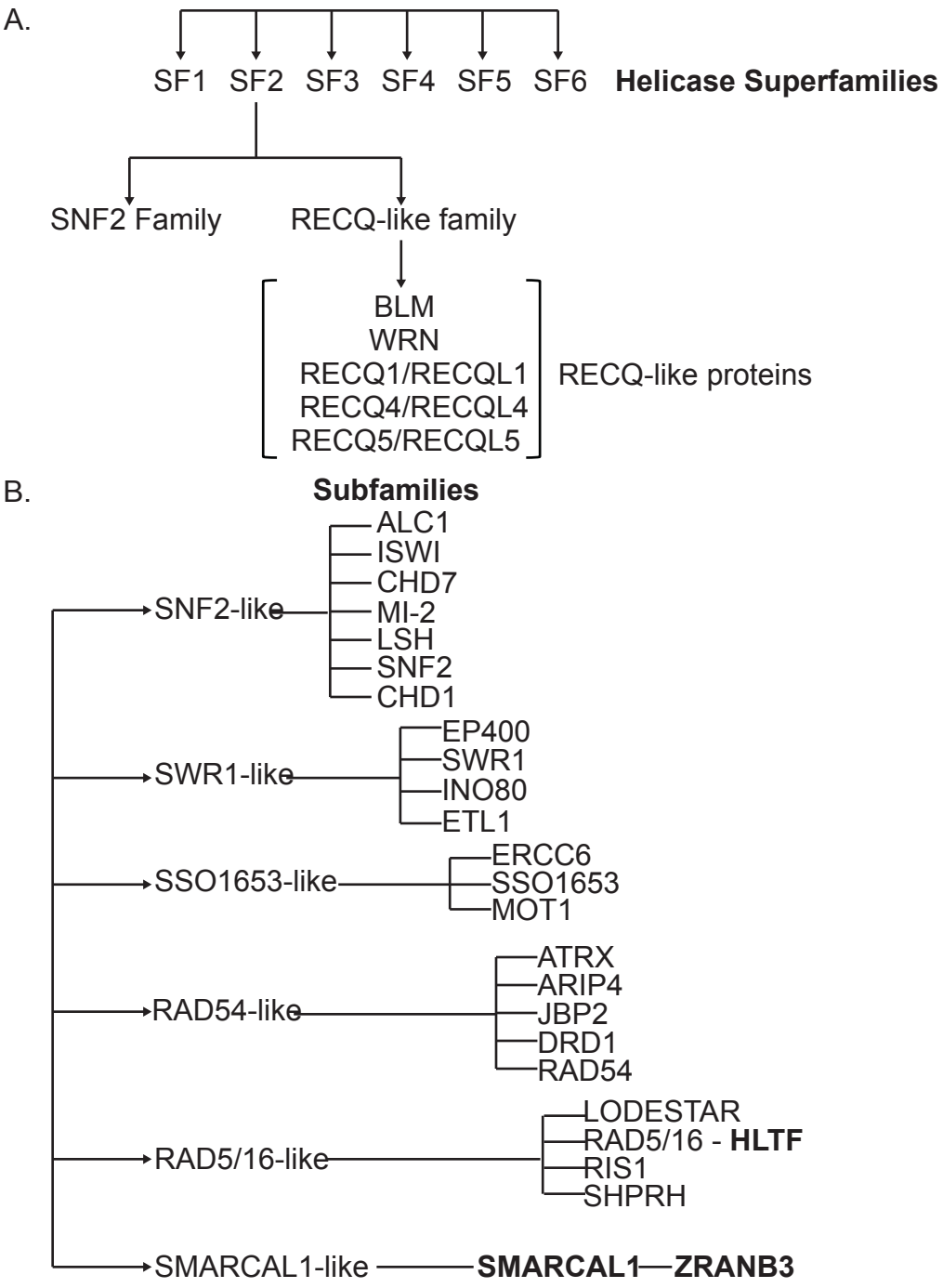


Figure 1.3 Helicase superfamilies and SNF2 subfamily classification trees. (A). Helicase superfamily clasification. Helicases are divided into six superfamilies based on sequence motif similarites of the RECA-like ATPase domain. (B) 24 subfamilies of the SNF2 family of proteins. This dissertation specifically focuses on the SMARCAL1-like subfamily and the RAD5/16 subfamily member HLTF. (A) and (B) were adapted from Singleton *et al* (27) and Flaus *et al* (37) respectively.

in replication fork stability and DNA repair (28). Collectively, RECQ helicases bind structured DNA intermediates present at replication forks and hydrolyze ATP molecules to motor their 3'-5' helicase activities (29). The enzymatic core of the RECQ helicase is the dual lobed ATPase domain and the RQC domain (RECQ C-terminal domain) that functions as a DNA binding domain and couples ATP hydrolysis to unwinding of structured DNA substrates (29). WRN is the only family member that contains a 3'-5' exonuclease domain (30). Mutations in *BLM* and *WRN* are associated with Bloom Syndrome and Werner Syndrome respectively; mutations in *RecQ4L* are linked to causing Rothmund–Thomson Syndrome (RTS), RAPADILINO Syndrome, and Baller–Gerold Syndrome (29). Although these syndromes are phenotypically diverse, most individuals develop cancers (28,29).

BLM and WRN catalyze ATP-dependent fork regression and fork restoration activities to model replication fork and regressed (Holiday junction) DNA substrates *in vitro* (Figure 1.2) (31,32). In these studies, a mutated variant inactivating the exonuclease domain in WRN was used in fork remodeling assays to prevent digestion of the DNA substrates. Subsequent work showed that WT (wild type) WRN preferentially regresses replication fork substrates containing an 11-13 nucleotide gap on the leading strand. For substrates that lacked a gap on the leading strand, WRN digested the nascent leading strand oligonucleotide to generate a gap on the leading strand before regressing the substrate (33). RECQ1 lacks fork regression activity yet restores regressed forks into replication fork structures (34). Replication forks in U2OS cells treated with 25nM CPT regress and RECQ1 was shown to facilitate fork restoration (34). RECQ5β, the longest RecQ5 isoform, catalyzes RPA stimulated fork reversal activity to replication fork substrates (35). RECQ4 is the least understood RECQ-like family member. Similar to WRN, BLM, RECQ1 and RECQ5, RECQ4 acts as an ATP-

dependent DNA helicase; there are no biochemical data determining whether RecQ4 catalyzes fork reversal or fork restoration activity to a replication fork or a regressed substrate *in vitro* (20,36).

The SNF2 Protein Family

Members of the SNF2 (sucrose nonfermenting 2) family of proteins harbor a dual lobed RECA type ATPase domain that is most conserved with the ATPase domain within the budding yeast chromatin remodeler SNF2P (37). The SNF2 family is further classified into 24 subfamilies (Figure 1.3B). Despite sharing a large ATPase domain, these subfamilies are functionally diverse suggesting that the conserved ATPase domain has been adapted or paired with accessory domains to accommodate multiple activities. For example, members of the SNF2 family of proteins include the nucleosome positioner ISWI (Imitation SWI), the RAD54 recombinase, TATA box binding protein displacement factor MOT1, the budding yeast RING finger-DNA remodeler RAD5, and the annealing helicase SMARCAL1 (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a like 1) (Figure 1.3B).

HLTF

HLTF (Helicase-like transcription factor) is a DNA-dependent ATPase, member of the RAD5 superfamily of SNF2 proteins and is a RAD5 homologue in that it also contains a RING finger ubiquitin ligase domain (38). In budding yeast, RAD5 functions as an ubiquitin ligase by promoting PCNA polyubiquitination (polyUB), a signal for an error free form of DDT called template switching. In yeast, sequential ubiquitination steps create PCNA polyUB molecules. The yeast RAD6-RAD18 E2-E3 ubiquitin conjugating and ligase complex mediates K63 linked, PCNA monoUB, which recruits factors for translesion synthesis; the RAD5-MMS2-UBC13 E2-E3 ubiquitin conjugating and ligase

complex then attaches K63 linked ubiquitin chains to the K164 monoubiquitinated residue on PCNA to achieve poly-UB PCNA molecules (38). Yeast cells treated with DNA damaging agents generate both mono and polyUB forms of PCNA; however, highly mutagenic yeast strains lacking RAD5 accumulate monoUB PCNA molecules (38). Also, efficient *in vitro* reconstitution of polyubiquitinated PCNA requires RAD5 (39). Although template switching is not as well characterized in mammals, the proteins that comprise the yeast E2 ubiquitin conjugating and E3 ubiquitin ligating complexes required to generate polyUB PCNA molecules are conserved in mammalian cells (38). HLTF and a distant RAD5 mammalian homolog SHPRH (SNF2 histone linker PHD RING helicase) are required for *in vivo* and *in vitro* PCNA polyubiquitination (40-43). Cells lacking either HLTF or SHPRH are more sensitive to DNA damage, and have elevated levels of chromosomal rearrangements and DNA breaks; these phenotypes are consistent with those observed in RAD5 null yeast strains (38,44).

Unlike SHPRH, HLTF and RAD5 are also DNA-dependent ATPase enzymes that catalyze *in vitro* fork reversal to replication fork structures (45-47). Also, HLTF and RAD5 contain a highly conserved HIRAN (HIP116 Rad5p N-terminal) domain (38). The HIRAN domain is evolutionarily conserved and present in prokaryotic and eukaryotic proteins (48). Based on intensive *in silico* analyses, the HIRAN domain was originally hypothesized to be a DNA binding domain (48). Three independent studies proved this to be true for the HIRAN domain in HLTF (49-51). The HIRAN domain is a 6-beta sheet, 2 alpha helices beta barrel that adopts an oligonucleotide binding (OB) type fold, similar to the OB DNA binding domains in RPA (49-51). Hishiki *et al*. showed that the HIRAN domain recognizes the 3' end of DNA and binds single stranded and double stranded DNA substrates (50). Kile *et al*. crystalized the HIRAN domain bound to single stranded DNA, and showed using NMR, electrophoretic mobility shift assays (EMSA) to various

DNA substrates and pull down assays that the HIRAN domain specifically recognizes and binds the 3' hydroxyl group of single stranded DNA substrates (51). Achar *et al*. crystallized the HIRAN domain with a single stranded DNA contaminant and showed using EMSA that the HIRAN domain binds single stranded DNA and replication fork substrates lacking single stranded regions; they did not test for 5' or 3' binding specificity (49).

The HIRAN domain is dispensable for HLTF DNA-dependent ATPase, translocase and ubiquitin ligase activities but is required for *in vitro* fork regression activity (49,51). Also, fiber-labeling experiments showed that the HIRAN domain is necessary for restarting stalled replication forks (51). Taken together, the HIRAN domain confers fork remodeling activity to the SNF2 ATPase domain and functions in HLTF mediated fork restart in cells. Whether HLTF is actually remodeling stalled forks in cells is unknown.

SMARCAL1

Coleman *et al*. were the first to clone the human and mouse *SMARCAL1* genes (52). SMARCAL1 is a member of the SMARCAL1-like family of SNF2 proteins (Figure 1.3B) (37,52). SMARCAL1 is highly conserved in that SMARCAL1 orthologs are present throughout metazoans including *Drosophila melanogaster*, *Xenopus laevis* and *Caenorhabditis elegans*. However no known SMARCAL1 ortholog exists in yeast (52,53). SMARCAL1 is a DNA-dependent ATPase that localizes to stalled replication forks by binding to RPA. SMARCAL1 functions to restart stalled replication forks (54-58). SMARCAL1 deficient cells are sensitive to replication stress and contain high levels of γH2AX (H2AX phosphorylated on serine 139) induction in the absence of DNA damage and replication stress; this result indicates that SMARCAL1 generally functions in

stabilizing replication forks and genome maintenance (54,58). SMARCAL1 overexpression activates the DNA damage response and phosphorylation of SMARCAL1 by ATR is thought to temper its function at stalled replication forks (59). These data suggest that SMARCAL1 function at stalled replication forks is regulated to prevent uncontrolled fork remodeling that may cause fork collapse (59). Recently, SMARCAL1 was shown to function to maintain telomere stability (60,61). The mechanism through which SMARCAL1 localizes to telomeres and how it functions to protect telomere ends is unknown.

SMARCAL1 binds forked DNA and DNA overhangs, which are substrates that are present at stalled replication forks. SMARCAL1 on its own catalyzes *in vitro* fork regression and fork restoration activities by re-annealing complementary DNA strands (53,62). SMARCAL1 binding to RPA directs its branch migration activity to specifically regress replication fork substrates mimicking a stalled leading strand and restore regressed substrates containing a longer leading nascent strand back to a normal replication fork (62). SMARCAL1 DNA binding and fork remodeling activities require its HARP2 domain, which is discussed in more detail in Chapter 3 (53,63). Briefly, the HARP2 domain acts as a DNA binding domain that dictates SMARCAL1 binding to structured DNA substrates and thereby imparts fork remodeling activity onto the ATPase domain (63).

ZRANB3

ZRANB3 (Zinc finger RAN binding protein 2-type containing 3) is a member of the SMARCAL1-like subfamily of proteins and is also a DNA-dependent ATPase enzyme (64). ZRANB3 localizes to sites of DNA damage through its PIP-box (PCNA interacting protein box), NZF (NPL4 zinc finger) and APIM (ALKB homolog 2 PCNA-

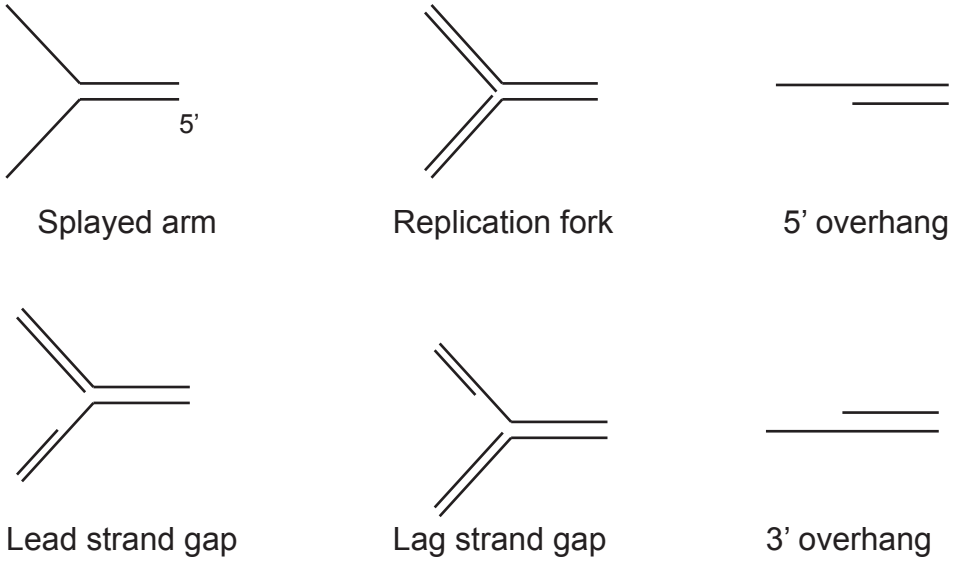
interaction motif) motifs. These motifs bind specifically to K63 linked poly-UB PCNA molecules located at stalled replication forks and DNA damage sites (65,66). ZRANB3 deficient cells display hallmark genome instability phenotypes in that cells are sensitive to replication stress and DNA damaging agents (HU, CPT, MMS) (65-67). ZRANB3 deficient cells treated with CPT and MMC have increased sister chromatid exchanges, which is indicative of unregulated recombination events between sister chromatids at stalled forks. ZRANB3 deficient cells are also defective in restarting stalled replication forks (65,67).

Similar to SMARCAL1, ZRANB3 specifically binds structured DNA *in vitro* such as forked and DNA overhang substrates (Figure 1.4) (64). Also, ZRANB3 catalyzes *in vitro* fork regression and fork restoration reactions to replication and regressed fork DNA substrates, respectively (62,65). As SMARCAL1 and unlike the RECQ helicases, ZRANB3 catalyzes branch migration reactions by re-annealing complementary strands (64).

ZRANB3 is the only protein identified in vertebrates to contain an HNH endonuclease domain (64,66). The ZRANB3 endonuclease domain cleaves structured DNA substrates in an ATP-dependent manner (66). ZRANB3 most efficiently cleaves forked DNA substrates with single stranded splayed arms (Figure 1.4-splayed fork substrate) (66). ZRANB3 endonuclease specificity is discussed in more detail in Chapter 5.

Currently, ZRANB3 is thought to function by binding to polyUB PCNA molecules assembled at stalled leading strand forks and facilitating template switching mediated fork restart by helping to remove the obstruction that caused the stall (65,66,68). This

ZRANB3 DNA Substrates



DNA substrates that ZRANB3 does not bind

Figure 1.4 ZRANB3 DNA binding specificity. Similar to SMARCAL1, ZRANB3 binds branched and DNA substrates containing a single stranded/double stranded junction. ZRANB3 does not bind double stranded or single stranded DNA substrates (Yusufzai et al) (64).

model has yet to be tested. How ZRANB3 couples its fork remodeling and endonuclease activities to help restart stalled replication forks is unknown.

RECG and UVSW

RECG and UVSW are well-characterized SF2 type- fork remodelers in *E. coli* and T4 bacteriophage, respectively, and also contain structure specific DNA binding domains (69,70). RECG, an SF2 helicase in *E.coli*, is the first enzyme identified to regress replication fork structures into Holliday junction substrates (71) . The solved crystal structure of full length RECG bound to a splayed arm substrate with an annealed nascent lagging strand oligonucleotide revealed that the ATPase domain interacts with the duplex parental strands. The “wedge” domain is positioned at the branch point of the two arms of the fork substrate and contains grooves that only accommodate single stranded DNA, and through which the two template arms pass. RECG is thought to catalyze fork regression by pulling the template strands through the grooves in the wedge domain. This movement is hypothesized to unwind the duplex parental strands and displace the nascent leading and lagging strands; the close proximity and complementarity of the nascent strands facilitate strand annealing (70).

UVSW, a SF2 helicase in T4 bacteriophage, is a RECG functional homolog in that UVSW complements the low viability phenotype of RECG null *E.coli* strains (71). UVSW catalyzes fork regression activity *in vitro* and is thought to function in recombination mediated repair of damaged DNA in T4 bacteriophage (72). Also UVSW facilitates the switch between early and late stages of T4 replication by unwinding R-loops (DNA-RNA duplexes) that prevent origin firing and DNA replication (72). UVSW contains an arginine and aromatic amino acid rich loop in the N terminal half of the protein which functions as a structure specific DNA binding domain (69).

Substrate Recognition Domains

As discussed, RECQ-like helicases, the SNF2 family of enzymes, RECG and UVSW are specialized to bind and act on structured DNA substrates. The RECQ-like helicases notably, RECQ1, BLM and WRN, require their RQC domain to bind DNA. The HIRAN domain is a DNA binding domain within HLTF; the HARP2 domain is necessary for robust SMARCAL1 DNA binding. The HARP2 domain and the SNF2 ATPase domains as a unit are sufficient in catalyzing fork regression activity, as are the HIRAN and ATPase domains of HLTF (49,53,63). The RQC, HIRAN, HARP, wedge and the N-terminal DNA binding domain in UvsW are functionally similar in that these domains dictate the specific DNA substrate that will be remodeled by the ATPase domain. Consequently, these proteins contain substrate recognition DNA binding domains that directly regulate how these enzymes function in mammalian cells (Figure 1.5).

Given the similarity shared with SMARCAL1, HLTF, and the RECQ-like family of helicases, I hypothesized that ZRANB3 contains a substrate recognition domain that binds structured DNA substrates and is necessary for its *in vitro* DNA binding, fork remodeling and endonuclease activities. This study identified a DNA binding domain within ZRANB3 that is necessary for all ZRANB3 *in vitro* enzymatic activities (Chapter 4). Given these data, the substrate recognition domain in ZRANB3 may dictate ZRANB3 function in cells.

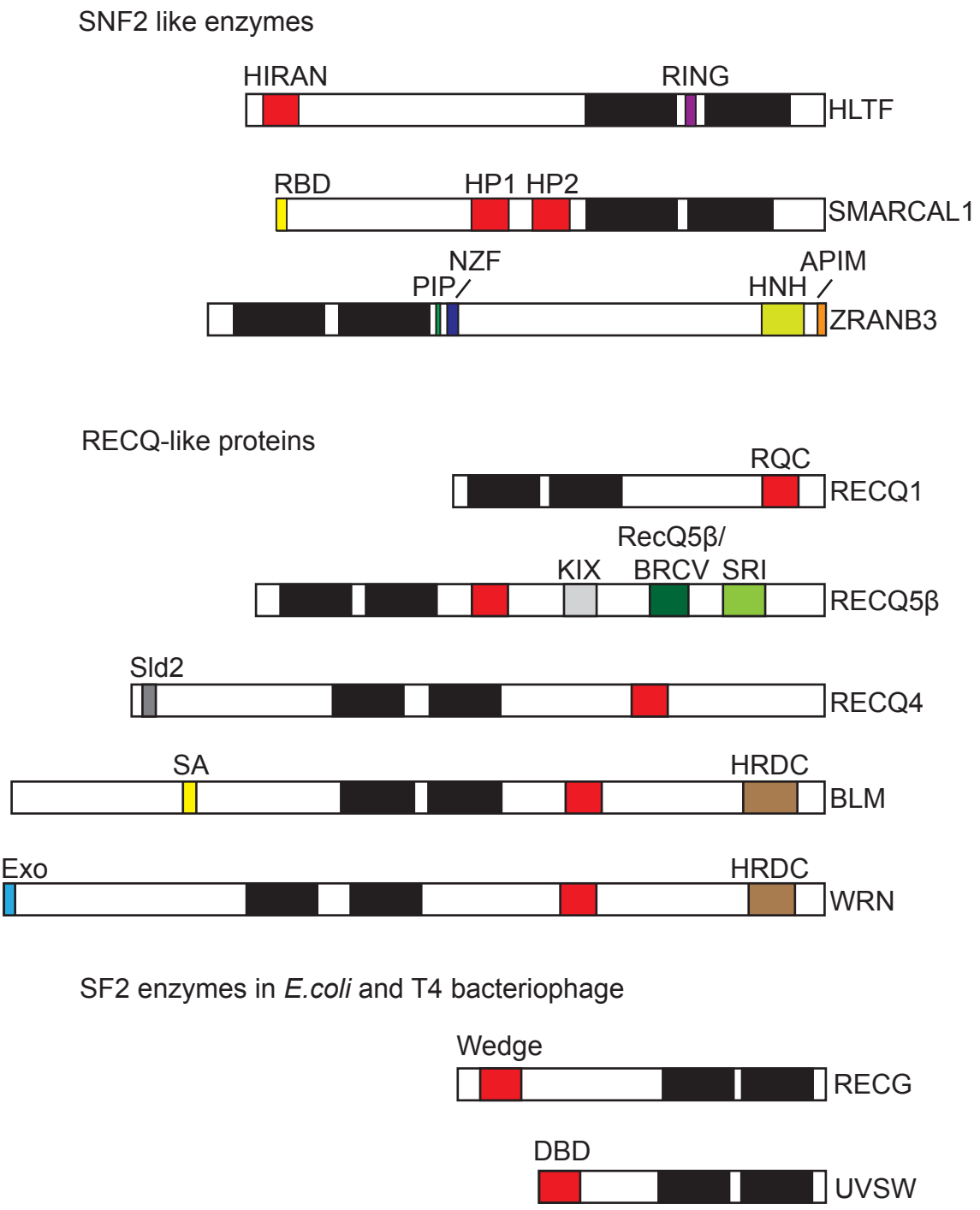


Figure 1.5 Protein architecture of the SNF2, RECQ-like and SF2 prokaryotic helicases. Predicted and confirmed DNA binding/substrate recognition domains are depicted in red. The conserved ATPase domain is depicted in black. The RECQ-like proteins depictions are adapted from Croteau *et al*.

CHAPTER II

MATERIALS AND METHODS

Cloning

pCB159 (pENTR3C – (-ATG) SMARCAL1 WT + 11x wobble) and pENTRE201 ZRANB3 from Open Biosystems (AM392702) were used to clone point, insertion, deletion and truncation mutations into SMARCAL1 and ZRANB3. Standard Gateway cloning allowed expression of WT and mutant SMARCAL1 and ZRANB3 recombinant protein using pDC1012 (pDEST-LPCX-FLAG-GW) and pRG09 (pDEST26-His-Flag-GW) expression vectors. Primers for point, deletion, insertion and truncation mutants were designed by QuickChange Primer Design software from Agilent Technologies (Table 2.1 A-B). All primers used in this study were ordered from Integrated DNA Technologies (IDT). Standard site directed mutagenesis methodology was used to generate all SMARCAL1 and ZRANB3 point, insertion and deletion mutations. The flexible linker sequences to create the minimal enzymatic unit construct were obtained from Chen *et al.* (Table 2.1B) (73). pBG101 (N-terminal GST tag, PreScission protease site) was used to express the ZRANB3 substrate recognition domain construct in bacteria cells.

Recombinant Protein Expression and Purification

Overexpression vectors for SMARCAL1 and ZRANB3 were transfected into HEK293T cells using standard PEI (polyethylenimine) transfection methodology. HEK293T cells were transfected as follows: 2.5 x 10^6 cells were plated onto a 10cm tissue culture dish 24 hours before intended transfection. For each sample, two 10cm tissue culture dishes of HEK293T cells were plated. The following day and for each dish, 4µg of expression plasmid, 100µL DMEM (Dulbecco's modified eagle medium) without FBS (Fetal bovine serum) and 24µL of 1mg/mL PEI was combined and incubated for 10-

Table 2.1A SMARCAL1 primers used to make point mutants.

SMARCAL1		
MUTATION	**PRIMER SEQUENCE (5'-3')**	**Notes/ Diagnosis**
F279S	tcctgacaccaagacgtggaactcgagcatgaatgactat	XhoI
E377Q	ccaggaagtggagctttctcctgcaggagcacagtaaactaattgc	PstI
H379P	aagtggagctttctcttggaagagccttcgaaactaattgcaaaggtgcgctgcc	BstBI
S488A	ggtggtgccatccgcggtgcgcttcacct	SacII
K555E	gatgaatctcacttcctcgagaacagtaggactgcccgc	XhoI
S859P	ctggcagaagccgggcttccggagaccaatttttcaga	BspEI & PciI
R645H	gcagtcatgctgcgtcacctcaagtccgacg	NarI

Table 2.1B ZRANB3 primers used to make point, deletion and truncation mutants. The gRNA seqeunces used to create the CRISPR U2OS ZRANB3 null cell lines cr 35 and 38 are also listed.

MUTATION	PRIMER SEQUENCE (5'-3)	Notes
ΔNZF	ctttcctgtagagggcactcctcaaggcagtg	
ΔHNH	ccaggggaaggacatttcagacaagctaaggaaaga	
ΔAPIM	gcatggatcagacatcacaaagtaagacccagctttct	
Δ712-818	aaacaccaaaaattgagaaagaagacacaaagcaacaaaccaaacaaaattg	
TAA stop codon insert	cgaacatcacacgatttttggtaaagaagtaagacccagctttcttgtacaaa	
1-501	ggggaccactttgtacaagaaagctgggtcttaactgtcatttggagtccaagcttc	
1-650	ggggaccactttgtacaagaaagctgggtcttaactgccttgaggagtctcac	
Δ651-720	gagactcctcaaggcagtaagagttcagacactttg	
Δ712-794	caccaaaaattgagaaagaagacgccatgaagcaaaggataatcag	
Δ795-859	cgagaatggagtagtctaactaggccacggga	
Amp 1-etc	ggggacaagtttgtacaaaaaagcaggcttcaccatgcctagggttcataacataaaaaag	Used for BP Gateway truncation mutants
1-859	ggggaccactttgtacaagaaagctgggtcttaggactcctttgtgatcagacg	
1-869	ggggaccactttgtacaagaaagctgggtcttaaagttttttgtgaaaggatcccgtgg	
1-879	ggggaccactttgtacaagaaagctgggtcttatagaaatgggacacaggctcc	
1-889	ggggaccactttgtacaagaaagctgggtcttaagtgagatctgcctggactgtg	
1-1008	ggggaccactttgtacaagaaagctgggtcttagtttcttatcatttcatttagctgttc	
S89A	taatagtggtcccttcggcgctgaggtacccttggac	HhaI
R377Q	gatagatggaagtgtttcatcttcagaacaaatacatctggttaatcagtttcaaaag	
K163D	tggatgaatcacactacatggactcgagaaatgcaactcgcagcag	XhoI
L760A/D761A/ I762A	aacagatgagctgtaatttcattcccgcggctgcaaaattagacctttgggaagatttac	SacII
W790A/S791A/ S792A	ctcactgattttgagatttgttcgagaagcggccgctctaactgccatgaagcaaaggat	NotI
F757A/ P759A	ctatactaaggatggaaaacagatgagctgtaatgccattgctctggatat caaattaga	EcoRV
L770A/P771A	tctggatataaaattagacctttgggaagatgcggccgcaagctttcagct gaaac	NotI
K748A/D749A/G750A	aatactgaccggatccacatctatactgcggctgcaaaacagatgagctgt aatttcatt	BamHI
$(GGGGS)_3$ - Top	gcgccggaggaggaggatcaggaggaggaggatcag gaggaggaggatcagcc	
$(GGGGS)_3$ - Bottom	ggctgatcctcctcctcctgatcctcctcctcctgatcc tcctcctccg	
720-869 cloning into GST expression vector pBG101	For (gagaggaattcgtggaagagttcagacactttgc)	Ligated using 5' EcoRI and 3' XhoI sites
	Rev (gagagctcgagttaaagtttttttgtgaaaggatcccgtg)	
gRNAs	**gRNAs used to generate ZRANB3 CRISPR U2OS cell lines (PAM sequence is in bold type)**	**Notes**
35 and 38 clones	(1) agctttgctcttagtctgtc**agg (**pTB08)	
	(2) ttttttatgttatgaaccct**agg (**pTB09)	

15 minutes at room temperature. During incubation, the old media from the HEK293T cells was replaced with 10mL of fresh media. After incubation, the combined expression plasmid, PEI and DMEM without FBS were added dropwise to cells while swirling the plate. The cells were incubated in transfection media for 24 hours. After incubation, the transfection media was removed, cells were trypsonized and pooled into a 15cm tissue culture dish. After 24 hours, the cells were harvested by scraping using a cell lifter tool, pooled into a 15mL conical tube and washed twice with 4mL of 1x PBS (Phosphate Buffer Saline) (cells were pelleted between washes by centrifuging for 1 min at 1000 RPMs). Cell pellets were frozen with liquid nitrogen and stored at -80°C until further use.

Recombinant SMARCAL1 protein was purified as follows: Frozen pellet was lysed in lysis buffer (20mM Tris pH 8.0, 150mM NaCl, 1mM EDTA, 0.5% IGEPAL, 5 µg/mL leupeptin, 5 µg/mL aprotinin, 0.2mM PMSF, 1:100 sodium vanadate; 2.5mL lysis buffer/15cm dish) on ice for 30 minutes. FLAG-SMARCAL1 from cleared cell lysate (cells were centrifuged for 30 minutes in a chilled 4°C centrifuge at 4000 RPM) was immunoprecipitated with 15µL of packed and washed FLAG bead slurry (F2426, EZ View Red Anti-FLAG M2 Affinity gel; 30 µL of bead slurry was washed once with lysis buffer to remove storage solution) for 4 hours rotating at 4°C. Beads were washed three times with 3mL lysis buffer, twice with 3mL LiCl buffer (10mM HEPES pH 7.9, 0.3M LiCl, 20% glycerol, 0.2mM EDTA, 0.1% Triton X-100, 1mM DTT, 0.2mM PMSF, 1.5mM $MgCl_2$, 5µg/mL aprotinin, 5µg/mL leupeptin) and twice with the elution buffer (20mM HEPES pH 7.6, 20% glycerol, 0.1M KCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 1mM DTT, 0.2mM PMSF, 0.01% IGEPAL). All lysis and elution buffer washes were conducted as follows: Beads were rotated for 3 minutes at 4°C and centrifuged at 4000RPM for 3 minutes to remove buffer. For the LiCl buffer washes, the tubes were inverted 4-6 times with 3mL LiCl buffer, centrifuged at max speed and the buffer was immediately aspirated

from the beads. Bound protein was eluted off FLAG beads by incubating beads with 80µL of elution buffer containing 300µg/mL FLAG peptide (F3290 Sigma), for 1.5 hours at 4°C or on ice. Beads were agitated every 10 minutes. To determine protein concentration, 10µL of eluted protein was resolved on a 10% SDS PAGE gel using known BSA standards (range from 1mg/mL – 0.0625mg/mL). The gel was then stained with SimplyBlue™ SafeStain (novex by life technologies catalog # LC6065) and analyzed for protein abundance using the Odyssey gel analysis system. BSA standards were diluted in the corresponding lysis buffer.

HEK293T cells were transfected with ZRANB3 expression vectors, harvested and stored using the same protocol as previously stated for transfecting and expressing SMARCAL1 constructs. Recombinant ZRANB3 protein from HEK293T cells was purified as follows: Frozen pellet was lysed in lysis buffer for 40 minutes on ice (20mM Tris pH 7.5, 150mM NaCl, 1mM DTT, 0.2mM PMSF, 5µg/mL aprotinin, 5µg/mL leupeptin, 0.1% Triton X-100, 0.1mM EDTA; 2.5mL/15cm dish). Cleared cell lysate (cells were centrifuged for 30 minutes in a chilled 4°C centrifuge at 4000 RPM) was incubated with FLAG-M2 beads (Sigma F2426, EZ View Red Anti-FLAG M2 Affinity gel) for 4 hours at 4°C. Beads were washed twice with 3mL lysis buffer, once with 3mL LiCl buffer (10mM HEPES pH 7.9, 0.3M LiCl, 20% glycerol, 0.2mM EDTA, 0.1% Triton X-100, 1mM DTT, 0.2mM PMSF, 1.5mM $MgCl_2$, 5µg/mL aprotinin, 5µg/mL leupeptin), and twice with 3mL elution buffer (20mM HEPES pH 7.9, 0.1M KCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 20% glycerol, 0.01% IGEPAL, 1mM DTT, 0.2mM PMSF, 5µg/mL aprotinin, 5µg/mL leupeptin). Washes were conducted as previously stated in transfecting and expressing SMARCAL1 constructs. Bound protein was eluted off FLAG beads by incubating beads with 80µL of elution buffer containing 300µg/mL FLAG peptide (F3290 Sigma), for 1.5 hours at 4°C or on ice. Protein concentrations were determined as previously described.

Dr. Miaw-Sheue Tsai at the Lawrence Berkeley National laboratory supplied the recombinant FLAG-ZRANB3 insect cell pellets. Recombinant FLAG-ZRANB3 protein from insect cells was purified as follows: A 0.4L insect cell pellet was thawed on ice for 45 minutes, washed twice with 10mL of 1x PBS, and resuspended in 10mL of lysis buffer (20mM Hepes pH 7.6, 0.5M KCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 20% Glycerol, 0.01% IGEPAL, 1mM DTT, 0.2mM PMSF, 1µg/mL leupeptin, 1µg/mL aprotinin, 0.5mM Benzamidine). Cells were homogenized 40 times with a dounce homogenizer on ice for 30 minutes (1 stroke/ 45 seconds). 50µL of packed, washed beads (washed 100µL of EZ View Red Anti-FLAG M2 affinity gel slurry with lysis buffer) was incubated with cleared cell lysate (the lysate was transferred to a 30mL Sorvall centrifuge compatible tube and centrifuged for 40 minutes at 10,000g) for 4 hrs at 4°C, rotating, to immunoprecipitate FLAG-ZRANB3. FLAG beads were then washed twice with lysis buffer and then twice with the elution buffer (20mM HEPES pH 7.9, 0.1M KCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 20% glycerol, 0.01% IGEPAL, 1mM DTT, 0.2mM PMSF, 5µg/mL aprotinin, 5µg/mL leupeptin). FLAG-ZRANB3 bound beads were incubated with 100µL elution buffer containing 200µg/mL FLAG peptide (F3290 Sigma) and rotated for 30 minutes at 4°C. The elution step was repeated twice. Protein concentration was determined using BSA standards as previously described.

GST-720-869 and GST-720-869 L760A/D761A/I762A constructs were expressed and purified from ArcticExpress *E.coli* cells as follows: ArcticExpress *E.coli* cells were grown at 37°C and upon reaching an $O.D_{600}$, protein expression was induced with 1mM IPTG and cells were grown at 16°C overnight. The cell pellet was solubilized in lysis buffer (25mM Tris pH 8.0, 50mM NaCl, 0.1mM EDTA, 5% glycerol, 1mM DTT, 0.1mM PMSF, 5µg/mL aprotinin, 5µg/mL leupeptin) and lysed by sonication. Triton X-100 was added to reach a final concentration of 1% and the lysate was incubated on ice for 30

minutes. Following high-speed centrifugation, the lysate was incubated with 200µL of washed GST beads for 4 hours at 4°C (prior to immunoprecipitation, the beads were washed twice with 1ml of lysis buffer; for washes, beads were centrifuged at 1000 RPM for 1 minute). Afterwards, the beads were washed three times with lysis buffer containing 1% Triton X-100. Protein was eluted with 200µL elution buffer (75mM Tris pH 8, 15mM glutathione, 0.1mg/mL leupeptin; equal volumes of elution buffer to bead volume), and dialyzed overnight at 4°C (dialysis buffer: 20mM Tris pH 8.0, 1mM DTT, 0.1mM EDTA, 20% glycerol, 5µg/mL leupeptin, 5µg/mL aprotinin, 0.2mM PMSF). Dialyzed samples were applied to a heparin column and eluted with increasing concentrations of KCl (50mM, 75mM, 150mM, 300mM). Fractions containing the desired protein were combined and concentrated using a Millipore 10,000 MWCO protein concentration filter.

Immunoblotting

Table 2.2 describes the use of antibodies for western blotting.

5' ^{32}P labeling and annealing oligonucleotides to make splayed arm (forked DNA), double stranded and single stranded DNA substrates

The following reaction was used to label the 5' end of oligonucleotides with γ-^{32}P-ATP: 3µL γ-^{32}P - ATP (BLU502Z250UC EasyTides® Adenosine 5' triphosphate), 1µL of oligonucleotide at 10µM, 2µL of 10X NEB T4 Polynucleotide Kinase buffer, 1µL NEB T4 Polynucleotide Kinase, and 13µL DNase/RNase free water were incubated in a 37°C water bath for 2 hours. Labeled oligonucleotide was purified using a GE G-25 column (GE Healthcare catalog# 27-5325-01) (vortex column right side up and upside down to properly mix resin, centrifuge column for 1 minute at 2.8 RPM to remove

Table 2.2: Antibodies used for immunoblotting

NAME	DESCRIPTION OF USE
FLAG-M2 FLAG ANTIBODY (F1804 Sigma) species: mouse	1:500, 5% Milk, 2 hours 4°C
SMARCAL1 909 (Open Biosystems Custom) species: rabbit	1:1000, 5% Milk, 2 hours 4°C
ZRANB3 BETHYL (A303-033A) species: rabbit	1:500, 5% Milk, 2 hours 4°C
GAPDH (Millipore/Chemicaon MAB374) species: mouse	1:10,000, 5% Milk, 2 hours 4°C

storage buffer, add labeled oligonucleotide reaction to resin, centrifuge for 2 minutes at 2.8 RPM and collect in fresh eppendorf tube). Final concentration of the labeled oligonucleotide is approximately 0.5μM.

The following procedure was used to anneal oligonucleotides to make the labeled splayed arm and double stranded DNA substrate: In a PCR tube combine 4μL of 5' ^{32}P labeled oligonucleotide, 2μL unlabeled oligonucleotide at 2μM, 2μL of 10X SSC buffer (10X SSC: 1.5M NaCl, 150mM Sodium Citrate pH 7.0) and 12μL DNase/RNase free water. Run the "RBanneal" program on a PCR machine to anneal oligonucleotides together. The final concentration of the splayed arm or double stranded DNA substrate is approximately 100nM. The substrates were stored at -80°C until further use.

RBanneal Program

- (1) 95.0°C for 30 seconds.
- (2) 95.0°C for 40 seconds (-3°C after every cycle)
- (3) 97.0°C for 40 seconds (-3°C after every cycle)
- (4) 96.0°C for 40 seconds (-3°C after every cycle)
- (6) Repeat steps (2)-(4) 24 additional times
- (5) 4.0°C hold

ATPase Assay

Purified FLAG-SMARCAL1 and FLAG-ZRANB3 proteins were used at a final concentration of 5nM for each reaction. Components for a typical reaction are as follows: 0.24μL γ- ^{32}P-ATP (BLU502Z250UC EasyTides ® Adenosine 5' triphosphate) 1μL of unlabeled ATP at 1μM, 1μL unlabeled splayed arm substrate stock at 12.5nM and 50nM or 20nM and 80nM (use 1μL of ATPase buffer for 0nM DNA substrate sample), 1μL of FLAG-SMARCAL1 or FLAG-ZRANB3 at 50nM and 6.76μL of ATPase Buffer (20mM

HEPES pH 7.6, 0.1M KCl, 5mM $MgCl_2$, 3% Glycerol, 0.25mg/mL BSA, 0.05mM DTT, 0.01% IGEPAL). The reaction was incubated in a 37°C water bath for 30 minutes. To stop the reaction, eppendorf tubes were placed on ice. 1µL of each reaction was spotted on a Cellulose PEI plate (Selecto Scientific, Flexible TLC Plates, catalog no: 10078, 20x20cm 100 Micron Cellulose PEI) and the plate was air dried for 5 minutes. To conduct thin layer chromatography, the plate was placed in a glass chamber containing the ATPase assay thin layer chromatography buffer (1M Formic Acid, 0.5M LiCl; to allow buffer to equilibrate within the chamber pour buffer into glass chamber multiple hours before incubating the plate,) and incubated for 45 minutes. After allowing the plate to dry for a minimum of 2 hours, it was exposed to a phosphorimager screen for 3-5 minutes and resolved using a Biorad phosphorimager scanner. "ATPase Activity (%)" was determined by calculating the percent of hydrolyzed ATP from total ATP (100* (hydrolyzed/ (hydrolyzed + un-hydrolyzed)).

To make the splayed arm substrate used in ATPase assay combine 3µL of the HJ1 oligonucleotide at 100µM, 3µL of the HJ4 oligonucleotide at 100µM, 2µL of 0.5M NaCl, and 12µL DNase/RNase free water together in a PCR tube and use a PCR machine to run the "RBanneal" program to anneal oligonucleotides together (Table 2.4).

DNA Binding Assay

FLAG-SMARCAL1 and FLAG-ZRANB3 DNA binding assays were achieved by incubating increasing amounts of purified protein with 10nM of a 5' ^{32}P labeled DNA substrate. A typical DNA binding reaction was conducted as follows: 2µL of protein at 0nM, 25nM, 50nM, 100nM (final protein concentration in reaction is 2.5nM, 5nM, 10nM; use 2µL of binding buffer for 0nM sample) and 2µL of the DNA substrate at 100nM was incubated in 16µL of the DNA binding buffer (20mM HEPES pH 7.9, 0.01% IGEPAL,

0.1M KCl, 5mM $MgCl_2$, 1% glycerol, 0.25 mg/mL BSA, 1mM DTT, 0.05M EDTA) for 30 minutes at room temperature. Following incubation, 15% Ficol (PM400 Sigma catalog# F4375) was added to the reaction to achieve a final concentration of 2.5%. Samples were resolved on a 5% polyacrylamide gel (1.25mL 40% 37.5:1 Acrylamide/Bis solution, 0.5mL 10X TBE, 8.25mL water, 200µL 10% APS, 20µL TEMED) in 1x TBE running buffer (10X TBE: 1M Tris, 0.9M Boric Acid, 20mM EDTA) at 40V for 180 minutes at 4°C. Immediately after making the gel and prior to loading the reaction, the gel was pre-ran at 60V for 30 minutes with 1X TBE.

After resolving the samples on the gel, the gel was dried overnight using cellophane, exposed for 2 hours to a phosphorimager screen, scanned and analyzed. "DNA binding %" was determined by calculating the percent of labeled DNA substrate bound to protein (100*(bound/(bound + unbound)).

Making the fork regression substrate (lead gap regression substrate)

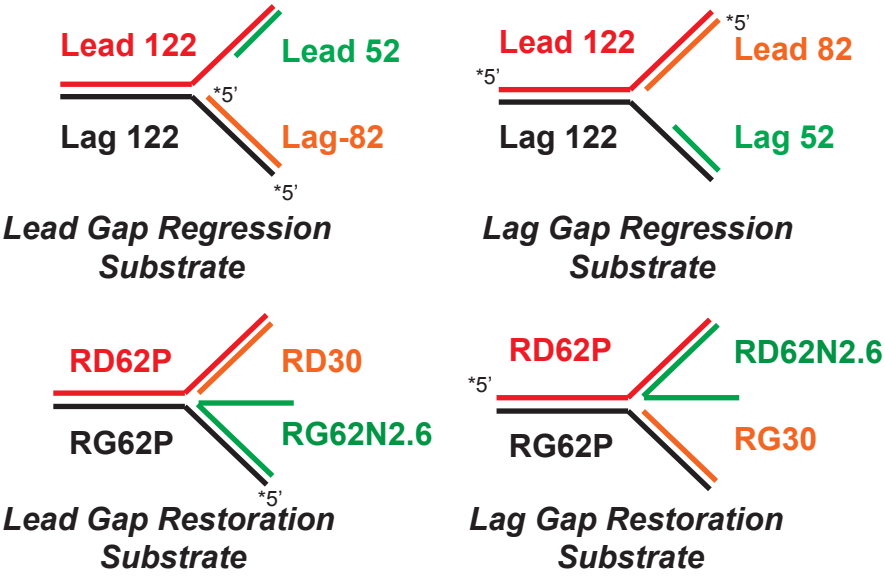
Although this procedure is briefly describe in Bétous *et al*. 2013, some changes have been made to decrease the preparation time without compromising the quality of the purified regression substrate (62). The following procedure describes the methodology to make the lead gap regression substrate (Table 2.3):

Step 1: Label

The four oligonucleotides that are used to make the lead gap regression substrate are lead 122, lead 52, lag 122 and lag 82 (Table 2.3). 5' label lag 122 and lag 82 each with γ-^{32}P-ATP by combining 10µL of a 10µM stock of lag 122 and Lag 82 with 4µL of 10X polynucleotide kinase buffer, 18µL of DNase/RNase free water (heat oligonucleotide, polynucleotide buffer and water together at 95°C for 2 minutes and place on ice to cool for 1 minute; this step removes DNA secondary structures), 2µL of

Table 2.3. Oligonucleotides used to assemble regression and restoration DNA substrates. Bolded and underlined sequences indicate mismatches

NAME OF OLIGOS FOR REGRESSION SUBSTRATES	SEQUENCE (5'-3')	LENGTH
LEAD 122	CGTGACTTGATGTTAACCCTAACCCTAAGATATCGCGT**T** **A**TCAGAGTGTGAGGATACATGTAGGCAATTGCCACGTGT CTATCAGCTGAAGTTGTTCGCGACGTGCGATCGTCGCT GCGACG	122
LAG 122	CGTCGCAGCGACGATCGCACGTCGCGAACAACTTCAGC TGATAGACACGTGGCAATTGCCTACATGTATCCTCACAC TCTGA**AT**ACGCGATATCTTAGGGTTAGGGTTAACATCAA GTCACG	122
LEAD 82	CGTCGCAGCGACGATCGCACGTCGCGAACAACTTCAGC TGATAGACACGTGGCAATTGCCTACATGTATCCTCACAC TCTGA	82
LAG 82	TCAGAGTGTGAGGATACATGTAGGCAATTGCCACGTGTC TATCAGCTGAAGTTGTTCGCGACGTGCGATCGTCGCTG CGACG	82
LEAD 52	CGTCGCAGCGACGATCGCACGTCGCGAACAACTTCAGC TGATAGACACGTGG	52
LAG 52	CCACGTGTCTATCAGCTGAAGTTGTTCGCGACGTGCGAT CGTCGCTGCGACG	52
NAME OF OLIGOS FOR RESTORATION SUBSTRATES	**SEQUENCE (5'-3')**	**LENGTH**
RD62P	TAGGCA**A**TTGCC**A**CGTGT**C**TATCA**G**CTGAA**GT**ACAAGCG CTGCACCCTAGGTCCGACGCTGC	62
RD62N2.6	GCAGCGTCGGACCTAGGGTGCAGCGCTTGT**TG**TTCAG**G** TGATA**C**ACACG**C**GGCAA**A**TGCCTA	62
RG62P	CGTCGCAGCCTGGATCCCACGTCGCGAACA**AC**TTCAG**C** TGATA**G**ACACG**T**GGCAA**T**TGCCTA	62
RG30	TGTTCGCGACGTGGGATCCAGGCTGCGACG	30
RD30	GCAGCGTCGGACCTAGGGTGCAGCGCTTGT	30
RG62N2.6	TAGGCA**T**TTGCC**G**CGTGT**G**TATCA**C**CTGAA**CA**TGTTCGC GACGTGGGATCCAGGCTGCGACG	62



T4 polynucleotide kinase, 6μL of γ-^{32}P-ATP and incubating for 2 hours in a 37°C water bath. Purify labeled lag 122 and lag-82 with a GE-25 column (method stated in "*5' ^{32}P labeling and annealing oligonucleotides to make splayed arm, double stranded and single stranded DNA substrates"* section). Proceed to "Step 2: Anneal".

Step 2: Anneal

Anneal the unlabeled leading half of the leading strand template as follows: combine 14μL of lead 122 at 10μM, 14μL of lead 52 at 10μM, 4μL of 10X SSC, 8μL DNase/RNase free water and split reaction into 2 PCR tubes, each containing 20μL. Run the "RBanneal" program to anneal oligonucleotides together.

Anneal the labeled lagging half of the lead gap regression substrate as follows: From "Step 1: Label" combine 25μL of labeled lag 122, 37.5μL of labeled lag 82, 8μL 10X SSC buffer, 9.5μL DNase/RNase free water together, split into 4 tubes, each with 20μL and run the "RBanneal" program. After assembling unlabeled leading half and labeled lagging half, proceed to "Step 3: Assemble".

To make the parental strand products of the regression reaction combine 5μL of labeled lag 122 (~2.5μM) with 7.5μL of unlabeled lead 122 from a 2.5μM stock, 2μL 10X SSC, and 5.5μL DNase/RNase free water into a PCR tube and run the "RBanneal" program to anneal oligonucleotides together. To make the nascent strand product of the regression reaction combine 5μL of labeled lag 82 (~2.5μM) with 7.5μL of unlabeled lead 52, 2μL 10X SSC and 5.5μL DNase/RNase free water into a PCR tube and run the "RBanneal" program to anneal oligonucleotides together. The parental strand and nascent strand products are approximately at ~0.625μM. The annealed parental strand product will be used in a later step to determine the concentration of the gel purified lead gap regression substrate.

Step 3: Assemble

Assemble the lead gap regression substrate as follows: Combine 78μL of the labeled lagging half with 39μL of unlabeled leading half and 13μL of 10X annealing buffer (10X annealing buffer: 400mM Tris pH 7.5, 50mM $MgCl_2$, 20mM DTT, 1mg/mL BSA, 200mM KCl). Place eppendorf tube in 37°C water bath, switch off water bath and incubate for 30 minutes or until water temperature reaches 30°C. Store at -80°C until further use.

Step 4: Purification

Make a 5% polyacrylamide gel (1.25mL 37.5:1 40% Acrylamide/Bis solution, 0.25mL 10X TBE, 8.5 mL water, 200μL 10% APS, 20μL TEMED) and pre-run for 30 minutes at 60V in 1X TBE. During the pre-run, thaw the assembled lead gap regression substrate (if stored at -80°C) and add 12μL of 3X stop dye (for 1mL 3X stop dye combine 200μL of 6X DNA loading dye, 655μL of 50% glycerol, 100μL of 0.5M EDTA and 45μL of 20% SDS). After pre-run, wash out wells, load 20μL of assembled substrate into multiple lanes and run gel for 60 minutes at 100V. Make sure to also run a small amount of the parental strand product to use as a size marker. After running the gel, wrap the gel in saran wrap or an antibody bag. Expose the gel to film for 2 minutes, resolve film, and use as a template to determine the portion of the gel that corresponds to the assembled lead gap regression substrate. Cut out portion of gel containing lead gap regression substrate with a new razor, and place gel piece into a clipped 3,500 Kd MWCO snakeskin bag, add 2 mL of 0.25X TBE and clip the top. Fill a designated Ethidium Bromide free agarose gel electrophoresis chamber with 0.25X TBE, submerge snakeskin bag containing gel piece in chamber on the side designated as the anode and electrophorese for 80 minutes at 80V. Transfer buffer in snakeskin bag into a fresh

eppendorf tube (prevent transferring any gel pieces and if this happens, filter the liquid through FACS filter paper). Concentrate sample to ~150μL using a 0.5mL Amicon Ultra 4, 10,000 MWCO column (follow directions in pamphlet for details on how to use this column).

Step 5: Concentration Determination

To determine the concentration of the purified substrate, use the parental strand product (~0.625μM) as the standard. Make an 8% polyacrylamide gel (1mL 19:1 40% Acrylamide/Bis solution, 1mL 10X TBE, 7mL water, 200μL 10% APS, 20μL TEMED) and pre-run for 60V at 30 minutes in 1X TBE. During pre-run, make serial dilutions of the parental strand product to achieve 0.312μM, 0.156μM and 0.078μM stocks. Add 2μL of each stock to 8μL of DNase/RNase free water and 5μL of 3X stop dye. Also add 2μL of purified substrate to 8μL of DNase/RNase free water and 5μL of 3X stop dye. After pre-run load all 15μL of standards and purified substrate onto the gel and resolve for 60 minutes at 80V. To analyze gel, wrap the wet gel in saran wrap or place in a plastic antibody bag, expose gel to phosphorimager screen for 1 hour (minimum) and scan using the BioRad phosphorimager.

To create the lag gap regression substrate, follow the same procedure as making the lead gap substrate, but in step 1, label the lead 122 and lead 82 oligonucleotides instead. Anneal labeled lead 122 and labeled lead 82 together to create the labeled leading half, and anneal unlabeled lag 122 and unlabeled lag 52 together to create unlabeled lagging half. Make the parental strand product by annealing labeled lead 122 with unlabeled lag 122 and the nascent strand product by annealing labeled lead 82 with the unlabeled lag 52. Use the parental strand product as the standard to determine the final concentration of the lag gap regression substrate. Maintain the volume and stock

concentrations of all the oligonucleotides used to make the lead gap regression substrate.

Making the fork restoration substrate (lag gap restoration substrate)

Although this procedure is briefly describe in Bétous *et al*. 2013, some changes have been made to decrease the preparation time without compromising the quality of the purified restoration substrate (62). The following procedure describes the methodology to make the lag gap restoration substrate (Table 2.3):

Step 1: Label

The four oligonucleotides that are used to assemble the lag gap restoration substrate are RD62P, RD62N2.6, RG62P, RG30. 5' label RD62P with γ- ^{32}P-ATP by combining 4μL of 50μM RD62P, 4μL of 10X Polynucleotide kinase buffer, 18μL of DNase/RNase free water (heat buffer, water and RD62P at 95°C for 2 minutes and incubate on ice for 1 minute; this step removes DNA secondary structure) 10μL γ-^{32}P-ATP, and 4μL of Polynucleotide Kinase and incubating at 37°C for 2 hours in a water bath. Purify labeled RD62P with a GE-25 column (method stated in "*5' ^{32}P labeling and annealing oligonucleotides to make splayed arm, double stranded and single stranded DNA substrates"* section). The concentration of labeled RD62P is ~5μM. Proceed to "Step 2: Anneal and Assemble".

Step 2: Anneal and Assemble

From step 1, combine 35μL of labeled RD62P, 5μL unlabeled RG62P at 50μM, 5μL of 10X SSC, and 5μL of DNase/RNase free water into two PCR tubes with 25μL in each tube (RBanneal step 1). Simultaneously, to make the products of the restoration reaction (which will also be used to determine the concentration of the purified

substrate), combine 3.5μL of labeled RD62P, 2μL of 10X SSC, 5μL of unlabeled RD62N2.6 at 50μM, and 9.5μL of DNase/RNase free water into a PCR tube. Run the "RBanneal" program to anneal oligonucleotides together. The concentration of the product is ~0.875μM.

Take 47μL from "RBanneal step 1" and combine with 5μL of RG30 at 100μM, 1μL of 10X SSC and 6μL of DNase/RNase free water and split into 3 PCR tubes with 20μL in each tube. Run the "RBanneal" program to anneal oligonucleotides together (RBanneal step 2). Take 58μL of "RBanneal step 2" and combine with 8μL of unlabeled RD62N2.6 at 50μM, 8μL of 10X annealing buffer (same buffer as described in making regression substrate) into an eppendorf tube and place tube in 37°C water bath, switch off water bath and incubate for 30 minutes or until water bath temperature reaches 30°C. Store at -80°C until further use.

Step 3: Purification

Follow the same procedure as outlined in the "Step 4: Purification" section to make the lead gap regression substrate.

Step 4: Concentration determination

Follow the same procedure as outlined in the "Step 5- Concentration determination" section to make the lead gap regression substrate. Using 0.875μM stock of reaction product, make serial dilutions to achieve 0.437μM, 0.218μM, 0.109μM stocks to use as standards.

To make the lead gap restoration substrate, label RG62P. For "RBanneal step 1", use RD62P as the unlabeled oligonucleotide. For "RBanneal step 2", use unlabeled RD30. For the final annealing step, use unlabeled RG62N.6. To make the products of a

lead gap restoration reaction anneal labeled RG62P with unlabeled RG62N2.6. Maintain the volume and stock concentrations of all of the oligonucleotides used to make the lag gap restoration substrate.

Fork Regression Assay

A standard regression assay was conducted as follows: Combine 2μL of 30nM protein (3nM final), 2μL of 30nM regression substrate (3nM final), 16μL of reaction buffer (40mM Tris pH 7.5, 100mM KCl, 5mM $MgCl_2$, 100μg/mL BSA, 2mM ATP unlabeled) and incubate in a 37°C water bath for 30 minutes. Resolve reaction products on an 8% polyacrylamide gel (2mL 19:1 40% Acrylamide/Bis solution, 1mL 10X TBE, 7mL water, 200μL 10% APS, 20μL TEMED; always pre-run gel for 30 minutes at 60V before resolving products) in 1x TBE at 80V for 80 minutes at room temperature. To analyze gel, wrap wet gel in saran wrap or place in a plastic antibody bag, expose to phosphorimager screen for 1 hour and scan screen using a Biorad phosphorimager.

Fork Restoration Assay

A standard restoration assay was conducted as follows: Combine 2μL of 30nM protein (3nM final), 2μL of 30nM restoration substrate (3nM final), 16μL of reaction buffer (40mM Tris pH 7.5, 100mM KCl, 5mM $MgCl_2$, 100μg/mL BSA, 2mM unlabeled ATP) and incubate in a 37°C water bath for 30 minutes. Resolve reaction products on an 8% polyacrylamide gel (2mL 19:1 40% Acrylamide/Bis solution, 1mL 10X TBE, 7mL water, 200μL 10% APS, 20μL TEMED; always pre-run gel for 30 minutes at 60V in 1X TBE before resolving products) in 1x TBE at 80V for 80 minutes at room temperature. Wrap wet gel in saran wrap or place in a plastic antibody bag, expose to phosphorimager screen for 1 hour and scan screen using Biorad phosphorimager.

Making the replication fork substrate used in DNA binding and nuclease assays

The following procedure describes the methodology to make the replication fork substrate (Table 2.4):

Step 1: Label

The four oligonucleotides that are used to assemble the replication fork substrate are 48-Remy, 49-Remy, 50-Remy, 51-Remy. Label the 5' end of 48-Remy with γ-^{32}P-ATP by combining 4μL of 50μM oligonucleotide, 4μL of 10X polynucleotide kinase buffer, 18μL DNase/RNase free water (heat buffer, water and 48-Remy at 95°C for 2 minutes and incubate on ice for 1 minute; this step removes DNA secondary structures), 10μL γ-^{32}P-ATP, 4μL of polynucleotide kinase and incubating in a 37°C water bath for 2 hours. Purify labeled 48-Remy with a GE-25 column (method stated in "*5' ^{32}P labeling and annealing oligonucleotides to make splayed arm, double stranded and single stranded DNA substrates"* section). The concentration of labeled 48-Remy is ~5μM. Proceed to "Step 2: Anneal and Assemble".

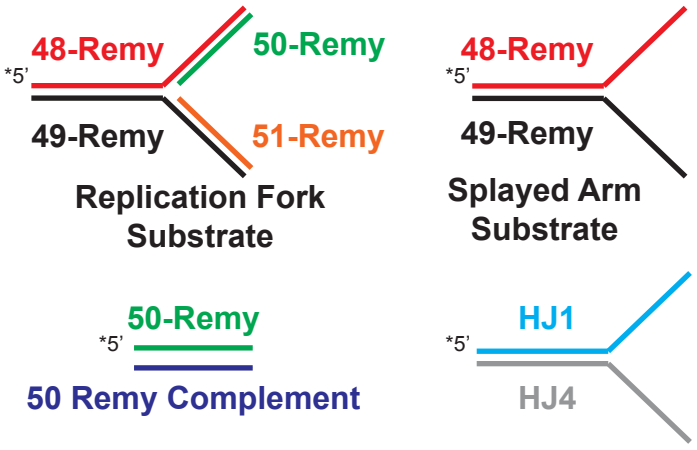
Step 2: Anneal and Assemble

Take 35μL of labeled 48-Remy and combine with 5μL of unlabeled 49-Remy at 50μM, 5μL 10X SSC and 5μL DNase/RNase free water. Split into two PCR tubes containing 25μL in each tube and anneal oligonucleotides using the "RBanneal" program (RBanneal step 1). The splayed arm product is ~3.5μM. Make sure to keep ~3μL of the splayed arm product to use as a standard to determine the final concentration of the purified substrate.

Take 47μL of "RBanneal step 1" and combine with 5μL of unlabeled 50-Remy at 50μM, 1μL 10X SSC and 7μL of DNase/RNase free water. Split into 3 PCR tubes each

Table 2.4 Oligonucleotides used to assemble replication fork, double stranded and splayed arm DNA substrates used for DNA binding, ATPase and nuclease assays.

NAME OF OLIGOS FOR REPLICATION FORK SUBSTRATE	SEQUENCE (5'-3')	LENGTH
48-REMY	ACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACATCTT TGCCCACCTGCAGGTTCACCC	60
49-REMY	TCGATAGTCGGATCCTCTAGACAGCTCCATGTAGCAAGG CACTGGTAGAATTCGGCAGCGT	61
50-REMY	GGGTGAACCTGCAGGTGGGCAAAGATGTCC	30
51-REMY	CATGGAGCTGTCTAGAGGATCCGACTATCGA	31
NAME OF OLIGOS FOR SPLAYED ARM SUBSTRATE	**SEQUENCE (5'-3')**	**LENGTH**
48-REMY	As above	60
49-REMY	As above	61
NAME OF OLIGOS FOR dsDNA SUBSTRATE	**SEQUENCE (5'-3')**	**LENGTH**
50-REMY	AS ABOVE	30
REMY 50 COMPLEMENT	GGACATCTTTGCCCACCTGCAGGTTCACCC	30
NAME OF OLIGOS FOR ATPase ASSAY	GGACATCTTTGCCCACCTGCAGGTTCACCC	30
HJ1	GTCGCTGCCCTAATCTGGCTTGCTAGGACATCTTTGCCG AGGTAGACCCG	50
HJ4	CGGATAGTCACTACACAGCATGTCCTAGCAAGCCAGATTA GGGCAGCGAC	50



containing 20µL. To anneal, use the “RBanneal” program (RBanneal step 2). Make sure to keep 3µL of this reaction to use as a size marker in the purification step.

Take 58µL of “RBanneal step 2” and combine with 8µL of unlabeled 51-Remy at 50µM, 8µL of 10X annealing buffer and 6µL of DNase/RNase free water into an eppendorf tube and place tube in 37°C water bath, switch off water bath and incubate for 30 minutes or until water bath temperature reaches 30°C. Store at -80°C until further use.

Step 3: Purification

Follow the same procedure as outlined in the “Step 4: Purification” section to make the lead gap regression substrate.

Step 4: Concentration determination

Follow the same procedure as outlined in the “Step 5: Concentration determination” section to make the lead gap regression substrate. Using the splayed arm intermediate generated in “RBanneal step 1” make dilutions to achieve 0.625µM, 0.312µM, 0.156µM, 0.078µM stocks and use as standards.

Nuclease Assay

I followed the same procedure described in Weston *et al*. to conduct the nuclease assays (66). Briefly, nuclease assays, under native conditions, to a splayed arm DNA substrate were conducted as follows: 2µL of 30nM ZRANB3 (3nM final) was incubated with 2µL of 100nM labeled splayed arm substrate (10nM final) for 30 minutes at 37°C in 16µL of nuclease buffer (50mM Tris pH 7.5, 50mM NaCl, 5mM $MgCl_2$, 2mM ATP, 0.1mg/mL BSA). The products were resolved on a 10% polyacrylamide gel (2.5mL 19:1 40% Acrylamide/Bis solution, 1mL 10X TBE, 6.3 mL water, 200µL 10% APS, 20µL

TEMED), in 1x TBE at 80V for 80 minutes at room temperature. Prior to loading the samples, the gel was pre-ran at 60V for 30 minutes in 1X TBE.

The nuclease assay to test lag gap regression reaction products was conducted as follows: Conduct a fork regression assay using a lag gap regression substrate. Add 20μL of 2X formamide buffer (95% formamide, 0.025% Bromophenol Blue, 0.025% Xylene Cyanol, 5mM EDTA) and heat at 95°C for 10 minutes. Resolve 20μL of the reaction on a 10% polyacrylamide 8M Urea gel (2.5mL 40% 29:1 Acrylamide/Bis solution, 4.8g Urea, 1mL 10X TBE, fill to 10mL with deionized water; heat in a 65°C water bath for 15 minutes and stir or until the Urea is completely dissolved and then add 33μL of 30% APS and 4μL TEMED) in 1X TBE at 120V for 2 hours or until the dye reaches half-way down the gel. Pre-run the gel for 30 minutes at 80V before loading samples. Clear each gel well before loading each sample. To analyze gel, wrap wet gel in saran wrap or place in a plastic antibody bag, expose to phosphorimager screen for 1 hour and scan screen using Biorad phosphorimager. Drying gel with plastic cellophane will cause the Urea to crash out of solution.

Running large gels to resolve nuclease products

These experiments were conducted with the manager of Dr. Osheroff's lab, Jo Ann Byl. Assemble gel casting chamber as follows: If the glass plates are new or have been used multiple times, soak each plate in 2M NaOH for 1 hour. Wash plates with water and thoroughly clean both sides of glass plates with glass cleaning soap. After wiping each plate dry, clean each side with Windex, then 100% ethanol. To the smaller plate, add 0.5mL Sigmacote (SL2-25ML Sigma) to the side of the plate that will make contact with the gel, coat the whole plate using a Kimwipe and allow to air dry for 10 minutes. Add another coat of Sigmacote to glass and allow to air dry for 10 minutes.

Assemble, glass, spacer and gasket, as instructed and set aside. Pour a 10% polyacrylamide 7.5M Urea gel and allow gel to polymerize overnight (25mL 40% 19:1 Acrylamide/Bis solution, 10mL 10X TBE, 30mL formamide, 45 grams Urea; heat at 65°C for 15 minutes and stir or until Urea is completely dissolved; cool down until it reaches room temperature, then add 625µL 10% APS and 62.5µL TEMED). Prior to loading samples, pre-run the gel at 40 Watts for 30 minutes. Run large gels in 1X TBE at 40 Watts for 2 hours or until dye front migrates halfway down the gel.

Creation, propagation and validation of the ZRANB3$^{-/-}$ mouse embryonic fibroblasts cell lines

The ZRANB3 $^{-/-}$ mouse embryonic fibroblast cell line was generously given to us by Dr. Christine Eischen. The *ZRANB3* gene disruption was generated from OST216789 OmniBank embryonic stem cells containing a neomycin resistance cassette inserted in the intron between Exon 8 and Exon 9 of *ZRANB3*. The ZRANB3$^{-/-}$ MEFS were confirmed using RT-PCR and western blotting (Figure 6.2B). The ZRANB3$^{-/-}$ MEFS were typically passaged at 1:3 when they reached 60% confluency. The MEFs were propagated in DMEM plus 10% FBS, 2mM glutamine, and 0.1mM nonessential amino acids (1% final).

Creation, propagation and validation of the U2OS ZRANB3 CRISPR cell lines

ZRANB3 CRISPR cell lines were generated using the procedure outlined on the ZHANG lab website: http://www.genome-engineering.org/crispr/?page_id=23 (74). Briefly gRNAs were cloned into the pSpCas9(BB)-2A-Puro Cas9 expression vector. U2OS cells were transfected and screened for editing using the following procedure: In each well of a 6-well dish, plate 2.5 x 10^5 U2OS cells (make sure to plate one or two extra wells of cells to use as the mock plate for Puromycin selection). 24 hours after

plating cells, dilute 1µg of each gRNA vector into an eppendorf tube containing 20µL of OptiMEM. Add 4.8µL of PEI at 1mg/mL, mix by flicking tube and incubate for 15 minutes at room temperature. During incubation remove old media from cells and add 2mL of fresh media (DMEM 10% FBS). After incubation, add transfection components to cells and incubate for 24 hours. Select for transfected cells by adding 2µL of puromycin at 1mg/mL to fresh media (1µg/mL final). Continue to select cells for 48 hours or until the cells on the control plate die. Plate surviving cells onto a 10cm tissue culture dish at a density to allow for single colony growth. Allow cells to form colonies; this will take 2-3 weeks. Once colonies have grown, pick and place individual colonies into a 24 well dish. Once the cells are confluent, passage 10% of the cells and pellet the remaining 90% for genomic DNA isolation using the Wizard gDNA purification kit (Promega catalog# A1120). Screen clones by PCR. Continue to passage cells that show evidence of editing. Validate protein absence using immunoblotting.

Colony Forming Assay

The following procedure was used to conduct colony forming assays using the U2OS ZRANB3 CRISPR cell lines and the ZRANB3$^{-/-}$ MEFs.

U2OS ZRANB3 CRISPR cells

Day 1

1. Use the following chart to plate 1mL of the indicated amount of cells/mL, in triplicate onto a 60mM tissue culture dish.

[MMS]	WT U2OS	Null (Clone 35)
0 mM MMS	200	200
0.1 mM MMS	500	500
0.2 mM MMS	500	500
0.3 mM MMS	1000	2000
0.4 mM MMS	2000	2000
0.5 mM MMS	2000	2000

Day 2

2. Carefully remove media and treat the U2OS cells with 2mL of the indicated concentration of MMS in the chart for 1 hour.
3. After treatment, immediately remove MMS and add 5mL of media (10% FBS DMEM) and place dishes in incubator. Make sure to clean and fill water bath to the top with fresh milliQ water.

Day 4 or 5

4. Assess cells using a light microscope for colony growth.

Day 12 (or until colonies are large enough to stain)

5. Stain dishes with Methylene blue dye (2% Methylene Blue, 50% Methanol), by pouring off media and incubating colonies with ~2 mL of dye for 2 minutes. After incubation, pour used dye into a collection container and gently wash dish with colonies under tap water. This dye can be reused multiple times.
6. Allow plates to dry overnight before counting colonies.

ZRANB3$^{-/-}$ Mouse Embryonic Fibroblasts

Day1

Use the following chart to plate 1mL of the indicated amount of cells/mL, in triplicate onto a 60mM dish.

[MMS]	WT MEFs	ZRANB3$^{-/-}$
0 mM MMS	200	200
0.1 mM MMS	200	200
0.25 mM MMS	200	200
0.5 mM MMS	500	500
0.75 mM MMS	500	500
0.1 mM MMS	500	500

Day 2

1. Carefully remove media and treat MEFs with 2mL of the indicated concentration of MMS in the chart for 1 hour.
2. After treatment, immediately remove MMS and add 5mL of MEF media (10% FBS, 2mM Glutamine, 0.1mM non-essential amino acids solution) and place dishes in incubator. Make sure to clean and fill water bath to the top with fresh milliQ water.

Day 3 or 4

3. Assess cells using a light microscope for colony growth. These cells proliferate quickly.

Day 7 or 8

4. Stain dishes with Methylene blue dye (2% Methylene Blue, 50% Methanol), by pouring off media and incubating cells with ~2 mL of dye for 2 minutes. After incubation, pour used dye into a collection container and gently wash dish with cells under tap water. This dye can be reused multiple times.
5. Allow plates to dry overnight before counting colonies.

Sister Chromatid Exchange Assay (U2OS cells)

I used the following procedure to create metaphase spreads of U2OS cells to detect sister chromatid exchange events.

Day 1:

1. Plate 3mL of 100,000 cells/mL into a 60mM dish (300,000 cells total).

Day 2:

2. 24 hours after plating cells, add BrdU to reach a final concentration of 10μM (the BrdU stock is at 10mM therefore add 3μL of stock directly to dish). Swirl dishes to evenly distribute BrdU and wrap dishes in aluminum foil to shield BrdU from light. Take note of time when BrdU was added.

   Note: After BrdU incubation, shield cells and slides from light using aluminum foil for the remaining of the procedure.

Day 4:

3. After 48 hour BrdU incubation, add colcemid to 150ng/mL final concentration and incubate for 1.5 hours (the colcemid stock is at 10μg/mL therefore add 45μL of stock to dish).

   Note: At this step, warm 75mM KCl in 37°C water bath. Also, put a box of glass slides into a -20°C refrigerator.

4. Pool media (floating cells) and attached cells into a 15mL conical tube and wash with 1X PBS. Spin down cells at room temperature at 180g for 8 minutes.
5. Aspirate supernatant, leaving 0.2mL of cells, and resuspend cells by gently flicking the bottom of the tube.
6. Add 6mL of pre-warmed 75mM KCl while gently vortexing tube. Add the first 1mL of 75mM KCl dropwise using a plastic bulb dropper or glass pipette and rubber bulb. Incubate for 16 minutes in a 37°C incubator.

   Note: At this step make 15mL of 3:1 methanol: glacial acetic acid fixative solution.

7. Add four drops of fixative, gently invert, and spin down cells at 180g for 8 minutes.
8. Aspirate supernatant leaving 0.2mL. Resuspend pellet by gently flicking the bottom of the tube and add 5mL of fixative, dropwise for the first 1mL while gently vortexing. Incubate for 20 min at 4°C. Spin down cells at 180xg for 8 minutes.
9. Aspirate the supernatant, leaving 0.2mL. Repeat step 7 twice.
10. Resuspend the pellet in a small volume of fixative (0.2mL), or until the cell suspension looks cloudy.
11. Drop cells onto prechilled slides using a P1000 pipette. Drop cells from a distance of 6 inches above the slide. Make sure that drops do not overlap. Dry slides at a slant and shield from light.
12. Allow slides to air dry in the dark for 2-3 days.

Day 6 or 7:

13. Stain cells with 0.1mg/mL acridine orange for 5 minutes at RT. Do this by pipetting 150µL of acridine orange onto a piece of parafilm and placing slide on top. Make 0.1mg/mL stock of acridine orange using deionized water and store in the dark. This stock is stable and can be used for multiple months.
14. Wash slides under running dH2O water for 2 minutes. Do this by placing slides in a coplin jar and placing it under a slow stream of deionized water.
15. Incubate slides for 1 minute in Sorenson Buffer (use same procedure as acridine orange incubation) (0.1M Na_2 $2HPO_4$ pH 6.8, 0.1M NaH_2PO_4). Mount slides using 150µL of Sorenson Buffer.
16. View immediately using a fluorescent microscope.

CHAPTER III

IMPACT OF THE HARP AND ATPase DOMAINS ON SMARCAL1 FUNCTION

Introduction

Initial biochemical characterization of SMARCAL1

Protein fractions containing a N-terminal truncated, mouse SMARCAL1 protein purified from bacteria showed modest DNA-dependent ATPase catalysis to a M13 single stranded phage DNA substrate; Coleman *et al*. provided the first evidence of SMARCAL1 DNA-dependent enzymatic activity (52). Further *in vitro* analysis showed that SMARCAL1 specifically binds forked, overhang and gapped DNA substrates and does not bind single or double stranded DNA substrates (53,75). SMARCAL1 DNA binding specificity is also reflected in its DNA stimulated ATPase activity in that it catalyzes robust ATP hydrolysis when bound to splayed arm DNA substrates and displays negligible ATP hydrolysis activity in the presence of single stranded and double stranded DNA substrates (53,75). Discrepancy in ATPase assay results between Coleman *et al*. and subsequent results from Yusufzai *et al*. and Bétous *et al*. that showed that SMARCAL1 catalyzes DNA-dependent ATPase activity to forked, gapped and overhang DNA substrates, is most probably due to the use of M13 single stranded DNA as a substrate for this reaction, which can fold into secondary DNA structures that SMARCAL1 is capable of binding and catalyzing ATP hydrolysis. Furthermore, SMARCAL1 catalyzes annealing helicase and fork remodeling activity in that it re-anneals complementary DNA strands (53,75).

In addition to having a highly conserved ATPase domain, SMARCAL1 contains a RPA binding domain (RBD) and tandem HARP domains in the N-terminal half of the protein. The HARP1 and HARP2 domains are approximately 70 amino acids in length

and share significant sequence similarity with one another, although they lack sequence similarity with other functionally established domains present in other proteins classified within the SNF2 subfamilies (53,63). A combined ΔHARP1-ΔHARP2 deletion mutant binds a splayed arm DNA substrate, is active as a DNA-dependent ATPase yet lacks annealing helicase activity (76). These results suggested that the HARP domains are dispensable for DNA binding and primarily function to help catalyze annealing helicase activity. Shortly after the Ghosal *et al*. (76) publication, Bétous *et al*. showed that a ΔHARP1-ΔHARP2 deletion mutant and full length SMARCAL1 containing point mutations within the HARP2 domain failed to bind a splayed arm DNA substrate (53). Bétous *et al.* and Ghosal *et al.* observed opposing results regarding whether the HARP domains are necessary for SMARCAL1 to bind its DNA substrates. The Ghosal *et al*. study did not include a necessary ATPase-dead negative control to ensure that the purification of the ΔHARP1-ΔHARP2 SMARCAL1 deletion mutant was not contaminated with other DNA-dependent ATPase enzymes, which would yield a positive result.

SMARCAL1 function in cells

SMARCAL1 localizes to sites of stalled replication forks by binding to RPA (54-57). SMARCAL1 phosphorylation is regulated by the DNA damage response kinases ATR, ATM and DNA-PK (54,57). Recently, Couch *et al*. hypothesized that the mechanism of fork collapse in cells lacking ATR function may be partially due to excess SMARCAL1 regression of stalled replication forks into Holliday junctions that can be cleaved by the structure specific nuclease SLX4 and undergo resection by the exonuclease CtIP. Specifically, ATR phosphorylates SMARCAL1 on S652 in a damage-dependent manner, and this modification is hypothesized to temper SMARCAL1 function at stalled replication forks by preventing unregulated fork regression that promotes SLX4 and CTIP mediated fork cleavage, resection and fork collapse (59). This effect is specific

to SMARCAL1 in that knockdown of functionally similar fork remodelers, notably HLTF, ZRANB3 and BLM, did not decrease the creation of nascent strand single stranded DNA in the absence of ATR function which is indicative of continued replication fork processing by SLX4 and CtIP. The ability of SMARCAL1 to bind RPA at stalled forks and binding DNA are both required for damaged dependent ATR mediated phosphorylation of SMARCAL1 in cells undergoing replication stress. As previously stated, SMARCAL1 deficient cells are more sensitive to HU, aphidicolin and CPT induced replication stress and have inherently higher levels of γH2AX signaling; therefore SMARCAL1 function in cells is specifically regulated by ATR to ensure genome integrity during DNA replication. Importantly, the HARP domains are required for robust damage-dependent ATR mediated phosphorylation of SMARCAL1.

Furthermore, SMARCAL1 functions to maintain telomere integrity (60,61). SMARCAL1 depleted cells accumulate telomere dysfunction-induced foci (TIFs) and replication-dependent extrachromosomal DNA circles (C-circles), which are directly attributed to telomere instability (60). SMARCAL1 DNA translocase activity is required to prevent C-circle formation and RPA binding is not, since the R764Q ATPase dead mutant is incapable of complementing this phenotype in SMARCAL1 deficient cells and a RPA binding mutant was proficient in suppressing this phenotype (60). SMARCAL1 deficiency in cells utilizing the alternative lengthening of telomere mechanism to elongate telomere ends show a significant increase in telomere foci size that is attributed to telomeric DNA clustering as a consequence of double strand break formation at telomere ends (61). In this context, SMARCAL1 prevents DNA breaks and replication stress at telomeres in ALT cells and promotes ALT mediated telomere lengthening. Collectively, SMARCAL1 promotes global replication stability.

How SMARCAL1 functions to maintain telomere integrity is unknown. As previously stated, SMARCAL1 binding to RPA regulates its fork remodeling activity *in vitro* and deleting the RBD fails to complement the γH2AX signaling response in SMARCAL1 deficient cells (54). Therefore, RPA binding is integral in SMARCAL1 function to maintain fork stability while replicating non-telomeric DNA but is negligible for SMARCAL1 function at maintaining telomere integrity. This strongly suggests that SMARCAL1 localizes to telomeres using an alternative mechanism than RPA and functions at telomeres in an RPA binding independent manner. SMARCAL1 localization and retention at telomere is more than likely facilitated by binding to a telomere specific protein. Furthermore, since SMARCAL1 binds its DNA substrates in a sequence independent manner, the HARP domains may also function to localize and retain SMARCAL1 at telomeres.

SMARCAL1 and disease

Schimke immune-osseous dysplasia (SIOD) is a rare disease linked to SMARCAL1 deficiency (77). Individuals with SIOD suffer from renal failure, growth retardation, short stature, spondyloepiphyseal dysplasia, and immunodeficiency (77). In some cases, SIOD patients contract cancer such as non Hodgkin lymphoma (Epstein Barr virus positive or negative), osteosarcoma, and an undifferentiated carcinoma of the sinus (78,79). Severe forms of SIOD cause death at an early age (77). Half of the documented cases of SIOD contain biallelic mutations within *SMARCAL1*; the remaining half contains either monoallelic or no detectable *SMARCAL1* mutations (77,80). Furthermore, SIOD patient derived fibroblasts have increased γH2AX levels, indicative of elevated DNA damage, which is rescued by exogenous SMARCAL1 expression (81). Therefore, SMARCAL1 activity is highly important in cells and analyzing how point

mutations linked to causing SIOD impacts SMARCAL1 activity will help us understand how SMARCAL1 functions to bind DNA substrates and catalyze fork remodeling.

Given the published biochemical data characterizing SMARCAL1 enzymatic activity, its DNA binding specificity and the discrepancies between the Ghosal *et al*. and Bétous *et al*. work, I wanted to study how SMARCAL1 functions as an annealing helicase on a molecular level. I hypothesized that *the HARP and ATPase domains contribute in DNA binding and provide functional specificity for SMARCAL1***.** To test this, I subjected SIOD point mutants within the HARP domains, and highly conserved residues present in the SNF2 ATPase domain to protein expression, DNA binding, and DNA-dependent ATPase assays. This assessment attempted to establish how the domains in SMARCAL1 impact its DNA binding and enzymatic activities *in vitro* and in cells.

Results

Biochemical characterization of SIOD associated point mutants

I compiled a list of point mutations present in SMARCAL1 associated with severe and mild forms of SIOD and mapped them on the primary structure of the protein (Figure 3.1). Most point mutants mapped to the ATPase domain. S859P, a novel point mutant from a child that was treated at Vanderbilt University, is located at the C-terminal end of the ATPase domain. F279S mapped to the HARP1 domain and mutations E377Q and H379P mapped to the HARP2 domain. F279 and E377 are highly conserved (Figure 3.2A); H379 is conserved through mammals (Figure 3.2B). The list of point mutants associated with SIOD were obtained from Boerkoel *et al*. (77), Clewing *et al*. (E377Q) (80), Lücke *et al*. (F279S) (82), Lücke *et al*. (F279S) (83), Carroll *et al*. (S859P) (78), and the SMARCAL1 SIOD database established by Piirilä *et al*. (84).

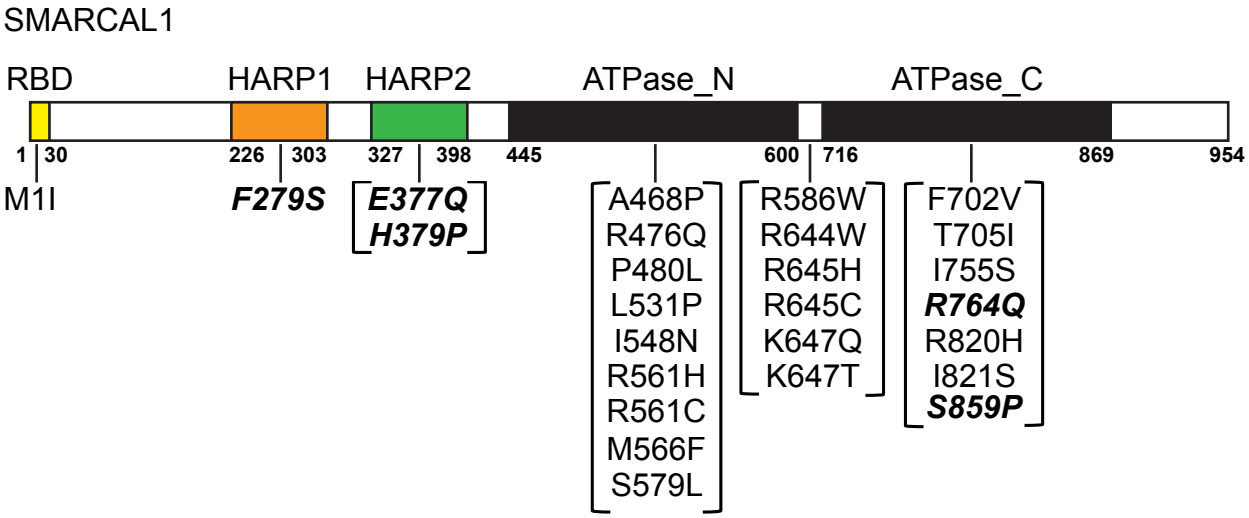
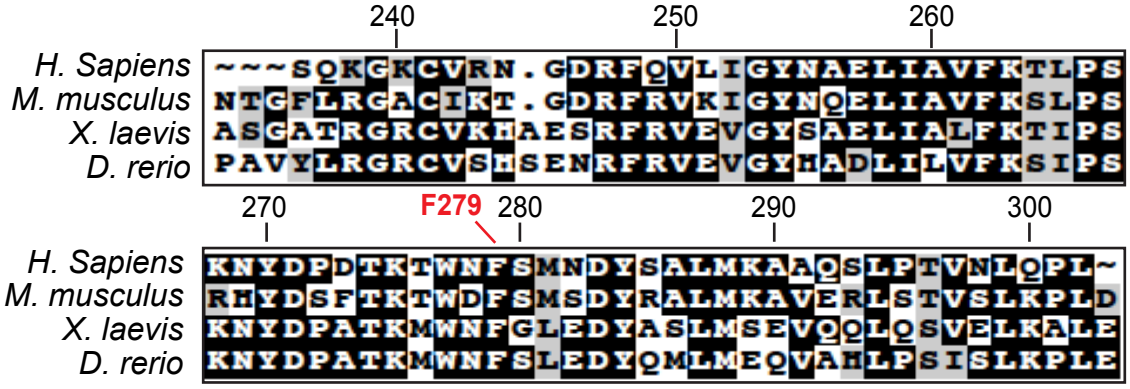


FIGURE 3.1 SIOD linked point mutants mapped onto SMARCAL1. Bolded and italized mutants were studied. Numbers indicate the N and C terminal boundaries of the RPA binding domain (RBD), HARP1, HARP2, ATPase_N and ATPase_C domains.

A. HARP1 Sequence Alignment



B. HARP2 Sequence Alignment

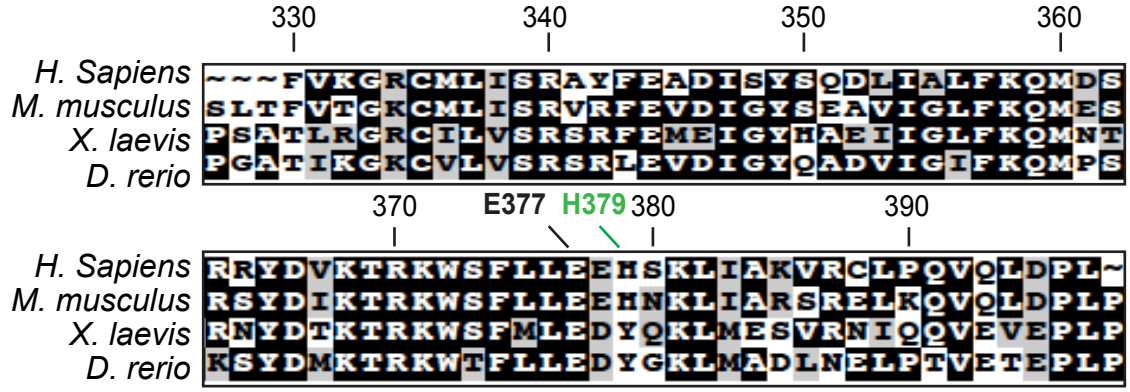


FIGURE 3.2 F279 and E377 are highly conserved and H379 is less conserved. Sequence alignments for the (A) HARP1 and (B) HARP2 domains of SMARCAL1 orthologs. Amino acid numbering is based on human SMARCAL1.

To test whether the point mutants within the HARP domains impact SMARCAL1 activity, I expressed recombinant protein harboring these mutations and conducted ATPase assays to a splayed arm DNA substrate (Figure 3.3). H379P and F279S did not express protein to the same extent as the other mutants and I did not test them in subsequent enzymatic assays. Mutant E377Q is as active as a DNA-dependent ATPase as WT SMARCAL1 (Figure 3.3.).

S859P is highly conserved and is located in close proximity to the C-terminal end of the ATPase domain (Figure 3.1 and Figure 3.4A). Compared to WT SMARCAL1, S859P catalyzes DNA-dependent ATPase activity to a lesser extent (Figure 3.4B).

Biochemical characterization of highly conserved ATPase domain residues

The *Sulfolobus solfataricus* Rad54 (SsoRad54) ATPase domain makes direct contacts with double stranded DNA (85). Similar to SMARCAL1, the SsoRad54 ATPase domain contains the same signature motifs defining the SNF2 family of chromatin remodelers (85). To determine whether the ATPase domain in SMARCAL1 contributes to binding its DNA substrates, I tested point mutants within this domain that are conserved with the SsoRAD54 ATPase domain in a DNA binding assay to a splayed arm DNA substrate (Figure 3.5A and Figure 3.5B). S488A binds a splayed arm DNA substrate with a similar affinity to WT SMARCAL1 and K555E has reduced DNA binding affinity (Figure 3.5B).

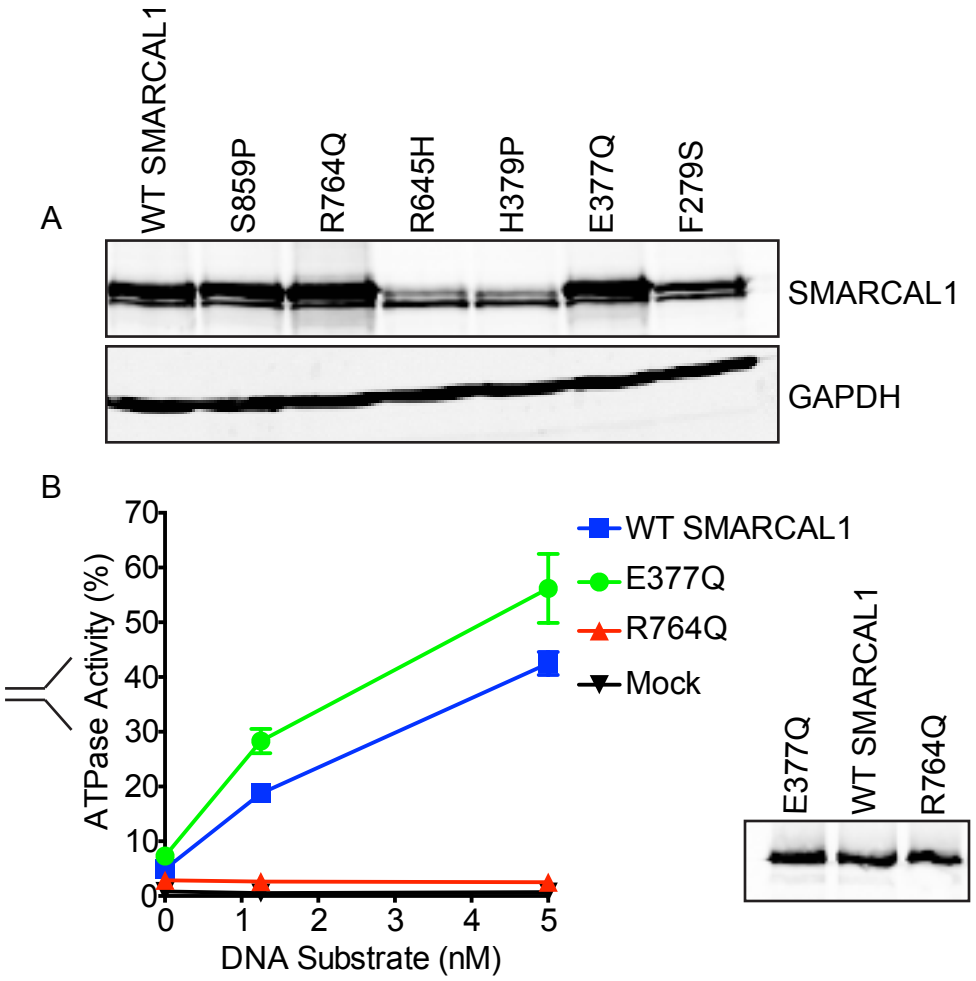


FIGURE 3.3 SIOD point mutant E377Q is as active as a DNA-dependent ATPase as WT SMARCAL1. (A) 293T cells were transfected with FLAG-SMARCAL1 cDNA expression vectors containing indicated SIOD mutations, lysed, and protein expression was analyzed using immunoblotting. (B) Purified FLAG-SMARCAL1 protein containing indicated SIOD point mutants were tested in an ATPase assay to 0nM, 1.25nM and 5nM of a splayed arm DNA substrate. The R764Q and Mock data points overlap. Immunoblotting of purified protein dilutions was conducted to ensure equal loading of protein used in the ATPase assay. Error bars represent the standard deviation of the mean for three experiments.

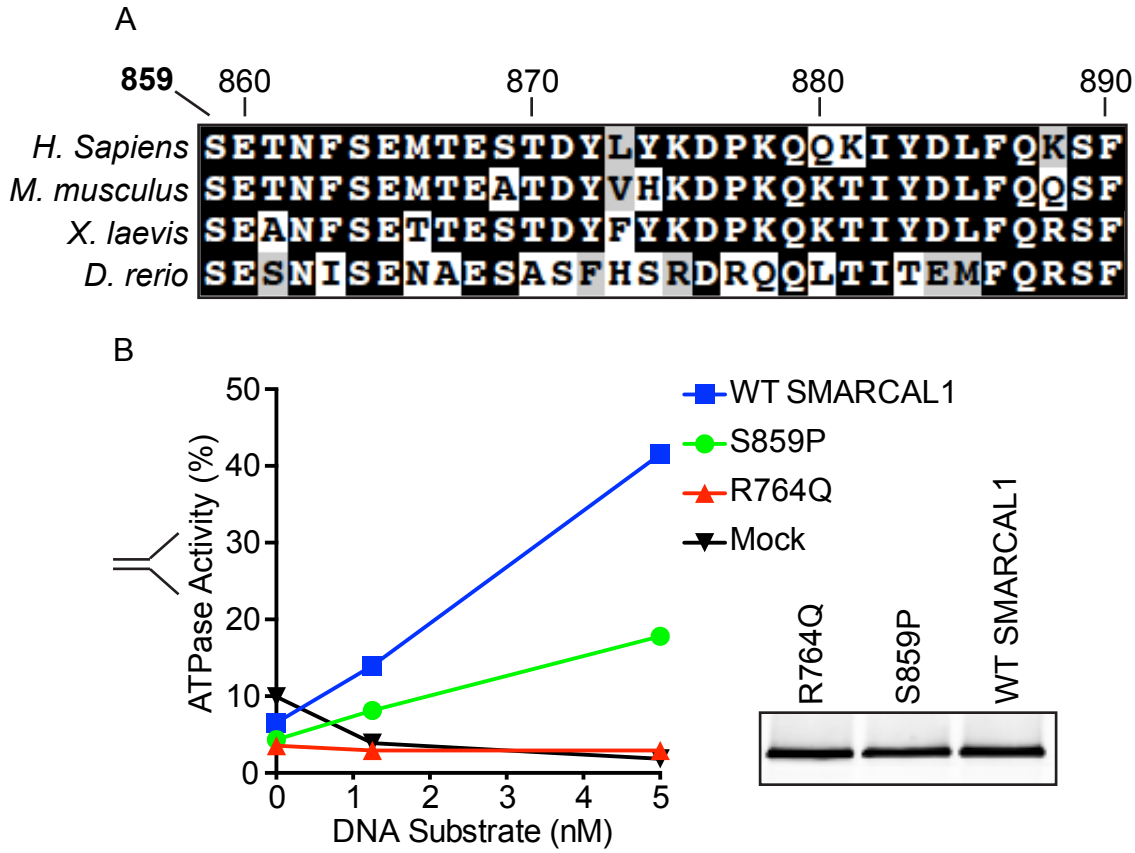


FIGURE 3.4. SIOD point mutant S859P is highly conserved and less active as a DNA-dependent ATPase as WT SMARCAL1 (A) Sequence alignment of *M. musculus, X. laevis and D. rerio* SMARCAL1 orthologs to human SMARCAL1 residues 859-890. Residue S859 is indicated in bold print. (B) Purified FLAG-SMARCAL1 S859P, wt SMARCAL1 and R764Q were tested in an ATPase assay to 0nM, 1.25nM and 5nM of a splayed arm DNA substrate. Immunobloting of purified protein dilutions was conducted to ensure equal loading of protein used in the ATPase assay. Error bars represent the standard deviation of the mean for three experiments. Amino acid numbering is based on human SMARCAL1.

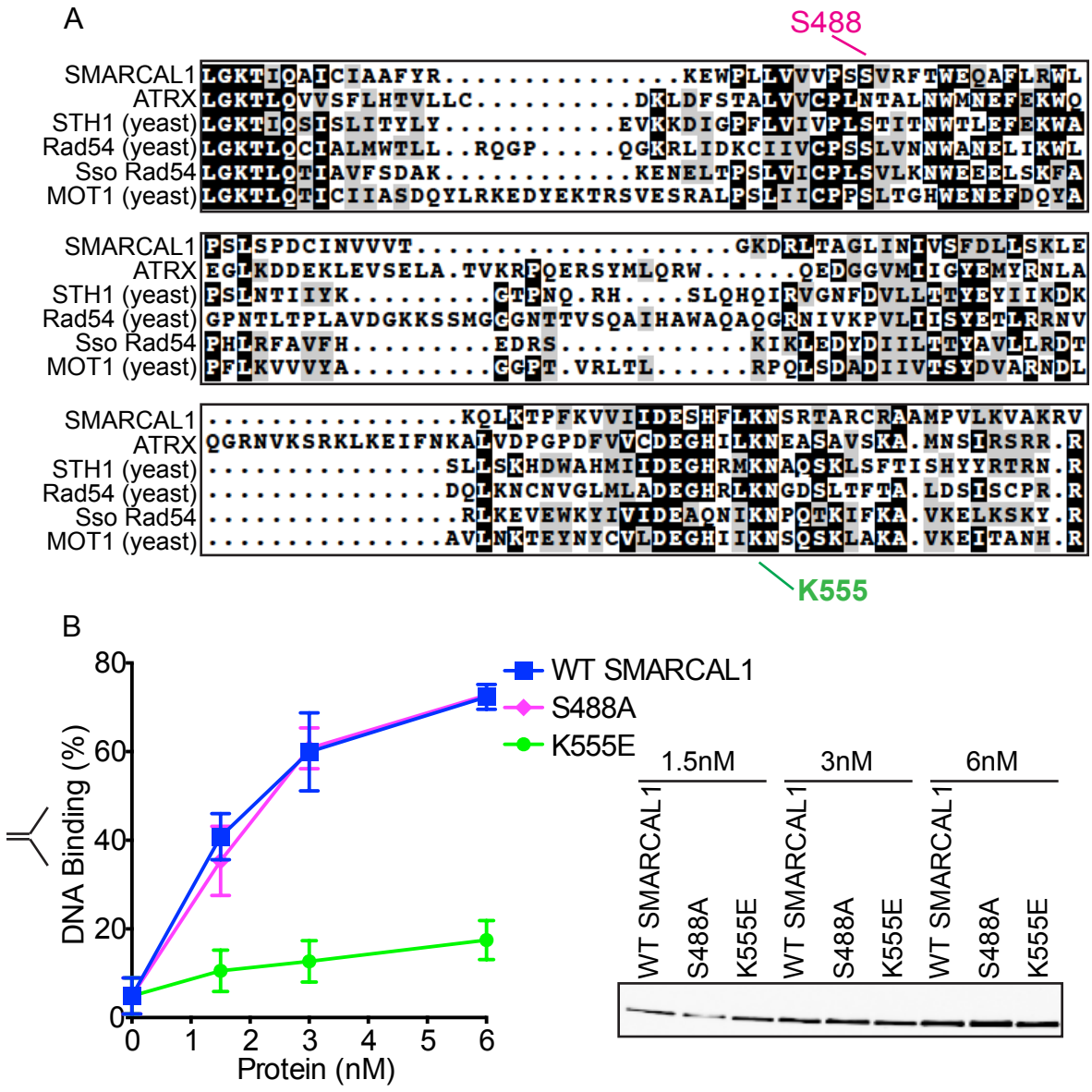


FIGURE 3.5 K555 contributes to SMARCAL1 DNA binding. (A) Human SMARCAL1, ATRX, yeast STH1, yeast Rad54, *Sulfolobus solfataricus* Rad54, and yeast MOT1 alignment of the conserved SNF2 ATPase domain (recreated from Dürr *et al*). (B) 1.5nM, 3nM and 6nM of purified WT SMARCAL1, S488A and K555E were tested in a DNA binding assay to a splayed arm DNA substrate. Immunobloting of purified protein dilutions was conducted to ensure equal loading of protein used in DNA binding assay. Error bars represent the standard deviation of the mean for three experiments.

Discussion

Analysis of SIOD associated F279S, H379P, E377Q, and S859P SMARCAL1 point mutants

Using sequence conservation and characterizing the activity of point mutants associated with SIOD, I identified residues that contribute to SMARCAL1 DNA binding, DNA-dependent ATPase activity and protein stability. SIOD point mutants located in the HARP1 (F279S) and HARP2 (H379P) domains created unstable recombinant protein, thereby making it difficult to continue further analysis. When mapped to the solved crystal structure of the mouse HARP1 domain (63) and predicted structure of human HARP2, F279 is oriented towards the center of the structure (Figure 3.6A). Mutating this residue may destabilize the HARP1 domain and thereby full length SMARCAL1 and cause a decrease in expression levels of recombinant F279S mutant proteins in cells. Interestingly, F279S is a reoccurring SIOD mutation that was identified in 6 patients that comprised 5 unrelated families; two patients are brothers containing identical heterozygous mutations in their SMARCAL1 alleles (F279S - paternal allele; E848X, where X denotes a non-sense mutation - maternal allele), one patient is of an unrelated boy also containing heterozygous SMARCAL1 mutations (H379P, F279S) and another patient has the SMARCAL1 genotype F279S and P113X (80,83,86). The medical reports of the remaining 2 patients were not listed (80,84). Residue H379 in the HARP2 domain is also positioned towards the center, which may also compromise protein stability (Figure 3.6B).

Residue E377Q did not interfere with SMARCAL1 ATPase activity. Based on the structural similarities of the HARP2 domain and the mismatch recognition domain (MRD) in MSH6, and how the MRD contacts DNA, Mason *et al.* tested mutant D361P/E377N,

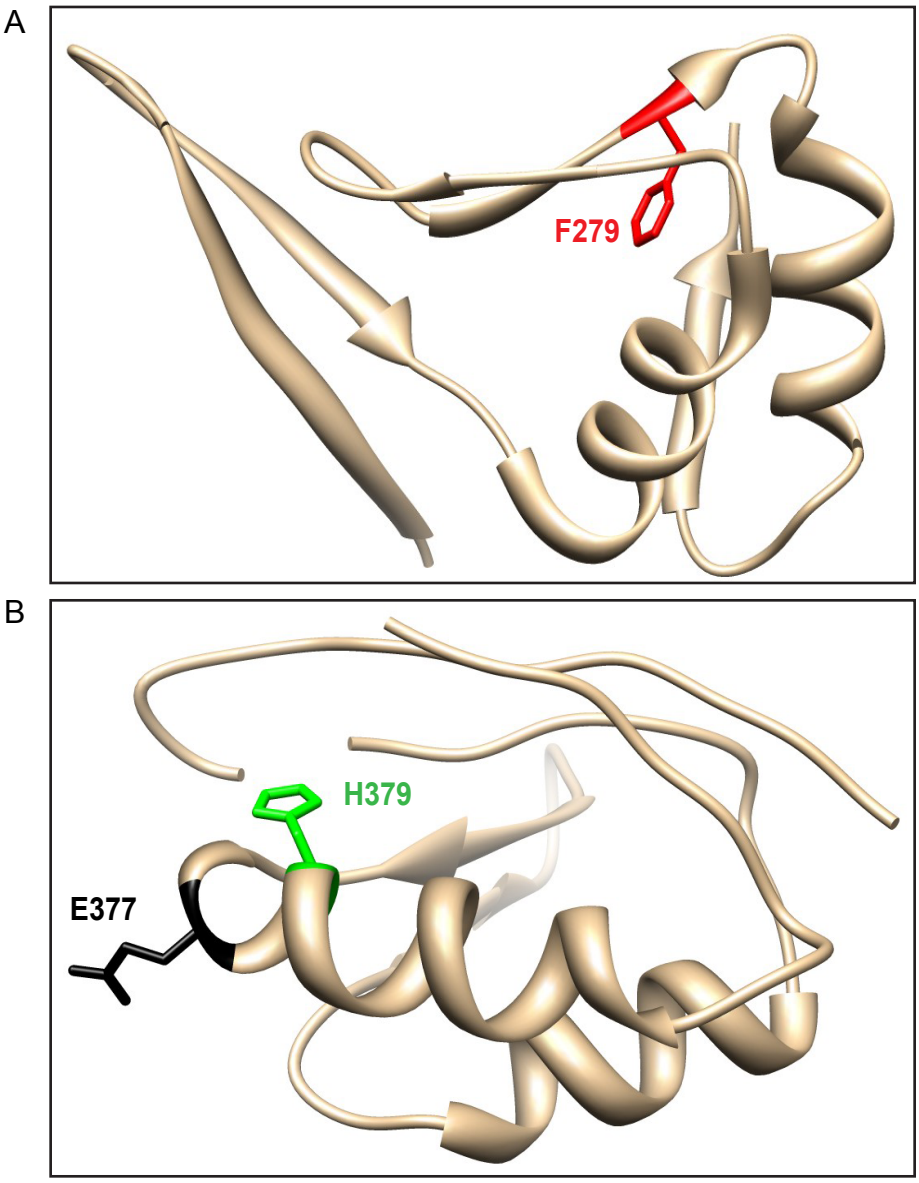


FIGURE 3.6 F279, E377, and H379 modeled onto the mouse HARP1 structure and predicted structure of HARP2 domains. (A) F279, indicated in red, is modeled onto the published crystal structure of the mouse HARP1 domain (PDB 4O66). (B) Residues E377 and H379 are modeled onto the predicted human HARP2 structure. E377 and H379 are indicated in black and green respectively.

made in the catalytic domain (cd) of SMARCAL1 (HARP2-ATPase domain), in DNA binding assays to various structured DNA substrates, ATPase and fork regression assays and observed reduced DNA binding, ATPase and fork regression activities (63).

Mason *et al*. did not test the E377Q SIOD associated mutation in the context of the SMARCAL1 cd or full length protein in DNA binding and fork remodeling assays. Therefore I cannot make strong conclusions between the D361P/E377N and my results to explain why the full length E377Q mutant retained robust DNA-dependent ATPase activity and the catalytic domain D361P/E377N mutant had reduced DNA-dependent ATPase activity. However, mutating a residue in the HARP2 domain in the context of the catalytic domain, may have a more pronounced affect than mutating this same residue in the context of the full length protein, since the HARP1 domain may functionally compensate in DNA binding in the absence of a functional HARP2 domain. This may explain why I did not observe a decrease in DNA-dependent ATPase activity with the single E377Q point mutant in the context of the full length protein (Figure 3.3B). As further support, ΔHARP1 has decreased binding affinity to a splayed arm and gapped DNA substrates (53). These results suggest that the HARP1 domain contributes to DNA binding to some extent or may be involved in binding specific DNA structures. Another possibility is that the dual D361P/E377N mutations are more effective at debilitating the HARP2 DNA binding ability. Therefore, mutating both residues in the context of the full length protein may yield a similar result as the catalytic domain D361P/E377N mutant.

An explanation for why the S859P SIOD point mutant has reduced DNA dependent ATPase activity is unclear. Although S859P is highly conserved in SMARCAL1 orthologs (Figure 3.4), this residue lacks sequence conservation to other SNF2 chromatin remodelers. I did not test S859P in a DNA binding assay therefore decreased DNA binding affinity cannot be ruled out as a possibility. If this mutant is also

defective in DNA binding, this would provide further evidence that the SMARCAL1 ATPase domain contributes to binding DNA.

The SMARCAL1 ATPase domain binds DNA

Based on the structure, residues S502 and K568 in the catalytically dead SsoRad54 binds the 3'-5' and the 5'-3' strand of a double stranded DNA substrate, respectively (85). No additional biochemical analysis was conducted to directly test either residue in full length SsoRAD54. I observed a severe decrease in DNA binding for the SMARCAL1 K555E mutant and observed no difference for the S488A mutant. This suggests that the ATPase domain in SMARCAL1 contacts DNA. Specifically how the SMARCAL1 ATPase domain is oriented on DNA is unknown. Based on published structures of other SF2 superfamily ATPase domains, the SMARCAL1 ATPase domain is speculated to bind and translocate on the duplex parental strand at the junction (37).

Concluding remarks

Although this project did not result in a published body of work, the techniques I learned and the collaborations I made with Dr. Aaron Mason, a post doctoral trainee in Dr. Brandt Eichman's lab at the time, greatly helped in my ZRANB3 project (Chapter 4). Furthermore the S859P ATPase assay result (Figure 3.4) is published and I am a co-author on that paper (78).

CHAPTER IV

IDENTIFICATION OF A SUBSTRATE RECOGNITION DOMAIN IN ZRANB3

Introduction

Genomic replication is a highly challenging task. The DNA replication machinery must precisely duplicate billions of base pairs while tolerating a multitude of obstacles including damaged DNA, collisions with transcriptional machineries, unusual DNA structures and other difficult to replicate sequences (9). Many of these obstacles stall replication forks and activate replication stress responses that stabilize and restart persistently stalled forks. These mechanisms include fork remodeling to regress replication forks into a chicken foot DNA structure (19,20). Fork regression may facilitate DNA repair or template switching to bypass the obstruction (20).

Several members of the SNF2 family of DNA-dependent ATPases including SMARCAL1, HLTF and ZRANB3 are replication stress response proteins that catalyze fork remodeling including fork regression (45,53,65). The replication stress response is essential to complete replication accurately. Therefore, defects in this response cause human disease (9). For example, bi-allelic loss of function mutations in *SMARCAL1* cause Schimke immuno-osseous dysplasia (SIOD) (77). *HLTF* is silenced in colorectal cancer and *ZRANB3* is mutated in endometrial cancers suggesting that both may be tumor suppressors (87,88).

The enzymatic activities of SMARCAL1 and HLTF are dependent on a SNF2 ATPase motor domain and a substrate recognition domain (SRD) that is thought to mediate binding to specific structures at stalled replication forks. The SRD of SMARCAL1 is its HARP2 domain, which is required for SMARCAL1 binding to branched DNA structures as well as DNA-dependent ATPase and fork regression activities

(53,63). The HARP domain is structurally related to the damage recognition domain of the XPB helicase (Xeroderma pigmentosum, complementation group B; ERCC3) and the mismatch recognition domain of MSH6 (MUTS homolog 6) (63). The SRD in HLTF is its HIRAN domain, which is unrelated in sequence and structure to the HARP domain and interacts with the exposed 3' end of small DNA flaps (50,51,63). The HIRAN domain is also important for HLTF mediated fork regression activity (49,51). In both SMARCAL1 and HLTF, mutations in the HARP or HIRAN domains interfere with their ability to bind DNA and catalyze fork remodeling (50,51,53,63).

Yuan *et al*. reported that ZRANB3 contains a domain similar in sequence to the HARP domains of SMARCAL1 (67). However, they reported that deletion of this putative SRD domain inactivates its strand annealing activity without interfering with DNA binding or DNA-dependent ATPase activity (67). Given the apparent differences in the reported activities of the SMARCAL1 HARP and ZRANB3 HARP-like domains, we revisited the requirements for ZRANB3 to bind DNA, hydrolyze ATP and catalyze fork remodeling. We define a ZRANB3 SRD that is essential for all three functions and define the ZRANB3 minimal enzymatic unit for fork remodeling as containing only the SNF2 ATPase domain and its SRD.

Results

Region 720-869 is highly conserved and necessary for ZRANB3 ATPase activity

ZRANB3 is a DNA-dependent ATPase in the same SNF2 family as SMARCAL1 and HLTF (37). SMARCAL1 and ZRANB3 are both annealing helicases that re-anneal complementary DNA strands (64,75) and catalyze replication fork remodeling reactions (53,65). The SMARCAL1 HARP2 domain is required for SMARCAL1 to bind DNA, hydrolyze ATP, anneal DNA, and remodel replication forks (53,63).

A previous study concluded that a region encompassing amino acids 712-818 in ZRANB3 contains a HARP-like domain that is required for ZRANB3 annealing helicase activity (67). However, unlike the SMARCAL1 HARP domain, the HARP-like domain was reported to be dispensable for DNA binding and ATPase activity. Fork remodeling was not tested. Due to the striking functional differences between the HARP and HARP-like domains, we revisited whether the HARP-like domain of ZRANB3 really shares similar functional properties to the SMARCAL1 HARP domains. We purified wild type (WT) and Δ712-818 ZRANB3 (Figure 4.1A) and tested their ability to bind a splayed arm DNA substrate and hydrolyze ATP. In contrast to the previously published findings, purified Δ712-818 ZRANB3, which lacks the HARP-like domain, failed to bind a splayed arm DNA substrate (Figure 4.1B). It also lacked DNA-stimulated ATPase activity (Figure 4.1C). In contrast, WT ZRANB3 displayed both DNA-binding and DNA-dependent ATPase activity.

It is unclear whether deleting amino acids 712-818 generates a protein that is properly folded. Since important amino acids in the SMARCAL1 HARP domain have already been identified (53,63), we attempted to generate a sequence alignment between the SMARCAL1 HARP domains and the ZRANB3 HARP-like domain to identify critical amino acids for mutagenesis. However, we were unable to find sufficient sequence similarity to generate a useful alignment. Therefore, using evolutionarily conserved regions of ZRANB3 as a guide, we designed and tested various deletion, truncation and point mutants to determine regions within the protein that are necessary for DNA-dependent ATPase activity (Figure 4.2A). Deletion of the NZF and APIM motifs, which bind polyubiquitinated PCNA (65), did not impair ATPase activity (Figure 4.2B). Deletion of the HNH nuclease domain caused a modest but reproducible decrease in activity (Figure 4.2B). Deletion of amino acids 651-720 also yielded an active enzyme

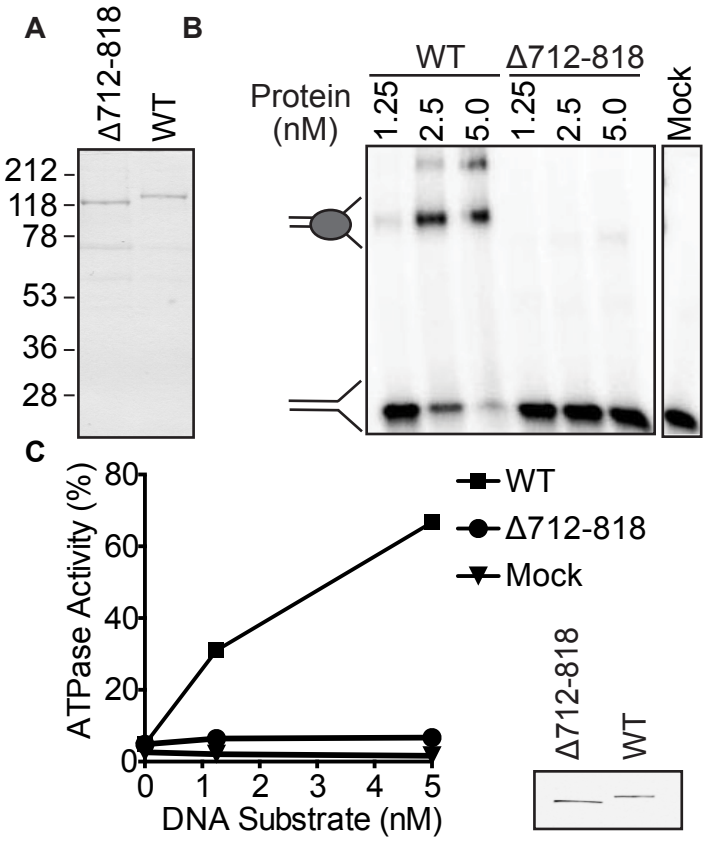


FIGURE 4.1 Δ712-818 ZRANB3 does not bind DNA and cannot hydrolyze ATP. (A) Coomassie stained SDS-PAGE gel of purified wild-type (WT) and Δ712-818 ZRANB3. (B) Δ712-818 ZRANB3 and WT ZRANB3 were incubated with a splayed arm DNA substrate. To assess DNA binding, samples were resolved on a polyacrylamide gel and visualized by autoradiography. A representative experiment is shown. (C) Δ712-818 ZRANB3 and WT ZRANB3 were incubated with a splayed arm substrate and ATPase activity was measured. The mean and standard deviation from three experiments are shown. In most cases the standard deviation is smaller than the symbol size. The inset is an anti-FLAG immunoblot of the purified proteins.

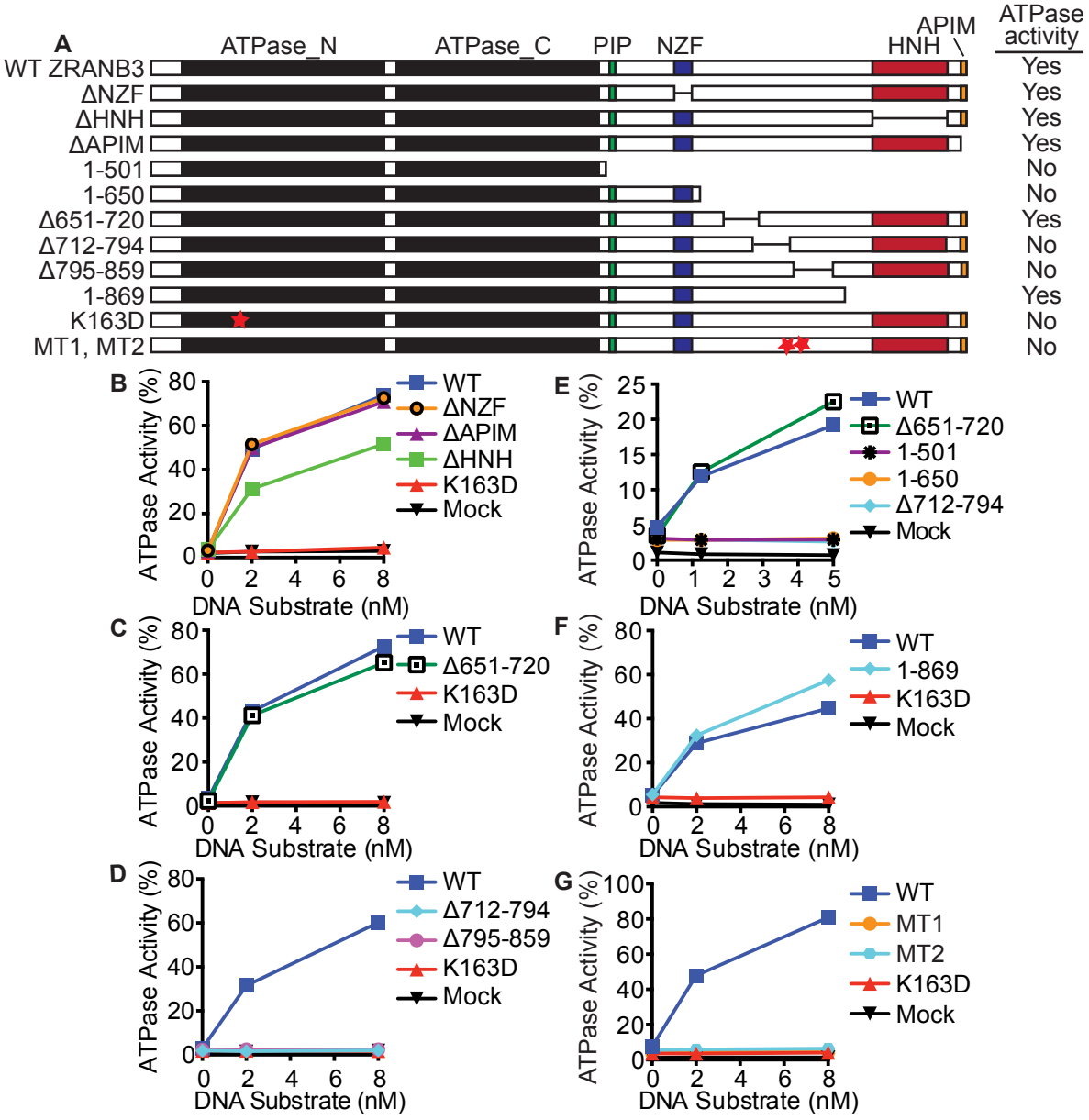


FIGURE 4.2 Amino acids 721-869 are necessary for ZRANB3 DNA-dependent ATPase activity. (A) Schematic and summary of results for the various deletion, truncation and point mutants. Purified (B) ΔNZF, ΔHNH, ΔAPIM motif deletion mutants; (C) Δ651-720 deletion mutant; (D) Δ712-794 and Δ795-859 deletion mutants; (E) 1-501 and 1-650 truncation mutants; (F) 1-869 truncation mutant; and (G) triple mutants L760A/D761A/I762A (MT1) and W790A/S791A/S792A (MT2) were incubated with a splayed arm DNA substrate and ATPase activity was measured. A representative experiment (of at least two replicates) is shown for each mutant.

(Figure 4.2C). In contrast, ZRANB3 Δ712-794 and ZRANB3 Δ795-859 were both inactive (Figure 4.2D).

Likewise, C-terminal deletion constructs containing only the ATPase domain (ZRANB3 1-501) or the ATPase domain, PIP and NZF motifs (ZRANB3 1-650) were also inactive (Figure 4.2E). However, a protein consisting of amino acids 1-869 was as active as the wild type protein (Figure 4.2F). In all cases, the active proteins required DNA for ATP hydrolysis. Thus, the ZRANB3 ATPase domain requires an accessory domain that likely includes amino acids 721-869 for activity.

Amino acids 720-869 contain most, but not all, of the residues previously described to make-up the HARP-like domain. It is relatively highly evolutionarily conserved compared to flanking regions of ZRANB3 (Figure 4.3). However, our sequence alignment failed to detect significant similarity with the HARP domains of SMARCAL1. We also compared the known secondary structure of the HARP domain to the predicted secondary structure of this ZRANB3 region. While there is some similarity, the ZRANB3 domain contains a large insertion that is predicted to be alpha helical. Mutations in highly conserved amino acids within this helix and in other highly conserved amino acids in this region (MT1: L760A/D761A/I762A and MT2: W790A/S791A/S792A) inactivate the protein (Figure 4.2G). These data confirm that this region is necessary for DNA-dependent ATPase activity. Based on this data as well as additional information (see below) we designate amino acids 720-869 of ZRANB3 as a substrate recognition domain (SRD).

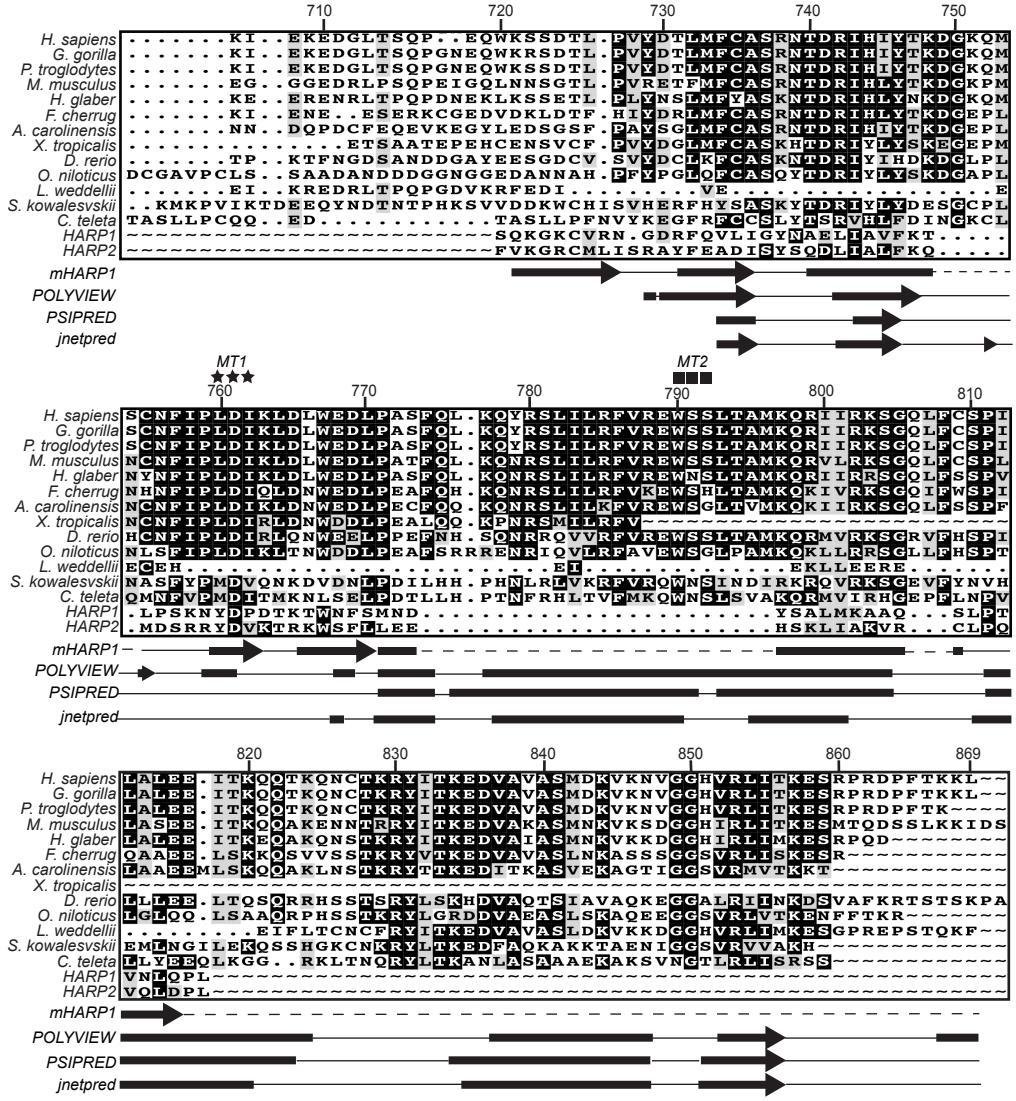


FIGURE 4.3 ZRANB3 amino acids 720-869 are highly conserved but have minimal similarity to the HARP domains of SMARCAL1. Amino acid sequence numbering corresponds to human ZRANB3 isoform 2. Secondary structure of mouse HARP1 (based on PDB ID 4O66), and predicted secondary structure of ZRANB3 generated from POLYVIEW, PSIPRED and JNETPRED are depicted. MT1 (L760A/D761A/I762A) residues are indicated as stars and MT2 (W790A/S791A/S792A) residues are indicated as squares.

The ZRANB3 SRD is sufficient to impart DNA binding, ATPase and fork remodeling activities to the ATPase domain

Since the ZRANB3 ATPase domain by itself is not active, we tested whether addition of the SRD via a flexible linker (Figure 4.4A) is sufficient to impart DNA-dependent activity (73). Indeed, ZRANB3 1-501~720-869 is active in the presence of DNA although its activity is modestly decreased compared to wild-type ZRANB3 (Figure 4.4B and Figure 4.4C). Consistent with its DNA-dependent ATPase activity, ZRANB3 1-501~720-869 is capable of binding complex DNA substrates that mimic a replication fork (Figure 4.4D). Similar to WT ZRANB3, ZRANB3 1-501~720-869 also catalyzes fork regression and fork restoration reactions, whereas an ATPase-deficient mutant (K163D) is inactive in these assays (Figure 4.5A and Figure 4.5B).

Incorporation of the MT1 mutations into this minimal enzymatic unit (1-501~720-869-MT1) inactivates its ATPase, DNA binding, and fork remodeling activities (Figure 4.6A-4.6D). Thus, we conclude that amino acids 720-869 of ZRANB3 encode a SRD that is necessary and sufficient to impart DNA binding, ATPase and *in vitro* fork remodeling activities onto the ZRANB3 motor domain. Furthermore, this analysis defines the minimal enzymatic unit of ZRANB3 capable of catalyzing fork remodeling as containing amino acids 1-501 and 720-869.

The ZRANB3 SRD is required for structure-specific endonuclease activity

In addition to catalyzing fork remodeling reactions, ZRANB3 was reported to act as an ATP-dependent, structure-specific endonuclease that nicks the duplex DNA of a splayed arm substrate (66). Endonuclease activity required both the HNH and ATPase domains (66). Thus, we hypothesized that the SRD domain of ZRANB3 may also be

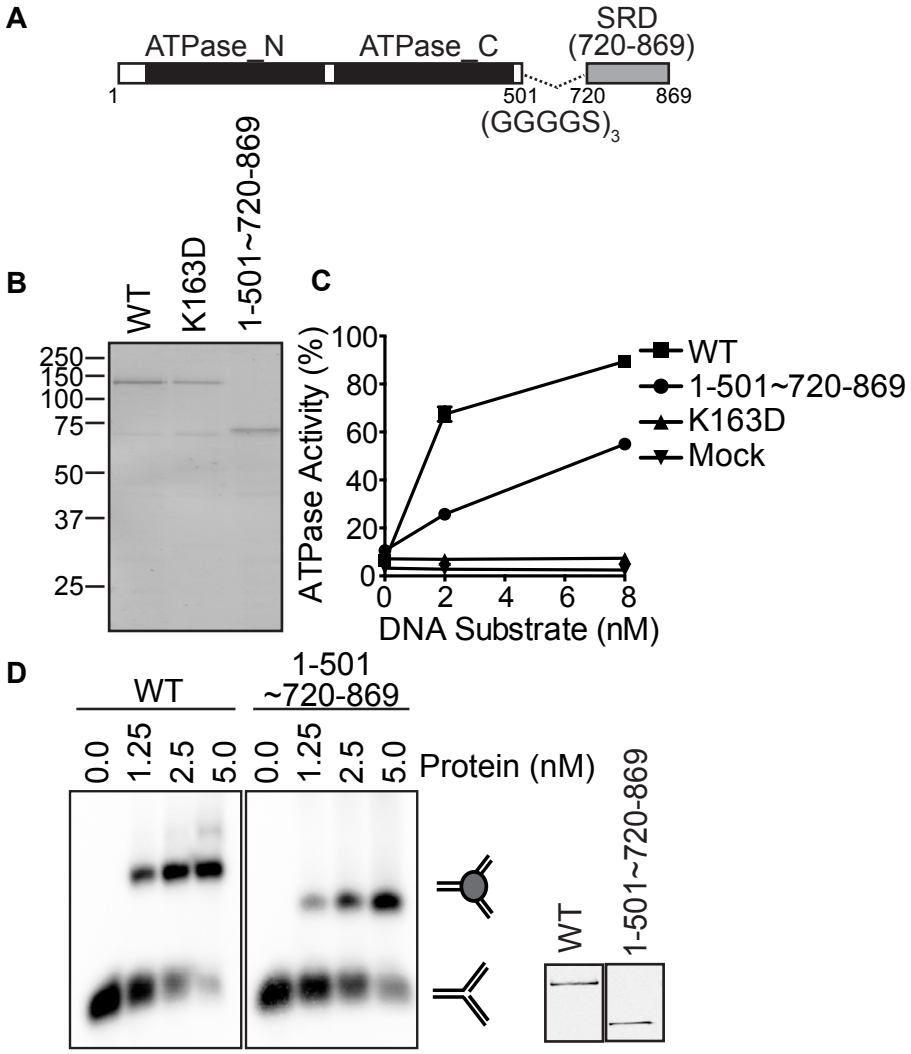


FIGURE 4.4 ZRANB3 1-501~720-869 represents a minimal enzymatic unit that retains DNA binding, ATPase, and fork remodeling activities (A) Schematic of ZRANB3 1-501~720-869. (B) Coomassie-stained SDS-PAGE gel of WT, K163D and ZRANB3 1-501~720-869. (C) ATPase activity of ZRANB3 1-501~720-869 to a splayed arm DNA substrate. The mean and standard deviation for three experiments are shown. In most cases the standard deviation is smaller than the symbol size. (D) DNA binding activity of ZRANB3 1-501~720-869 to a replication fork mimicking DNA substrate. Samples were resolved on a poly-acrylamide gel and visualized by autoradiography. The inset is an anti-FLAG immunoblot of the purified proteins.

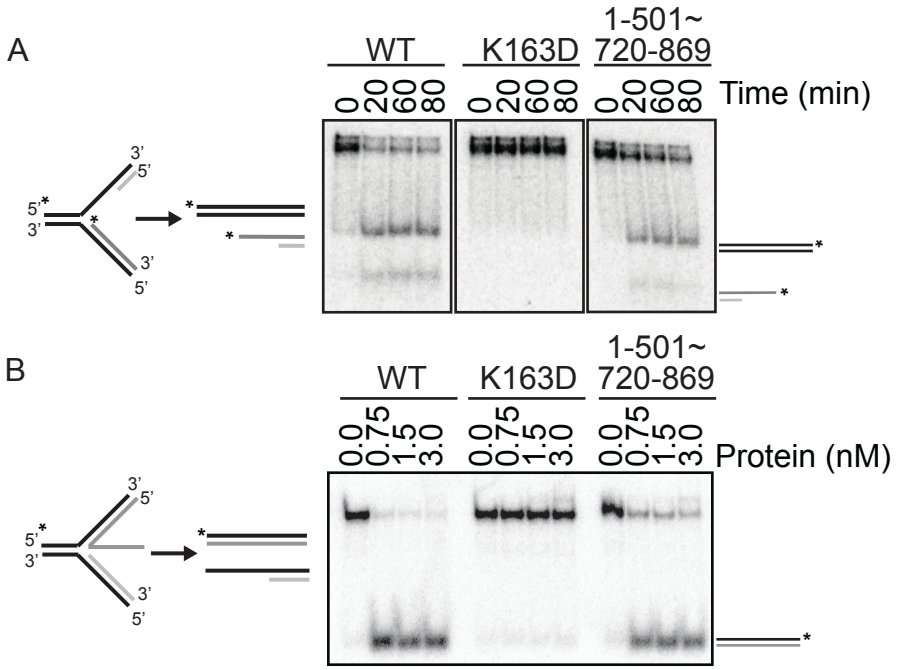


FIGURE 4.5 ZRANB3 1-501~720-869 retains fork remodeling activities (A) Purified WT ZRANB3, K163D and ZRANB3 1-501~720-869 were incubated with a model stalled fork DNA substrate for increasing times. Reaction products were separated by gel electrophoresis and visualized by autoradiography. (B) Increasing amounts of purified WT ZRANB3, K163D and ZRANB3 1-501~720-869 were incubated with a model fork restoration substrate for 30 minutes at room temperature. Reaction products were separated by gel electrophoresis and visualized by autoradiography.

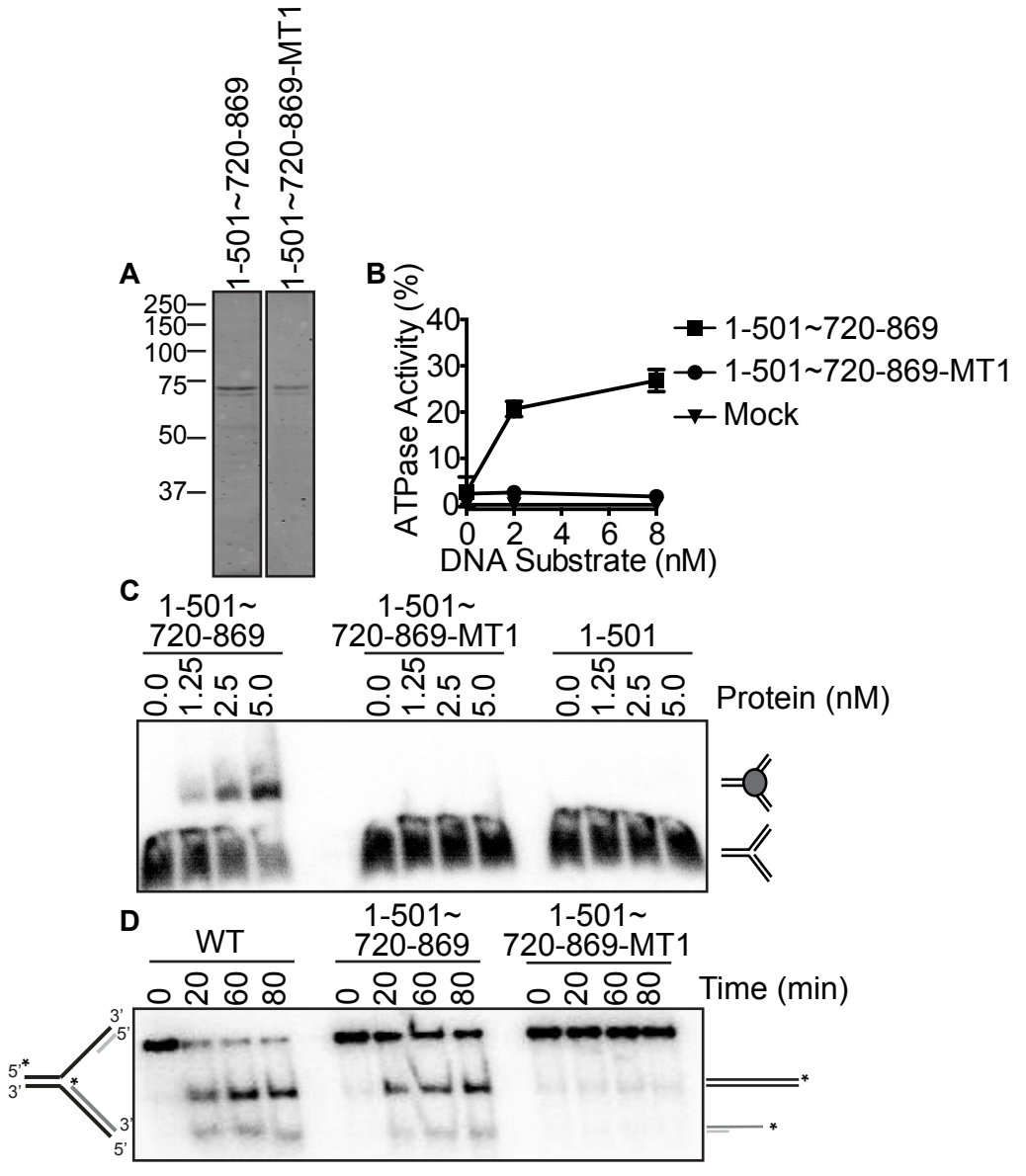


FIGURE 4.6 The substrate recognition domain is necessary and sufficient to impart DNA binding, ATPase, and fork remodeling activities to the ZRANB3 ATPase domain (A) Coomassie stained SDS-PAGE gel of purified 1-501~720-869 and 1-501~720-869 containing MT1 mutations. (B) ATPase activity of ZRANB3 1-501~720-869 and 1-501~720-869-MT1. The mean and standard deviation for three experiments are shown. (C) DNA binding of 1-501~720-869, 1-501~720-869-MT1, and 1-501 to a replication fork mimicking DNA substrate. Products were separated by gel electrophoresis and visualized by autoradiography. (D) Fork regression activity of full length ZRANB3, 1-501~720-869 and 1-501~720-869-MT1 to a model stalled fork DNA substrate for increasing times. Reaction products were separated by gel electrophoresis and visualized by autoradiography

required for its nuclease activity. Indeed, mutations in the SRD inactivate nuclease activity (Figure 4.7A).

The ZRANB3 SRD binds DNA

To test if amino acids 720-869 in ZRANB3 actually contain a DNA binding domain, we purified recombinant GST-720-869 from *E. coli* (Figure 4.8A). Like full-length ZRANB3, GST-720-869 is not capable of binding either single-stranded or double-stranded DNA (Figure 4.8B). However it can bind a splayed arm substrate, albeit with reduced affinity compared to full-length ZRANB3 (Figure 4.8C and Figure 4.8D). Incorporating the MT1 mutations into either GST-720-869 or full-length ZRANB3 greatly reduced their ability to bind the splayed arm DNA substrate (Figure 8C and Figure 8D).

Overall these results indicate that ZRANB3 amino acids 720-869 contains a domain that is both necessary and sufficient to impart substrate-selective DNA binding and enzymatic activity to the ZRANB3 ATPase domain. Thus, it acts as a SRD similar to the HARP domain of SMARCAL1 and the HIRAN domain of HLTF.

Discussion

In this study, we identified a structure recognition domain (SRD) in ZRANB3 that binds branched DNA substrates and confers DNA-dependent ATPase and fork remodeling activity to its SNF2-type motor domain. The SRD is also required for structure-specific endonuclease activity. A minimal enzymatic unit, containing only the SRD and the SNF2 ATPase domains, retains similar fork remodeling activities as the full-length protein. Thus, these data suggest that ZRANB3 shares a similar mechanism of action as SMARCAL1 and HLTF and support the idea that structure recognition domains impart fork remodeling activities onto the motor domains of these proteins.

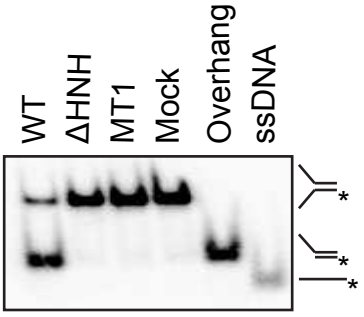


FIGURE 4.7 The substrate recognition domain is necessary for ZRANB3 nuclease activity. Nuclease assay of WT, MT1, and ΔHNH ZRANB3 proteins to a splayed arm substrate. Overhang DNA and ssDNA were used as size markers. Products were separated by gel electrophoresis and visualized by autoradiography.

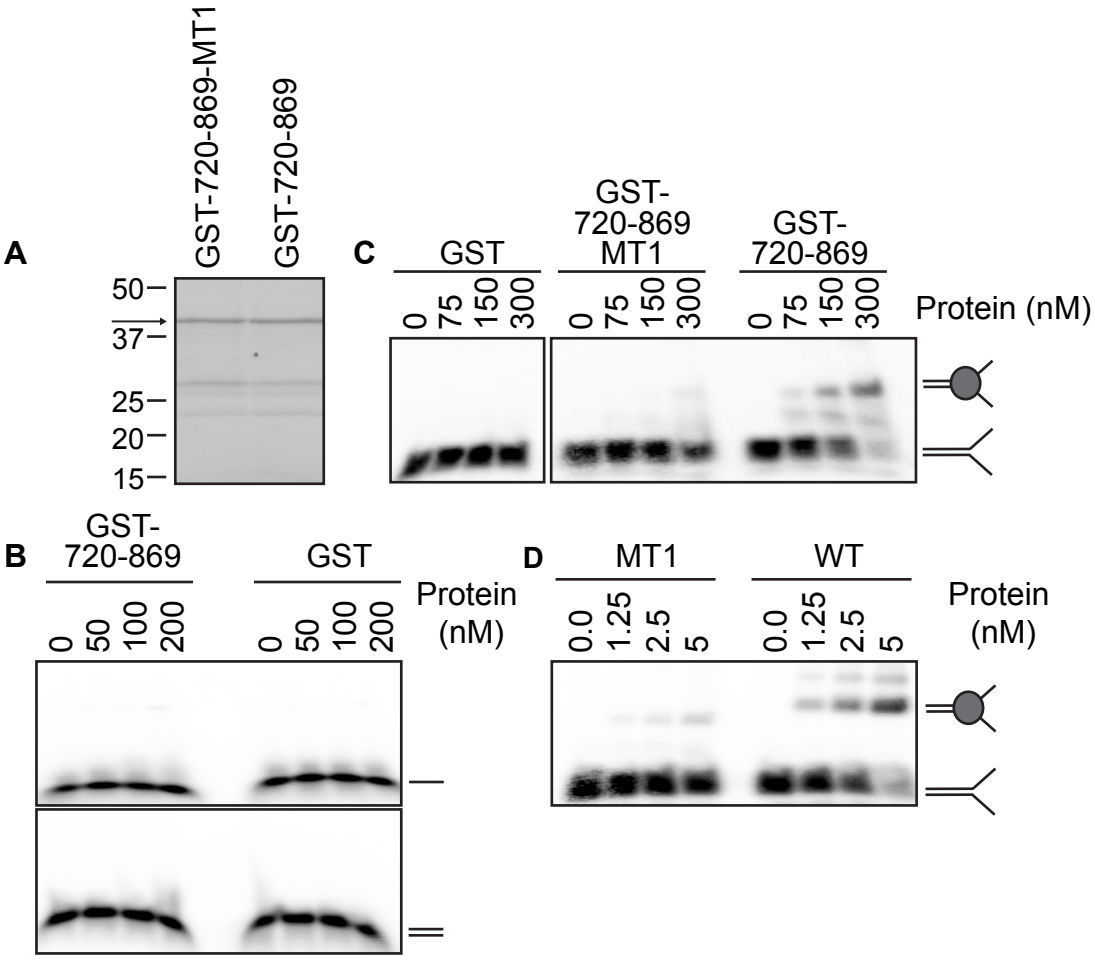


FIGURE 4.8 ZRANB3 amino acids 720-869 encode a structure-specific DNA binding domain. (A) Coomassie-stained SDS-PAGE gel of GST-720-869-MT1 and GST-720-869. (B and C) DNA binding of GST-720-869, GST-720-869-MT1, and GST proteins to (B) single or double-stranded DNA or a (C) splayed arm DNA substrate was assessed by a electrophoretic mobility shift assay. (D) DNA binding assay of full-length WT and MT1 proteins to a splayed arm DNA substrate. Representative experiments are shown.

Why cells express several different fork remodeling proteins that catalyze similar reactions *in vitro* is unknown. The exact DNA structure that is formed at a stalled replication fork in a cell is also unknown. Presumably, the stalled fork adopts a multitude of structures dependent upon the nature of the obstacle. This heterogeneity may warrant the need for several fork remodeling enzymes with substrate recognition domains that interact and bind to different DNA forms present at a stalled and/or regressed fork.

The high-resolution structures of the SRDs of HLTF and SMARCAL1 (HIRAN and HARP domains respectively) have been determined by X-ray crystallography (50,51,63). The HIRAN domain structure includes DNA, and explains its binding preference for duplex DNA with a short 3' single-stranded DNA overhang (51). The HARP domain structure did not include DNA, but it resembles domains in other proteins that bind distorted DNA structures (63). SMARCAL1 prefers to bind DNA structures that contain at least five nucleotides of ssDNA (53), and point mutants in the HARP2 domain impair the ability of the SMARCAL1 catalytic domain to bind branched DNA structures (53,63). Whether the HARP domain recognizes the fork junction itself is unknown. Our data indicate that the ZRANB3 SRD does bind forked DNA on its own although with significantly lower binding affinity than when it is attached to the motor domain. Most likely the SRD recognizes the fork structure and the motor binds the duplex DNA. As it hydrolyzes ATP, it can act to displace the nascent strands while re-annealing the parental strands. The SRD may act at the junction to facilitate this reaction.

Our data are inconsistent with the results from Yuan and colleagues (67). While the SRD we identified overlaps their HARP-like domain, we find that ZRANB3 is unable to bind forked DNA structures without this region and also lacks DNA-dependent ATPase activity. Also, we found that mutations in the SRD inactivated ZRANB3 endonuclease activity as would be predicted if the SRD were required for DNA binding. We do not

know why Yuan and colleagues were able to observe both DNA binding and ATPase activity in their mutant protein; however, we note that other mammalian DNA-dependent ATPases could have contaminated their protein purifications.

This study extends our understanding of how ZRANB3 operates as a fork remodeling enzyme and determines the necessary components to carry out its enzymatic activities. Future high-resolution structural analyses of the ZRANB3 and SMARCAL1 proteins bound to DNA will be useful to better understand how their SRDs provide specificity to their fork remodeling activities.

CHAPTER V

CHARACTERIZATION OF ZRANB3 ENDONUCLEASE ACTIVITY

Introduction

Nuclease Overview

Nucleases function during DNA replication to process lagging strand Okazaki fragments and during the replication stress response to fix damaged, mismatched, and highly toxic double stranded DNA breaks. Nucleases cleave phosphodiester bonds, generating a 5' terminal phosphate group and a 3' terminal hydroxyl group. Exonucleases cleave nucleotides present at DNA ends in a 5'-3' or 3'-5' direction and endonucleases cleave internal phosphodiester bonds. Interestingly, some nucleases function as dual endo- and exo- nucleases; this functional overlap enables participation in multiple DNA replication and repair processes (89).

HNH domain

The HNH endonuclease domain has been identified in over 1000 proteins throughout the prokaryotic and eukaryotic life kingdoms (90,91). Crystal structures generated from the bacteriophage homing endonuclease IHmu-I and polypeptide bacteria toxin ColE7, show that the HNH endonuclease domain adopts a consensus "ββα-metal" motif (92). As the name suggests, the HNH domain contains a conserved histidine-asparagine-histidine sequence motif. During an endonuclease reaction, the metal ion is positioned to stabilize the creation of the phosphoanion once the phosphodiester bond is cleaved. The conserved asparagine and C-terminal most histidine residue coordinates the bound metal ion. The N-terminal histidine residue acts

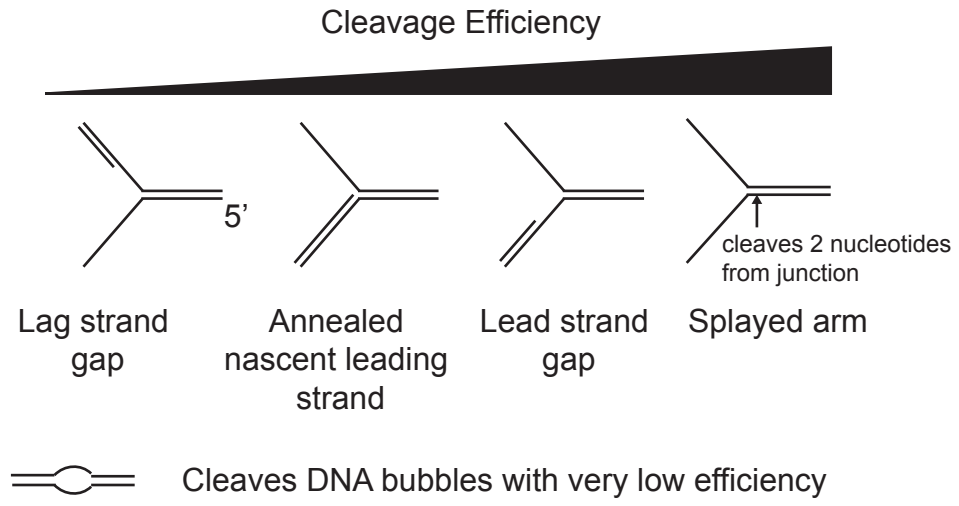
as a general base and creates an active hydroxyl group that initiates the endonuclease reaction (92).

ZRANB3 is a structure specific endonuclease

Yusufzai and collaborators were the first to report that ZRANB3 contains a conserved HNH endonuclease domain (64). To determine whether ZRANB3 acts as an endonuclease, they tested endonuclease activity to isolated genomic DNA from *E.coli* and used KpnI, a sequence specific endonuclease also containing a HNH endonuclease domain, as a positive control. No detectable endonuclease activity was observed in the reactions containing ZRANB3. They concluded that endonuclease activity possibly requires specific reaction conditions. Multiple years later, Weston *et al*. showed that ZRANB3 acts as a structure specific ATP-dependent endonuclease that cleaves the parental leading strand two nucleotides within the duplex DNA portion of the annealed template strands (Figure 5.1). Specifically, ZRANB3 optimally cleaves a splayed arm DNA substrate, and cleaves with less efficiency, substrates containing a lagging strand gap. ZRANB3 fails to cleave substrates containing a nascent lagging strand annealed at the junction (Figure 5.1). Weston *et al*. hypothesized that ZRANB3 endonuclease activity functions to remove damaged DNA on the parental leading strand template that cause replication fork stalling (66). This model has yet to be tested and is further discussed in Chapter 6.

In the presence of ATP, ZRANB3 acts as a fork reversal and endonuclease enzyme *in vitro*. This raises a major issue of whether ZRANB3 catalyzes both fork remodeling and endonuclease activity in cells and how these two activities are regulated. To date, no studies have been done to determine the impact of the HNH endonuclease and fork remodeling activities on ZRANB3 function in cells. To begin to answer these

A Endonuclease Substrates



B Substrates that ZRANB3 fails to cleave

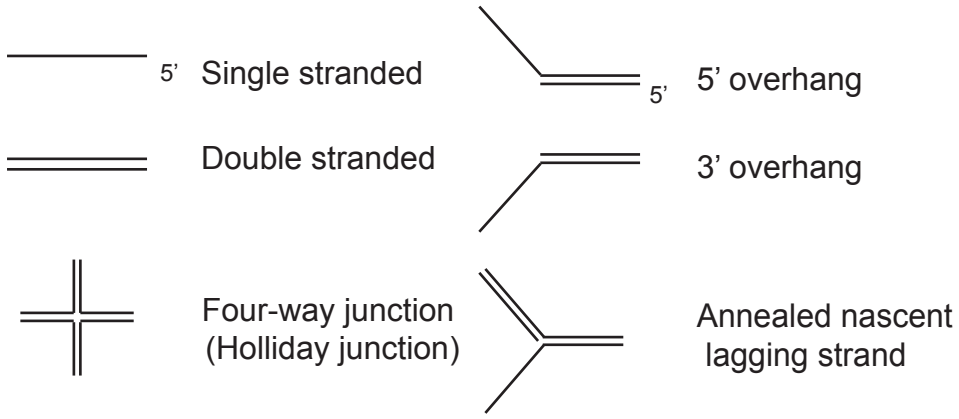


Figure 5.1 ZRANB3 Endonuclease and non-endonuclease substrates. (A) As determined by Weston *et al*, (66) ZRANB3 cleaves, with increasing efficiency, lag strand gap, annealed nascent leading strand, lead strand gap, and splayed arm substrates. (B) DNA substrates that ZRANB3 fails to cleave.

questions further *in vitro* characterization of ZRANB3 endonuclease activity in the context of fork remodeling is required. In this chapter, I re-validated the endonuclease results obtained by Weston *et al.* and developed assays to detect endonuclease products that form in fork remodeling reactions. Hopefully this data will serve as a foundation for a ZRANB3 project to further characterize its endonuclease and fork reversal activity *in vitro*, and support future work to discern how these activities contribute to its functional purpose in cells.

Results

ZRANB3 cleaves splayed arm and lag gap regression substrates

Weston *et al.* and Yusufzai *et al.* tested and observed opposing ZRANB3 endonuclease results (64,66). Therefore, before proceeding with my studies, I tested endonuclease activity following the procedure outlined in the Weston *et al.* study (66). In my hands, ZRANB3 catalyzes ATP-dependent endonuclease activity to a splayed arm substrate (Figure 5.2A-B). ZRANB3 cleaves a splayed arm substrate under fork remodeling buffer conditions (Figure 5.3 A-B). Therefore assays testing fork remodeling activity to endonuclease compatible DNA substrates may contain endonuclease products (Figure 5.3 A-B). To detect fork remodeling and endonuclease products within the same reaction I tested ZRANB3 endonuclease activity to a lag strand gap fork regression substrate and resolved the reaction under denaturing conditions that yield single stranded DNA substrates. Based on where ZRANB3 was shown to cut a splayed arm substrate, if ZRANB3 cleaves the lag strand gap substrate I expect to resolve, under denaturing conditions, single stranded 36 nucleotide long products in a regression reaction (Figure 5.4A). I detected a single stranded product slightly larger than 30 nucleotides (Figure 5.4B). This suggests that every published ZRANB3 fork

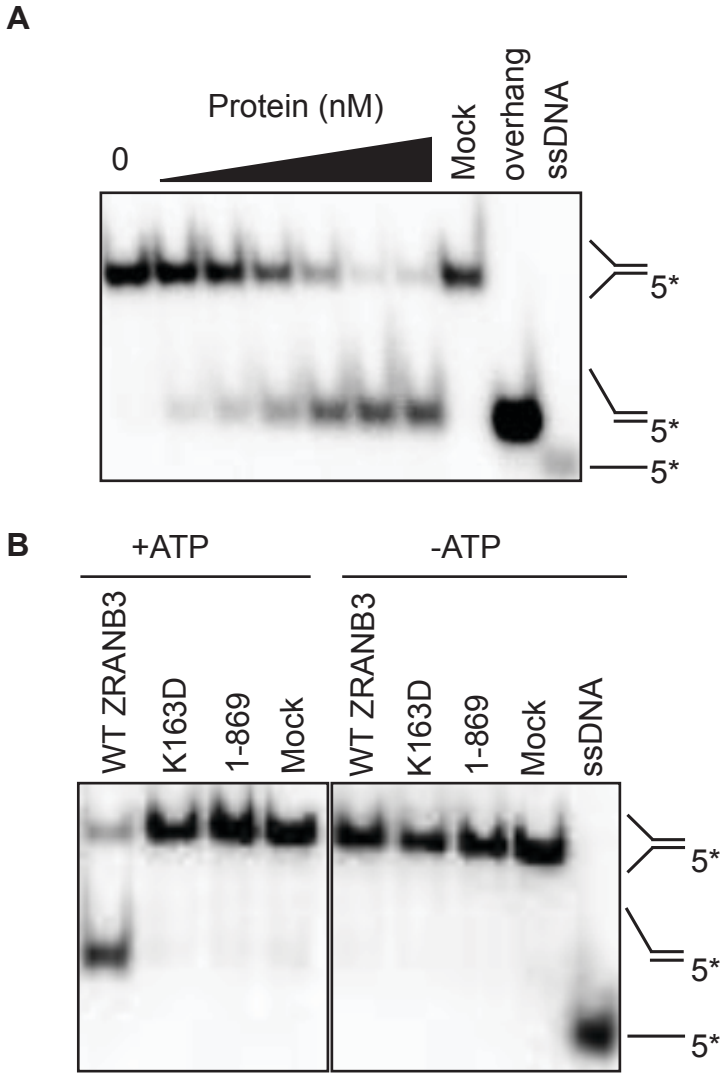


Figure 5.2 ZRANB3 endonuclease activity validation. (A) Increasing concentrations of purified ZRANB3 was incubated with 10uM of a 5' ^{32}P labeled splayed arm substrate. (B) Purified WT, K163D, 1-869 were incubated with 10uM of 5' ^{32}P labeled splayed arm substrate in the presence and absence of 2mM ATP. Products were resolved on an 8% polyacrylamide gel under native conditions.

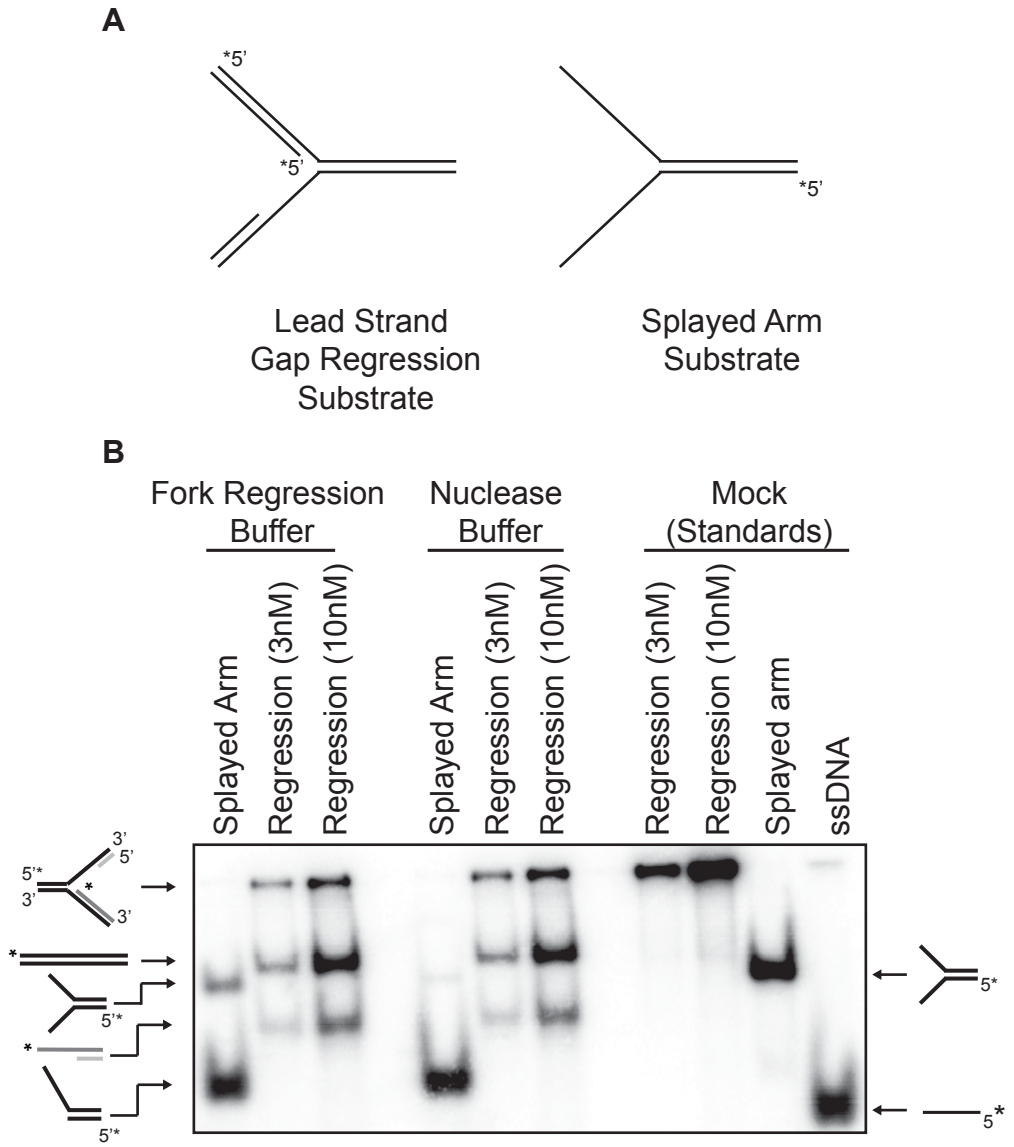


Figure 5.3 ZRANB3 endonuclease activity test in standard fork regression assay conditions. (A) Lead strand gap substrate and splayed arm substrate used in (B). (B) To test whether ZRANB3 catalyzes endonuclease activity to a splayed arm substrate under standard fork regression assay conditions, 3nM ZRANB3 was incubated with either a 5' ^{32}P labeled splayed arm (10nM) or a ^{32}P labeled lead strand gap regression substrate (3nM, and 10nM) in reaction buffer used to conduct a fork regression assay. On the left side of the graph is a depiction of the lead gap replication fork substrate, regression products, splayed arm nuclease substrate and the nuclease product. To the right of the graph is a depiction of the splayed arm and single stranded DNA markers. Products were resolved on a 8% polyacrylamide gel under native conditions.

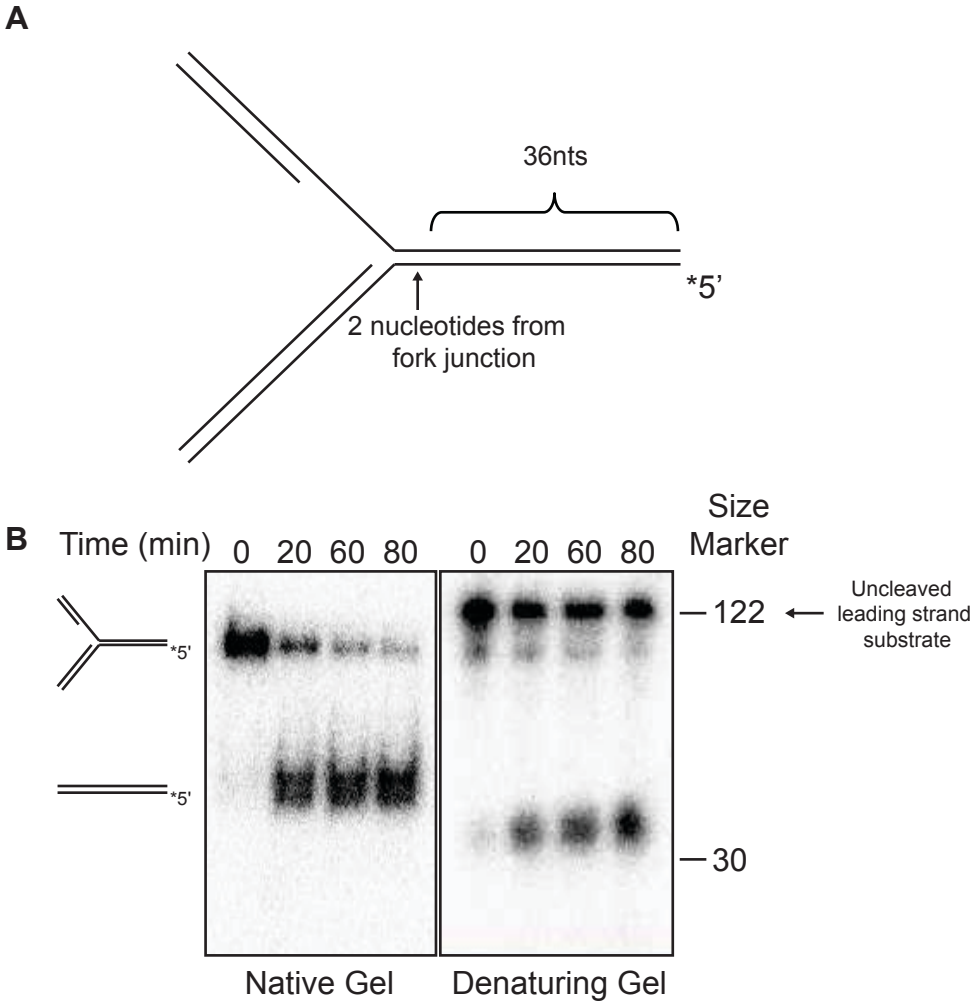


FIGURE 5.4 ZRANB3 cleaves a lag strand gap regression substrate. (A) Depiction of where ZRANB3 would cleave a lag strand gap regression substrate. The expected size of the nuclease product is approximately 36 nucleotides (nts) in length. (B) Purified ZRANB3 was incubated with a lag strand gap regression substrate, for increasing times. Half of the reaction was resolved under native conditions (left) and the other half was resolved under denaturing conditions (right) The arrow points to the uncleaved parental strand oligonucleotide.

regression assay to a compatible DNA substrate may contain products formed due to its endonuclease activity.

ZRANB3 may cleave DNA substrates at more than one location

ZRANB3 was proposed to cut a forked DNA substrate once (66). To test whether the endonuclease product formed is of one size, I resolved the products formed from a regression reaction to a lag strand gap substrate on a large 30 x 40 cm (13.7 x 15.7 inch) 10% polyacrylamide 8M Urea gel. The oligonucleotides used to assemble the lag strand gap regression substrate were PAGE purified to ensure that the oligonucleotides used to assemble the lag strand gap substrate are of the same length. The products of the regression reaction contained several oligonucleotides between 30-50 nucleotides in length (Figure 5.5). Weston *et al.* tested ZRANB3 endonuclease activity to DNA substrates that cannot be regressed since the oligonucleotides used to assemble the splayed fork with an annealed nascent leading or lagging oligonucleotides are not complimentary. Therefore, the products formed from the lag strand gap regression/endonuclease reaction in Figure 5.5 may be due to its ability to regress the substrate. To test this, I assembled and purified a splayed arm substrate with an annealed nascent leading strand and performed an endonuclease assay to this substrate and a splayed arm substrate (the substrate originally tested in the Weston *et al.* publication) and resolved the products on a large 10% polyacrylamide 8M Urea gel. Surprisingly, in reactions with either the splayed arm substrate alone or the splayed arm with the annealed nascent leading strand oligonucleotide, contained products of multiple sizes (Figure 5.6).

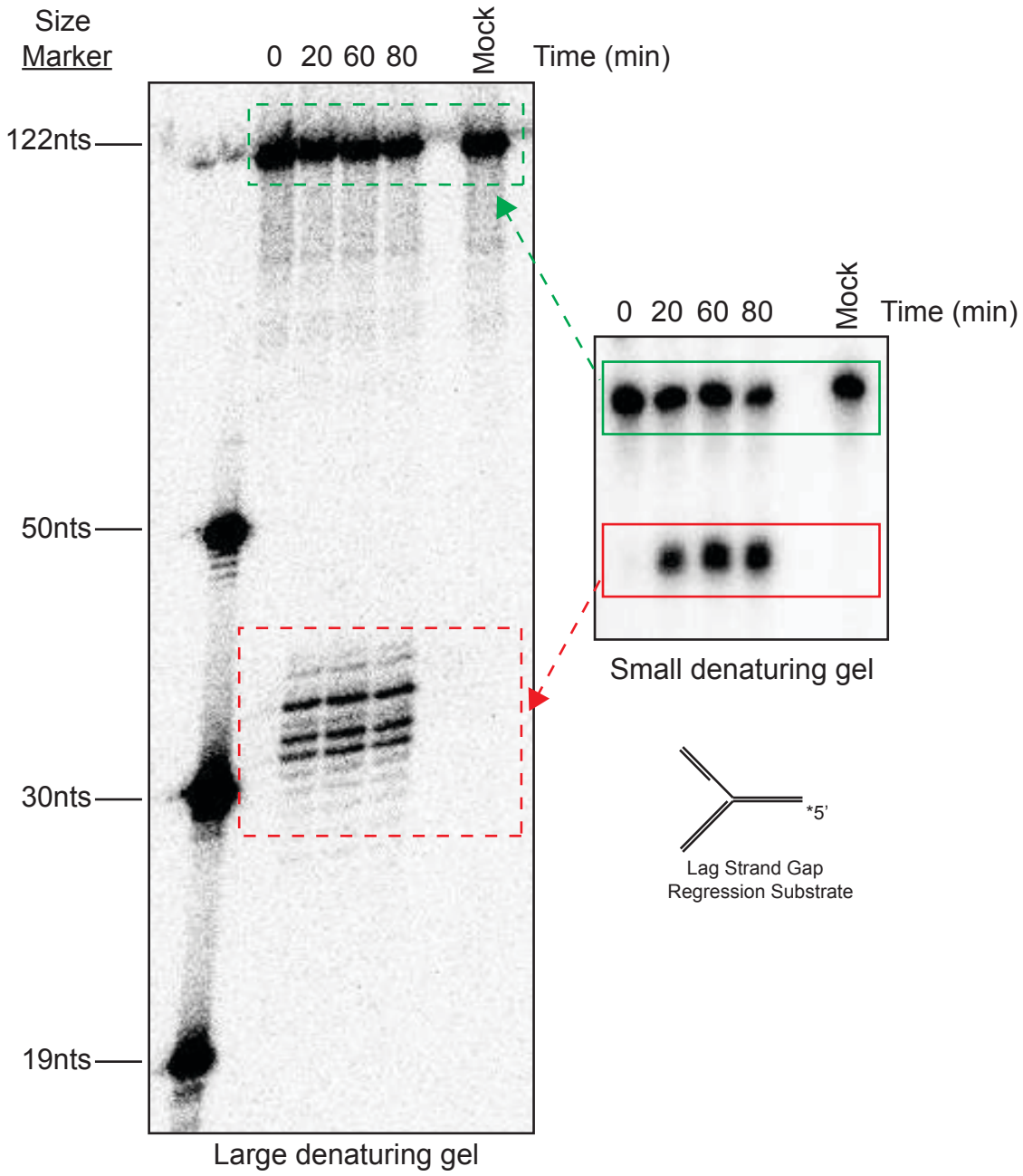


FIGURE 5.5 Fork regression/endonuclease assay yeilds several oligonucleotides of various lengths. Purified ZRANB3 was incubated for various times with a lag strand gap regression substrate and the products were resolved on a small and large 10% polyacrylamide 8M Urea gel.

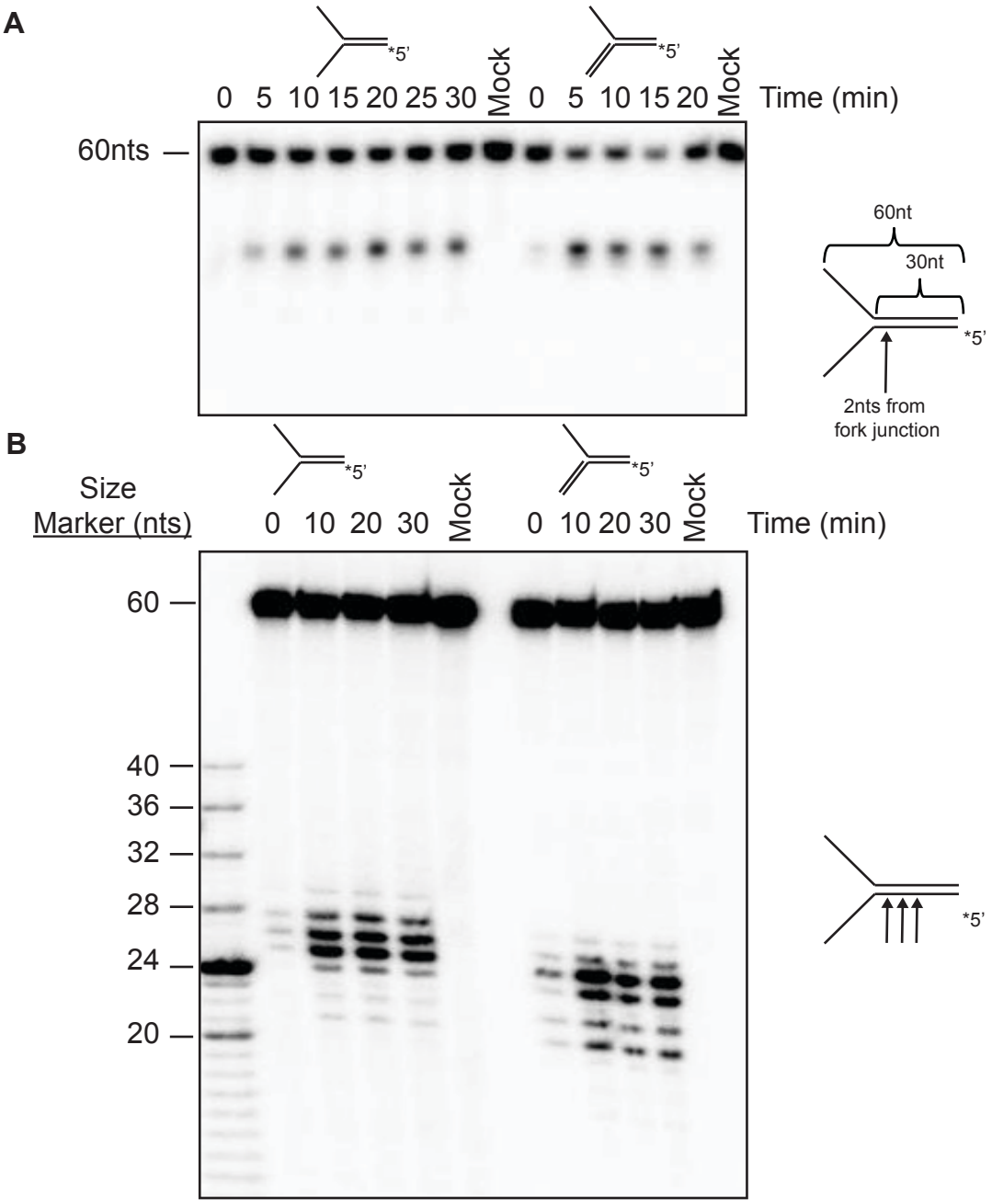


FIGURE 5.6 DNA substrates that ZRANB3 cannot regress yeild nuclease products of multiple lengths. (A) Endouclease assay to a splayed arm substrate with an annealed nascent leading strand and a splayed arm substrate. To the right of the gel is a depiction of where ZRANB3 is hypothesized to cleave. The nuclease reaction should yeild products aproximately 28nts in size. Products are resolved on a small 10% polyacrylamide 8M Urea gel under denaturing conditions (B) Products from (A) resolved on a large 10% polyacrylamide 8M Urea gel under denaturing conditions. To the right of the gel is a depiction of a splayed arm substrate and where ZRANB3 is cutting to obtain products less than 28 nucleotides in size.

Discussion

ZRANB3 most efficiently catalyzes endonuclease activity to substrates containing singled stranded DNA at the junction, in other words, substrates lacking annealed nascent leading and lagging strand oligonucleotides (66). *E.coli* genomic DNA is circular and contains one origin of replication, and upon origin firing two replication forks are created that contain transient single stranded DNA at the fork junction due to helicase mediated unwinding of the parental strands (1). Presumably, ZRANB3 can catalyze endonuclease activity at the replication fork junction. Therefore, one of many possibilities to explain why Yusufzai *et al.* failed to detect nuclease activity is that the isolated genomic DNA from *E.coli* used in their study may have contained low levels of DNA being replicated at the time it was harvested that also contained single stranded DNA regions adjacent to the fork junction (64). More generally, the lack of compatible endonuclease substrates in the isolated genomic DNA from bacteria is a plausible reason for why they failed to detect endonuclease activity.

Analysis of proteins containing a SNF2 ATPase and HNH domains

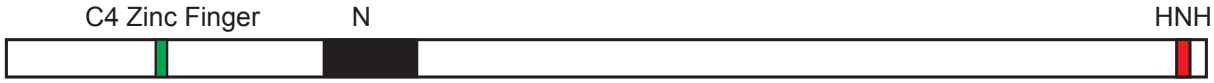
Similar to what was observed by Weston *et al.*, ZRANB3 endonuclease activity requires ATP, which is quite curious especially since other proteins containing HNH domains, like the restriction endonuclease Kpn1, do not require ATP to catalyze endonuclease activity (64,93). There are 53 different types of protein architecture groups encompassing the HNH family of proteins and of those, 5 groups of proteins contain a SNF2-like ATPase domain (Figure 5.7) (94). A part from the two groups containing ZRANB3-like orthologs, the other 3 groups are of functionally uncharacterized proteins. The uncharacterized proteins in *Emiliania huxleyi*, an ocean dwelling plankton and *Toxoplasma gondii*, a single celled parasite, only contain the N-terminal SNF2 domain.

Blind cave fish (ZRANB3-like; lacks NZF motif): 1092 Amino Acids

Human (ZRANB3): 1079 Amino Acids

Emiliania huxleyi (Uncharacterized Protein): 1086 Amino Acids

Neospora caninum (Uncharacterized Protein): 3651 Amino Acids



Toxoplasma gondii (Uncharacterized Protein): 3776 Amino Acids

Figure 5.7 HNH and SNF2 domain containing proteins obtained from the Pfam protein collection database. Depicted are domain layouts of representative proteins from groups identified to contain both an HNH and a SNF2-like ATPase domain. “N” and “C” denote the N and C terminal halves of the ATPase domain. Proteins and domains are drawn to scale. The APIM and PIP motifs are not annotated in the Pfam database for human ZRANB3 and the Blind cave fish ZRANB3-like protein.

The HNH domain containing proteins in *Neospora caninum*, a domestic and farm animal specific parasite, and *Toxoplasma gondii* have a topoisomerase DNA binding C4 zinc finger domain. Mining and studying these proteins may help explain why the HNH domain in ZRANB3 requires ATP to cleave DNA (91,94,95). Whether ZRANB3 requires ATP binding or ATP hydrolysis to catalyze its endonuclease activity is unknown and requires further analysis. Presumably, ATP binding and/or hydrolysis facilitates a conformational change within the protein which may allow the HNH domain to access and cleave the DNA substrate. Another hypothesis is that ATP hydrolysis also alters the structure of the DNA substrate making it more compatible for cleavage.

Analysis of Weston et al. study

Since ZRANB3 requires ATP to catalyze both fork regression and endonuclease activities, it is unsurprising that both regression and endonuclease products are produced within a fork regression assay. The products of a regression assay are annealed oligonucleotides and only under denaturing conditions was I able to detect the endonuclease products. ZRANB3 is incapable of cleaving DNA substrates containing a nascent lagging strand oligonucleotide annealed at the junction which suggests that, in a cell, ZRANB3 may not cleave a regressed stalled fork (66). Therefore, I did not test endonuclease activity to a regressed chicken foot substrate. ZRANB3 cleaves at multiple locations regardless of whether it is capable of remodeling the substrate or not (Figure 5.5 and Figure 5.6). However, the endonuclease products are approximately the size of what I would expect if ZRANB3 cleaves in the location characterized by Weston *et al.* (two nucleotides 5' from the junction on the leading parental strand) (66).

The endonuclease activity observed by Weston *et al*., specifically the analysis to determine the size of the cleaved product, and their denaturing gels testing

endonuclease activity to a splayed arm substrate clearly shows that the product of those endonuclease reactions are composed of several oligonucleotides of various lengths (66). Since they did not resolve the endonuclease products on a large gel, it is difficult to assess the sizes of the predominant endonuclease products and to determine whether the size of those products are consistent with the splayed arm endonuclease results in Figure 5.6B. Weston *et al*. did not report whether they used PAGE purified DNA oligonucleotides to assemble the substrates tested in their studies. Therefore I cannot rule out the possibility that the endonuclease products of various lengths is due to testing DNA substrates assembled with differently sized oligonucleotides.

Concluding Remarks

The endonuclease reaction in Figure 5.4B did not go near completion as did the fork regression reaction, which may suggest that ZRANB3 is more efficient at regressing than nicking replication fork substrates. The lag gap regression substrate used to detect both regression and endonuclease reaction products contains an annealed nascent lagging strand and ZRANB3 cannot cleave replication fork substrates containing an annealed nascent strand at the junction. Therefore, once ZRANB3 regresses the lag strand gap substrate to the point where the nascent lagging strand reaches the junction, ZRANB3 may no longer be able to cleave the substrate. This may also explain why there is no gradual increase in the endonuclease product overtime.

The predominant endonuclease products formed in the reaction containing the splayed arm substrate with an annealed nascent leading strand oligonucleotide is smaller in size than the endonuclease products obtained in the splayed arm substrate samples (Figure 5.6B). This suggests that ZRANB3 cleaves 5’ from the location that yielded the products in the splayed arm samples. The smaller endonuclease products

may be due to the nascent leading strand oligonucleotide obstructing the site where ZRANB3 cleaves.

In Figure 5.6 A-B, it is difficult to discern differences in the amount of product formed over time in the endonuclease reactions testing either a splayed arm substrate or a splayed arm substrate with an annealed nascent leading arm. Therefore, these experiments will need to be repeated and designed to properly assess the kinetics of its endonuclease activity.

These findings along with the Weston *et al*. study are truly interesting and further characterize ZRANB3 endonuclease activity. More work is required to determine (1) how ZRANB3 catalyzes its endonuclease activity, (2) ZRANB3 endonuclease activity efficiency when compared to its fork regression activity, (3) the substrate limitations of ZRANB3 endonuclease activity, (4) the relevance of ZRANB3 endonuclease activity in resolving stalled replication forks, and (5) how the endonuclease and fork remodeling activities are coordinated. In chapter 6, I outline a series of experiments to further elucidate how ZRANB3 functions as an endonuclease *in vitro* and how its endonuclease and fork remodeling activities impact its function in cells.

CHAPTER VI

DISCUSION AND FUTURE DIRECTIONS

ZRANB3 overview

ZRANB3 contains a functionally conserved substrate recognition domain and binds structured DNA present at stalled replication forks. ZRANB3 binds polyUB PCNA molecules at stalled replication forks, which implicates ZRANB3 functioning in template switching mediated bypass and restart of stalled forks (Figure 6.1) (68). Furthermore, ZRANB3 contains a highly conserved, ATP-dependent, structure specific endonuclease domain. How ZRANB3 couples all of its enzymatic activities to promote fork restart is poorly understood and requires more research to elucidate its actual function in cells (68).

ZRANB3 and disease

SMARCAL1 deficiency causes SIOD, a rare developmental and cancer predisposing disease (77). *HLTF* is silenced in 43% of primary colon cancer tumors; re-expressing HLTF in colon cancer cell lines decreased their proliferative ability suggesting that HLTF acts as a tumor suppressor (87). Bloom, Werner and Rothmund-Thomson syndromes are rare diseases caused by deficiencies in BLM, WRN and RECQ4 respectively. Bloom Syndrome patients suffer from dwarfism, type II diabetes, infertility, and can manifest non-Hodgkins lymphoma, leukemia, breast, colon and skin cancers (28). Werner Syndrome patients suffer from premature aging; develop cataracts, osteoporosis and various bone cancers. Rothmund-Thomson syndrome causes skin atrophy, short stature and primarily bone cancers (28). This is insurmountable evidence that chromatin remodelers are very important in maintaining genome stability and preventing disease.

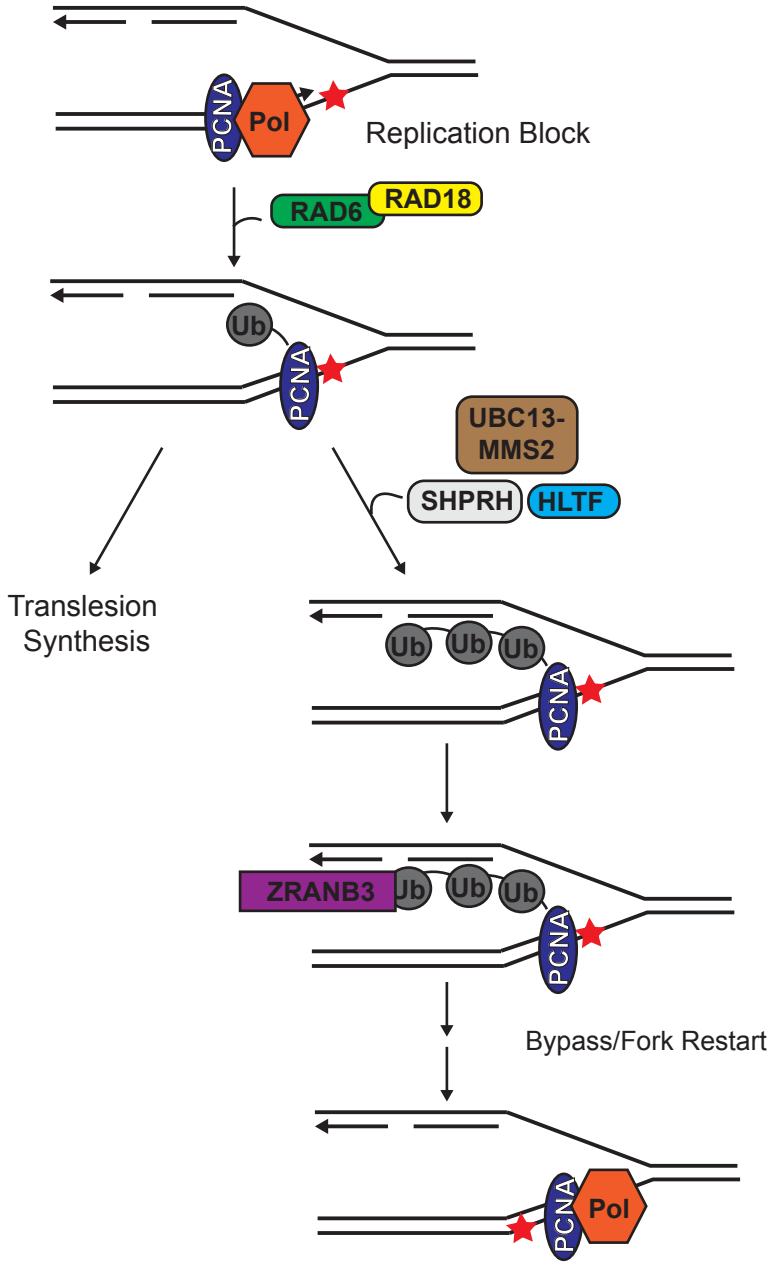


Figure 6.1 A model for ZRANB3 function during the template switching pathway for restarting stalled replication forks. This model was adapted from Mailand *et al*. (68). A replication block (depicted as a red star) stalls the leading strand replisome. Rad6/Rad18 E2-E3 ubiquitin conjugating and ligase complex creates monoUB PCNA molecules at the stalled replication fork. Depending on the nature of the block, translesion synthesis or template switching mediated bypass of the block may be utilized. For template switching in mammalian cells, the monoUB PCNA molecule is polyubiqutinated by the E3 ubiquitin ligase complex Ubc13/Mms2-SHPRH-HLTF. ZRANB3 binds polyUB PCNA and facilitates fork remodeling to bypass the block.

To identify genes associated with cancers that are less frequently mutated yet represent a large portion of mutated genes in tumors (96), Lawrence and collaborators analyzed and compared exome sequencing data from nearly 5000 genes spanning 21 tumor types to sequences obtained from normal tissue samples (88). *ZRANB3* was identified as a potential tumor suppressor gene that is significantly mutated in endometrial cancers (Figure 6.2) (88). In this study, 6 of the 11 ZRANB3 mutations that were identified in the tumor of patients with endometrial cancer are missense mutations located in the ATPase domain. Mutants T66A and K340T affect conserved residues within the SMARCAL1 and SsoRad54 ATPase domains and are specifically located in motifs I and IV respectively (85). As stated previously, the ATPase domain within the SF1 and SF2 superfamily of helicases contain 7 signature motifs involved in ATP binding and transferring the energy obtained from hydrolyzing ATP molecules into helicase and translocase reactions (97). Therefore, these mutants may negatively impact the functionality of the ATPase domain.

A nonsense mutation within the PIP-box motif at residue R523, if expressed in cells, would create a truncated protein containing only the ATPase domain (Figure 6.2). Truncation mutants 1-501 and 1-650 fail to catalyze DNA-dependent ATPase activity and 1-501 is incapable of binding DNA (Figure 4.2 and Figure 4.6). Therefore, I would expect a prematurely truncated 1-523 mutant to also lack DNA binding and ATPase activities. The remaining variants include a splice site mutant at F514, a C-terminal nonsense mutant at residue R947, a missense mutant at residue D1020Y and a frame shift mutant at residue L1040. A thorough analysis is necessary to determine whether ZRANB3 is truly implicated in cancer.

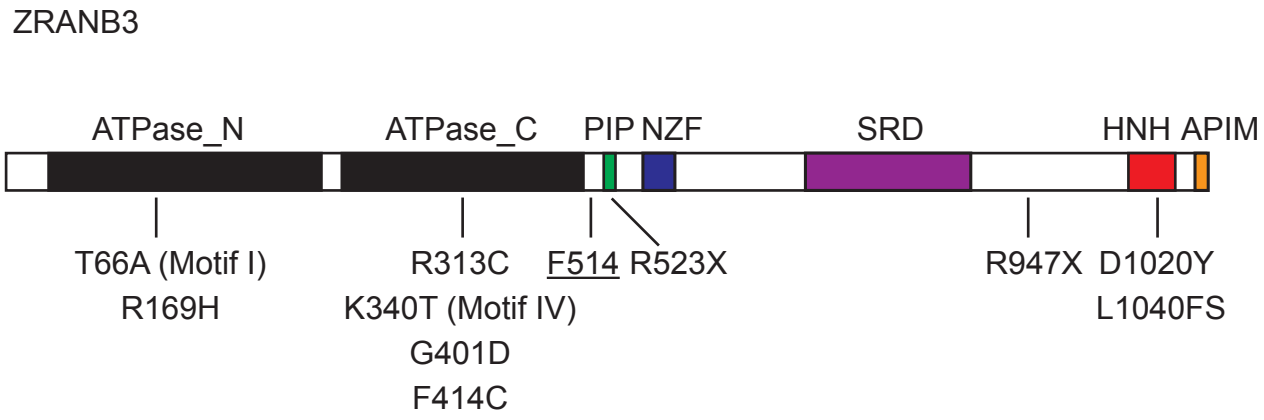


Figure 6.2 Depiction of ZRANB3 mutations identified in endometrial cancers. Mutant T66A is located in motif I of the N-terminal half of the ATPase domain and K340T is located in motif IV of the C-terminal portion of the ATPase domain. F514 is a splice site mutant; R532X and R947X are nonsense mutants; L1040FS is a frame shift mutant.

Substrate recognition domain function in cells

Fiber labeling experiments analyzing replication restart of ZRANB3 deficient cells unequivocally showed a significant decrease in replication fork restart after damage (65,67). Also, dual ΔPIP-ΔAPIM and ΔNZF deletion mutants did not complement the replication restart phenotype in ZRANB3 deficient cells (65). ZRANB3 deficient cells are sensitive to MMS, CPT, HU and have increased sister chromatid exchange events (65-67). ZRANB3 colocalizes with γH2AX at DNA damage induced laser stripes (65). These data suggest that ZRANB3 indeed functions at damaged and/or stalled replication forks marked with polyUB PCNA; how ZRANB3 functions in fork restart is unknown.

There is minimal evidence determining whether ZRANB3 fork remodeling activity impacts its function in cells. Yuan *et al.* showed that the HARP-like deletion mutant was unable to complement the cell sensitivity phenotype observed with treating ZRANB3 deficient cells with HU (67). Even though this result is what I would expect for a ΔSRD mutant, making conclusions based solely on a large deletion mutant is inaccurate. Therefore, testing point mutants that significantly reduce ZRANB3 DNA binding affinity to its DNA substrates in complementation assays assessing cell viability, fork restart and sister chromatid exchange events will yield more definitive and reliable results.

To test whether the substrate recognition domain impacts ZRANB3 function in cells, Dr. Cory Holland, a post doctoral trainee, generated ZRANB3 null U2OS cell lines using CRISPR methodology and we acquired ZRANB3$^{-/-}$ MEFS from Dr. Christine Eischen to develop complementation systems to test point mutants within the substrate recognition domain (Figure 6.3). In my hands, short-term cell proliferation assays assessing sensitivity of the ZRANB3 null MEFs to CPT, cisplatin, and MMS and the U2OS ZRANB3 null cell lines to CPT, cisplatin and HU showed no difference in cell

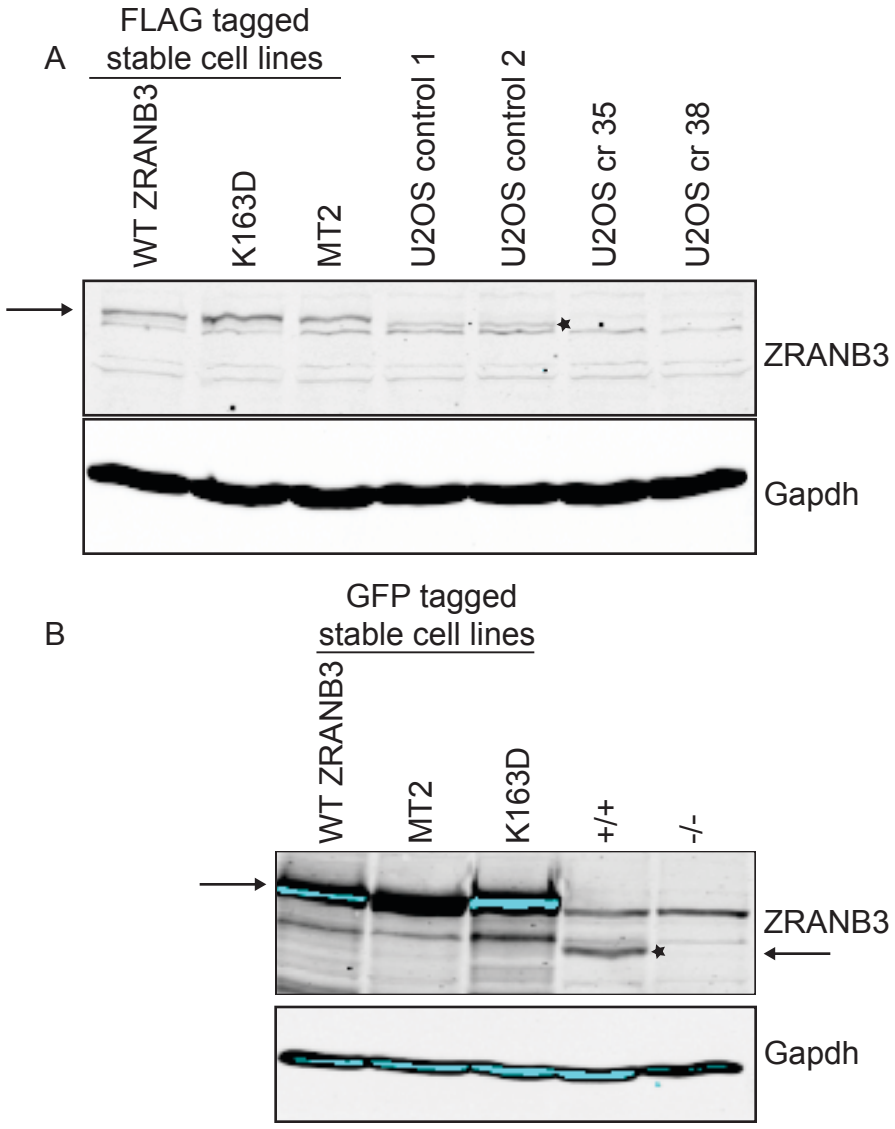


Figure 6.3 Immublots to validate U2OS ZRANB3 CRISPR null cell lines and ZRANB3$^{-/-}$ MEFs. (A) Protein isolated from U2OS cr 35 cells stably expressing WT, K163D or MT2 ZRANB3 protein, U2OS cell control 1, U2OS control 2 and CRISPR cell lines 35 and 38 were used to probe for ZRANB3 expression. The arrow pointing to the right indicates the stably expressing FLAG tagged ZRANB3 protein and the star indicates endogenous ZRANB3 in the U2OS control cell lines. GAPDH was used as a loading control. (B) Protein from MEFs isolated from ZRANB3$^{-/-}$ and WT mice were used to probe for ZRANB3 expression. The arrow pointing to the right indicates the stably expressing GFP tagged ZRANB3 protein which is ~150kDa in size. The ZRANB3 mouse ortholog is ~120kDa in size.

sensitivity between the WT and ZRANB3 null cells. However, testing long term cell proliferation using a colony forming assay showed a marked sensitivity of the ZRANB3$^{-/-}$ MEFs and the ZRANB3 null U2OS cell lines to MMS, which is consistent with what was observed by Weston *et al*. (Figure 6.4). Also, we observed an increase in sister chromatid exchange events in the ZRANB3 null U2OS cell lines treated with CPT when compared to WT U2OS cells; these results are consistent with what was observed by Ciccia *et al.* (65). Although these results were promising, we were unable to complement the cell growth and sister chromatid exchange phenotypes by exogenously expressing WT ZRANB3 in the ZRANB3 null MEFs and U2OS cells lines. As a means to combat this setback, one saying rings true, “less is more”. The stably expressing GFP-ZRANB3 and the FLAG-ZRANB3 constructs in the ZRANB3$^{-/-}$ MEFs and ZRANB3 CRISPR cell lines are expressed at much higher levels than the endogenous ZRANB3 protein and this may interfere with its function in cells. Therefore an alternative expression vector containing a weaker promoter is necessary to lower the expression of ZRANB3 to levels comparable to the endogenous protein. Once a reliable complementation system is created, it will serve as a powerful tool to test the substrate recognition domain mutants.

As stated previously, ZRANB3 localization to stalled replication forks is required for fork restart (65). Also, ZRANB3 retention at DNA damage induced laser stripes is dependent on protein components required for creating polyUB PCNA such as RAD18 and UBC13 (65). ZRANB3 is robustly retained at DNA damage induced laser stripes in the absence of the deubiquitinase (DUB) USP1 (65). USP1 depletion was also shown to enhance the interaction of ZRANB3 with polyUB PCNA molecules (65). These data suggest that USP1 may act to deubiquitinate polyUB PCNA molecules and ZRANB3 localization and retention at stalled forks is due to polyUB PCNA. Therefore, to properly test ZRANB3 function in cells, all complementation systems must ensure that PCNA

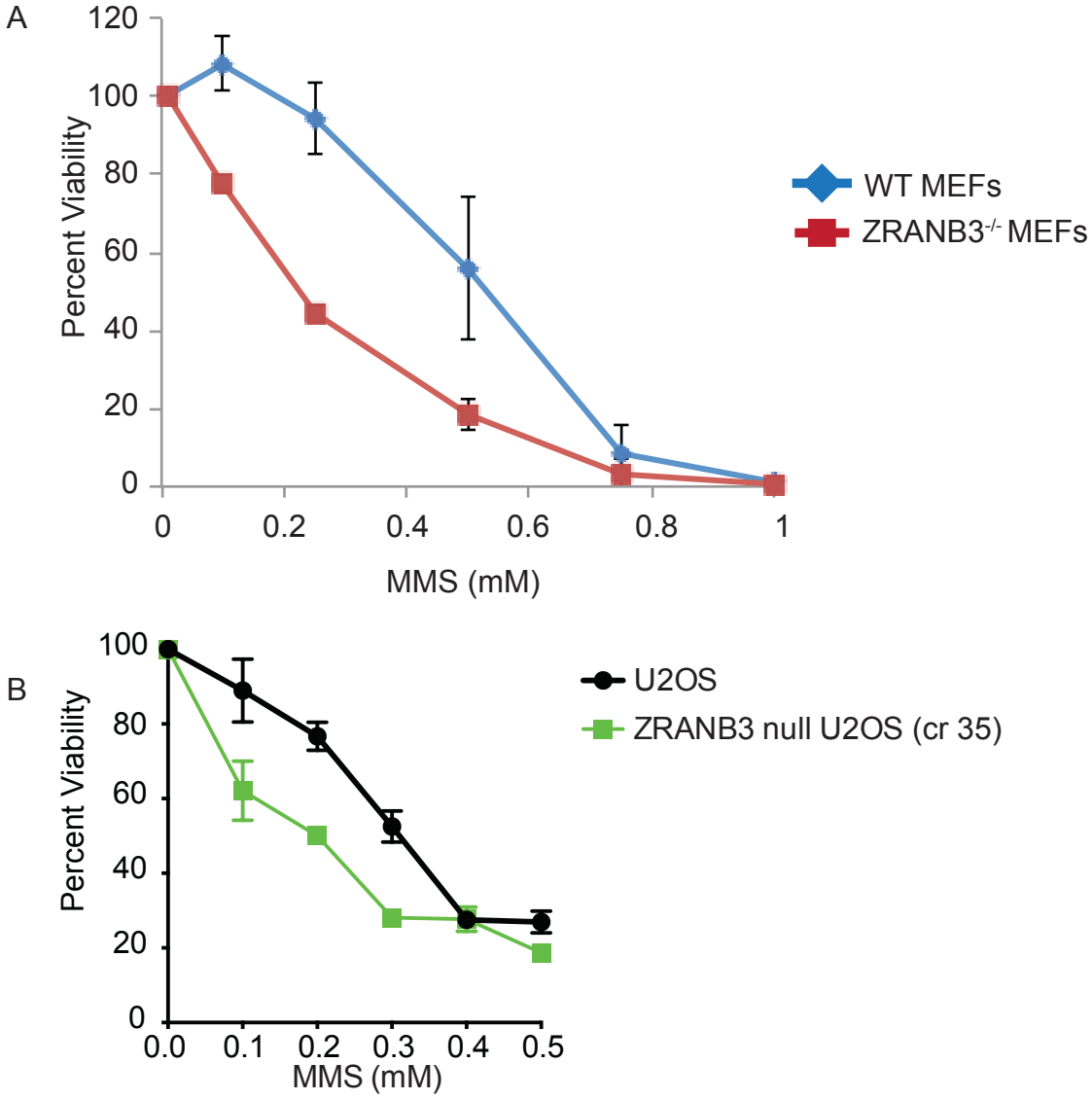


Figure 6.4 Colony forming assay analysis. (A) WT MEFs and ZRANB3$^{-/-}$ MEFs were treated with indicated MMS concentrations in a colony forming cell proliferation assay. (B) U2OS and ZRANB3 null CRISPR U2OS cell lines were treated with indicated MMS concentrations in a colony forming cell proliferation assay.

polyUB is induced under the experimental conditions that will be utilized. We did not check for PCNA polyUB induction in my experiments therefore I cannot rule this out as a reason for why our complementation studies did not work as expected. However, I am confident that future work elucidating the function of the substrate recognition domain using the cell lines and other reagents we have generated will further our understanding of how its DNA binding and fork remodeling activity impacts ZRANB3 function in cells.

Functional relevance of the HNH endonuclease domain

As previously stated in Chapter 5, ZRANB3 contains an evolutionarily conserved, structure specific, endonuclease domain (66). Further *in vitro* and *in vivo* characterization is necessary to build an understanding of how the endonuclease domain affects ZRANB3 function in cells. In the presence of ATP, ZRANB3 catalyzes both endonuclease and regression activities *in vitro* (Figure 5.4). Interestingly, ZRANB3 does not cleave the substrate at one location as originally predicted by Weston *et al.* (Figure 5.5 - Figure 5.6). Therefore, mapping the location where ZRANB3 cleaves and determining whether ZRANB3 catalyzes both regression and cleavage reactions on the same substrate will provide a better understanding of the substrate that ZRANB3 cleaves at a stalled fork in cells.

One method that can be used to map where ZRANB3 cleaves is by comparing the sizes of cleavage products obtained from incubating purified ZRANB3 protein with a splayed arm substrate containing a 5' ^{32}P label on the leading strand oligonucleotide and a splayed arm substrate containing a 3' ^{32}P label on the leading strand oligonucleotide. A method to determine whether ZRANB3 acts to regress and cleave a regression substrate simultaneously is to conduct a time course nuclease assay to a lag strand gap regression substrate that is assembled with a 3' ^{32}P labeled leading strand parental

oligonucleotide. If over time the labeled cleaved product decreases in size, this indicates that as ZRANB3 regresses the substrate, it is cleaving the substrate at the same time.

A clear answer to whether the HNH domain is functionally relevant in cells is lacking. Yuan *et al.* concluded that since a C-terminal deletion mutant lacking the last 80 amino acids, that contains the HNH and the APIM motif, was unable to localize to HU induced RPA2 foci and complement the sensitivity phenotype to treating cells with HU, the HNH domain may be involved in localizing ZRANB3 to sites of stalled forks. However, at the time, they did not know that ZRANB3 contains an APIM motif located C-terminal to the HNH domain. Therefore it is difficult to discern whether the APIM or the HNH domain is responsible for this phenotype. Furthermore, Weston *et al.* observed that the HNH domain may partially function to localize ZRANB3 to UV induced DNA damage sites (66). These observations are further complicated by the Ciccia *et al.* work in that the PIP, NZF and APIM motifs are sufficient in localizing ZRANB3 to laser induced DNA damage sites. Therefore, it is unclear whether the HNH domain functions to facilitate localization and/or cleave structured DNA at stalled forks.

I favor the Weston *et al.* model proposing that the HNH domain functions in cells to cleave the leading template strand of a stalled replication fork to mediate the removal of an obstructive base (Figure 6.5) (66). Following cleavage, the stalled fork is reversed and the DNA flap containing the damaged base could be removed by the flap endonuclease FEN1. Fork reversal provides a template for a DNA polymerase to synthesize the gap generated by FEN1 activity. Ligation of the newly synthesized DNA with the leading strand template precedes fork restoration and creation of a normal fork. Creating a complementation system using the fiber labeling methodology outlined in Ciccia *et al.* and testing HNH point mutants that inactivate the endonuclease activity is one way to test this hypothesis. Results from this and other analyses elucidating the

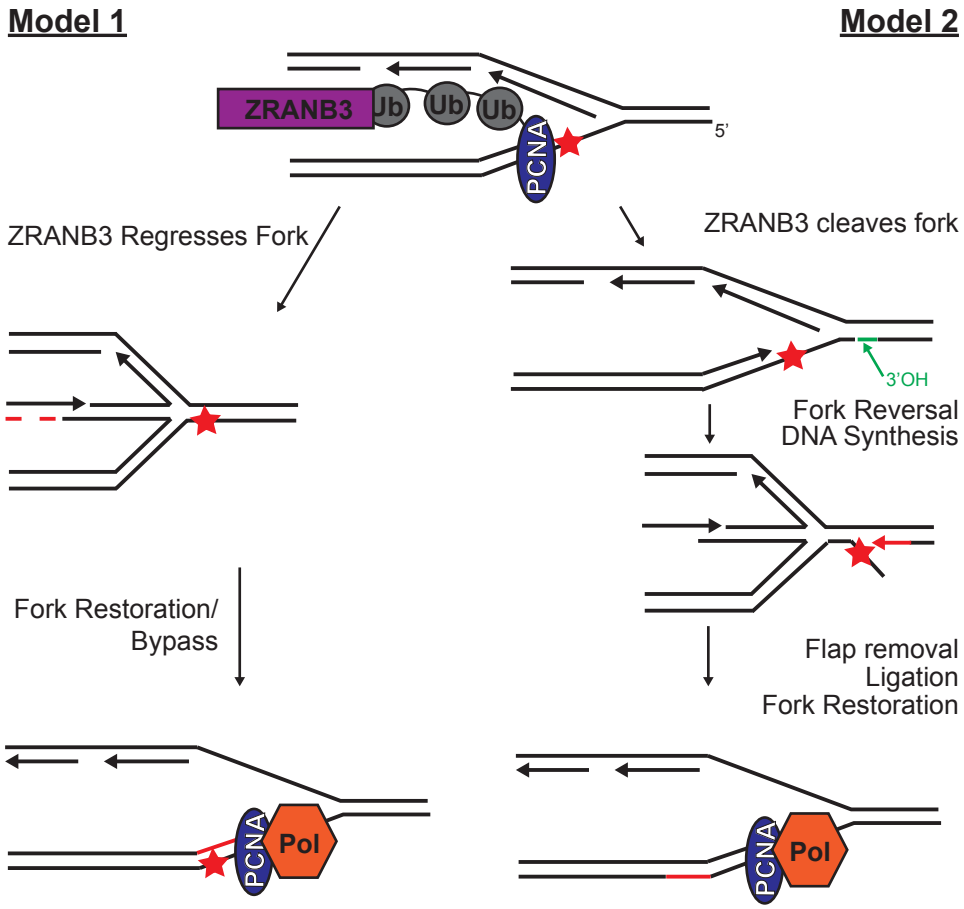


Figure 6.5 Two models of ZRANB3 activity at a stalled replication fork to promote fork restart. Model 1 depicts a scenario where ZRANB3 catalyzes fork regression to a stalled leading strand replication fork. Fork regression enables replication past the damage using the nascent lagging strand as the template strand. Fork restoration creates a normal fork whereby replication can restart. Model 2 depicts an alternative form of bypass, reported by Weston *et al.* (66), that incorporates ZRANB3 endonuclease activity. First, ZRANB3 cleaves the fork two nucleotides downstream of the junction, which generates a 3'- OH end. Upon fork reversal, the 5' flap is cleaved off by Fen1 endonuclease, and the 3'-OH is extended by a DNA polymerase up to the newly created 5' phosphate group. DNA Ligase I seals the nick and the regressed fork is restored to a normal fork.

function of the HNH domain in cells will provide a clearer understanding of how ZRANB3 endonuclease activity functions in replication restart.

Regulation of ZRANB3

Unregulated fork remodeling catalyzed on stalled replication forks will induce replication stress and may cause fork collapse. SMARCAL1 overexpression induces pan nuclear γH2AX, a marker for DNA damage (54). ATR phosphorylation of SMARCAL1 is thought to prevent aberrant replication fork remodeling that causes deleterious nuclease processing by SLX4 and CtIP and fork collapse (59). Therefore, unregulated replication fork cleavage and remodeling by ZRANB3 maybe equally deleterious. To test this, we can determine whether ZRANB3 overexpression causes a similar pan nuclear γH2AX phenotype as SMARCAL1 overexpression.

Regulating the function of a protein is not limited to post translational modifications. Similar to how SMARCAL1 binding to RPA directs SMARCAL1 to regress and restore specific replication fork and regressed substrates, ZRANB3 *in vitro* enzymatic activity may be regulated by its interaction with PCNA. Therefore testing WT ZRANB3 and PIP-NZF-APIM triple mutants in fork regression, fork restoration and nuclease activities to DNA substrates assembled with polyUB PCNA molecules may offer an insight to whether binding polyUB PCNA does indeed regulate ZRANB3 enzymatic activities.

Structural characterization of the substrate recognition domain

The HARP2 domain in SMARCAL1 adopts a fold similar to the mismatch recognition domain of MUTS and the damage recognition domain of XPB (63). The HIRAN domain contains a RPA OB type DNA binding domain (51). Structural characterization of the ZRANB3 substrate recognition domain may identify a conserved

fold shared with other DNA binding domains and provide an answer to how the ZRANB3 SRD interacts with DNA.

Concluding Remarks

This body of work accomplished multiple feats. I identified a novel DNA binding domain within ZRANB3, developed methodologies to test and further characterize ZRANB3 endonuclease activity, and created reagents to further investigate ZRANB3 function in cells. Hopefully, my work will serve as a foundation for future *in vivo* studies on ZRANB3.

REFERENCES

1. O'Donnell, M., Langston, L., and Stillman, B. (2013) Principles and Concepts of DNA Replication in Bacteria, Archaea, and Eukarya. *Cold Spring Harbor Perspectives in Biology* **5**

2. Ciccia, A., and Elledge, S. J. (2010) The DNA Damage Response: Making It Safe to Play with Knives. *Molecular cell* **40**, 179-204

3. De Bont, R., and van Larebeke, N. (2004) Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis* **19**, 169-185

4. Yeeles, J. T. P., Deegan, T. D., Janska, A., Early, A., and Diffley, J. F. X. (2015) Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**, 431-435

5. Moldovan, G.-L., Pfander, B., and Jentsch, S. (2007) PCNA, the Maestro of the Replication Fork. *Cell* **129**, 665-679

6. Shiloh, Y., and Ziv, Y. (2013) The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* **14**, 197-210

7. Smith, G. C. M., and Jackson, S. P. (1999) The DNA-dependent protein kinase. *Genes & Development* **13**, 916-934

8. Cimprich, K. A., and Cortez, D. (2008) ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Biol* **9**, 616-627

9. Zeman, M. K., and Cimprich, K. A. (2014) Causes and consequences of replication stress. *Nat Cell Biol* **16**, 2-9

10. Barlow, J. H., Faryabi, Robert B., Callén, E., Wong, N., Malhowski, A., Chen, Hua T., Gutierrez-Cruz, G., Sun, H.-W., McKinnon, P., Wright, G., Casellas, R., Robbiani, Davide F., Staudt, L., Fernandez-Capetillo, O., and Nussenzweig, A. (2013) Identification of Early Replicating Fragile Sites that Contribute to Genome Instability. *Cell* **152**, 620-632

11. Helmrich, A., Ballarino, M., Nudler, E., and Tora, L. (2013) Transcription-replication encounters, consequences and genomic instability. *Nat Struct Mol Biol* **20**, 412-418

12. Helmrich, A., Ballarino, M., and Tora, L. (2011) Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Molecular cell* **44**, 966-977

13. Bochman, M. L., Paeschke, K., and Zakian, V. A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat Rev Genet* **13**, 770-780

14. Yekezare, M., Gómez-González, B., and Diffley, J. F. X. (2013) Controlling DNA replication origins in response to DNA damage – inhibit globally, activate locally. *Journal of Cell Science* **126**, 1297-1306

15. Mourón, S., Rodriguez-Acebes, S., Martínez-Jiménez, M. I., García-Gómez, S., Chocrón, S., Blanco, L., and Méndez, J. (2013) Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nat Struct Mol Biol* **20**, 1383-1389

16. Waters, L. S., Minesinger, B. K., Wiltrout, M. E., D'Souza, S., Woodruff, R. V., and Walker, G. C. (2009) Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. *Microbiology and Molecular Biology Reviews : MMBR* **73**, 134-154

17. Petermann, E., Orta, M. L., Issaeva, N., Schultz, N., and Helleday, T. (2010) Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. *Molecular cell* **37**, 492-502

18. Hanada, K., Budzowska, M., Davies, S. L., van Drunen, E., Onizawa, H., Beverloo, H. B., Maas, A., Essers, J., Hickson, I. D., and Kanaar, R. (2007) The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nat Struct Mol Biol* **14**, 1096-1104

19. Petermann, E., and Helleday, T. (2010) Pathways of mammalian replication fork restart. *Nat Rev Mol Cell Biol* **11**, 683-687

20. Neelsen, K. J., and Lopes, M. (2015) Replication fork reversal in eukaryotes: from dead end to dynamic response. *Nat Rev Mol Cell Biol* **16**, 207-220

21. McGlynn, P., and Lloyd, R. G. (2002) Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol* **3**, 859-870

22. Lopes, M., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C. S., and Foiani, M. (2001) The DNA replication checkpoint response stabilizes stalled replication forks. *Nature* **412**, 557-561

23. Sogo, J. M., Lopes, M., and Foiani, M. (2002) Fork Reversal and ssDNA Accumulation at Stalled Replication Forks Owing to Checkpoint Defects. *Science* **297**, 599-602

24. Hu, J., Sun, L., Shen, F., Chen, Y., Hua, Y., Liu, Y., Zhang, M., Hu, Y., Wang, Q., Xu, W., Sun, F., Ji, J., Murray, Johanne M., Carr, Antony M., and Kong, D. (2012) The Intra-S Phase Checkpoint Targets Dna2 to Prevent Stalled Replication Forks from Reversing. *Cell* **149**, 1221-1232

25. Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., Neelsen, K. J., Fachinetti, D., Bermejo, R., Cocito, A., Costanzo, V., and Lopes, M. (2012) Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nat Struct Mol Biol* **19**, 417-423

26. Zellweger, R., Dalcher, D., Mutreja, K., Berti, M., Schmid, J. A., Herrador, R., Vindigni, A., and Lopes, M. (2015) Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. *The Journal of Cell Biology* **208**, 563-579

27. Singleton, M. R., Dillingham, M. S., and Wigley, D. B. (2007) Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annual Review of Biochemistry* **76**, 23-50

28. Hickson, I. D. (2003) RecQ helicases: caretakers of the genome. *Nat Rev Cancer* **3**, 169-178

29. Croteau, D. L., Popuri, V., Opresko, P. L., and Bohr, V. A. (2014) Human RecQ Helicases in DNA Repair, Recombination, and Replication. *Annual Review of Biochemistry* **83**, 519-552

30. Huang, S., Li, B., Gray, M. D., Oshima, J., Mian, I. S., and Campisi, J. (1998) The premature ageing syndrome protein, WRN, is a 3[prime][rarr]5[prime] exonuclease. *Nat Genet* **20**, 114-116

31. Machwe, A., Xiao, L., Groden, J., and Orren, D. K. (2006) The Werner and Bloom Syndrome Proteins Catalyze Regression of a Model Replication Fork. *Biochemistry* **45**, 13939-13946

32. Machwe, A., Karale, R., Xu, X., Liu, Y., and Orren, D. K. (2011) The Werner and Bloom Syndrome Proteins Help Resolve Replication Blockage by Converting (Regressed) Holliday Junctions to Functional Replication Forks. *Biochemistry* **50**, 6774-6788

33. Machwe, A., Xiao, L., Lloyd, R. G., Bolt, E., and Orren, D. K. (2007) Replication fork regression in vitro by the Werner syndrome protein (WRN): Holliday junction formation, the effect of leading arm structure and a potential role for WRN exonuclease activity. *Nucleic Acids Research* **35**, 5729-5747

34. Berti, M., Chaudhuri, A. R., Thangavel, S., Gomathinayagam, S., Kenig, S., Vujanovic, M., Odreman, F., Glatter, T., Graziano, S., Mendoza-Maldonado, R., Marino, F., Lucic, B., Biasin, V., Gstaiger, M., Aebersold, R., Sidorova, J. M., Monnat, R. J., Lopes, M., and Vindigni, A. (2013) Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nat Struct Mol Biol* **20**, 347-354

35. Kanagaraj, R., Saydam, N., Garcia, P. L., Zheng, L., and Janscak, P. (2006) Human RECQ5β helicase promotes strand exchange on synthetic DNA structures resembling a stalled replication fork. *Nucleic Acids Research* **34**, 5217-5231

36. Capp, C., Wu, J., and Hsieh, T.-s. (2010) RecQ4: the Second Replicative Helicase? *Critical Reviews in Biochemistry and Molecular Biology* **45**, 233-242

37. Flaus, A., Martin, D. M. A., Barton, G. J., and Owen-Hughes, T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Research* **34**, 2887-2905

38. Unk, I., Hajdú, I., Blastyák, A., and Haracska, L. (2010) Role of yeast Rad5 and its human orthologs, HLTF and SHPRH in DNA damage tolerance. *DNA Repair* **9**, 257-267

39. Parker, J. L., and Ulrich, H. D. (2009) Mechanistic analysis of PCNA poly-ubiquitylation by the ubiquitin protein ligases Rad18 and Rad5. *The EMBO Journal* **28**, 3657-3666

40. Motegi, A., Liaw, H.-J., Lee, K.-Y., Roest, H. P., Maas, A., Wu, X., Moinova, H., Markowitz, S. D., Ding, H., Hoeijmakers, J. H. J., and Myung, K. (2008) Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 12411-12416

41. Motegi, A., Sood, R., Moinova, H., Markowitz, S. D., Liu, P. P., and Myung, K. (2006) Human SHPRH suppresses genomic instability through proliferating cell nuclear antigen polyubiquitination. *The Journal of Cell Biology* **175**, 703-708

42. Unk, I., Hajdú, I., Fátyol, K., Hurwitz, J., Yoon, J.-H., Prakash, L., Prakash, S., and Haracska, L. (2008) Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination. *Proceedings of the National Academy of Sciences* **105**, 3768-3773

43. Masuda, Y., Suzuki, M., Kawai, H., Hishiki, A., Hashimoto, H., Masutani, C., Hishida, T., Suzuki, F., and Kamiya, K. (2012) En bloc transfer of polyubiquitin chains to PCNA in vitro is mediated by two different human E2–E3 pairs. *Nucleic Acids Research* **40**, 10394-10407

44. Motegi A, L. H., Lee KY, Roest HP, Maas A, Wu X, Moinova H, Markowitz SD, Ding H, Hoeijmakers JH, Myung K. (2008) Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks *Proceedings of the National Academy of Sciences of the United States of America* **105**, 12411-12416

45. Blastyák, A., Hajdú, I., Unk, I., and Haracska, L. (2010) Role of Double-Stranded DNA Translocase Activity of Human HLTF in Replication of Damaged DNA. *Molecular and Cellular Biology* **30**, 684-693

46. Blastyák, A., Pintér, L., Unk, I., Prakash, L., Prakash, S., and Haracska, L. (2007) Yeast Rad5 Protein Required for Postreplication Repair Has a DNA Helicase Activity Specific for Replication Fork Regression. *Molecular cell* **28**, 167-175

47. Machado, L. E. F., Pustovalova, Y., Kile, A. C., Pozhidaeva, A., Cimprich, K. A., Almeida, F. C. L., Bezsonova, I., and Korzhnev, D. M. (2013) NMR Structure Note: PHD Domain from Human SHPRH. *Journal of biomolecular NMR* **56**, 393-399

48. Iyer, L. M., Babu, M., and Aravind, L. (2006) The HIRAN Domain and Recruitment of Chromatin Remodeling and Repair activities to Damaged DNA. *Cell Cycle* **5**, 775-782

49. Achar, Y. J., Balogh, D., Neculai, D., Juhasz, S., Morocz, M., Gali, H., Dhe-Paganon, S., Venclovas, Č., and Haracska, L. (2015) Human HLTF mediates postreplication repair by its HIRAN domain-dependent replication fork remodelling. *Nucleic Acids Research* **43**, 10277-10291

50. Hishiki, A., Hara, K., Ikegaya, Y., Yokoyama, H., Shimizu, T., Sato, M., and Hashimoto, H. (2015) Structure of a Novel DNA-binding Domain of Helicase-like Transcription Factor (HLTF) and Its Functional Implication in DNA Damage Tolerance. *Journal of Biological Chemistry* **290**, 13215-13223

51. Kile, Andrew C., Chavez, Diana A., Bacal, J., Eldirany, S., Korzhnev, Dmitry M., Bezsonova, I., Eichman, Brandt F., and Cimprich, Karlene A. (2015) HLTF's Ancient HIRAN Domain Binds 3′ DNA Ends to Drive Replication Fork Reversal. *Molecular cell* **58**, 1090-1100

52. Coleman, M. A., Eisen, J. A., and Mohrenweiser, H. W. (2000) Cloning and Characterization of HARP/SMARCAL1: A Prokaryotic HepA-Related SNF2 Helicase Protein from Human and Mouse. *Genomics* **65**, 274-282

53. Bétous, R., Mason, A. C., Rambo, R. P., Bansbach, C. E., Badu-Nkansah, A., Sirbu, B. M., Eichman, B. F., and Cortez, D. (2012) SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication. *Genes & Development* **26**, 151-162

54. Bansbach, C. E., Bétous, R., Lovejoy, C. A., Glick, G. G., and Cortez, D. (2009) The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes & Development* **23**, 2405-2414

55. Yusufzai, T., Kong, X., Yokomori, K., and Kadonaga, J. T. (2009) The annealing helicase HARP is recruited to DNA repair sites via an interaction with RPA. *Genes & Development* **23**, 2400-2404

56. Yuan, J., Ghosal, G., and Chen, J. (2009) The annealing helicase HARP protects stalled replication forks. *Genes & Development* **23**, 2394-2399

57. Postow, L., Woo, E. M., Chait, B. T., and Funabiki, H. (2009) Identification of SMARCAL1 as a Component of the DNA Damage Response. *The Journal of Biological Chemistry* **284**, 35951-35961

58. Ciccia, A., Bredemeyer, A. L., Sowa, M. E., Terret, M.-E., Jallepalli, P. V., Harper, J. W., and Elledge, S. J. (2009) The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. *Genes & Development* **23**, 2415-2425

59. Couch, F. B., Bansbach, C. E., Driscoll, R., Luzwick, J. W., Glick, G. G., Bétous, R., Carroll, C. M., Jung, S. Y., Qin, J., Cimprich, K. A., and Cortez, D. (2013)

ATR phosphorylates SMARCAL1 to prevent replication fork collapse. *Genes & Development* **27**, 1610-1623

60. Poole, L. A., Zhao, R., Glick, G. G., Lovejoy, C. A., Eischen, C. M., and Cortez, D. (2015) SMARCAL1 maintains telomere integrity during DNA replication. *Proceedings of the National Academy of Sciences* **112**, 14864-14869

61. Cox, Kelli E., Maréchal, A., and Flynn, Rachel L. (2016) SMARCAL1 Resolves Replication Stress at ALT Telomeres. *Cell reports* **14**, 1032-1040

62. Bétous, R., Couch, F. B., Mason, A. C., Eichman, B. F., Manosas, M., and Cortez, D. (2013) Substrate-selective repair and restart of replication forks by DNA translocases. *Cell reports* **3**, 1958-1969

63. Mason, A. C., Rambo, R. P., Greer, B., Pritchett, M., Tainer, J. A., Cortez, D., and Eichman, B. F. (2014) A structure-specific nucleic acid-binding domain conserved among DNA repair proteins. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 7618-7623

64. Yusufzai, T., and Kadonaga, J. T. (2010) Annealing helicase 2 (AH2), a DNA-rewinding motor with an HNH motif. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20970-20973

65. Ciccia, A., Nimonkar, A. V., Hu, Y., Hajdu, I., Achar, Y. J., Izhar, L., Petit, S. A., Adamson, B., Yoon, J. C., Kowalczykowski, S. C., Livingston, D. M., Haracska, L., and Elledge, S. J. (2012) The ZRANB3 translocase associates with poly-ubiquitinated PCNA to promote fork restart and limit recombination after replication stress. *Molecular cell* **47**, 396-409

66. Weston Ria, H. P., and Dragana Ahel. (2012) ZRANB3 is a structure-specific ATP-dependent endonuclease involved in replication stress response. *Genes & Development* **26**, 1558-1572

67. Yuan J, G. G., Chen J. (2012) The HARP-like domain-containing protein AH2/ZRANB3 binds to PCNA and participates in cellular response to replication stress. *Molecular cell* **10**, 410-421

68. Mailand, N., Gibbs-Seymour, I., and Bekker-Jensen, S. (2013) Regulation of PCNA-protein interactions for genome stability. *Nat Rev Mol Cell Biol* **14**, 269-282

69. Sickmier, E. A., Kreuzer, K. N., and White, S. W. (2004) The Crystal Structure of the UvsW Helicase from Bacteriophage T4. *Structure* **12**, 583-592

70. Singleton, M. R., Scaife, S., and Wigley, D. B. (2001) Structural Analysis of DNA Replication Fork Reversal by RecG. *Cell* **107**, 79-89

71. Atkinson, J., and McGlynn, P. (2009) Replication fork reversal and the maintenance of genome stability. *Nucleic Acids Research* **37**, 3475-3492

72. Carles-Kinch, K., George, J. W., and Kreuzer, K. N. (1997) Bacteriophage T4 UvsW protein is a helicase involved in recombination, repair and the regulation of DNA replication origins. *The EMBO Journal* **16**, 4142-4151

73. Chen, X., Zaro, J., and Shen, W.-C. (2013) Fusion Protein Linkers: Property, Design and Functionality. *Advanced drug delivery reviews* **65**, 1357-1369

74. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013) Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science (New York, N.Y.)* **339**, 819-823

75. Yusufzai, T., and Kadonaga, J. T. (2008) HARP Is an ATP-Driven Annealing Helicase. *Science* **322**, 748-750

76. Ghosal, G., Yuan, J., and Chen, J. (2011) The HARP domain dictates the annealing helicase activity of HARP/SMARCAL1. *EMBO Reports* **12**, 574-580

77. Boerkoel CF, T. H., John J, Yan J, Stankiewicz P, Rosenbarker L, André JL, Bogdanovic R, Burguet A, Cockfield S, Cordeiro I, Fründ S, Illies F, Joseph M, Kaitila I, Lama G, Loirat C, McLeod DR, Milford DV, Petty EM, Rodrigo F, Saraiva JM, Schmidt B, Smith GC, Spranger J, Stein A, Thiele H, Tizard J, Weksberg R, Lupski JR, Stockton DW. (2002) Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia. *Nature Genetics* **30**, 215-220

78. Carroll, C., Badu-Nkansah, A., Hunley, T., Baradaran-Heravi, A., Cortez, D., and Frangoul, H. (2013) Schimke Immunoosseous Dysplasia Associated with Undifferentiated Carcinoma and a Novel SMARCAL1 Mutation in a Child. *Pediatric blood & cancer* **60**, E88-E90

79. Baradaran-Heravi, A., Raams, A., Lubieniecka, J., Cho, K. S., DeHaai, K. A., Basiratnia, M., Mari, P.-O., Xue, Y., Rauth, M., Olney, A. H., Shago, M., Choi, K., Weksberg, R. A., Nowaczyk, M. J. M., Wang, W., Jaspers, N. G. J., and Boerkoel, C. F. (2012) SMARCAL1 Deficiency Predisposes to Non-Hodgkin Lymphoma and Hypersensitivity to Genotoxic Agents In vivo. *American journal of medical genetics. Part A* **158A**, 2204-2213

80. Clewing, J. M., Fryssira, H., Goodman, D., Smithson, S. F., Sloan, E. A., Lou, S., Huang, Y., Choi, K., Lücke, T., Alpay, H., André, J.-L., Asakura, Y., Biebuyck-Gouge, N., Bogdanovic, R., Bonneau, D., Cancrini, C., Cochat, P., Cockfield, S., Collard, L., Cordeiro, I., Cormier-Daire, V., Cransberg, K., Cutka, K., Deschenes, G., Ehrich, J. H. H., Fründ, S., Georgaki, H., Guillen-Navarro, E., Hinkelmann, B., Kanariou, M., Kasap, B., Kilic, S. S., Lama, G., Lamfers, P., Loirat, C., Majore, S., Milford, D., Morin, D., Özdemir, N., Pontz, B. F., Proesmans, W., Psoni, S., Reichenbach, H., Reif, S., Rusu, C., Saraiva, J. M., Sakallioglu, O., Schmidt, B., Shoemaker, L., Sigaudy, S., Smith, G., Sotsiou, F., Stajic, N., Stein, A., Stray-Pedersen, A., Taha, D., Taque, S., Tizard, J., Tsimaratos, M., Wong, N. A. C. S., and Boerkoel, C. F. (2007) Schimke immunoosseous dysplasia: suggestions of genetic diversity. *Human Mutation* **28**, 273-283

81. Bansbach, C. E., Boerkoel, C. F., and Cortez, D. (2010) SMARCAL1 and replication stress: An explanation for SIOD? *Nucleus* **1**, 245-248

82. Lücke, T., Kanzelmeyer, N., Baradaran-Heravi, A., Boerkoel, C. F., Burg, M., Ehrich, J. H. H., and Pape, L. (2009) Improved outcome with immunosuppressive monotherapy after renal transplantation in Schimke-immuno-osseous dysplasia. *Pediatric Transplantation* **13**, 482-489

83. Lücke, T., Billing, H., Sloan, E. A., Boerkoel, C. F., Franke, D., Zimmering, M., Ehrich, J. H. H., and Das, A. M. (2005) Schimke-immuno-osseous dysplasia: New mutation with weak genotype–phenotype correlation in siblings. *American Journal of Medical Genetics Part A* **135A**, 202-205

84. Piirilä, H., Väliaho, J., and Vihinen, M. (2006) Immunodeficiency mutation databases (IDbases). *Human Mutation* **27**, 1200-1208

85. Dürr, H., Körner, C., Müller, M., Hickmann, V., and Hopfner, K.-P. X-Ray Structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase Core and Its Complex with DNA. *Cell* **121**, 363-373

86. Kilic, S. S., Donmez, O., Sloan, E. A., Elizondo, L. I., Huang, C., André, J.-L., Bogdanovic, R., Cockfield, S., Cordeiro, I., Deschenes, G., Fründ, S., Kaitila, I., Lama, G., Lamfers, P., Lücke, T., Milford, D. V., Najera, L., Rodrigo, F., Saraiva, J. M., Schmidt, B., Smith, G. C., Stajic, N., Stein, A., Taha, D., Wand, D., Armstrong, D., and Boerkoel, C. F. (2005) Association of migraine-like headaches with Schimke immuno-osseous dysplasia. *American Journal of Medical Genetics Part A* **135A**, 206-210

87. Moinova, H. R., Chen, W.-D., Shen, L., Smiraglia, D., Olechnowicz, J., Ravi, L., Kasturi, L., Myeroff, L., Plass, C., Parsons, R., Minna, J., Willson, J. K. V., Green, S. B., Issa, J.-P., and Markowitz, S. D. (2002) HLTF gene silencing in human colon cancer. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 4562-4567

88. Lawrence MS, S. P., Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495-501

89. Marti, T. M., and Fleck, O. DNA repair nucleases. *Cellular and Molecular Life Sciences CMLS* **61**, 336-354

90. Hsia, K.-C., Chak, K.-F., Liang, P.-H., Cheng, Y.-S., Ku, W.-Y., and Yuan, H. S. (2004) DNA Binding and Degradation by the HNH Protein ColE7. *Structure* **12**, 205-214

91. Huang, H., and Yuan, H. S. (2007) The Conserved Asparagine in the HNH Motif Serves an Important Structural Role in Metal Finger Endonucleases. *Journal of Molecular Biology* **368**, 812-821

92. Stoddard, B. L. (2005) Homing endonuclease structure and function. *Quarterly Reviews of Biophysics* **38**, 49-95

93. Saravanan, M., Bujnicki, J. M., Cymerman, I. A., Rao, D. N., and Nagaraja, V. (2004) Type II restriction endonuclease R.KpnI is a member of the HNH nuclease superfamily. *Nucleic Acids Research* **32**, 6129-6135

94. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman, A. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* **44**, D279-D285

95. Mehta, P., Katta, K., and Krishnaswamy, S. (2004) HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily. *Protein Science : A Publication of the Protein Society* **13**, 295-300

96. Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., and Kinzler, K. W. (2013) Cancer Genome Landscapes. *Science* **339**, 1546-1558

97. Bhattacharyya, B., and Keck, J. L. (2014) Grip it and rip it: Structural mechanisms of DNA helicase substrate binding and unwinding. *Protein Science : A Publication of the Protein Society* **23**, 1498-1507