REPLICATION FORK REMODELING BY

HELICASE-LIKE TRANSCRIPTION FACTOR

By

Diana Andrea Chavez

Dissertation

Submitted to the Faculty of the

Graduate School at Vanderbilt University

in partial fulfillment of requirements

for the degree of

DOCTOR OF PHILOSOPHY

in
Biological Sciences
August 10, 2018
Nashville, Tennessee

Approved:

Todd Graham, Ph.D.

Brandt F. Eichman, Ph.D.

Walter Chazin, Ph.D.

David Cortez, Ph.D.

Carl Johnson, Ph.D.

Lauren Jackson, Ph.D.

To my loving husband,

For all your unwavering support and encouragement.

ACKNOWLEDGEMENTS

I am very grateful for the continuous support and guidance from my mentor, Dr. Brandt Eichman. Thank you for introducing me to the exciting field of DNA repair and for sharing your passion for science with me. I would also like to thank all of my committee members. Drs. Walter Chazin, Dave Cortez, Todd Graham, Lauren Jackson, and Carl Johnson, I am so grateful for the time and expertise you shared with me. Thank you for pushing me to be a better scientist and for helping me make my project a success.

I would also like to thank the past and present members of the Eichman lab. Thank you for making the lab a fun place to be and for all the science and not science-related discussions that have made this journey enjoyable. I will carry great memories from my time in the Eichman lab and at Vanderbilt with me always.

Finally, I would like to thank all of my friends and family. I would not be here without your support. Mom and Dad, thank you for all your love and for encouraging me to always follow my dreams. Carlos and Alex, thank you for always being there. I know I can always count on you for anything. Ruben, thank you for following me to Nashville so that I could chase my dream. Thank you for sharing my failures and triumphs and for pushing me to not give up when things seemed bleak. You have kept me sane through grad school and have been a great blessing.

There are so many that have encouraged and pushed me to do my best throughout my years at Vanderbilt. I am truly thankful to all of you; I would not be where I am without your support.

TABLE OF CONTENTS

Chapter

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Å   Angstrom

APIM   AlkBH2-PCNA interacting motif

AT-hooks   Asparagine/aspartate-hooks

ATM   Ataxia-Telangiectasia Mutated

ATP   Adenosine tri-phosphate

ATR   Ataxia-Telangiectasia Rad3-related

BER   Base excision repair

CD   Circular dichroism

Cdc45   Cell-division control protein 45

Chd1   Chromodomain helicase DNA binding protein 1

CHROMO   Chromatin organization modifier

CldU   5-Chloro-2′-deoxyuridine

DDT   DNA damage tolerance

DNA   Deoxyribonucleic Acid

dsDNA   Double-stranded DNA

DSBs   Double-strand breaks

dNTPs   Deoxyribonucleotides

DTT   dithiothreitol

EDTA   Ethylenediaminetetraacetic acid

Edu   5′-ethynyl-2'-deoxyuridine

EMSAs   Electorphoretic mobility shifts assays

FAM   Fluorescein

FRET   Fluorescence energy transfer

GINS   Go-ichi-ni-san complex

HAND   Histone and DNA interaction

HARP   HepA related protein

HEAT   Huntingtin, EF3, PP2A, TOR1

HEPES   (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)

HIRAN   HIP116/Rad5 N-terminal

HLTF   Helicase-Like Transcription Factor

HSA   Helicase/SANT-associated

HRV3C   Human Rhinovirus 3C

HU   Hydroxyurea

IdU   5-Iodo-2′-deoxyuridine

iPOND   Isolation of protein on nascent DNA

IsWI	Imitation switch
kDa	Kilodalton
LB	Luria broth
MBP	Maltose binding protein
MCM	Mini chromosome maintenance proteins
MMR	Mismatch repair
MMS2	Methyl methane sulfonate sensitivity protein 2
NER	Nucleotide excision repair
Ni-NTA	Nickel nitrilotriacetic acid
NZF	Npl4 zinc finger
OB-fold	Oligonucleotide/oligosaccharide-binding fold
PAZ	Piwi-Argonaute-Zwille
PEG	polyethylene glycol
PIP	PCNA interacting protein
PCNA	Proliferating Cell Nuclear Antigen
PDB	Protein data bank
PMSF	phenylmethylsulfonyl fluoride
PPS	PreScission protease

PriA   Priming protein A

RFC   Replication factor C

RING   Really interesting new gene

RNA   Ribonucleic Acid

RPA   Replication protein A

rNTPs   Ribonucleotides

Rad5   Radiation sensitivity protein 5

Rad6   Radiation sensitivity protein 6

Rad18   Radiation sensitivity protein 18

Rad51   Radiation sensitivity protein 51

RecA   Recombination protein A

SANT   Swi3, Ada2, N-Cor, and TFIIIB

SAXS   Small-angle x-ray scattering

SDS-PAGE   sodium dodecyl sulfate polyacrylamide gel electrophoresis

ssDNA   Single-stranded DNA

SHPRH   SNF2 histone-linker PHD-finger helicase

SLIDE   SANT-like ISWI domain

SMARCAL1   SWI-SNF related, matrix associated, actin-like, regulator of chromatin,

            subfamily A-like 1

SmpB	Small protein B
SnAC	Snf ATP Coupling
SNF2	Sucrose Non-Fermenting 2
SRDs	Substrate recognition domains
SUMO	Small ubiquitin-like modifier
TBE	TRIS, Boric acid, EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch Virus
TLS	Translesion Synthesis
tRNA	Transfer RNA
TRIS	tris(hydroxymethyl)aminomethane
UBC13	Ubiquitin conjugating protein 13
UvsW	UV-repair and recombination protein
VRR	Viral replication and repair
ZRANB3	Zinc finger RANBP2-type containing 3

CHAPTER I

Introduction

DNA replication stress and stress responses

DNA replication is a fundamental biological process necessary for all life. It is therefore highly regulated and entails the coordination of many specialized proteins to ensure the accurate and complete duplication of a genome. These proteins, which together make up a replication fork or replisome, assemble at a licensed origin and fire bidirectionally to copy DNA. DNA replication consists of the unwinding of a DNA duplex by the 11-subunit Cdc45-MCM-GINS helicase (composed of cell-division control protein 45 (Cdc45), minichromosome maintenance proteins (MCM) and DNA replication complex (GINS) to generate two single-stranded DNA strands. Complementary (daughter or nascent) strands are generated by Proliferating Cell Nuclear Antigen (PCNA)-associated polymerases. The polymerase α-primase complex synthesizes the RNA-DNA primer that is required by both polymerase ε and δ for synthesis of nascent leading and lagging strands (Figure 1.1) (Burgers & Kunkel, 2017). Due to the polarity of DNA synthesis, one nascent strand is generated continuously, (leading strand) by polymerase ε, and the other is synthesized discontinuously (lagging strand) by polymerase δ in segments referred to as Okazaki fragments.
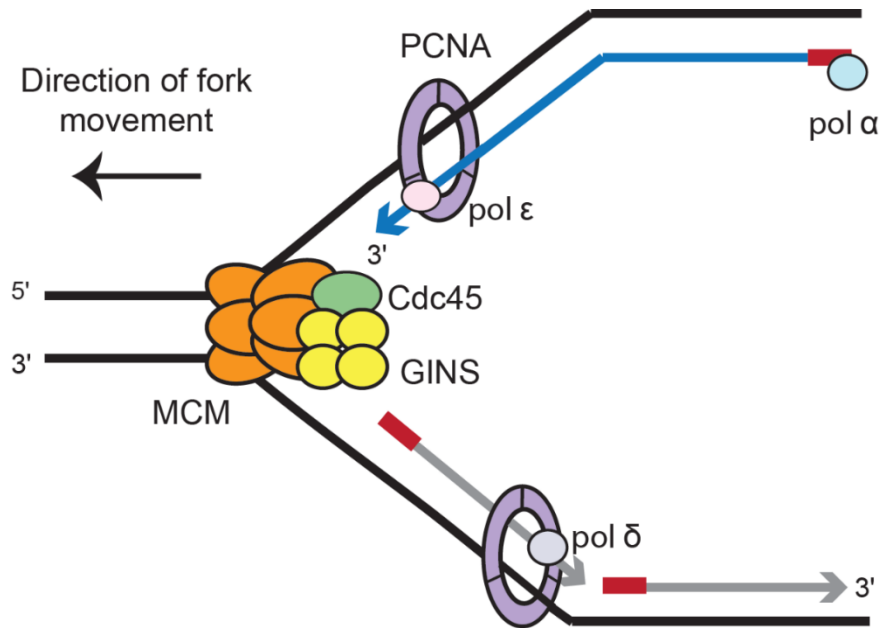
Figure 1.1 Architecture of a replication fork. Diagram of a replication fork. The CMG helicase unwinds the parental DNA strands (black), and polymerases synthesize the leading (blue) and lagging (gray) strands. RNA-DNA primers are shown in red.

Replication stress

During DNA replication, the moving replication fork may encounter obstacles that can lead to its transient stalling or slowing down. This phenomenon is known as replication stress and can result in a variety of negative effects on genomic stability. (Cortez, 2015). There are many sources of replication stress, including DNA damage, ribonucleotide incorporation, collisions with transcription machinery, RNA-DNA hybrid structures, repetitive sequences, and miss-regulation of replication initiation.

One commonly recognized source of replication stress is the presence of unrepaired DNA lesions. In addition to blocking DNA synthesis, DNA lesions threaten DNA integrity. Exogenous factors, such as ultra-violet rays from sunlight, can result in a

total of 10^5 DNA lesions per cell per day (Ciccia & Elledge, 2010) . In addition, ionizing radiation (IR), can lead to oxidation of DNA bases which result in DNA breaks. Moreover, chemical alkylating agents and cross-linking agents, often used in chemotherapy, lead to alkylated DNA bases and can result in intra- and inter-strand crosslinks respectively (Ciccia & Elledge, 2010). Endogenous factors can also lead to DNA lesions. Bases can undergo spontaneous depurination and deamination reactions than can lead to abasic bases and transversion mutations (Ciccia & Elledge, 2010). Reactive oxygen species produced as a byproduct of cell metabolic reactions, such as oxygen singlets and hydroxyl radicals, can also lead to a variety of modified DNA bases (Cadet & Wagner, 2013).

In addition to DNA lesions, ribonucleotide incorporation can also generate replication stress. Failure of DNA polymerases to discriminate between ribonucleotides (rNTP) and deoxyribonucleotides (dNTP) can result in the misincorporation of rNTPs in DNA and result in increased sensitivity of DNA to cleavage due to the higher reactivity of the 2'OH in the RNA ribose sugar. Although DNA polymerases can remove rNTPs in DNA using their proofreading function, the rates of misincorporated rNTPs in normal cells can be quite high and therefore pose a threat to DNA integrity (Potenski & Klein, 2014).

Another source of stress that can block progression of the replication machinery are DNA-protein crosslinks. These DNA-protein crosslinks serve as large adducts that are caused by the trapping of covalent reaction intermediates of a number of DNA processing enzymes, such as topoisomerases and DNA methyltransferases, or by the trapping of proteins in close proximity to DNA. Trapping of these enzymes on DNA is caused by a number of agents that result from endogenous metabolic processes, such

as the reactive aldehyde species generated from alcohol metabolism, or from exogenous agents that are commonly used in chemotherapeutics (Brooks & Theruvathu, 2005; Stingele *et al*, 2017).

Replication stress is also caused by a loss of RNA processing factors in cells. Unavailability of RNA processing factors can affect transcription rates and cause prolonged association of the transcriptional machinery with DNA. This can produce collisions between the transcriptional machinery and a replication fork and generate topological stress as the two approach each other. Loss of RNA processing factors can also lead to the formation of aberrant DNA-RNA hybrid structures, such as R-loops, where nascent RNA synthesized by RNA polymerase invades DNA, displacing one of the DNA strands. These DNA-RNA hybrid structures preferentially form in G-C rich regions, are more stable than duplex DNA and not only impede a moving replication fork, but cause DSBs (Aguilera & García-Muse, 2012; Zeman & Cimprich, 2014).

In addition to encountering DNA-RNA hybrid structures, the replication machinery can also encounter DNA sequences that are difficult to replicate. DNA sequence repeats, such as $(GAA)_n$ and $(CNG)_n$ and GC-rich $(CGG)_n$ $(C_4GC_4GCG)_n$, make up about 30% of the human genome and can form unusual secondary structures, including hairpins, triplexes and quadruplexes that block progression of the replication machinery. These repetitive sequences can also lead to replication slippage, and cause compaction or expansion of the repeat sequences (Bochman *et al*, 2012; Zeman & Cimprich, 2014; McMurray, 2010).

Finally, improper regulation of replication initiation can limit replication factors and cause slow fork progression and replication stress. These factors include nucleotide

pools, replication machinery components, and DNA packaging components such as histones and histone chaperones. Chromatin compaction, and other natural process that limit DNA accessibility to the replication machinery can also be a source of stress for a replication fork (Zeman & Cimprich, 2014).

Replication stress responses

Stalled replication and a failure of the cell to respond to replication stress can lead to fork collapse, or the inability of a fork to continue DNA synthesis. Fork collapse is detrimental to a cell and results in genomic instability, chromosomal rearrangements, mutations, cell death, and a number of human diseases (Zeman & Cimprich, 2014; Cortez, 2015). To avoid such deleterious effects, cells have developed various stress-response mechanisms that rely on the recruitment and coordination of many protein factors to stabilize, repair and restart stalled replication forks.

Upon encountering replication stress, a replicative polymerase may stall and uncouple from the moving helicase, generating stretches of ssDNA that are coated with replication protein A (RPA) (Cortez, 2015; Zeman & Cimprich, 2014). This, in conjunction with its proximity to a stalled fork, signals for the recruitment of additional proteins, which results in ataxia-telangiectasia Rad3-related (ATR)-dependent checkpoint activation. ATR is a kinase that acts as a key coordinator of DNA damage response mechanisms and phosphorylates many substrates to induce checkpoint activation, leading to suppression of late origin firing and inhibition of cell cycle progression to provide

additional time for resolution of replication stress and fork rescue (Cortez, 2015; Zeman & Cimprich, 2014).

Another set of stress response mechanisms are the DNA damage tolerance (DDT) pathways (Figure 1.2), whose activation relies on the post-translational modification of PCNA (Friedberg, 2005). Monoubiquitination of PCNA on lysine-164 by the E2-E3 ligase complex, Rad6-Rad18, and activates the error-prone translesion synthesis (TLS), where specialized polymerases that can incorporate nucleotides using a damaged DNA template are recruited through interactions with PCNA. Even though the exact mechanism for specific recruitment of these TLS polymerases is not known, they are recruited according to their specificity and ability to bypass a lesion. Most of these polymerases are error-prone, although some like polymerase η can bypass UV-induced thymidine dimers yet maintain accuracy (Choe & Moldovan, 2017; Branzei & Psakhye, 2016; Kanao & Masutani, 2017).

Lysine-63 linked polyubiquitination of PCNA by the E2 complex, Mms2/Ubc13 and either helicase-like transcription factor (HLTF) or SNF2 histone-linker PHD-finger RING-finger helicase (SHPRH), which function as E3 ubiquitin ligases, promotes template switching. The template switching pathway uses an undamaged template for error-free synthesis past a replication blockage. One template switching mechanism is D-loop formation, where the strand across the damaged template usually invades and uses the newly synthesized sister strand as a template. Another mechanism in template switching is fork regression (Kanao & Masutani, 2017).

Fork regression, also known as fork reversal, is a DNA-remodeling reaction in which the parental strands are re-annealed and the two nascent daughter strands are

unpaired from their templates and repaired together to form a four-way junction (Figure 1.3). Fork reversal was first proposed as a branch migration mechanism for genetic recombination over 40 years ago (Hotchkiss, 1974). Soon after, fork reversal received experimental support and was proposed to account for four-arm replication forks observed by electron microscopy and heavy/heavy DNA using pulse labelling of human replicating cells (Higgins *et al*, 1976). It was first confirmed as a relevant mechanism for maintaining genomic stability in prokaryotes as a response by the T4 bacteriophage protein UvsW to stalled forks caused by leading strand lesions (Manosas *et al*, 2012; Nelson & Benkovic, 2010) and as a response to replication-transcription collisions in *E.coli* (De Septenville *et al*, 2012). In the last few years fork reversal has also proved to be a relevant mechanism in eukaryotes. Fork slowing and unusual replication intermediates, determined to be reversed forks, were observed as a result of mild type1 topoisomerase poisoning (Chaudhuri *et al*, 2012). In addition, reversed forks have been observed as a response to prolonged S-phase in mouse embryonic stem cells and as a global response to genotoxic treatments in human cells (Ahuja *et al*, 2016; Zellweger *et al*, 2015). Furthermore, fork reversal has been detected in a number of cultured human cells and in experiments using human plasmids with difficult to replicate repeat sequences, which suggests that fork reversal is a response to endogenous replication stress and occurs in unperturbed S phase. Collectively this recent evidence points to fork reversal as an important, regulated, transient response to many forms of replication stress that maintains genomic integrity (Follonier *et al*, 2013; Neelsen *et al*, 2013; Ray Chaudhuri *et al*, 2012; Chaudhuri *et al*, 2015; Zellweger *et al*, 2015).

Fork reversal presents several advantages for maintaining fork stability. First, it limits accumulation of ssDNA regions that are vulnerable to nuclease cleavage and degradation. Second, it positions the lesion back into the context of duplex DNA, permitting repair by other pathways such mismatch-repair (MMR), base-excision repair (BER), and nucleotide-excision repair (NER). Additionally, formation of the regressed arm places the two nascent strands together and allows for synthesis of one of the nascent strands using the other as a template. Last, recognition by ataxia-telangiectasia mutated (ATM), a key cell-cycle regulating kinase, of the newly formed DNA-end at the regressed arm resembling DNA-ends at double-stranded breaks (DSBs), may induce checkpoint activation. This checkpoint activation by ATM ensures repair of damage before DNA replication is completed (Neelsen & Lopes, 2015). Once the fork has been reversed and the damage repaired, the fork can be remodeled back and restored for replication to continue.
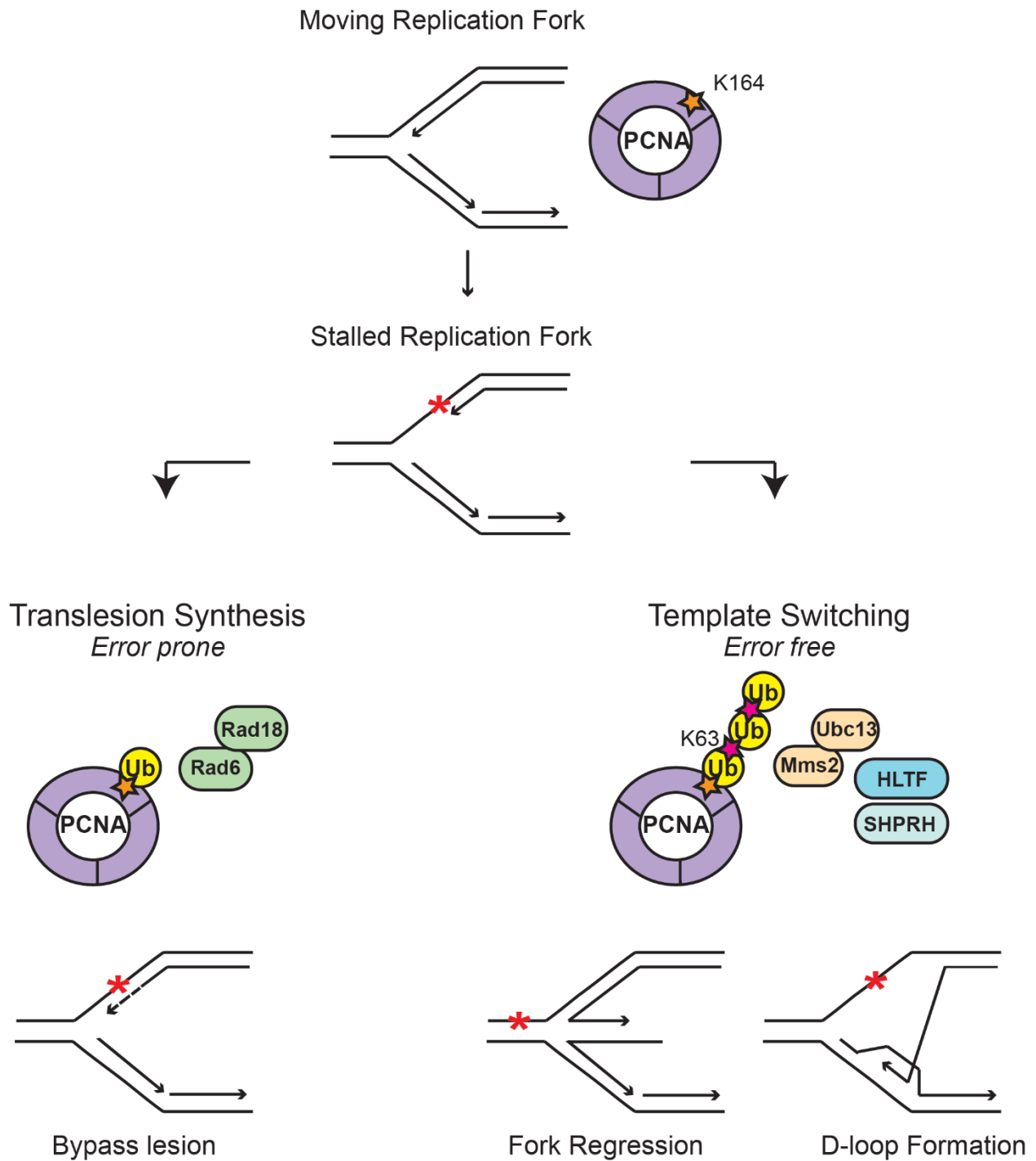
Figure 1.2 DNA damage tolerance pathways. As a result of fork stalling, PCNA is monoubiquitinated by Rad6/Rad18 to turn on translesion-synthesis and bypass the damage or polyubiquitinated by Mms2/Uc13 in complex with HLTF of SHPRH to turn on the template switching pathways which utilize fork regression or D-loop formation mechanisms.
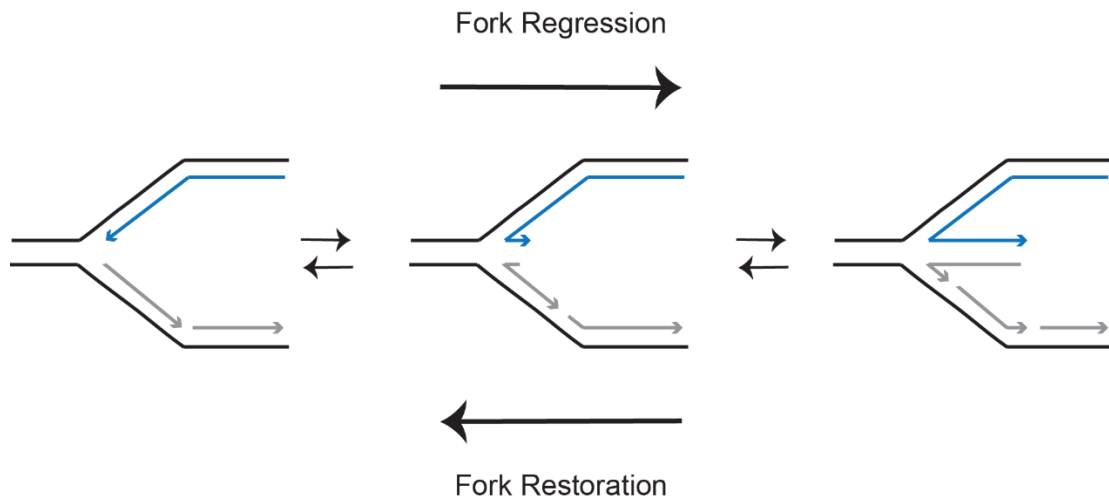
Figure 1.3 Fork remodeling reactions. In fork regression, the leading (blue) and lagging (gray) nascent strands are annealed together, forming a four-way junction. Pushing the regressed fork in the opposite direction results in fork restoration.

SNF2 fork remodelers

Many proteins that catalyze fork-remodeling reactions are classified as SNF2 proteins. These proteins are DNA-dependent ATPases that translocate on DNA to remodel nucleic acid structures and are involved in recombination, transcription, replication, and in repair. Enzymes in this family have a conserved two-lobed ATPase-motor domain, where each lobe has a RecA-like protein fold. There are 12 conserved motifs that line the interdomain pocket between the two lobes and based on the sequence conservation of these motifs, SNF2 proteins are classified into different subfamilies (Figure 1.4). Moreover, amino acids found in these motifs are essential for SNF2 function as they coordinate DNA, ATP, and Mg^{2+} binding.

Recent structures of SNF2 proteins have illustrated how these motors engage DNA. Interactions between the conserved motifs and the phosphate backbones along the

DNA minor groove allow the two ATPase lobes to engage both DNA strands and places the two ATPase lobes together. In addition, ATP binding by the motor results in a closed conformation that positions ATP in the interdomain pocket between the two lobes, near catalytic residues. This forms an active site for ATP hydrolysis. ATP binding and hydrolysis results in a series conformational changes that allows these enzymes to translocate on DNA, usually with a 3' to 5' directionality (Farnung *et al*, 2017; Flaus *et al*, 2006; Flaus & Owen-Hughes, 2011; Liu *et al*, 2017).

In addition to their ATPase-motor domain, SNF2 proteins have accessory domains that are essential for their function. For example, in the chromatin remodeler Chd1, CHROMO domains bind methylated lysines on histones to direct Chd1 to specific chromatin sites and regulates the activity of the SNF2 ATPase motor (Hauk *et al*, 2010). The HAND, SANT, and SLIDE domains in IsWI, another chromatin remodeler, bind chromatin and serve as nucleosome spacing modules (Bartholomew, 2014). Other domains that have been found in conjunction SNF2 motor domains include HEAT repeats, HSA domains, AT-hooks, and SnAC domains among others (Hopfner *et al*, 2012). Apart from their function in chromatin dynamics, SNF2 ATPases also have been characterized to catalyze fork remodeling reactions. Two of these proteins are SWI-SNF related, matrix associated, actin-like, regulator of chromatin subfamily A–like 1 (SMARCAL1) and zinc finger RANBP2 type–containing 3 (ZRANB3).

SMARCAL1 and ZRANB3 are fork remodeling proteins that play critical roles in stabilizing stalled forks during replication stress, as loss of either results in hypersensitivity to DNA damaging agents (Poole & Cortez, 2017). Both SMARCAL1 and ZRANB3 are recruited to stalled replication forks through protein-protein interactions. SMARCAL1, by
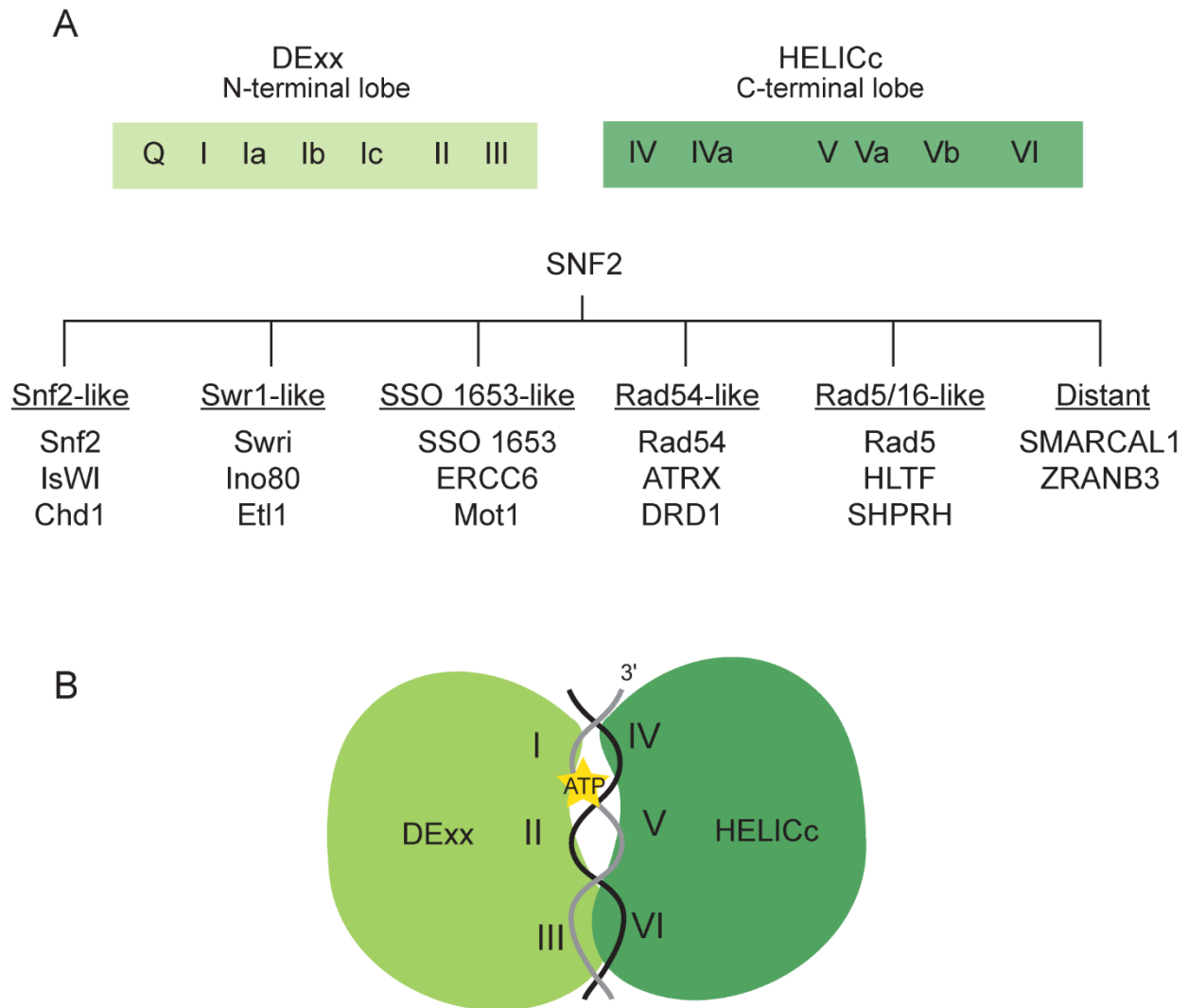
Figure 1.4 SNF2 family of ATPases. A. Domain organization of the SNF2 ATPase domain. Depicted are the conserved motifs in the DExx and HELICc ATPase lobes. SNF2 subfamilies are depicted below. Common members of each subfamily are also listed. **B.** Schematic of how ATPase lobes bind ATP and DNA. The conserved motifs are depicted within each lobe. The DNA tracking strand is shown in gray.

binding of its N-terminal domain to RPA, and ZRANB3 through interactions of its (PCNA interacting protein) PIP-box, AlkBH2-PCNA interacting motif (APIM) and Npl4 zinc finger (NZF) to polyubiquitinated PCNA (Poole & Cortez, 2017). In addition to these domains that are involved in protein-protein interactions, both proteins also have substrate-

recognition domains (SRDs) that provide specificity to DNA structures and are important for catalyzing fork regression in conjunction with the ATPase motor. The SMARCAL1 HARP domain binds branched DNA structures and is required for fork reversal (Bétous *et al*, 2013). ZRANB3's SRD binds fork-like DNA structures and is required for ATP hydrolysis and fork regression (Badu-Nkansah *et al*, 2016).

HLTF is a fork remodeler that promotes genomic stability

Another SNF2-family motor protein that has been characterized to catalyze fork regression is HLTF. Like SMARCAL1 and ZRANB3, it plays a critical role in promoting genomic stability as its inactivation leads to increased sensitivity to UV and MMS and also alters fork progression in cells under replication stress (Lin *et al*, 2011; Blastyák *et al*, 2010). Like other SNF2 ATPases, HLTF has additional domains to its SNF2 motor that are essential for its function (Figure 1.5). It has a HIP116/Rad5 N-terminal (HIRAN) domain, predicted to function as a SRD, and a conserved C3HC4-type RING domain positioned between the two SNF2 ATPase lobes. HLTF has a multi-faceted role in promoting genome stability through DDT mechanisms that requires both its RING and SNF2 motor domains (Minca & Kowalski, 2010; Ortiz-Bazán *et al*, 2014; Gangavarapu *et al*, 2006; Choi *et al*, 2015; Blastyák *et al*, 2010).

Through its RING domain, HLTF can act as an E3 ubiquitin ligase to polyubiquitinate PCNA as a response to replication stress *in vivo* (Motegi *et al*, 2008; Lin *et al*, 2011; Unk *et al*, 2008). It is able to interact with the ubiquitin conjugating enzyme

complexes, Rad6-Rad18 and Mms2-Ubc13 to form lysine-63 linked ubiquitin chains on PCNA (Motegi *et al*, 2008). Biochemical polyubiquitination assays have shown that HLTF
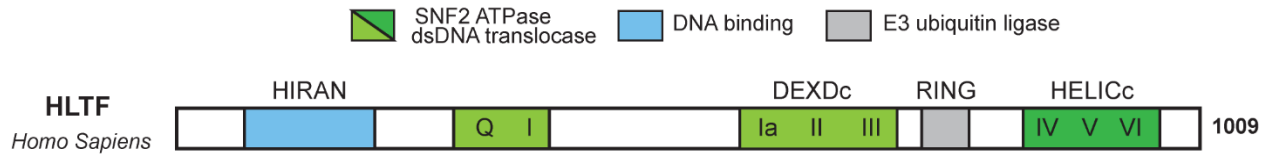
Figure 1.5 HLTF domain organization. Schematic of the domain organization of human HLTF. The conserved motifs within the SNF2 motor are also depicted.

modifies PCNA through an *en-bloc* transfer mechanism, where HLTF forms a ubiquitin chain on Ubc13 and transfers it to Rad6, which then is transferred to PCNA by Rad18 (Masuda *et al*, 2012). It has yet to be determined if this mechanism is relevant *in vivo*. PCNA polyubiquitination is necessary to promote the template switching pathway of DDT. It may also facilitate displacement of TLS polymerases, enable access of additional template switching factors, and act as a platform for recruitment of other proteins involved in fork stabilization, such as ZRANB3 (Ciccia *et al*, 2012; Saugar *et al*, 2014).

HLTF's SNF2 motor domain allows it to hydrolyze ATP to translocate on DNA and catalyze fork regression *in vitro*, by concertedly unwinding and annealing the nascent and parental strands (Blastyák *et al*, 2010). In addition, HLTF can clear proteins from a stalled fork, to allow fork remodeling and stabilization to occur. *In vitro* studies showed that HLTF removes PCNA, RPA, and replication factor C (RFC) from a fork substrate in an ATP-dependent manner (Achar *et al*, 2011). Furthermore, biochemical studies also showed that HLTF forms D-loops independently from Rad51. This study also suggests that unlike

Rad51, which is inhibited by RPA, HLTF can form D-loops when RPA is present on the invading ssDNA strand (Burkovics *et al*, 2014).

HLTF in disease

HLTF is important for cell viability. In mice, HLTF knockdown is semi-lethal, causes defects in the G2/M cell cycle transition, and results in brain and heart development defects. HLTF has also been shown to regulate several cellular processes including collagen biogenesis and angiogenesis, but it is not clear if this regulation is due to its function in promoting genomic stability or if its related to its function as a transcription factor (Helmer *et al*, 2013).

Interestingly, HLTF was first identified as a SV40 enhancer and HIV-1 promoter binding protein in human cells (Sheridan *et al*, 1995). Soon after, it was identified to bind the uteroglobin promoter in rabbit cells and was characterized as a transcription factor (Hayward-Lester *et al*, 1996). Many studies have characterized HLTF as a transcription factor and have concluded that HLTF binds to several promoters including the beta-globin enhancer, the tumor necrosis factor response element, the Gata4 promoter and the Hif-1a promoter (Dhont *et al*, 2016). Due to the nature of the techniques used in many of these studies, which did not test sequence-specific DNA binding, it is unclear if HLTF truly functions as a transcription factor that binds to specific promoters to regulate expression. Additionally, it is also unclear if this function is unrelated to its function in promoting genomic stability.

Due to its multi-faceted role in promoting genomic stability, it is no surprise that HLTF has been implicated in cancer progression. Alteration of HLTF expression, either by gene silencing or by expression of non-functional proteins leads to a variety of cancers. Although the mechanism is unknown, promoter hypermethylation leads to HLTF silencing in colon, gastric, uterine, and lung cancers. Conversely, HLTF is also upregulated in neck, head and cervix cancers (Dhont *et al*, 2016). These upregulated proteins include alternative spliced variants, resulting in truncated non-functional HLTF proteins (Hayward-Lester *et al*, 1996). The suppression of wild-type HLTF in these tumors suggests that HLTF is a driver of carcinogenesis, as its inactivation is seen in early stages of cancer and seen more prevalently in more advance stages of dysplasia. In addition, HLTF silencing is correlated with a poor prognosis for patient survival, disease recurrence and tumor aggressiveness (Dhont *et al*, 2016).

Scope of this work

This dissertation presents work characterizing the HIRAN domain of *H. sapiens* HLTF and explore how it impacts HLTF function. Chapter II describes the structural determination of HIRAN bound to a DNA substrate and biochemical characterization of its DNA-binding specificity. Chapter III further describes HIRAN's binding specificity, characterizing its binding to fork substrates, and shows how HIRAN promotes HLTF binding to a fork. This chapter also explores the role that HIRAN plays in promoting HLTF fork remodeling. Finally, Chapter IV discusses the implications of this study and discusses the ongoing work to structurally characterize HLTF. This work has expanded our knowledge of HLTF fork remodeling mechanisms and provides further understanding of

the action of fork remodeling enzymes.

CHAPTER II

The HIRAN Domain of HLTF Binds 3' DNA Ends to Drive Fork Regression

Introduction

Stalled replication forks and other downstream effects of replication stress are significant challenges to genome stability. Unrepaired DNA damage, secondary DNA structures, protein-DNA complexes, and nucleotide depletion can all lead to replication fork collapse and DNA breaks, and eventually to cell death and disease (Branzei and Foiani, 2010; Ciccia and Elledge, 2010; Zeman and Cimprich, 2014). Cells therefore have developed a number of mechanisms to respond to replication stress, bypass stalled replication forks, and repair damaged DNA. DDT pathways, for example, minimize fork stalling through bypass of replication blocks and leave resolution of the block for a later time (Branzei and Foiani, 2010; Saugar et al., 2014). In one form of DDT, TLS directly bypass lesions in what can be an error-prone process. In another form of DDT, an alternative template is used (template switching) to enable error-free bypass. Understanding the molecular pathways controlling these processes is key to understanding how cells maintain genomic integrity in the face of replication stress, and how errors in DNA damage response pathways lead to diseases such as cancer.

One versatile mechanism for overcoming stalled replication forks is fork reversal (also called fork regression). Fork reversal involves active remodeling of the stalled

replication fork, in which the three-armed fork is converted into a Holliday junction (HJ)-like structure by pairing the newly synthesized sister chromatids to form a fourth regressed arm. This structural change can have several protective effects on genomic integrity, such as repositioning fork-stalling lesions in the context of dsDNA to facilitate DNA repair (Neelsen and Lopes, 2015). Fork reversal creates an opportunity for template switching. This final process allows the indirect and error-free bypass of fork-blocking lesions using the undamaged sister chromatid. Genotoxic agents, repetitive DNA sequences, and oncogene-induced replication stress all lead to fork regression in mammalian cells, suggesting this process may be a common response to stalled forks (Follonier et al., 2013; Neelsen et al., 2013; Zellweger et al., 2015). Despite growing evidence for the importance of fork reversal in protecting the genome, and what may be the frequent use of this process in fork repair, how fork reversal occurs is not well understood.

Recent evidence suggests that molecular components of DDT might be involved in fork reversal. In yeast, error-free DDT is dependent on Rad5, a RING domain-containing ubiquitin ligase, which promotes the polyubiquitination of PCNA (Hoege et al., 2002). In mammalian cells, two Rad5-related proteins, SHPRH and HLTF, promote PCNA polyubiquitination (Motegi et al., 2008; 2006; Saugar et al., 2014; Unk et al., 2008; 2006). Rad5, HLTF, and SHPRH are all members of the SNF2 family of ATP-dependent DNA translocases involved in chromatin remodeling and DNA repair (Unk et al., 2010). Indeed, HLTF and Rad5 can directly catalyze replication fork reversal on model DNA substrates (Achar et al., 2011; Blastyák et al., 2010; 2007). In spite of these observations, it is not

known whether or how Rad5-related proteins mediate fork reversal in cells following replication stress.

Structural elements within Rad5-related proteins provide intriguing clues about their potential mechanism of action in fork reversal. Rad5, HLTF, and SHPRH share a domain structure that contains a ubiquitin ligase RING domain embedded within the SNF2 ATPase motor (Unk et al., 2010). In addition, HLTF and Rad5 contain an uncharacterized HIP116 Rad5 N-terminus (HIRAN) domain, which was identified fortuitously by sequence profile searches aimed at understanding the evolutionary classification of chromatin remodelers. HIRAN has been identified as a highly conserved domain from bacteria to humans. It can be found as a stand-alone protein in gram-positive and gram-negative bacteria and in eukaryotes it is coupled to DNA processing domains, including tyrosyl-DNA phosphodiesterase and viral replication and repair (VRR) nuclease domains, suggesting that the HIRAN domain may be integral to the function of these proteins. Based on domain architectures, conserved gene neighborhoods and functional interactions, HIRAN was predicted to be a DNA-binding domain that is important for recognition of DNA damage and stalled replication forks (Iyer *et al*, 2006). Consistent with this idea, deletion of the HIRAN domain in Rad8, the *S. pombe* ortholog of HLTF and Rad5, leads to DNA damage sensitivity (Ding & Forsburg, 2014).

In this study, we describe, in collaboration with the Cimprich lab, an important role for the HIRAN domain in driving replication fork regression by HLTF. Using biochemical, structural, and genetic approaches, we establish that the HIRAN domain recognizes 3′-ssDNA ends and directs HLTF to the 3′-end of the nascent leading strand to remodel replication forks. This requirement for the 3′-end is unique among factors involved in

replication fork reversal and the 3′-end binding activity appears to be a conserved activity of the ancient HIRAN domain. Lastly, we demonstrate this activity is required for fork progression in cells by showing that the 3′-end binding function of HIRAN affects the length of newly synthesized DNA fibers. Our findings indicate that the HIRAN domain is a substrate specificity factor for HLTF that dictates its biological activity, and thus provide important insights into the distinct mechanism by which HLTF recognizes and remodels replication forks.

Results

HIRAN recognizes the 3′ end of ssDNA

Even though HLTF is capable of binding and regressing a DNA fork, it is not known how it recognizes these fork structures (Blastyák et al., 2010). We were intrigued by the possibility that the evolutionarily conserved (Figure 2.1), yet functionally uncharacterized N-terminal HIRAN domain of HLTF is important for its DNA recognition and remodeling activities. To investigate if HIRAN functions to recognize a model fork, we first tested the DNA-binding ability of the HIRAN domain. We purified the HIRAN domain (residues 55-180) of human HLTF and tested its ability to bind ssDNA or dsDNA substrates using electrophoretic-mobility shift assays (EMSAs). Surprisingly, we found that the HIRAN domain bound the ssDNA substrate with strong affinity ssDNA (K_d = 13 ± 3 nM), but showed no detectable affinity for the dsDNA substrate. This result contradicts work where HIRAN was observed to bind to short dsDNA oligonucleotides containing both blunt ends and 5′ ends. It is possible that the duplex regions in these short oligonucleotides

denatured resulting in ssDNA regions for HIRAN to bind (Hishiki *et al*, 2015). To further probe this ssDNA specificity, we tested a series of ssDNA substrates where the 3' contained a hydroxyl (OH), a phosphate (PO_4), a fluorescein (FAM), or hydrogen (H). Interestingly, HIRAN only bound to the ssDNA that contained a hydroxyl at the 3' end. These results indicate that HIRAN interacts specifically with the 3' end of ssDNA.

Figure 2.1 Conservation of HIRAN DNA binding residues. A. HLTF domain schematic B. Sequence alignment of the HIRAN domain from several eukaryotic and prokaryotic HIRAN-containing proteins. Protein names refer to UniProt identifiers. Sequences were aligned using Clustal Omega and visualized using the ESPript 3.0 server (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi). Downward arrows and yellow shading indicate the HLTF HIRAN residues that interact with DNA (see also Figure 2.3), with the secondary structural elements of this region placed above them. The boxes with red text highlight additional regions of conservation.

Structure of HIRAN bound to ssDNA

To elucidate the molecular details of HIRAN's binding specificity, we determined a structure of HIRAN bound to a ssDNA substrate using X-ray crystallography. This model contains four HIRAN-DNA complexes in each asymmetric unit (asu) and was refined against data extending to 1.5 Å resulting in a crystallographic residual (R/R$_{free}$) of 15.2%/18.4% (Table 2.1). In each complex in the asu, 125 HIRAN residues and at least four 3' nucleotides are resolved. Our model reveals that HIRAN adopts a globular α/β architecture with an embedded modified oligosaccharide/oligonucleotide binding (OB)-fold (β2-β6) (Figure 2.3). OB-folds are a general nucleic acid binding platform that are present in many ssDNA binding proteins. Two loops between β1 and β2 (L12) and β2 and β3 (L23) extend from the OB-fold and with the β2-β3-β4/ αB surface form a binding pocket for the 3' end of ssDNA. This pocket is positively charged and accommodates two nucleotides that are well ordered at the 3' end. The 5' ends of the DNA are directed away from the pocket and fold (β2-β6). OB-folds are a general nucleic acid binding platform that are present in many ssDNA binding proteins. Two loops between β1 and β2 (L12) and β2 and β3 (L23) extend from the OB-fold and with the β2-β3-β4/ αB surface form a binding pocket for the 3' end protein surface by a loop between β5 and β6 (L56) and

participate in crystal contacts with the neighboring protein molecules. As a consequence, these 5' ends are highly flexible and their conformations are highly variable.

Comparison to other structures in the Protein Databank (PDB) revealed that the HLTF HIRAN domain is virtually identical to that of a predicted HIRAN domain within a protein of unknown function from *Lactobacillus plantarum* (PDB: 3K2Y, Figure 2.5). This similarity indicates that the HIRAN fold is conserved in organisms separated by more than

a billion years of evolution. The 3′-binding pocket in HIRAN is positively charged also strongly conserved, as indicated by mapping sequence homology of HIRAN domains from 150 proteins onto the crystal structure (Figure 2.4).
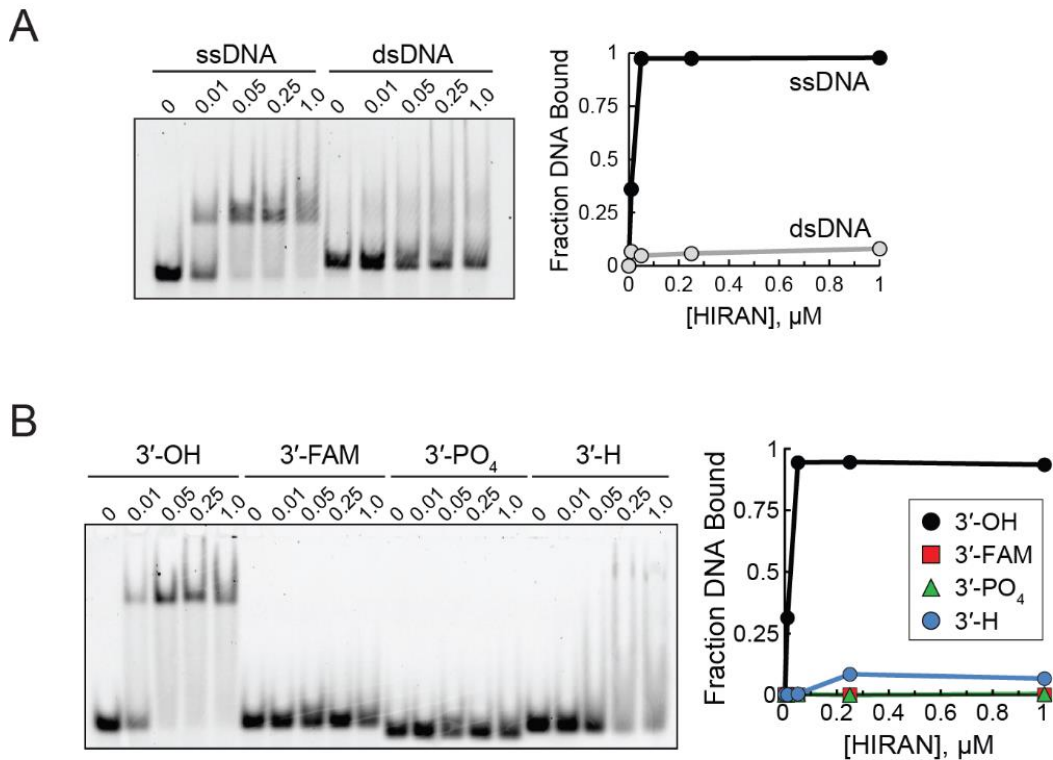
Figure 2.2 HIRAN binds 3´ end of ssDNA. A. EMSA of HIRAN binding to 5′-FAM labeled 40-mer ssDNA and dsDNA. Quantitation of the gel is shown on the right. **B.** EMSA of HIRAN binding to 5′-FAM-labeled dT_{10} oligonucleotides modified as shown at the 3′-ends. Quantification of the gel is shown on the right.

We also found that the HIRAN domain shares significant structural homology to small protein B (SmpB, PDB: 2CZJ), the tRNA-binding component of the bacterial transfer messenger RNA machinery (Bessho et al., 2007; Dong et al., 2002) (Figure 2.6). Interestingly, SmpB binds an internal segment of RNA using the same general
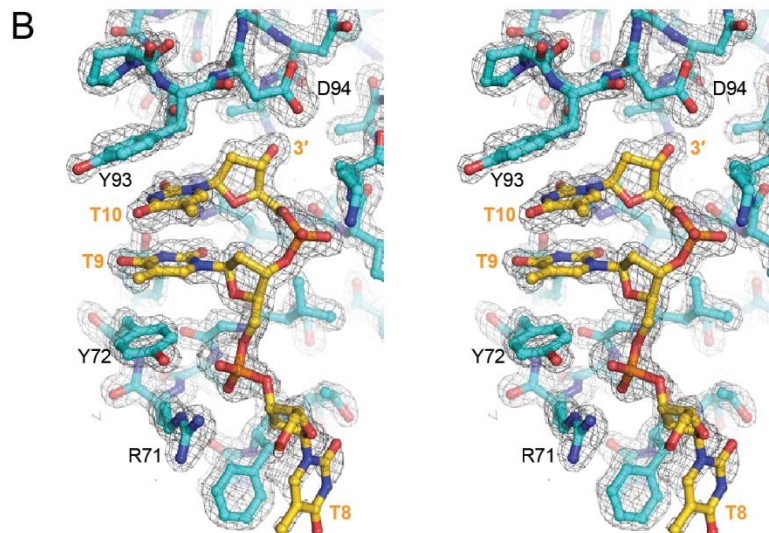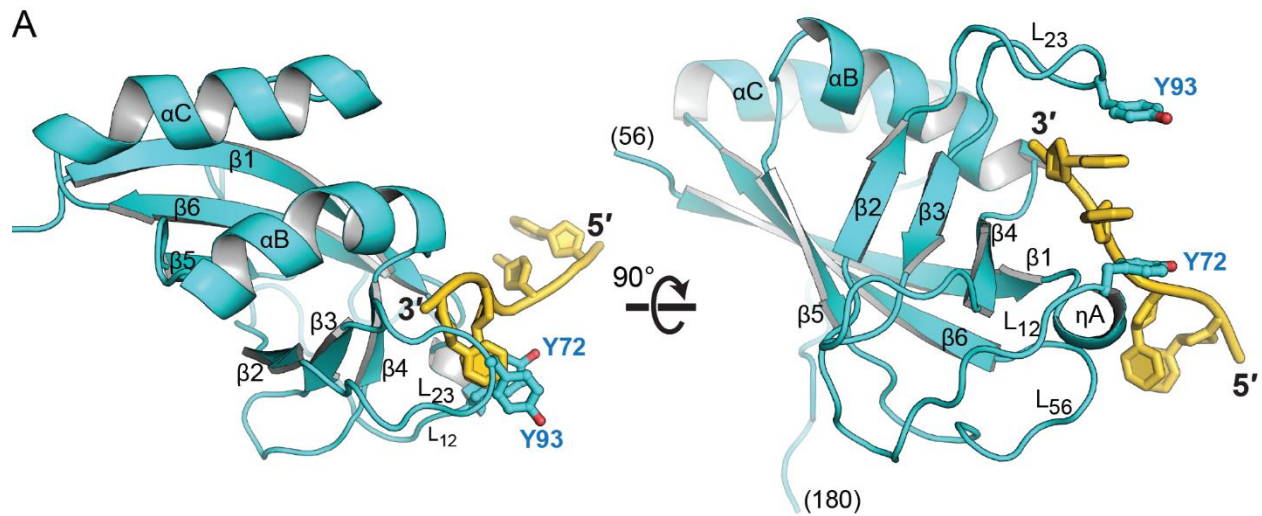
Figure 2.3 Structure of HLTF-HIRAN bound to DNA A. Crystal structure of the HLTF HIRAN domain bound to dT_{10} ssDNA (gold). **B.** Stereo-images of the refined HIRAN-ssDNA crystallographic model superimposed onto the 1.5 Å experimental SIRAS electron density map contoured at 1.5σ.

surface that HIRAN uses to bind ssDNA (Bessho et al., 2007), although the specific

HIRAN-DNA and SmpB-RNA contacts are distinct. Similarly, the 3′-end binding of the

HIRAN OB-fold is distinctly different than the manner in which RPA binds ssDNA

(Theobald et al., 2003).

Upon comparing the HIRAN architecture to other known 3′-end binding domains,

we found structural similarity to nucleic acid binding proteins from organisms throughout

evolution, ranging from the 3′-DNA binding domain (3′BD) of the bacterial PriA replication

restart helicase to the 3′- RNA binding PAZ domain of human Argonaute-1 (Ma et al.,

2004; Sasaki et al., 2007). Each of these end-binding domains use a topologically distinct

arrangement of β-strands to achieve a similar 3D architecture. Indeed, superposition of
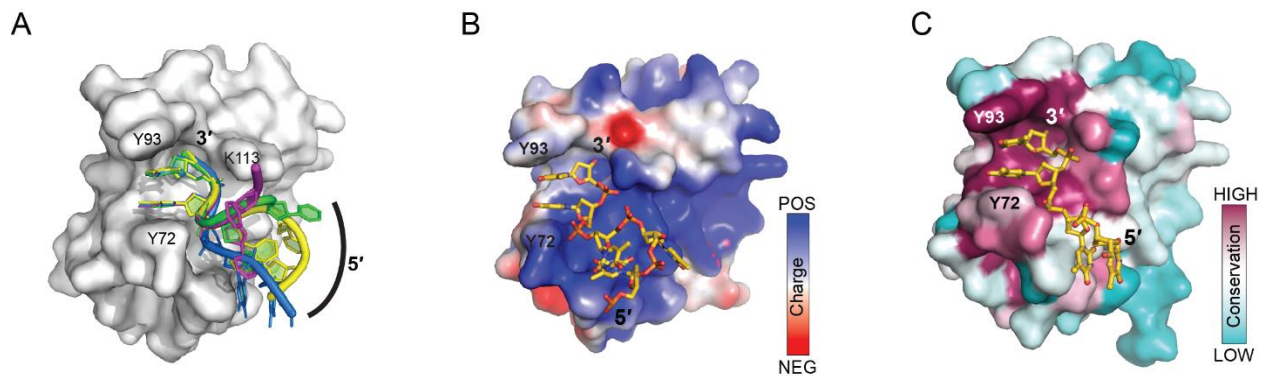
Figure 2.4 HIRAN DNA binding pocket. A. Superposition of the four HIRAN protein-DNA complexes in the asymmetric unit (RMSD = 0.378 Å for all atoms in residues 56-173). Protein is shown as a solvent accessible surface. The 2 nucleotides at the 3′-ends are bound in identical positions relative to the protein, whereas the conformations of the unbound 5′-ends are variable. **B.** Solvent-accessible surface of HIRAN colored by electrostatic potential **C.** Sequence conservation from 150 HIRAN orthologs mapped onto the solvent accessible surface of the HIRAN crystal structure using the ConSurf web server (http://consurf.tau.ac.il/).

Table 2.1 HIRAN-ssDNA X-ray data collection and refinement statistics

	Native	SeMet
Data Collection [a]		
Space group	$P2_1$	$P2_1$
Cell Dimensions		
a, b, c (Å)	61.03, 74.21, 66.18	60.87, 74.28, 65.93
a, b, g (°)	90.00, 113.68, 90.00	90.00, 113.76, 90.00
Wavelength	0.97872	0.97872
Resolution (Å)	100–1.50 (1.55–1.50)	100–1.92 (1.99–1.92)
R_{sym}	0.068 (0.455)	0.097 (0.507)
I/s I	27.2 (3.9)	20.3 (4.3)
Completeness (%)	99.8 (100)	100 (100)
Redundancy	7.6 (7.5)	7.6 (7.6)
Refinement		
Resolution (Å)	1.5	
No. reflections	86,121	
R_{work}/R_{free}	0.1528/0.1849	
No. atoms	4,933	
Protein/DNA	3,939/389	
Solvent	605	
B-Factors		
Protein/DNA	32.4/48.9	
Solvent	39.7	
Rms Deviations		
Bond lengths (Å)	0.006	
Bond angles (°)	0.973	

[a]Numbers in parentheses refer to data in the highest-resolution shell.
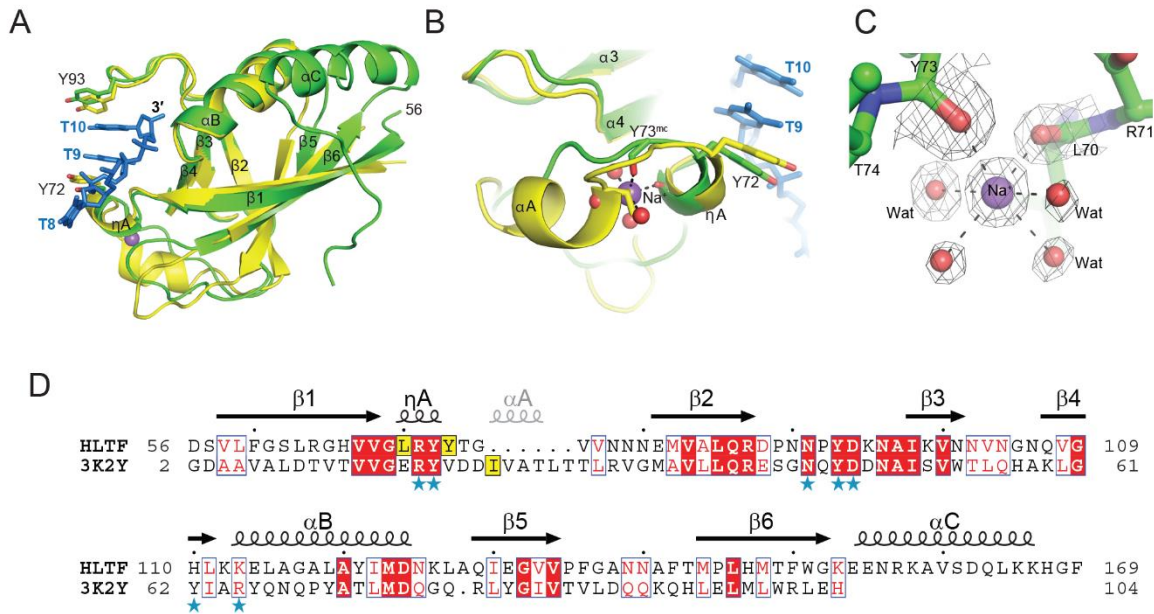
Figure 2.5 Comparison of the HIRAN domains from human HLTF and Lactobacillus plantarum. A. Human HLTF-HIRAN ssDNA (green) complex is shown superimposed onto the crystal structure of the Lactobacillus plantarum lp_0118 (PDB: 3K2Y, yellow) HIRAN domain. The RMSD for all backbone atoms between the two structures is 1.23 Å. **B.** Close-up of the most significant difference between the two HIRAN crystal structures. In the HLTF HIRAN domain, a sodium ion (Na+, purple sphere) is coordinated by 4 water molecules (red spheres) and carbonyl oxygens of residues L70 and Y73, which flank the 310-helix (ηA). This ion is absent in the bacterial protein, which instead contains an inserted alpha helix not present in HLTF and an isoleucine side chain (I22) located in the position of the sodium ion. **C.** 2Fo-Fc electron density contoured at 1.5σ superimposed onto the hexacoordinate sodium ion complex. **D.** Structure-based sequence alignment, showing invariant (red boxes) and conserved (red type) residues. Residues in the location of the sodium ion in HLTF-HIRAN are highlighted yellow. Blue stars denote DNA-binding residues in HLTF.

28

these structures places the nucleic acid binding surfaces in the same location relative to the β-sheet motifs (Figure 2.6). Thus, the HIRAN fold represents a general nucleic acid binding architecture that HLTF and other proteins have adapted to bind specifically to 3′-ends. We note that not all 3′-binding domains show this same architecture. For example, Flap endonuclease-1 (FEN-1) captures the 3′-end at a DNA nick using an α-helical domain with no structural resemblance to HIRAN, PriA-3′BD, or PAZ domains (Tsutakawa et al., 2011).

Analysis of our crystal structure revealed several interactions that form the basis of HIRAN's specificity for the 3' end of ssDNA. First, the DNA 3' hydroxyl group is nestled deep in the back of the pocket and interacts with the carboxyl side chain of D94 through hydrogen bonding. This explains the requirement for a free 3' OH for binding. Additionally, two nucleobases at the 3' end are stacked between two tyrosine side chains (Y72 andY93) that extend from loops L12 and L23. These two nucleobases also participate in Watson-Crick hydrogen bonding with Y73, N91, and H110. All of these interactions exclude dsDNA from binding to this pocket. Consistent with this, a recently published crystal structure (Hishiki *et al*, 2015) of HIRAN with dsDNA (PDB: 4XZF) showed the domain bound to two un-duplexed nucleotides at the 3' end, in a highly similar manner to our structure (RMSD = 0.51 Å for all atoms). Furthermore, the phosphate of the two 3' nucleotides are stabilized by electrostatic interaction with R71 and K113 side chains as well as a hydrogen bond from the Y72 hydroxyl group. Outside of the binding pocket, the third nucleotide from the 3' end is base-stacked against F142 from loop L56, which results a 90° kink. Thus, the two nucleotides in the binding pocket are stabilized by an extensive network of interactions. Thus, the two nucleotides at the 3′-end are stabilized by an

extensive network of interactions, and we observe no binding outside of the 3′-pocket. Notably, all of the residues in the DNA binding pocket are conserved in the *Lactobacillus* lp_0118 protein with the exception of H110, which in lp_0118 is a tyrosine capable of analogous hydrogen bonding with the DNA bases. This striking conservation indicates that 3′-binding is conserved in the bacterial HIRAN proteins.

Characterization of HIRAN DNA binding residues

To investigate the functional importance of the residues mediating the interactions between HIRAN and the 3'end of ssDNA, we purified mutant HIRAN domains containing substitutions at each of the residues identified above and tested the ability of these mutants to bind ssDNA using EMSAs. None of the substitutions significantly destabilized the protein, as determined by thermal denaturation profiles (Figure 2.7). All of the mutations diminished HIRAN's ability to bind a ssDNA substrate (Figure 2.8). The most significant reduction was seen when R71 was substituted to a glutamate, with 80% to 100% (depending on the buffer conditions Figure 2.7A vs 2.7B) reduction in binding. The K113E, N91A, D94A, and H110A mutations decreased ssDNA binding by three orders of magnitude. Substitution of the tyrosines involved in base-stacking to alanines (Y72A and Y93A) had a more modest effect and resulted in a 60% decrease in binding and a 15% decrease in binding (F142A). These results indicate that all residues lining the DNA binding pocket are critical for HIRAN binding to DNA, and that at least in our experimental conditions, the polar contacts to the phosphoribose backbone and nucleobases contribute more to the strength of the interaction than do the base stacking interactions.
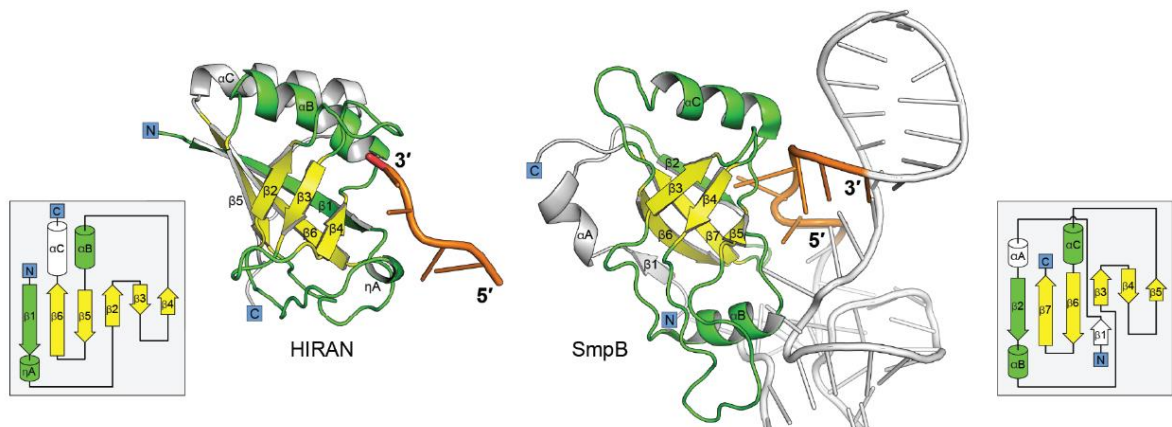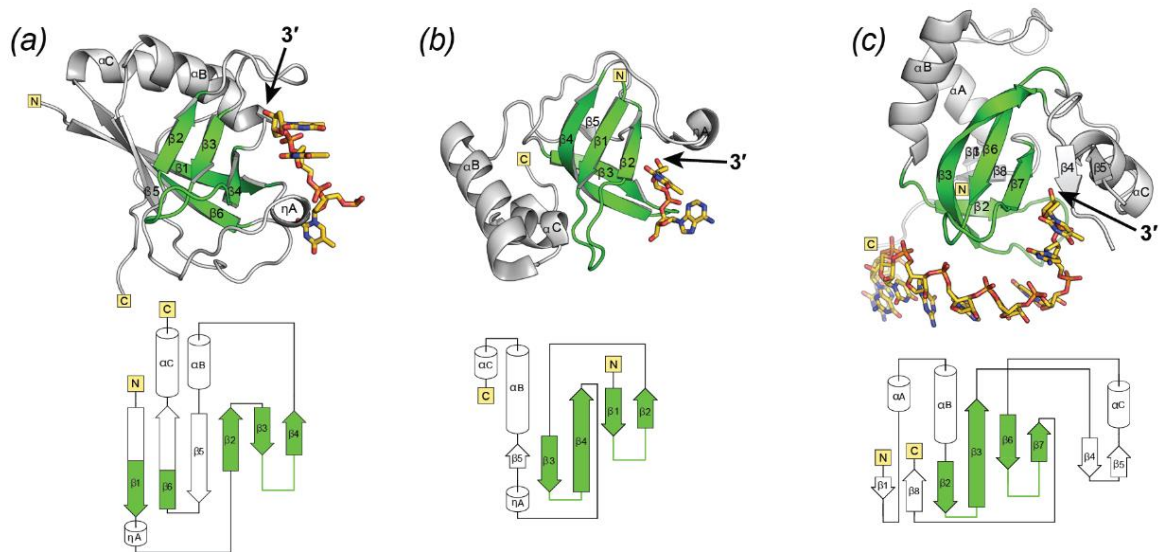
Figure 2.6 HIRAN has a general nucleic acid binding fold. A. Comparison of HLTF HIRAN-ssDNA and SmpB-tRNA (PDB: 2CZJ) structures. Homologous protein secondary structural elements are shown in green, OB-folds in yellow, and nucleic acid segments contacting the OB-fold in orange. Topology diagrams are shown on the side and colored according to the structures. **B.** Comparison of 3′-end binding domains. **(a)** HLTF HIRAN domain. **(b)** PriA 3′BD (PDB: 2DWM). **(c)** Argonaute-1 PAZ domain (PDB: 1SI2). Homologous secondary structural elements are colored green. Topology diagrams are shown below and colored according to the structures.

HIRAN is required for HLTF fork regression

The interaction of the HIRAN domain with ssDNA raised the possibility that this domain might facilitate HLTF-catalyzed DNA remodeling by directing its translocase activity to specific substrates *in vivo*. To test whether 3′-end binding by the HIRAN domain is required for HLTF's known activities, we purified a panel of full-length HLTF proteins with single mutations of the residues we found to compromise the HIRAN-ssDNA interaction. We also purified a D557A/E558A (DEAA) Walker B mutant previously shown to lack ATPase and fork regression activity (Blastyák et al., 2010; Gangavarapu et al., 2006).

In contrast to the DEAA mutant, all of the HIRAN mutants were proficient in DNA-dependent ATPase activity (Figure 2.9), indicating that DNA binding by HIRAN is not required for HLTF's ATPase activity and that the HIRAN mutations do not noticeably disrupt the functional integrity of HLTF. Next, we asked whether mutating the HIRAN domain affected the ability of HLTF to catalyze fork regression on a model fork structure. Strikingly, the HIRAN mutants had a reduced ability to reverse a model replication fork (Figure 2.9). The R71E and the Y72A/Y93A mutations caused the greatest defects in fork regression, whereas the D94A or H110A mutations had moderate effects. Together, these data indicate that DNA binding by the HIRAN domain is important for HLTF's fork regression activity through a mechanism other than impairment of ATP hydrolysis.
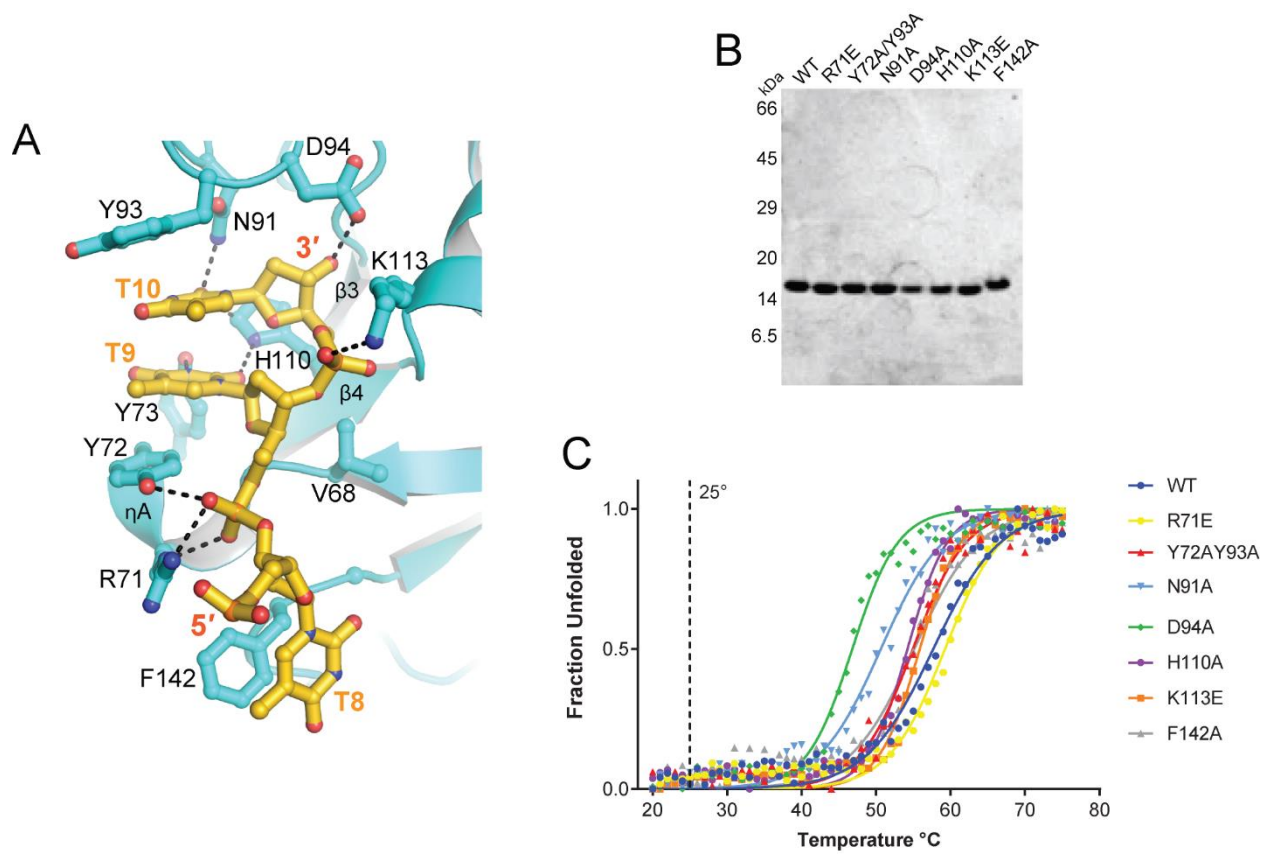
Figure 2.7 HIRAN DNA binding mutations. A. Up-close view of the residues in the HIRAN DNA-binding pocket. DNA is shown in yellow, and HIRAN in teal. **B.** SDS-PAGE gel of HIRAN WT and mutant proteins. **C.** Thermal denaturation of WT and mutant HIRAN monitored by circular dichroism. A dashed line is shown at the temperature that binding experiments were done.
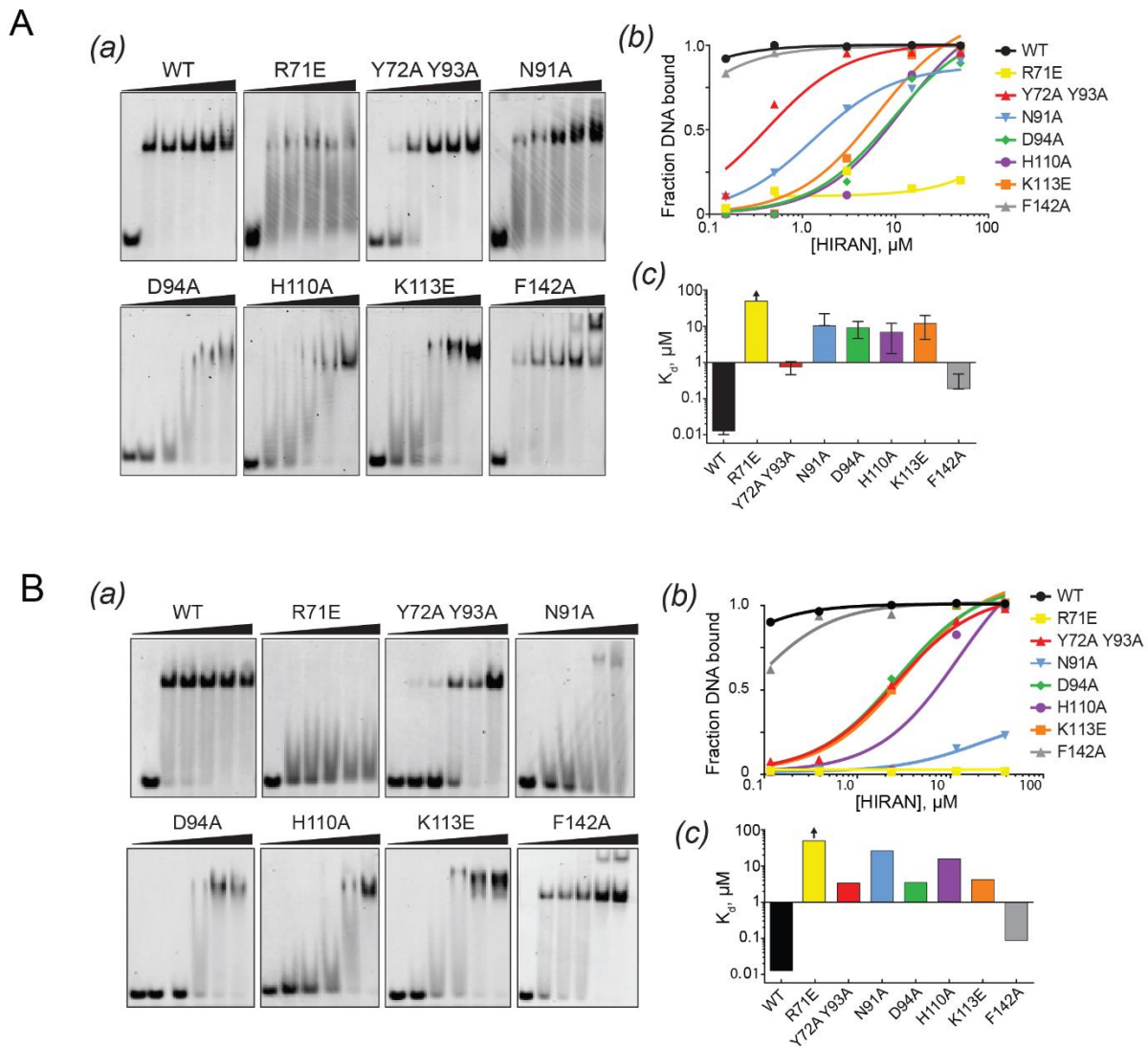
Figure 2.8 Residues in HIRAN pocket are important for DNA binding. HIRAN binding experiments performed in a **A.** HEPES buffer or **B.** TRIS buffer. **(a)** EMSAs of FAM-ssDNA in the presence of increasing concentrations of wild-type (WT) and mutant HIRAN proteins. Concentrations of protein were chosen to show the full range of binding for all mutants. **(b)** Quantitation of the data shown in panel E. Zero [HIRAN] is not plotted. **(c)** Dissociation constants (K_d) extracted from the binding isotherms shown in panel. Absolute K_d values for WT and F142A are approximations since the transition range was not defined. Values represent the average ± S.D. of three independent measurements for **A-(c).**

3'-DNA ends uniquely promote HLTF-dependent fork regression

Stalled replication forks contain a 3′-hydroxyl group on the nascent leading strand, and the requirement for the HIRAN domain in HLTF-mediated fork regression suggested that this group might be required for HLTF to promote fork regression. To test this hypothesis, we compared HLTF's ability to regress model replication forks containing either an unmodified (3′-OH) or phosphorylated (3′-PO$_4$) 3′-end on the nascent leading strand and compared this activity to that of SMARCAL1, another SNF2 protein shown to catalyze fork reversal *in vitro* (Bétous et al., 2012; Ciccia et al., 2012). We found that HLTF had reduced activity toward the 3′-PO$_4$-capped fork (Figure 2.10A). Strikingly, the 3′-PO$_4$-capped fork structure was efficiently reversed by SMARCAL1, and there was no difference in the ability of this enzyme to reverse the capped versus uncapped substrate. Together, these findings suggest the requirement of the free 3′-end of the nascent leading strand is specific to HLTF, and thus indicate that there is a difference in how HLTF recognizes and reverses stalled forks as compared to SMARCAL1.
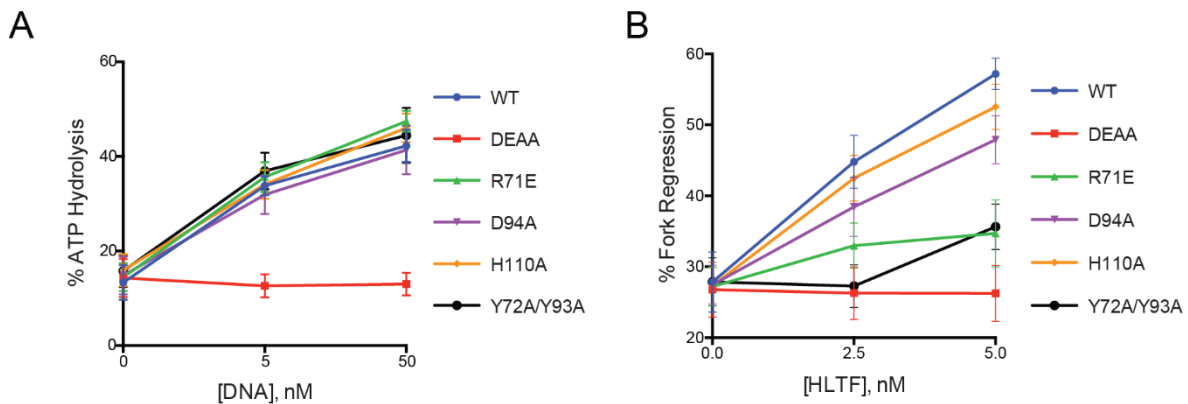
Figure 2.9 HIRAN is necessary for efficient HLTF fork regression. A. ATPase activity or **B.** fork regression activity of WT vs. ATPase-dead (DEAA) or HIRAN-mutant HLTF proteins incubated with the indicated amount of splayed-arm DNA (**A**) or HLTF proteins (**B**). Experiments were performed by Andrew Kile.
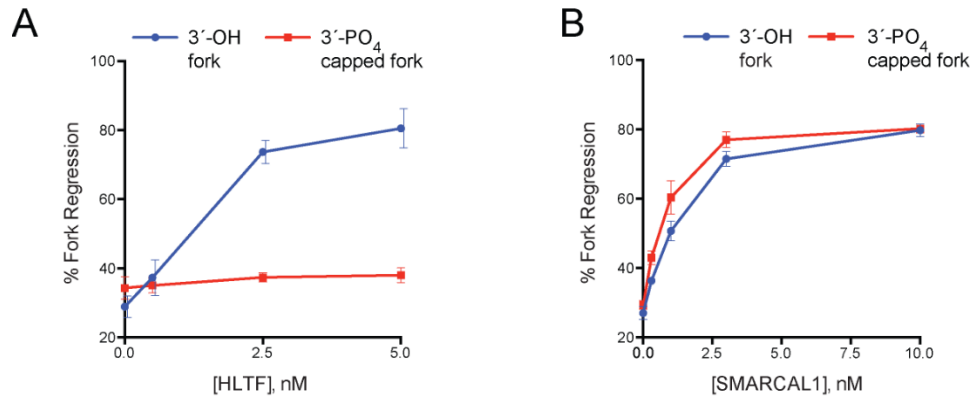
Figure 2.10 3´OH end promotes HLTF fork regression. Fork regression activity for **A.** HLTF or **B.** SMARCAL1 on model fork substrates with either a 3′-OH or a 3′-PO$_4$ end on the leading strand. The mean values ± SEM for three replicate experiments are shown in each graph. Experiments were performed by Andrew Kile.

A number of mammalian enzymes other than HLTF can also catalyze fork reversal *in vitro*, including SMARCAL1 (Bétous et al., 2012; Ciccia et al., 2012). To determine if SMARCAL1 is similarly affected by the 3′-PO$_4$-capped fork structure, we tested its activity using this as a substrate. Strikingly, the 3′-PO$_4$-capped fork structure was efficiently reversed by SMARCAL1, and there was no difference in the ability of this enzyme to reverse the capped versus uncapped substrate (Figure 2.10B). Together, these findings suggest the requirement of the free 3′-end of the nascent leading strand is specific to HLTF, and thus indicate there is a fundamental difference in how HLTF recognizes and reverses stalled forks as compared to SMARCAL1.

HLTF associates with the replication fork

Fork reversal by HLTF, if relevant *in vivo*, would be expected to occur on chromatin and at the replication fork. To determine whether HLTF is present at active replisomes, we used iPond (Isolation of Proteins on Nascent DNA) (Sirbu et al., 2011) to capture

nascent, EdU-labeled DNA and its associated proteins from proliferating cells. We found that HLTF is associated with EdU-labeled DNA immediately after the EdU pulse, but not 10-30 minutes after the EdU was washed out (Figure 2.11). Rad18, a RING-domain containing protein that interacts with HLTF (Motegi et al., 2008; Unk et al., 2008), also associated with nascent EdU-labeled DNA with similar kinetics, as did RPA and PCNA, two well-established markers of the replication fork. Together, these findings demonstrate that HLTF is a component of active replisomes, but not mature chromatin. Furthermore, they suggest that HLTF associates with DNA at a replication fork, even in the absence of exogenous DNA damage or replication stress.

The HIRAN domain is necessary for HLTF to slow replication forks upon dNTP depletion

Next, we wanted to determine whether HLTF function is required for replication fork remodeling in cells, and the contribution of the HIRAN domain to this function. Nucleotide depletion induced by hydroxyurea (HU) is commonly used to disrupt fork progression in eukaryotic cells and has recently been shown to promote fork reversal (Zellweger et al., 2015). As HU would also expose the nascent 3′-end of DNA, we reasoned that HLTF's fork reversal activity might be required under these conditions. To test this hypothesis, we monitored replication fork progression in the presence of 50 μM HU, a dose that reduces fork progression by approximately 45% as measured by fiber assay. (Jackson and Pombo, 1998). To disrupt HLTF function, we knocked out the *HLTF* gene in U2OS cells (Figure 2.12A-D). Using two *HLTF*-knockout lines and the parental line, we then monitored the effects of HLTF loss on replication fork progression using the fiber assay. We pulsed U2OS cells with IdU (30 min, red) followed by CldU (30 min, green)
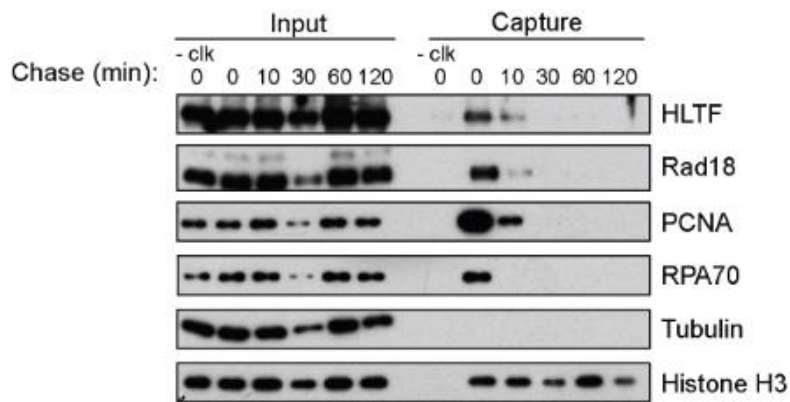
Figure 2.11 HLTF is enriched at replication forks. Western blot of nascent DNA-protein complexes found at replication forks isolated using iPOND. 293T cells were pulsed with EdU for 10 min and chased with thymidine for the time   shown. The -clk condition represents cells pulsed with EdU, and processed without biotin-azide in the click reaction step. Experiment was performed by Andrew Kile.

in the presence or absence of 50 μM HU, and then measured CldU tracks (Figure 2.12B).

Surprisingly, *HLTF*-knockout cells had substantially longer CldU tracks than did the

parental cells in the presence of HU (Figure 2.12C, D). However, the knockout of *HLTF*

had no effect in untreated cells (Figure 2.12C, D). These unexpected observations

suggest that HLTF functions at the stalled replication fork to restrain replication fork

progression and are consistent with the idea that HLTF has a fork remodeling activity that

slows fork progression under conditions of replication stress.

To determine if the fork progression phenotype is dependent on HLTF's HIRAN

domain, we compared the ability of wild-type HLTF and the HIRAN mutants to rescue this

phenotype in one of the *HLTF*-knockout lines. Expression of wild-type HLTF in the *HLTF*-

knockout line restored the CldU track length to wild-type levels (Figure 2.12E), but

expression of the HIRAN mutants, or the ATPase mutant did not restore CldU track length

to wild-type levels. Thus, the HIRAN domain is necessary for HLTF's function at the

replication fork. Taken together, these findings strongly suggest that HLTF constrains replication fork progression through HIRAN-mediated replication fork reversal.

Discussion

In this study, we present biochemical, structural and biological evidence that the HIRAN domain of HLTF is a 3′-ssDNA end-binding module important for replication fork reversal and proper replication fork progression following replication stress. Several lines of data support these conclusions. First, the HIRAN domain of HLTF binds tightly and specifically to the 3′-end of ssDNA. Second, residues lining the ssDNA binding pocket identified in the crystal structure of the HIRAN domain are critical for 3′-end binding in solution. Third, the ability of HLTF to efficiently regress model replication forks is dependent on a functional HIRAN domain and a free 3′-hydroxyl group on the nascent leading strand of the model fork structure. Finally, the HIRAN domain is required for HLTF to restrain replication fork progression *in vivo*. These findings provide crucial insights into the mechanism of replication fork regression by HLTF, and suggest that HLTF helps maintain genome stability by promoting replication fork reversal following replication stress.

The HIRAN domain was first identified as an evolutionarily conserved domain of unknown function, but its location within proteins containing DNA processing domains led to the prediction that it was needed to associate with damaged DNA or stalled replication forks (Iyer et al., 2006). However, thus far there has been no evidence to support this hypothesis. Our biochemical and structural data now show the HIRAN domain to be a *bona fide* DNA-binding domain with an unexpected 3′-end binding activity that is essential

39

for HLTF-dependent fork reversal. The strong structural similarity between human HLTF and *Lactobacillus* lp_0118 (Figure 2.5) HIRAN domains, and the conservation of the DNA binding residues in all of the known HIRAN sequences (Figure 2.1) implies that this 3′-end binding activity is a universal feature of HIRAN-containing proteins across all kingdoms of life in which they appear. The HIRAN structure is adapted from an OB-fold, a general nucleic acid binding motif most commonly associated with ssDNA binding but also known to interact with a variety of single- and double-stranded DNA and RNA structures. To our knowledge, however, the specific interaction with the end of the DNA is unique to the HIRAN OB-fold.

Possibly the most revealing aspect of the structure and 3′-binding function of the HIRAN domain is its similarity to the PriA-3′BD. Both domains are coupled to a superfamily II type ATPase motor (Bhattacharyya et al., 2014), and capping the 3′-end of the nascent leading strand in model fork structures disrupts HLTF-mediated fork regression and PriA-fork binding (Mizukoshi et al., 2003). The PriA-3′BD is proposed to orient the protein at a stalled fork to enable the helicase domain to unwind the nascent lagging strand duplex and create the ssDNA necessary for reloading of DnaB helicase and the replication restart primosome (Gabbai and Marians, 2010; Jones and Nakai, 1999).

In contrast, we propose that HLTF, which does not have strand unwinding activity, combines its dsDNA translocase activity (Blastyák et al., 2010) with HIRAN's 3′-end binding function to recognize and reverse stalled forks (Figure 2.12F). We hypothesize that HLTF binds to dsDNA on the unreplicated template ahead of the replication fork, and
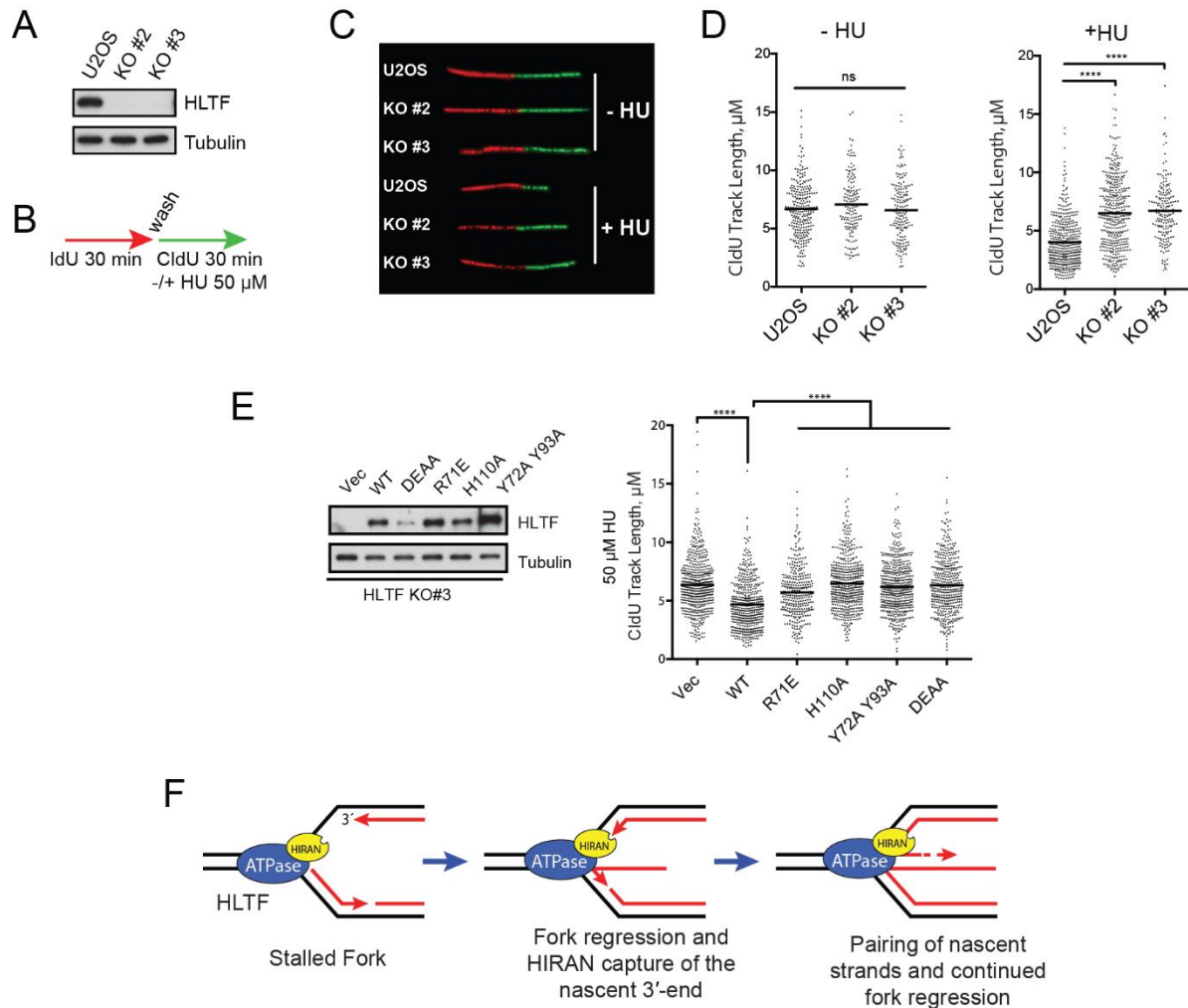
Figure 2.12 Loss of HLTF leads to longer DNA replication tracks upon depletion of nucleotide pools. A. Expression of HLTF in U2OS cells or CRISPR-generated HLTF knockout U2OS cells. B. Experimental setup. Cells were pulsed with IdU (30 min), then incubated with CldU and 50 μM hydroxyurea (HU) for 30 min. **C.** Representative IdU and CldU replication tracks in WT U2OS and HLTF-knockout clones. **D.** *Left*, dot-plot of CldU replication track lengths in the indicated cell lines. *Right*, dot-plot of CldU replication tracks in the indicated cell lines after treatment with 50 μM HU as in *B*. In both experiments the line represents mean. **** = p < 0.0001, by two-tailed nonparametric Mann-Whitney test. **E.** Left, western blots of lysates from HLTF KO #3 cells transfected with empty vector (vec), WT or mutant forms of HLTF. Right, dot plot of CldU replication track lengths in HLTF KO #3 cells transfected as on the left and treated with HU as in *B*. Line represents mean. **** = P < 0.0001 were calculated using two-tailed nonparametric Mann-Whitney test. **F.** Model of fork regression by HLTF, which utilizes the HIRAN domain to drive replication fork reversal. Experiments were performed by Andrew Kile.

uses its translocase activity to re-anneal the unwound template strands as it moves toward the stalled fork. Upon reaching the nascent leading strand, the translocase activity may destabilize the leading strand duplex allowing the HIRAN domain to capture the 3′-ssDNA end. This would facilitate annealing of the nascent strands and formation of a four-way HJ. The translocase activity of HLTF would then promote branch migration and further fork reversal. In support of this 3′-capture model, our crystal structure and biochemistry show that HIRAN is not able to bind a duplexed 3′-end.

Several additional lines of biochemical evidence support the proposed mechanism for fork reversal. First, HLTF binds dsDNA, a property that is likely associated with its SNF2 motor domain (Blastyák et al., 2010; Dürr et al., 2005; Singleton et al., 2007) and not its HIRAN domain. Furthermore, HLTF does not appear to have a preference for binding to different fork structures (Blastyák et al., 2010). Second, the ability of HLTF to reverse a model replication fork *in vitro* is significantly diminished when key DNA binding residues in the HIRAN domain are mutated, indicating that the HIRAN domain is needed for its biochemical activity. In contrast, the ATPase activity of HLTF is not affected by these mutations, indicating that ATPase activity is separable from 3′-end binding. Finally, model replication forks in which the leading strand 3′-OH has been capped are poor substrates for HLTF.

The HIRAN domain of HLTF may be analogous to the N-terminal HARP domain of SMARCAL1 in that both act as substrate recognition domains (Bétous et al., 2012; Mason et al., 2014). However, differences in the HIRAN and HARP structures as well as our finding that HLTF, but not SMARCAL1, requires a free 3′-end on the nascent leading strand during fork regression, suggest that these two translocases utilize different

mechanisms to recognize and/or remodel their substrates. The HIRAN domain may direct HLTF to forks where the 3′-end is exposed, whereas the HARP domains of SMARCAL1 likely promote fork reversal by recognizing a particular conformation of DNA at the branch point (Bétous et al., 2012; Mason et al., 2014). ZRANB3, another SWI/SNF2 translocase capable of fork reversal, also exhibits a substrate preference distinct from SMARCAL1, although the molecular basis of this preference is unknown (Bétous et al., 2013; Ciccia et al., 2012). These differences in substrate recognition could allow HLTF, SMARCAL1 and ZRANB3 to act on different types of stalled fork structures. By extension, these results may also indicate that many fork remodeling proteins exist in mammalian cells because forks stalled by distinct obstacles require several different modes of fork recognition and remodeling.

Our cellular data support an *in vivo* role for HLTF in promoting replication fork reversal. First, we show using iPond that HLTF and Rad18 associate specifically with nascent DNA in cells; a proteomic analysis utilizing a related approach also identified HLTF at replication forks (Alabert et al., 2014). These findings suggest that HLTF travels with the replication fork, which would enable it to respond rapidly to DNA damage or replication stress. Second, we also find that depletion or deletion of HLTF leads to longer replication tracts under conditions of replication stress (Figure 2.12D). HU dramatically increases the frequency of replication fork reversal in cells, and a defect in fork reversal could account for the effect of HLTF depletion on fork progression in the presence of HU (Zellweger et al., 2015). Although it is possible that other fork remodeling processes could lead to this phenotype, two other factors needed for fork reversal, RAD51 and PARP, are also required to slow the fork under conditions of genotoxic stress (Zellweger et al., 2015).

More importantly, the fork progression phenotype we observe is dependent on the HIRAN domain, which is required for fork reversal *in vitro*. We therefore hypothesize that the increased fork speed observed upon loss of HLTF is a result of loss of fork reversal.

Our findings raise interesting questions about the role of the HIRAN domain and HLTF in the cell and how HLTF may function together with other DDT and fork reversal proteins. For example, the HIRAN domain could help recruit HLTF to the replication fork in cells, which consequently might impact the functions of other proteins. Indeed, although HLTF's role in promoting PCNA polyubiquitination is shared with other ligases (Krijger et al., 2011; Unk et al., 2010), a decrease in this modification brought about by HLTF loss could reduce the recruitment of ZRANB3, which associates with polyubiquitinated PCNA (Ciccia et al., 2012). It will also be important to determine how SMARCAL1 and ZRANB3 affect fork progression under these conditions, and to investigate whether HLTF acts in the same pathway as these remodelers and as other proteins that restrain fork speed, such as RAD51 and PARP. Finally, the physiological effects of unrestrained replication fork progression are intriguing. HLTF's fork reversal activity may protect the replication fork and prevent the accumulation of mutations and genome instability. For instance, under conditions of nucleotide depletion, HLTF-mediated fork reversal could protect the fork by limiting ssDNA accumulation. Notably, *HLTF* is silenced in more than 40% of colon cancers, and its disruption promotes genome instability and intestinal carcinogenesis on the Apc$^{min/+}$ mutant background in mice (Sandhu et al., 2012; Unk et al., 2010). The Cancer Genome Atlas (TCGA) also indicates that *HLTF* amplification is observed in many cancers. We speculate that loss of HLTF could drive tumor progression by preventing a proper response to replication stress. Conversely, increased expression of HLTF may be

advantageous in cancers that need to tolerate elevated levels of replication stress. HLTF may therefore be an important vulnerability point for tumorigenesis.

In summary, characterization of HLTF's HIRAN domain and elucidation of its DNA-bound structure reveals surprising clues about HLTF function and the mechanism of fork reversal and could help pharmacological efforts to target HLTF activity in cancers. Moreover, the striking conservation of the ancient HIRAN domain throughout evolution and the role we have ascribed to it paves the way for functional studies on other uncharacterized HIRAN-domain containing proteins.

Materials and Methods

Protein purification.

The *S. cerevisiae* strain BJ2168 carrying pY6xHIS-GST plasmids encoding wild-type or mutant human *HLTF* cDNA were galactose-induced for 8-16 hr. at 30°C and these cells were frozen with liquid nitrogen. Yeast cells were lysed using a Retsch mixer mill in the presence of liquid nitrogen. Lysed cells were rehydrated in Buffer A500 (40 mM Tris, pH 7.5, 500 mM NaCl, 0.01% NP40, 0.1% Triton X-100, 10% glycerol, 1 mM TCEP-HCl), which also included 2 mM imidazole, 0.5 mM PMSF, and 1X Roche ETDA-free protease inhibitor cocktail. Lysates were cleared by centrifugation and then subjected to incubation with Ni-NTA resin (Life Technologies) overnight at 4°C. HLTF-bound Ni-NTA was washed with Buffer A500/25 mM imidazole and eluted with Buffer A500/500 mM imidazole. Elutions were diluted 10-fold in Buffer A500 and EDTA was added to 5 mM final concentration. The elutions were then incubated overnight at 4°C with glutathione

agarose (Pierce). The glutathione agarose-bound proteins were washed with Buffer A500/1 mM EDTA, Buffer A1000 (1 M NaCl)/1 mM EDTA, and Buffer A250 (250 mM NaCl)/1 mM EDTA. The beads were incubated overnight at 4°C with Turbo3C protease (A.G. Scientific), and cleaved proteins were concentrated using Vivaspin (GE Healthcare) 30K MWCO concentrators, and stored at -80°C.

The cDNA encoding the *H. sapiens* HLTF HIRAN domain (amino acids 55-180) was PCR amplified from a HLTF[1-180]/pGEX4T3 plasmid and cloned into pBG101 (Vanderbilt Center for Structural Biology) to produce a fusion protein with a cleavable N-terminal His_6-GST tag. The HIRAN/pBG101 plasmid was transformed into *E. coli* HMS174 cells and grown in Luria Broth (LB) at 37°C to an OD_{600} of 0.6. HIRAN expression was induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3 h at 37°C. Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris-HCl pH 7.8, 500 mM NaCl, and 10% glycerol), and lysed using an Avestin Emulsifier C3 homogenizer. The soluble fraction of the lysate was subjected to nickel affinity chromatography. His_6-GST-HIRAN protein was eluted in lysis buffer containing 500 mM imidazole, followed by cleavage of the His_6-GST tag by an overnight incubation with PreScission Protease at 4°C. The protein was further purified using Heparin Sepharose chromatography in 50 mM Tris-HCl pH 7.8, 5% glycerol, 2 mM DTT, 0.1 mM EDTA and a 0.05-1 M NaCl gradient, followed by size exclusion chromatography (Superdex S-200, GE Healthcare) in 20 mM Tris-HCl pH 7.8, 5% glycerol, 150 mM NaCl, 2 mM DTT and 0.1 mM EDTA. Mutant HIRAN proteins were generated by site-directed mutagenesis using the Quik-Change Kit (Stratagene). SeMet-labeled HIRAN was expressed in minimal media for 9 hr. at 37°C in *E. coli* HMS174 cells by metabolic inhibition of methionine

biosynthesis (Van Duyne et al., 1993). HIRAN mutants and SeMet-labeled HIRAN were purified in the same manner as wild-type HIRAN. Integrity of the purified HIRAN proteins were verified by circular dichroism (CD).

SMARCAL1 protein was a gift from D. Cortez, Vanderbilt University.

HIRAN crystallization and structure determination.

HIRAN was exchanged into 20 mM Tris pH 8.3, 100 mM NaCl, 2 mM DTT, and 0.1 mM EDTA buffer and mixed with dT_{10} oligonucleotide in a 1:1.2 (protein: DNA) ratio. Crystals were grown at 16°C by sitting-drop vapor diffusion by mixing 200 μM protein-DNA complex with an equal volume of mother liquor containing 25% PEG 4000, 250 mM sodium acetate trihydrate and 100 mM Tris-HCl pH 8.5 (native) or 24% PEG 8000, 100 mM magnesium acetate tetrahydrate, 200 mM ammonium acetate, 50 mM sodium cacodylate trihydrate pH 6.5 (SeMet). Crystals were flash frozen in the mother liquor supplemented with 28% PEG 4000, and 10% glycerol (native) or 15% glycerol (SeMet). X-ray diffraction data (Table S1) were collected at beamline 21-ID at the Advanced Photon Source and processed using HKL2000 (Otwinowski and Minor, 1997). Phase determination and density modification were carried out using autoSHARP (Vonrhein et al., 2007). Phases were determined from the positions of 16 selenium atoms in the asymmetric unit (asu) by single isomorphous replacement with anomalous scattering (SIRAS). After density modification, four protein-DNA complexes were readily identified in the resulting 1.5-Å electron density map. Successive rounds of model building and refinement were carried out using Coot (Emsley and Cowtan, 2004) and Phenix (Adams

et al., 2010). The protein-DNA model was subjected to simulated annealing, followed by multiple rounds of restrained coordinate and anisotropic B-factor refinement using a maximum likelihood target and combining experimental and refined phases. The refined model was validated using MolProbity as implemented in Phenix and contained all residues in allowed regions of the Ramachandran plot. Structure factors and coordinates were deposited in the Protein Data Bank under accession code PDB: 4S0N. Structural homology to other proteins was analyzed using PDBeFold (Krissinel and Henrick, 2004) and Dali (Holm and Sander, 1993).

Electrophoretic mobility shift assays (EMSAs).

EMSAs were performed with purified HIRAN protein and oligonucleotides end-labeled with 6-carboxyfluorescein (FAM). Single- versus double-stranded DNA was measured using *dna40s* and *dna40d* deoxyoligonucleotides (Table 2.2), and 3′-end binding was measured using single-stranded dT_{10} oligonucleotides containing either no modification, FAM, PO_4, or dideoxycytidine conjugated to the 3′-ends. Binding reactions were carried out with 25 nM FAM-DNA for 20 min at 25°C in 20 mM HEPES pH 7.6, 100 mM KCl, 0.5 mM DTT, 3% glycerol, 2 mM $MgCl_2$, 0.01% NP-40, and 250 µg/mL BSA. Reactions were spiked with native loading dye containing 70% glycerol and 0.1% bromophenol blue and loaded onto a 5% (29:1) acrylamide gel. Gels were run in 0.5X TBE for 30 min at 200 V and imaged with a Typhoon Trio variable mode imager at 532 nm excitation and 526 nm emission wavelengths. Data were quantified with GelAnalyzer and fit to a two-state binding model in GraphPad Prism.

Circular dichroism.

CD spectra were collected on a Jasco J-180 spectropolarimeter in a 1.0 mm path quartz cuvette containing 70 µM protein and 20 mM HEPES pH 7.6, 100 mM KCl, 0.5 mM DTT, 3% glycerol, 2 mM MgCl2. Thermal denaturation was monitored as an increase in molar ellipticity (θ) at 222 nM as temperature was increased from 20-75 ^{0} C in 1 °C steps. Data were normalized and plotted with GraphPad Prism. Melting temperatures (Tm) were derived by fitting the data using the equation $\theta = (1+e(Tm-T)/k)-1$, where k describes the cooperativity of the transition.

Fork regression.

Oligonucleotides (Table 2.2) AK203 (5′-FAM-labeled) with AK125, and AK183 with AK124 were separately annealed at a 1:1 ratio to 250 nM in 1X SSC (15 mM NaCitrate pH 7.0 and 150 mM NaCl) by decreasing the temperature (1.2°C /min) from 95°C to 4°C. For 3′-PO$_4$-capped fork structures, oligonucleotide AK206 replaced AK124. To create the fork structures, the products of the above reactions were annealed in a 1:1 ratio at 37°C for 30 min in 1X reaction buffer (40 mM Tris pH 7.65, 50 mM NaCl, 5 mM MgCl$_2$, 1 mM TCEP-HCl). 10-µl reactions were assembled on ice with 5 mM ATP, 5 nM fork substrate, and the indicated amount of protein. Following incubation at 37°C for the indicated amount of time, reactions were stopped by the addition of 10X stop buffer (4% SDS, 50 mM EDTA, 0.1% bromophenol blue) and incubated for 5 min. Samples were run on 6% 0.5X TBE gels and visualized using a Typhoon 9410 Imager (GE healthcare) and then

quantified using ImageJ. All experiments were performed at least three times, with the mean plotted ± SEM.

ATPase assays.

Splayed-arm DNA substrates were annealed as above using oligonucleotides AK122 and AK123. Ten microliter reactions were assembled in ATPase reaction buffer (20 mM Tris-HCl pH 7.65, 10mM KCl, 1mM MgCl$_2$, 10% glycerol, 60 µg/mL BSA) with 50 nM protein, 1 mM ATP trace-labeled with ^{32}P-γATP and the indicated amount of DNA. After incubation at 37°C for 30 min, 1 µL of each reaction was spotted to TLC PEI Cellulose F chromatography plates (Millipore), and the products were migrated in solvent (1 M formic acid, 0.25 M LiCl) for 30 min. The plates were imaged by radiography on a Typhoon 9410 (GE Healthcare). Quantification was performed with ImageJ. All experiments were performed at least three times, with the mean plotted ± SEM.

Isolation of proteins on nascent DNA (iPond).

Briefly, for each sample 1x10^8 293T cells were pulsed with 10 uM EdU for 10 min, followed by a thymidine chase for the indicated amount of time. Cells were fixed with 1% formaldehyde, collected, and stored at -80°C. Cells were permeabilized in 1X PBS with 0.25% Triton X-100 on ice. Biotin-azide (10 µM) was conjugated to EdU-labeled DNA-protein complexes by click chemistry for 1 hr. at room temperature. Cells were then lysed in RIPA buffer (100 mM Tris, pH7.5, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate) by sonication, and cleared by centrifugation. The lysates were incubated

with high-performance streptavidin sepharose (GE healthcare), washed in RIPA buffer, and boiled in 2X sample buffer.

Cell culture, RNA interference and plasmids.

U2OS and 293T cells were maintained in DMEM (Life Technologies) supplemented with 10% FBS, 2 mM L-glutamine, and penicillin/streptomycin in 5% CO_2 at 37°C. For add-back experiments, plasmids were transfected using Fugene6 (Promega) according to manufacturer's directions and experiments were performed after 48 hr. The non-targeting control siRNA luciferase (GL3; D-001400-01-20) and the previously-described synthetic siRNAs targeting HLTF (Lin et al., 2011) (HL4; D-006448-04-0005, HL5; D-006448-05-0005) were purchased from ThermoFisher. Transfections were performed using 25 nM siRNA with Dharmafect 1 (ThermoFisher) according to the manufacturer's directions. The yeast expression vector pY6xHIS-GST was generated by replacing the SacI-Hind III fragment of pEG(KT) (Pierce and Wendland, 2009) with a cassette containing a tandem 6xHistidine and Glutathione-S transferase tag followed by an HRV3C protease site and a multiple cloning site. *HLTF* cDNA was amplified by PCR and subcloned into the SalI-NotI restriction sites of the pY6xHIS-GST vector or with a FLAG-epitope into pcDNA3.1(+). Site-directed mutagenesis of *HLTF* was carried out using the QuikChange kit (Stratagene). The ΔHIRAN deletion mutant was generated by PCR-mediated in-frame deletion of DNA sequence corresponding to amino acids 55-180 of HLTF.

CRISPR/Cas9-mediated knockout of HLTF in U2OS cells.

The single-guide RNA (sgRNA) target site (GTTGGACTACGCTATTACACGGG) toward *HLTF* exon 2, downstream of the start ATG, was designed using the CRISPR design tool (http://crispr.mit.edu) and oligonucleotides encoding the exon 2 sgRNA were cloned into the pX330 plasmid (Addgene) (Ran et al., 2013). U2OS cells were transfected with the pX330-Exon2 sgRNA plasmid, and 48 hr. post-transfection the cells were diluted to allow for colony formation from single cells. Well-isolated single colonies were picked, expanded, and screened by western blot analysis for the loss of HLTF. The *HLTF* locus was then sequenced to verify the disruption of HLTF generated by the Cas9 nuclease, and RT-qPCR was used to determine *HLTF* mRNA transcript levels.

DNA fiber experiments.

For experiments using transfected DNA, U2OS cell lines were seeded in 6-well dishes at 1.5-2 x 10^5 cells/well. The next day, cells were transfected and allowed to grow another 48 hr. For experiments with untransfected DNA, U2OS cell lines were seeded at 5 x 10^5 cells/well 24 hr. before the experiment. On the day of the experiment, cells were pulsed with 25 µM IdU in fresh DMEM for 30 min, washed twice with DMEM, and DMEM that included 200 µM CldU was added with or without 50 µM hydroxyurea. After an additional 30 min, the cells were collected by trypsinization and used for DNA spreading. Briefly, cells were adjusted to 1 x 10^6 cells/mL in phosphate buffered saline and 4 uL were spotted to microscope slides. 10 µL lysis buffer (200 mM Tris-HCl pH 7.5, 50 mM EDTA, 0.5% SDS) was gently swirled with the cells and incubated for 2-3 min at room

temperature. The slides were tilted at a 15-20° angle to allow DNA fiber spreading, and then air dried and fixed in a solution of 3:1 methanol:acetic acid. DNA fibers were stained with antibodies to IdU and CldU and collected on a Zeiss AxioSkop 2. IdU and CldU track lengths were measured in ImageJ. Means were calculated and statistical analysis was carried out by nonparametric Mann-Whitney test or ANOVA as indicated in the figure legends.

Table 2.2 Oligodeoxynucleotides used in this Chapter II.

Name	Sequence
<u>EMSA</u>	
dna40s	FAM-d(CTCAGGACTCAGTTCGTCAGCCCTTGACAGCGATGGAAGC)
dna40d	d(GCTTCCATCGCTGTCAAGGGCTGACGAACTGAGTCCTGAG)
dT10-OH	FAM-d(TTTTTTTTTT)
dT10-FAM	d(TTTTTTTTTT)-FAM
dT10-PO$_4$	FAM-d(TTTTTTTTTT)-PO$_4$
dT10-H	FAM-d(TTTTTTTTTT)-ddC
<u>Fork Regression</u>	
AK124	d(AGCTACCATGCCTGCCTCAAGAATTCGTAA)
AK125	d(TTACGAATTCTTGAGGCAGGCATGGTAGCT)
AK183	d(GTTTTCCCAGTCACGACGATGCTCCGGTACTCCAGTGTAGGCATGTTACGAATTCTTGAGGCAGGCATGGTAGCT)
AK203	FAM-d(AGCTACCATGCCTGCCTCAAGAATTCGTAATATGCCTACACTG-GAGTACCGGAGCATCGTCGTGACTGGGAAAAC)
AK206	d(AGCTACCATGCCTGCCTCAAGAATTCGTAA)-PO$_4$

CHAPTER III

The HIRAN Domain of HLTF Positions the DNA Translocase Motor to Drive Efficient DNA Fork Regression

Introduction

The accurate and complete replication of DNA is crucial for maintaining genomic stability and for cell survival. Multiple forms of replication stress, including DNA damage, difficult to replicate sequences, DNA secondary structures, and protein-DNA and RNA-DNA complexes, inhibit progression of the replication fork (Zeman & Cimprich, 2014). Polymerase stalling can cause uncoupling of DNA synthesis and unwinding activities (Byun *et al*, 2005), leading to an accumulation of ssDNA that is vulnerable to nuclease cleavage or formation of aberrant DNA structures (Berti & Vindigni, 2016) . Stalled replication and a failure of the cell to respond to replication stress leads to genomic instability, chromosomal rearrangements, mutations, cell death, and a number of human diseases (Zeman & Cimprich, 2014). As a way to avoid the genomic instability associated with fork stalling and arrest, cells possess DDT pathways that maintain fork progression, facilitate replication restart, and promote the completion of DNA replication (Friedberg, 2005; Chang & Cimprich, 2009; Branzei & Psakhye, 2016). TLS involves specialized DNA polymerases that bypass DNA lesions in an error-prone manner, while the template switching pathway relies on an undamaged template for error-free synthesis past a replication blockage.

---

*The work in this chapter was published in Chavez DA, Greer BH, Eichman BF (2018). The HIRAN domain of helicase-like transcription factor positions the DNA translocase motor to drive efficient DNA fork regression. *J Biol Chem*, 293: 8484-94

One template switching mechanism by which stalled forks may be rescued or avoided is fork regression (also known as fork reversal), in which the two parental strands are re-annealed and the two nascent daughter strands are unpaired from their templates and repaired together to form a four-way junction (Figure 1A). Fork reversal has been observed as a response to Top1 poisoning in yeast, mouse, and human cells, to replication stress caused by a prolonged S-phase in mouse embryonic stem cells, and as a response to other forms of replication stress (Ray Chaudhuri *et al*, 2012; Ahuja *et al*, 2016; Zellweger *et al*, 2015; Follonier *et al*, 2013; Neelsen *et al*, 2013; Neelsen & Lopes, 2015). Reversed forks are important replication intermediates that maintain genomic integrity in several ways (Neelsen & Lopes, 2015). In addition to initiating template switching to allow for error-free DNA synthesis using the undamaged sister chromatid, fork regression may limit the accumulation of vulnerable DNA structures generated by polymerase stalling. Additionally, fork reversal may facilitate DNA repair through recombination pathways or by sequestering the lesion back into the context of duplex DNA.

HLTF serves to promote DDT in mammalian cells (Motegi *et al*, 2008; Unk *et al*, 2008; Lin *et al*, 2011). Inactivation of HLTF leads to increased UV and MMS sensitivity in cells and alters the progression of replicating forks under replication stress (Lin *et al*, 2011; Blastyák *et al*, 2010). Like its yeast homologue Rad5, HLTF functions as a RING E3 ubiquitin ligase that works together with Rad18 and Mms2-Ubc13 complexes to polyubiquitinate proliferating cell nuclear antigen (PCNA), activating the template switching pathway (Unk *et al*, 2008; Burkovics *et al*, 2014). HLTF contains a DNA-dependent ATPase motor domain related to the SWI/SNF2-family chromatin remodelers

that enables translocation on dsDNA and regression of model replication forks *in vitro* (Blastyák *et al*, 2010). HLTF's dsDNA translocation activity can displace proteins that might be found at a stalled replication fork from DNA (Blastyák *et al*, 2010; Achar *et al*, 2011) . HLTF has also been reported to form D-loops without the assistance of Rad51 *in vitro* (Ahcar *et al*, 2011). Both the ubiquitin ligase and ATPase activities of HLTF and Rad5 are required for DDT (Blastyák *et al*, 2010; Choi *et al*, 2015; Minca & Kowalski, 2010; Ortiz-Bazán *et al*, 2014).

Two other SNF2-related DNA translocases—SMARCAL1, and ZRANB3—have been identified to catalyze fork regression *in vitro* and to be important for DDT and maintenance of genomic stability (Neelsen & Lopes, 2015; Yusufzai & Kadonaga, 2008; Bansbach *et al*, 2009; Ciccia *et al*, 2009; Yuan *et al*, 2009; Yusufzai *et al*, 2009; Postow *et al*, 2009; Yusufzai & Kadonaga, 2010; Bétous *et al*, 2012; Ciccia *et al*, 2012; Yuan *et al*, 2012; Poole & Cortez, 2017). Inactivation of these proteins leads to a sensitivity to genotoxic agents, as well as defective fork restart, increased double-strand breaks, and sister chromatid exchange events in response to replication stress (Bansbach *et al*, 2009; Ciccia *et al*, 2009; Yuan *et al*, 2009; Yusufzai & Kadonaga, 2010; Ciccia *et al*, 2012; Yuan *et al*, 2012). Outside of their ATPase motor domains, HLTF, SMARCAL1, and ZRANB3 are distinguished by substrate recognition domains (SRDs) that impart specificity for a particular DNA structure found at a stalled fork (Poole & Cortez, 2017; Kile *et al*, 2015; Mason *et al*, 2014; Badu-Nkansah *et al*, 2016). Despite the importance of these enzymes to DDT, the mechanisms by which the motor and SRD domains catalyze fork remodeling and how this activity is tied to their cellular roles is not well understood.

HLTF contains a HIRAN (HIP116, Rad5p, N-terminal) domain that serves as the SRD through its ability to bind specifically to the 3′-end of ssDNA (Kile *et al,* 2015; Hishiki

et al, 2015; Achar *et al*, 2015). Crystal structures of HIRAN bound to DNA showed that HIRAN adopts a modified oligonucleotide/oligosaccharide (OB)-fold that clamps the two 3′-terminal nucleobases and forms a binding pocket for the 3′-hydroxyl group. Mutation of the interacting residues impaired DNA binding by HIRAN in solution (Kile *et al*, 2015; Hishiki *et al*, 2015; Achar *et al*, 2015). Importantly, site-directed incorporation of these mutations into full-length HLTF impaired *in vitro* fork regression but not ATPase, ubiquitin ligase, or translocation activities, and also prevented complementation of replication fork progression and UV sensitivity defects in HLTF-depleted cells (Kile *et al*, 2015; Achar *et al*, 2015). Moreover, deletion of HIRAN from HLTF recapitulates these defects. Depletion of the HLTF ortholog from human cells, *Arabidopsis* (Rad5A), and *S. pombe* (Rad8) caused sensitivity to DNA damaging agents and defects in DNA replication and cell growth, none of which were complemented by ΔHIRAN mutants (Achar *et al*, 2015; Ding & Forsburg, 2014; Kobbe *et al*, 2016). Thus, HIRAN's DNA 3'-end binding activity is essential to HLTF function.

Previous work established that the HLTF HIRAN domain serves as the SRD through its ability to bind specifically to the 3′-end of ssDNA (Kile *et al*, 2015; Hishiki *et al*, 2015; Achar *et al*, 2015). Crystal structures of HIRAN bound to DNA showed that HIRAN adopts a modified oligonucleotide/oligosaccharide (OB)-fold that clamps the two 3′-terminal nucleobases and forms a binding pocket for the 3′-hydroxyl group. Mutation of HIRAN's DNA binding residues impaired DNA binding by HIRAN in solution (Kile *et al*, 2015; Hishiki *et al*, 2015; Achar *et al*, 2015), and incorporation of these mutations into full-length HLTF impaired *in vitro* fork regression but not ATPase, ubiquitin ligase, or translocation activities. These mutations in full-length HLTF also prevented complementation of replication fork progression and UV sensitivity defects in HLTF-

58

depleted cells (Kile *et al*, 2015; Achar *et al*, 2015). Moreover, deletion of HIRAN from HLTF recapitulates these defects. Depletion of HLTF and HLTF orthologs *Arabidopsis* (Rad5A), and *S. pombe* (Rad8) caused sensitivity to DNA damaging agents and defects in DNA replication and cell growth, none of which were complemented by ΔHIRAN mutants (Achar *et al*, 2015; Ding & Forsburg, 2014; Kobbe *et al*, 2016). Thus, HIRAN's DNA 3'-end binding activity is essential to HLTF function.

Despite its importance to DDT, the mechanism by which HLTF remodels stalled forks, and the roles of HIRAN and motor domains in this process, is unknown. From our previous mutational analysis of HIRAN function in fork reversal and replication fork progression, we proposed a model in which recognition and remodeling of stalled forks by HLTF is accomplished through the combined activities of the dsDNA translocase motor and HIRAN 3′-end-binding domains (Kile *et al*, 2015). In this model, HLTF facilitates reannealing of parental strands by binding and translocating along the parental duplex DNA toward the stalled fork, eventually destabilizing the nascent duplexes, after which capture of the destabilized 3′-end of the nascent leading strand by the HIRAN domain would facilitate efficient unwinding of the leading arm and/or annealing of the nascent strands (Figure 3.1). This model has not been tested and it is not yet clear how HLTF engages a fork to catalyze fork remodeling. Specifically, we do not know how HLTF is positioned at a stalled replication fork, how the protein deals with ssDNA gaps that would be present on leading and lagging strands of stalled and normal forks or know if HIRAN can denature the 3'-end of the nascent leading strands. We tested these aspects of our model biochemically using HIRAN deletion mutants previously shown to impair HLTF function in cells (Achar *et al*, 2015; Ding & Forsburg, 2014; Kobbe *et al*, 2016) to better

understand how the HIRAN and ATPase domains engage fork structures to facilitate efficient fork regression.
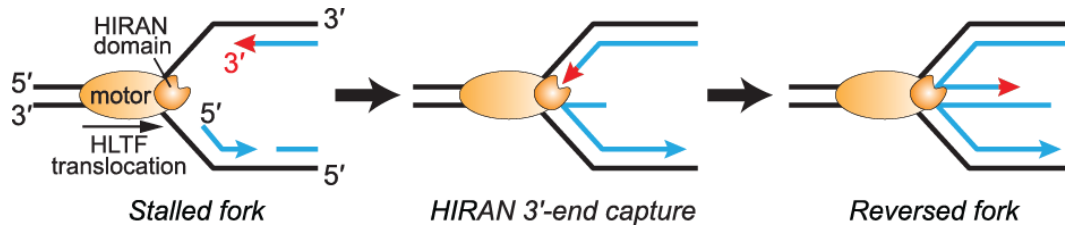
Figure 3.1 Model of fork regression reaction by HLTF.

Results

HIRAN specifically interacts with the 3'-end of the nascent leading strand

We previously showed that efficient fork regression by HLTF is dependent on a functional HIRAN domain and on a free 3′-hydroxyl group on the nascent leading strand of a model replication fork, which suggested a direct interaction between HIRAN and the 3′-end of the nascent leading strand (Figure 3.1) (Kile *et al*, 2015). Specifically, HLTF fork regression activity is reduced either by amino acid substitutions within the DNA binding cleft of HIRAN or by phosphorylation of the 3′-end of the nascent leading strand, both of which inhibit ssDNA binding by the HIRAN domain. Moreover, this reduction in fork regression when blocking the leading 3′-end is HLTF specific, as a 3′-phosphorylated substrate did not impair fork reversal by SMARCAL1 or T4 phage UvsW, which share a similar SRD unrelated to HIRAN (Kile *et al*, 2015). To verify that a direct interaction between HIRAN and the 3′-end of the nascent leading strand is responsible for the reduced activity of HLTF against a 3′-blocked substrate, we constructed a HIRAN-

UvsW^{motor} chimera protein by replacing UvsW's N-terminal SRD with HLTF's HIRAN domain and tested fork regression activity against substrates containing 3′-OH (unblocked) and 3′-PO$_4$ (blocked) nascent leading strands (Figure 3.2). In this assay, we monitor regression as a time-dependent formation of annealed template strands from a homologous fork containing a mismatch at the junction to prevent spontaneous regression. We previously showed that removal of UvsW's SRD from the motor domain abrogates fork regression activity, and that this activity can be partially restored by fusion of the structurally homologous SRD from SMARCAL1 onto the UvsW motor (Mason *et al*, 2014). Similarly, the HIRAN-UvsW^{motor} chimera showed only a modest reduction in activity against an unblocked substrate (Figure 3.2A and 3.2B), indicating that HIRAN acts independently of the motor domain to facilitate fork regression. However, whereas 3′-phosphorylation of the nascent leading strand had no effect on wild-type UvsW, blocking the 3′-end abrogated fork regression activity of the HIRAN-UvsW^{motor} chimera (Figure 3.2A and 3.2B). We therefore conclude that HIRAN serves as the SRD of HLTF through a specific interaction with the 3′-end of the nascent leading strand.

HIRAN captures unpaired 3'-ends from duplex DNA

Crystal structures of HIRAN showed that the protein clamps two stacked nucleobases from the 3′-terminal nucleotides in a conformation that precludes base pairing to a complementary strand (Achar *et al*, 2015; Hishiki *et al*, 2015; Kile *et al*, 2015). This specificity for ssDNA and HIRAN's association with the leading end of the nascent strand prompted us to explore whether HIRAN has the ability to denature the 3′-end on its own, or whether it facilitates fork regression by capturing the 3′-end of an already

denatured duplex. We tested the ability of HIRAN to bind to the 3′-end in the context of

dsDNA using an EMSA. Purified HIRAN domain was added to DNA substrates containing

0, 1, 2, or 3 unpaired 3′-nucleotides across from a 5′-overhang (Figure 2A). The 3′-end of
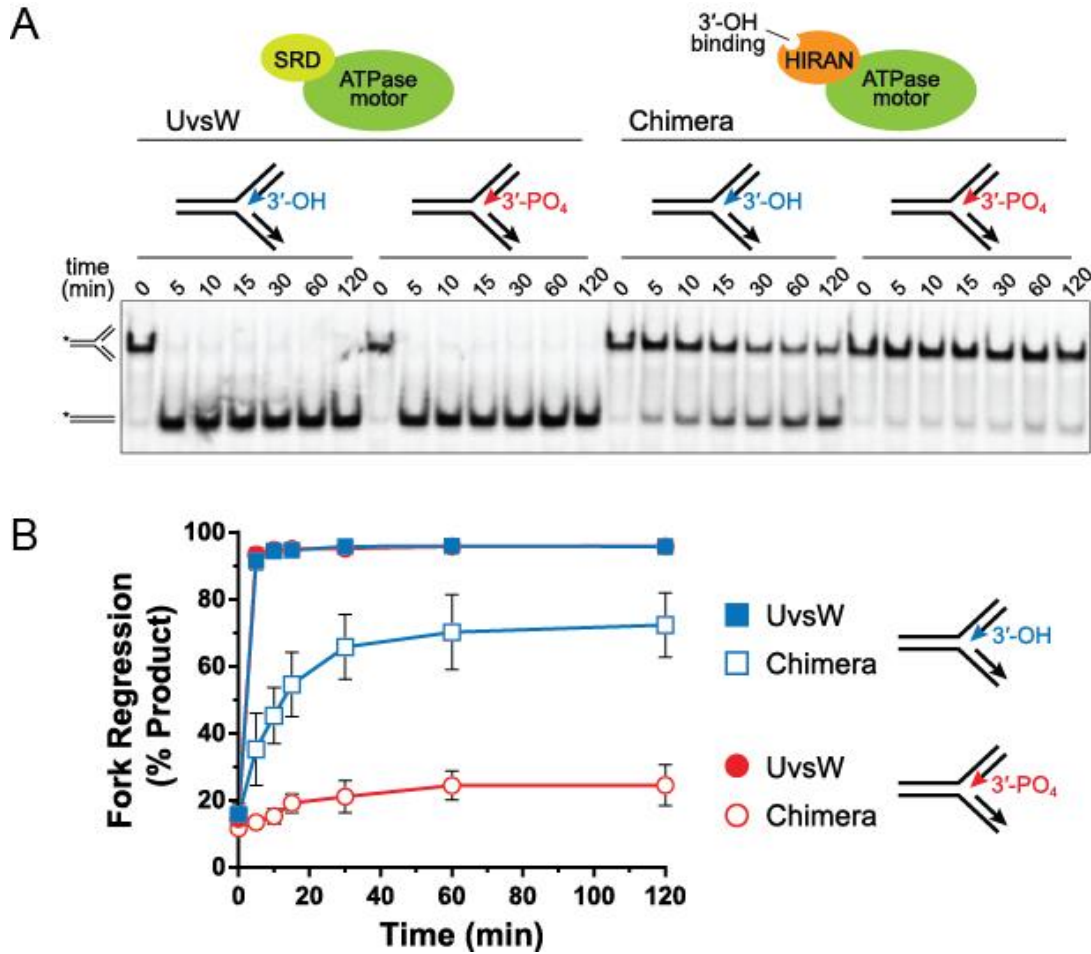
Figure 3.2 HIRAN specifically interacts with 3' end. A. Representative native gel showing time-dependent fork regression by wild-type UvsW and a HIRAN-UvsW^{motor} chimera against model fork substrates containing either a native 3′-OH or capped 3′-PO$_4$ on the nascent leading strand. ^{32}P-labels are indicated by asterisks. SRD, substrate recognition domain. **B.** Quantitation of data from three independent experiments (mean ± S.D.). Data generated by Briana Greer.

the overhang strand was capped with a fluorescein molecule to prevent any spurious

binding there (Kile *et al*, 2015). Consistent with the crystal structure, binding was only

observed to DNA that contained at least two unpaired 3′-nucleotides. No binding was

observed when the 3′-end was fully base-paired or frayed by only one nucleotide (Figure 3.3A), indicating that the HIRAN-DNA binding energy is not sufficient to denature the 3′-end away from the template strand.

We next examined the ability of HIRAN to bind to DNA ends in the context of a DNA fork structure. Since HIRAN cannot access a paired DNA end, we hypothesized that any apparent binding of HIRAN to a fork would occur only if the nascent leading 3′-end were frayed. To test this possibility, we performed an EMSA with purified HIRAN against fork structures containing a nascent leading strand that was either fully base paired or that contained two unpaired nucleotides at the 3′-end. Consistent with our previous results, binding was observed only to the fork with a frayed leading strand, and this binding was abolished by phosphorylation of the frayed end (Figure 3.3B). These data confirm that HIRAN only recognizes an unpaired 3′-end, even in the context of a fork. Moreover, we conclude that HIRAN has no affinity for the structure of the fork itself, inconsistent with a previous report (Achar *et al*, 2015). The binding to fork structures used in that study may have resulted from partially denatured or frayed fork substrates.

HIRAN enforces a specificity to HLTF at forks containing a frayed nascent leading strand

HLTF was previously shown to bind to various DNA junction structures with similar affinity, suggesting that the full-length protein lacks specificity for a particular fork (Blastyák *et al*, 2010). However, we hypothesized that the interaction between HIRAN and the 3′-end of a frayed nascent leading strand would impose a specificity to HLTF.
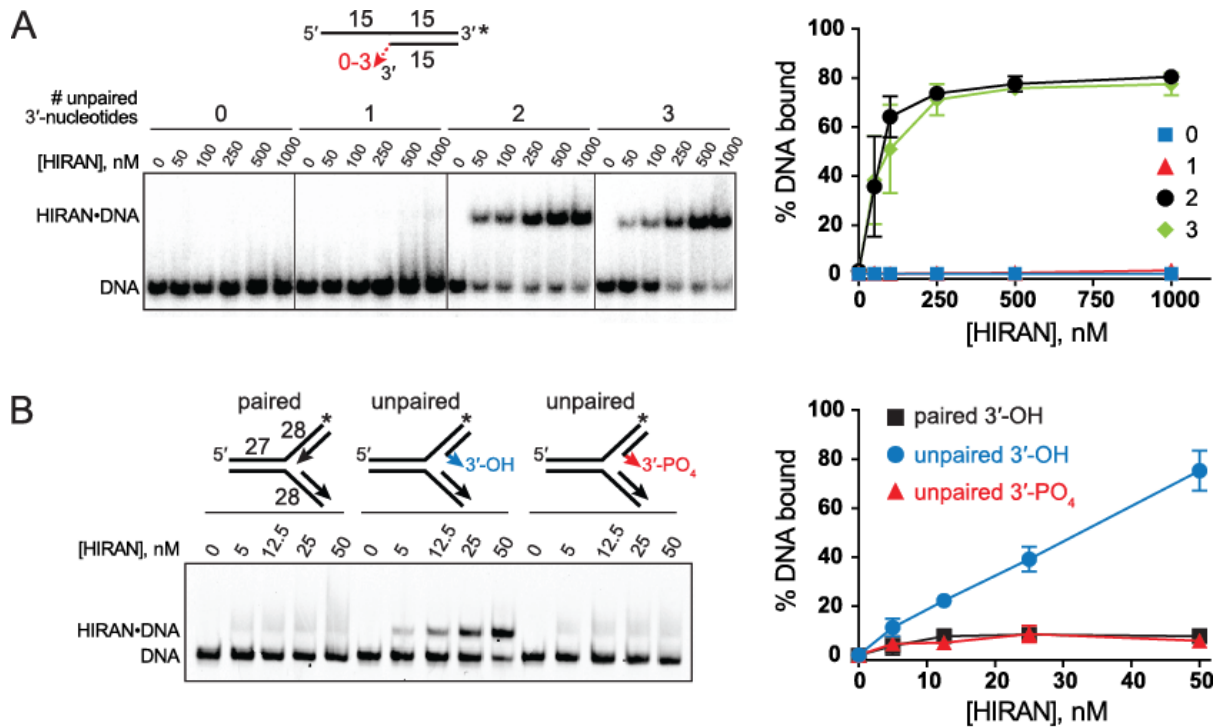
Figure 3.3 HIRAN binds unpaired DNA ends in the context of duplex DNA. EMSAs of purified HIRAN domain binding to DNA containing 0, 1, 2, or 3 unpaired 3′-nucleotides within dsDNA (**A**) or to DNA forks that contain either fully paired or two unpaired nucleotides at the 3′-end of the nascent leading strand (**B**). Numbers with the DNA schematics refer to the number of nucleotides, and asterisks represent the location of fluorescein tags. Quantitation of three experiments (mean ± S.D.) is shown to the right of a representative gel.

We tested this by comparing binding of HLTF and a deletion mutant lacking the HIRAN domain (ΔHIRAN) to fully paired and partially frayed fork structures (Figure 3.4). Consistent with the previous study (Blastyák *et al*, 2010), our EMSAs showed a protein-dependent appearance of 3-4 discrete bands corresponding to multiple protein-DNA complexes, likely containing one or more proteins bound per DNA since the dsDNA translocase motor domain would be capable of binding any of the three arms. However, the presence of the frayed nascent leading 3′-end resulted in a significant accumulation in the lowest mobility complex (band 1) and a concomitant decrease in the highest mobility

complex (bands 3+) for the full-length protein (Figure 3.4A), suggesting that this structural feature favors a stoichiometric HLTF-DNA complex. This increase in band 1 was not observed from the ΔHIRAN construct (Figure 3.4B), indicating that the HIRAN domain imparts a specificity to HLTF for the 3′-end of the nascent leading strand in these model replication forks. Consistent with this conclusion, removing the HIRAN domain resulted in an accumulation of the higher mobility complexes (bands 3+) relative to the others (Figure 3.4B).

HLTF binds to the parental duplex ahead of the fork and tracks with the 3′ leading end

Our model of fork regression by HLTF postulates that the motor domain translocates on the unreplicated parental duplex toward the fork while the HIRAN domain engages a displaced 3′-end of the nascent leading strand (Figure 3.1). To probe how HLTF is positioned at a DNA fork, we performed DNase footprinting on immobile forks containing two unpaired nucleotides at the 3′-end of the nascent leading strand to promote the stoichiometric 1:1 complex (Figure 3.5). We tested both full-length HLTF and the HLTF-ΔHIRAN mutant against two forks that differed by the presence (Figure 3.5A and 3.5C) or absence (Figure 3.5B and 3.5D) of a 10-nucleotide ssDNA gap on the leading template to alter the position of the nascent 3′-end with respect to the junction. Full-length HLTF showed a virtually identical pattern of nuclease protection on the lagging templates across the junctions of both forks that spanned 8-17 nucleotides along the parental duplexes and 4-5 nucleotides on the unwound arms behind the junction (Figure 3.5A and 3.5B). There was no protection of the leading strand of the parental duplex ahead of either fork, but instead was moderately sensitized to nuclease cleavage upon addition of HLTF.

Deletion of the HIRAN domain delocalized the protection on the lagging template away from the junction on both gapped and non-gapped forks (Figure 3.5A and 3.5B). Moreover, the ΔHIRAN mutant modestly protected the leading strand upstream of the ungapped fork (Figure 3.5B and 3.5D), providing additional evidence that the motor has reduced specificity for this fork without the HIRAN domain.
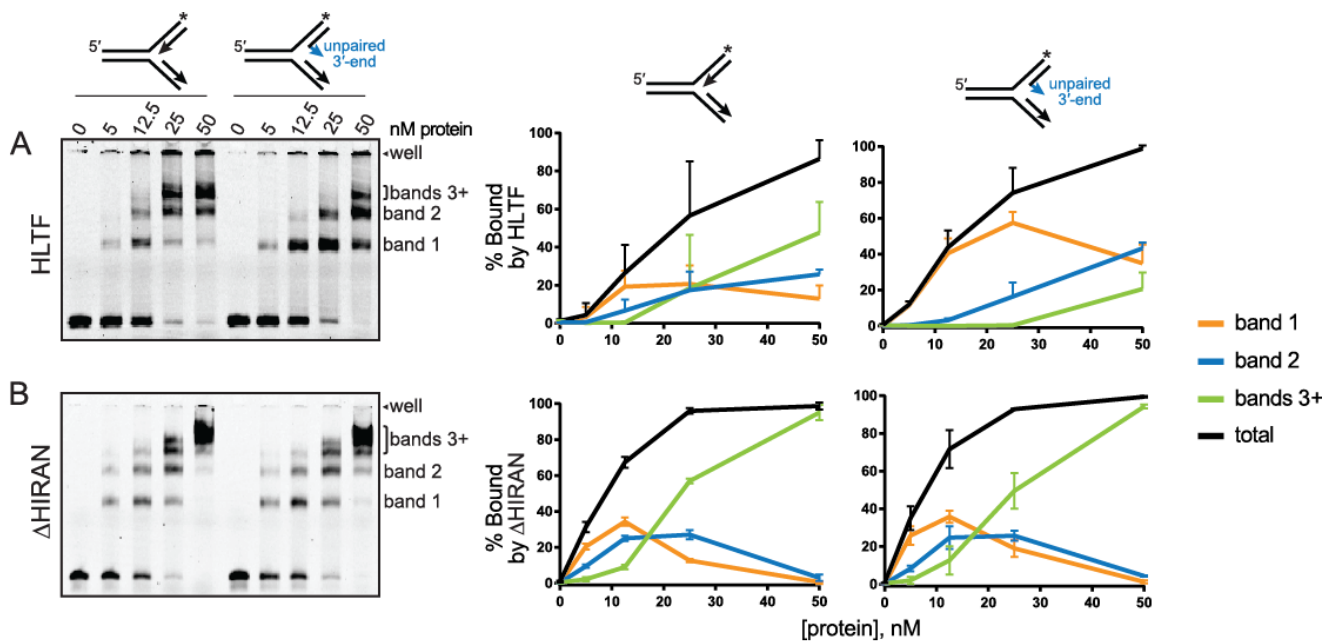
Figure 3.4 Effect of HIRAN on HLTF-fork binding. EMSAs of HLTF (**A**) or ΔHIRAN (**B**) binding to 25 nM DNA forks containing fully paired or two unpaired nucleotides at the 3′-end of the nascent leading strand. Bands representing discrete protein-DNA complexes are labeled to the right of each gel. Quantitation of the percentage of each shifted band and the total from three experiments (mean ± S.D.) is shown to the right of a representative gel. The DNA remaining in the well (topmost band of each gel) was not included in the quantification.
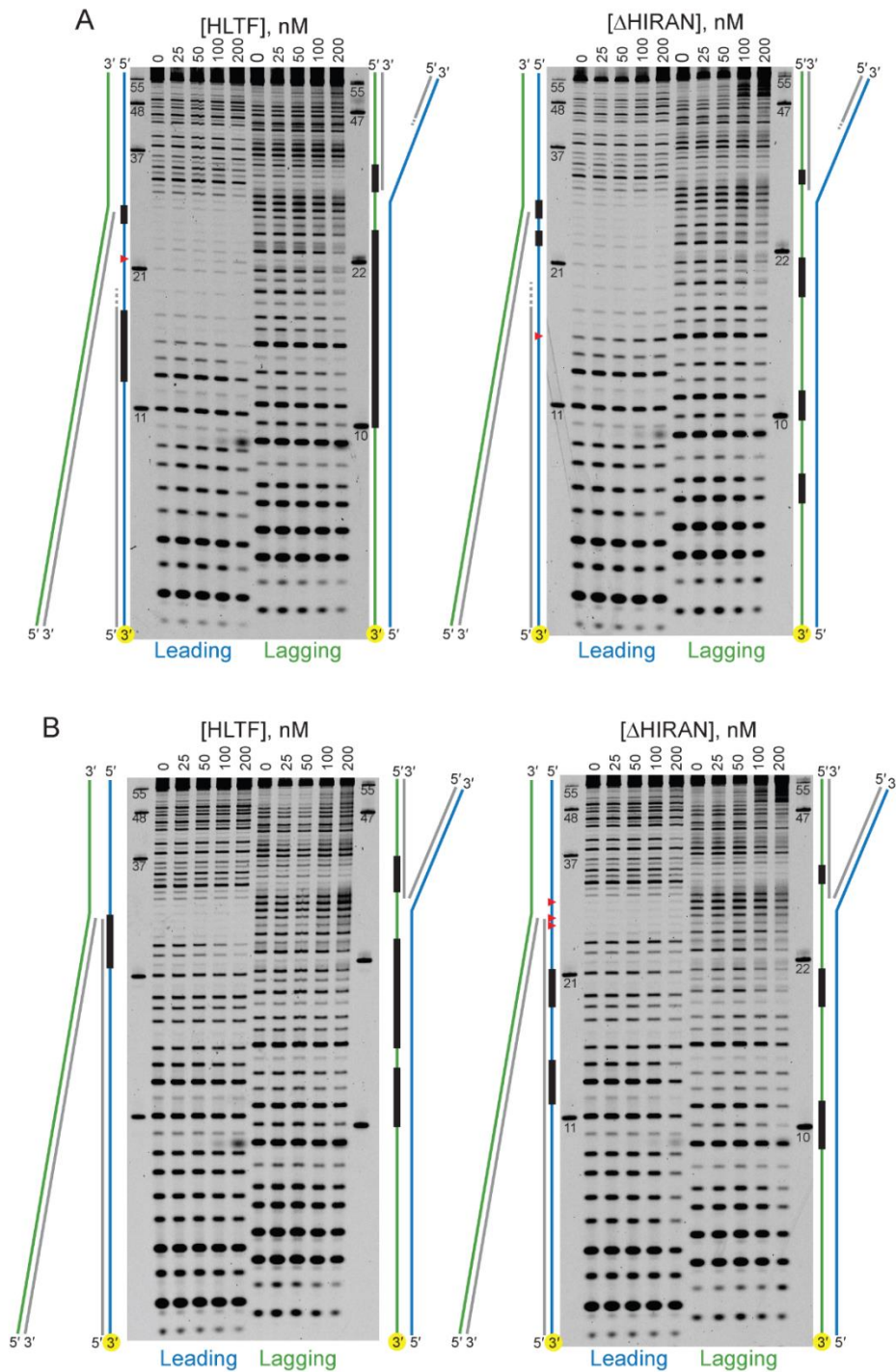
Figure 3.5 HLTF DNase footprint of HLTF bound to model forks. Representative footprinting gels of gapped (**A**) and ungapped (**B**) forks are shown. Either the leading template (left sides of gels) or the lagging template (right sides) strand was labeled with fluorescein, indicated by the yellow circle in the DNA schematics. DNA concentration was held fixed at 100 nM. Lanes represent increasing concentrations of HLTF. Only 0 and 200 nM HLTF lanes were used in quantification shown in Figure 3.6.

Protection of the leading template behind the junction tracked with the position of the 3′-end of the nascent strand in a HIRAN-dependent manner. On the ssDNA gapped fork, HLTF protection of a 4-5 nucleotide region and strong hypersensitivity of a single nucleotide was observed in the base-paired and ssDNA regions, respectively, of the leading template across from the nascent 3′-end (Figure 3.5A). Strikingly, this pattern of protection and sensitivity opposite the 3′-end was not evident in the ΔHIRAN mutant, consistent with HIRAN engagement with the frayed 3′-end of the nascent leading strand. When the nascent 3′-end abutted the junction in the non-gapped fork, HLTF protection of the leading template coincided with the position of the nascent 3′-end (Figure 3.5B and 3.5D). This pattern did not significantly change by deleting the HIRAN domain (Figure 3.5B and 3.5D). Interestingly, ΔHIRAN caused nuclease sensitivity directly at the non-gapped junction, although the significance of this is unclear considering that a small amount of sensitivity is also evident at the junction of the gapped fork when treated with HLTF and ΔHIRAN proteins (Figure 3.5A). Nonetheless, these results are consistent with HLTF positioned on the lagging strand ahead of the fork and show that engagement of the unwound templates behind the fork corresponds to the position of the stalled nascent strand on the leading arm and to the position of the junction on the lagging arm.

HIRAN facilitates initial formation of the 4-way junction

To investigate the manner in which remodeling of the leading strand by the HIRAN domain facilitates fork regression, we compared fork regression activities of HLTF and ΔHIRAN against forks containing ssDNA gaps on either the leading or lagging templates

immediately behind the junction (Figure 3.6). These substrates model replication forks that have stalled from impediments on the leading or lagging strands, and also provide a way to ascertain the importance of the HIRAN and motor domains in annealing and denaturing parental and nascent strands. The presence of nascent strands annealed right up to the junction present a barrier to initial formation of the 4-way junction from the 3-way fork (Manosas *et al*, 2013) and thus moving the ends of the nascent strands from the junction to create a ssDNA region in the template would lower the barrier to this transition.
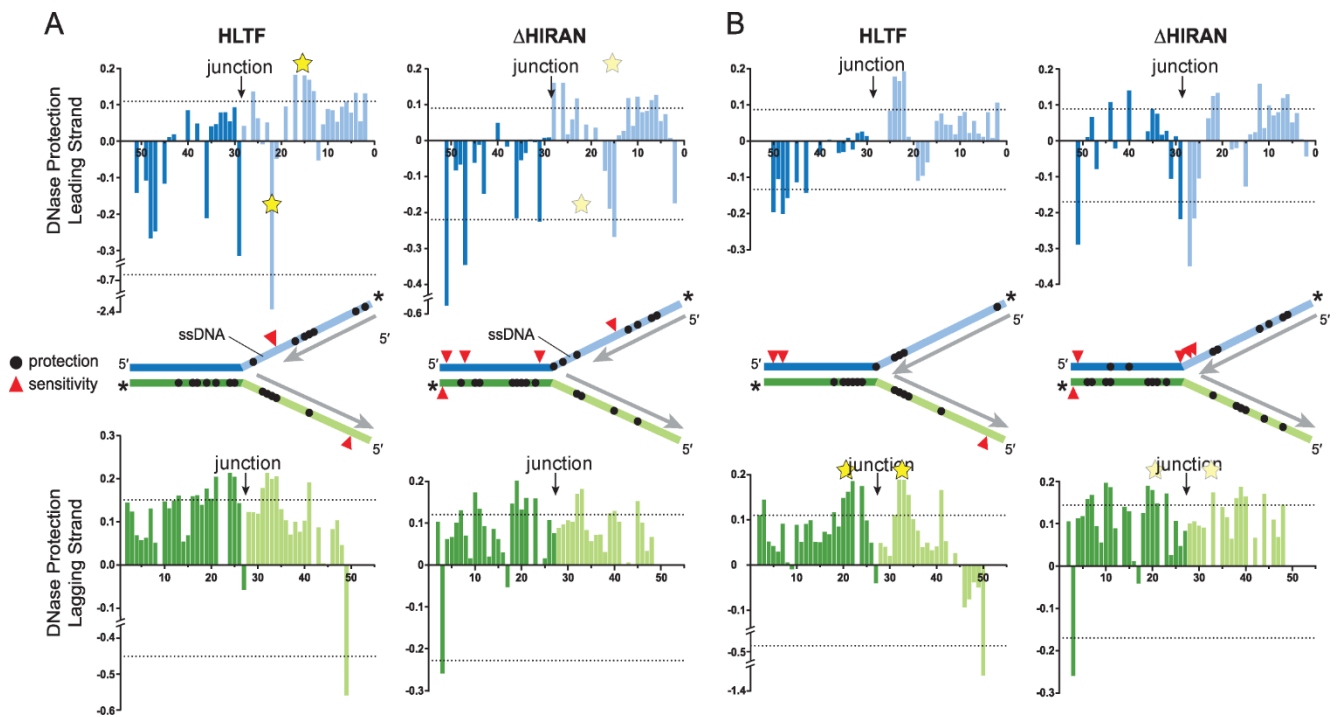
Figure 3.6 Quantitation of HLTF DNase footprinting. Quantitation of HLTF (wild-type and ΔHIRAN mutant) footprinting on gapped (**A**) and ungapped (**B**) fork substrates. Bar graphs plot the amount of nuclease protection (positive value) or sensitivity (negative value) as a function of nucleotide position relative to the fluorescein label along the leading (blue) or lagging (green) strand. Dashed lines signify 0.5 standard deviation from the mean protection/sensitivity. Yellow stars indicate significant differences in protection between HLTF and ΔHIRAN in each fork. The location of significant protection and sensitivity sites are shown as black circles and red triangles, respectively. Asterisks indicate the position of the fluorescein labels on each strand, and unlabeled nascent strands are shown in grey.
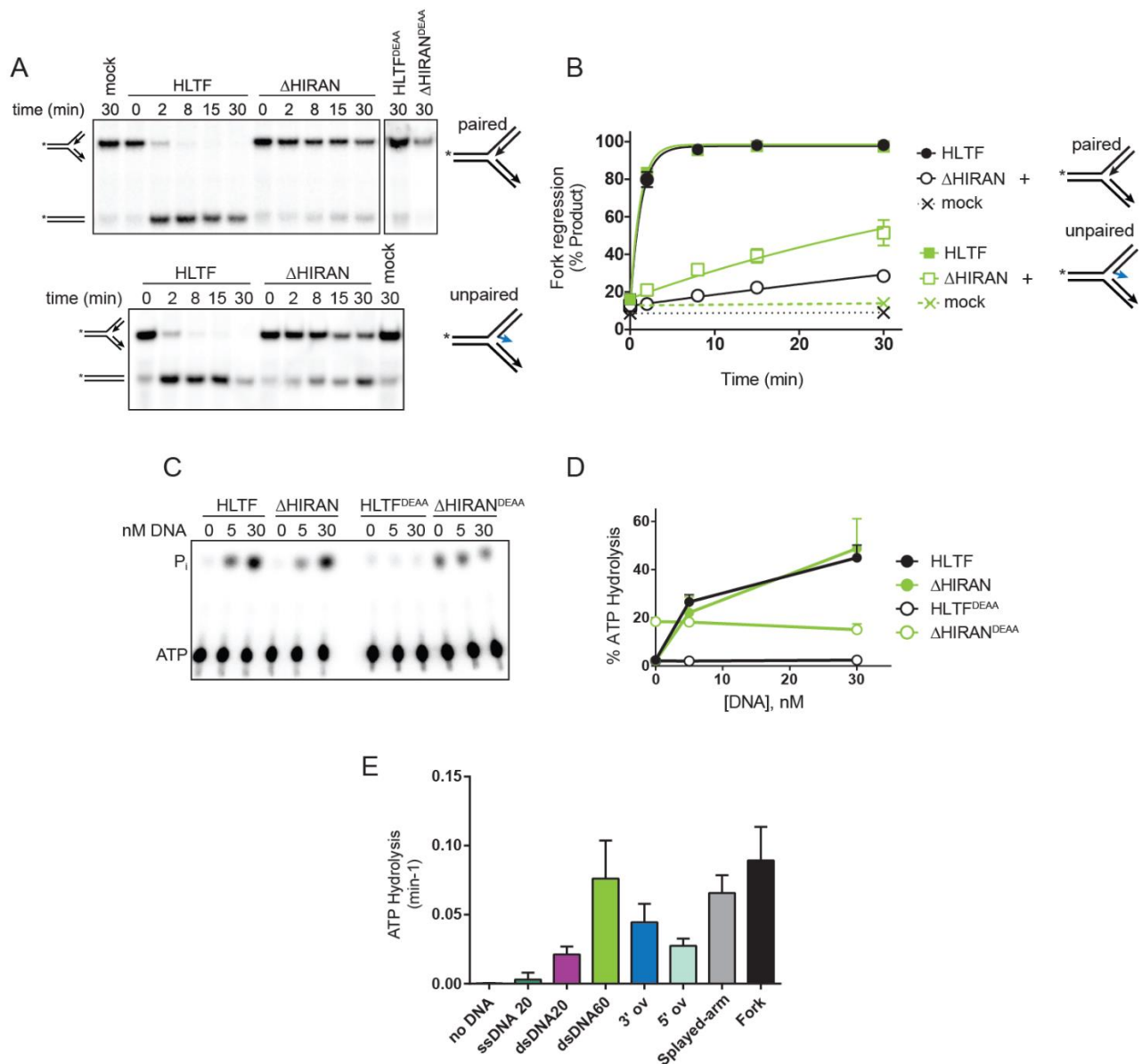
Figure 3.7 HIRAN is required for fork regression. A. Representative native gel of time-dependent fork regression by HLTF and ΔHIRAN against a fork substrate with fully paired nascent arms or containing 2-mismatched base pairs at the 3'end. Endpoints of reactions containing ATPase-dead HLTF^{DEAA} or ΔHIRAN^{DEAA} proteins, or of a mock reaction with no protein are shown as negative controls. **B.** Quantitation of three experiments (mean ± S.D.). **C.** Representative ATPase activities of HLTF, ΔHIRAN, or D557A/E558A mutants of each (HLTF^{DEAA}, ΔHIRAN^{DEAA}) in the presence of fork DNA. Shown is the phosphorimage of TLC plates containing ^{32}P-γATP substrate and inorganic ^{32}P product. **D.** Quantification of data (mean ± S.D.) from three independent measurements. The ΔHIRAN^{DEAA} preparation shows a contaminating ATPase activity not stimulated by DNA. **E.** ATP hydrolysis rates (min^{-1}) of HLTF incubated with different DNA substrates. Values are mean ± S.D., n=3.

70

We hypothesized that if the HIRAN domain played a role in denaturing the nascent arm, then deletion of the HIRAN domain would affect regression activity of a fully paired fork more than a ssDNA gapped fork. Indeed, it was previously shown that deletion of the HIRAN domain from HLTF abolished regression of a fully base-paired fork without affecting DNA-dependent ATPase or DNA translocation activities (Achar *et al*, 2015). Consistent with those results, under our experimental conditions in a standard regression assay using a model fork with both nascent strands abutted against a junction (paired fork), the ΔHIRAN mutant severely diminished HLTF's robust regression activity (Figure 3.6A and 3.6B). Deletion of HIRAN did not impair HLTF ATPase activity (Figure 3.6C and 3.6D), which coincides with previous observations that HLTF's ATPase activity is dependent on the amount of dsDNA, and is highly robust even if HIRAN is not engaged (Achar *et al*, 2015; Kile *et al*, 2015). Interestingly, we found that the ΔHIRAN mutant had greater activity against a fork containing an unpaired nascent 3′-end (unpaired fork) as compared to the fully paired fork (Figure 3.6A and 3.6B). We confirmed that the small amount of residual regression activity present in our purified ΔHIRAN preparation was attributed to the HLTF motor domain, as incorporation of an ATPase-dead D557A/E558A (DEAA) mutation abolished fork regression activity by both HLTF and ΔHIRAN (Figure 3.6A).

Incorporation of a 30-nucleotide ssDNA gap on either leading or lagging template did not significantly affect fork regression activity by HLTF as compared to the fully paired fork (Figure 3.7A and 3.7C). The ΔHIRAN mutant reduced HLTF activity against the gapped fork substrates, but this reduction of activity was strikingly weaker than that of the
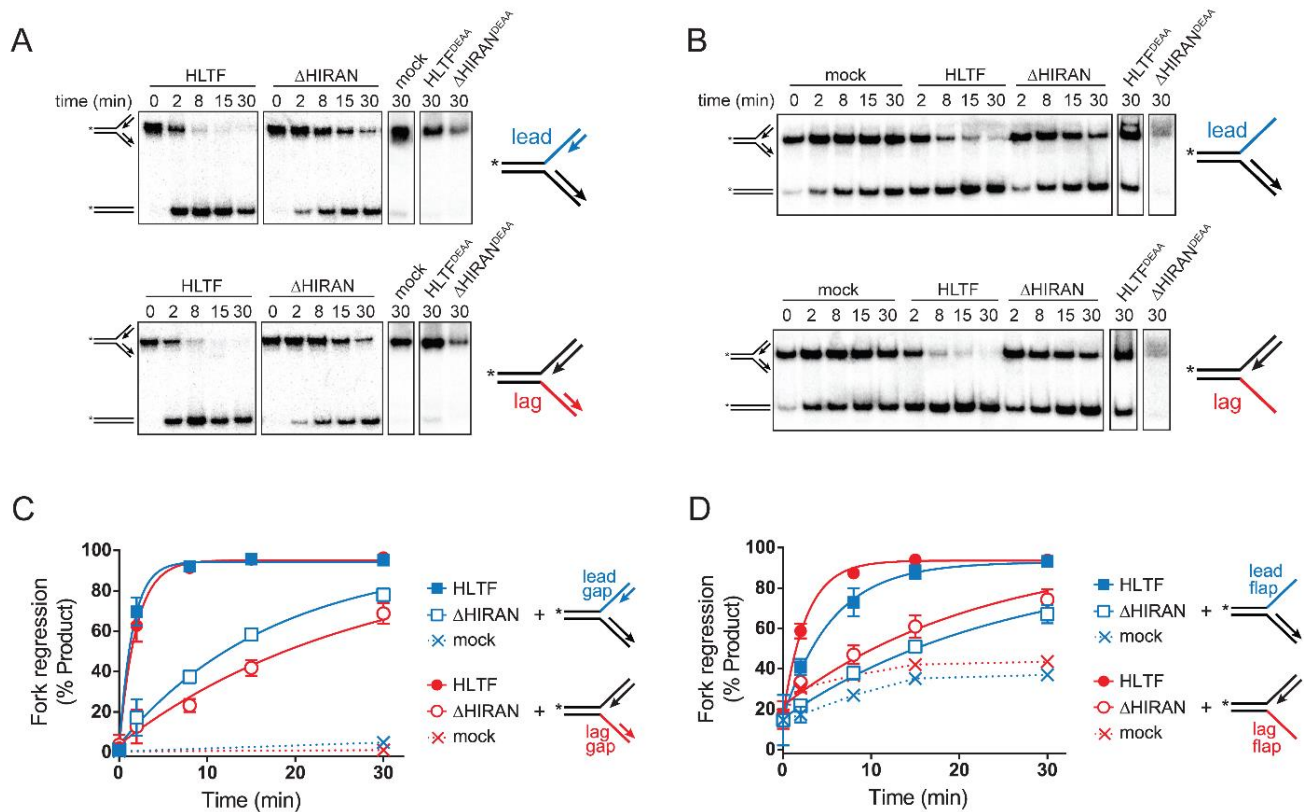
Figure 3.8 HIRAN is not required for regression of forks containing ssDNA at the junction. Representative native gels of time-dependent fork regression by HLTF and ΔHIRAN using forks in which the leading or lagging nascent strands have been shortened (**A**) or removed (**B**). Endpoints of reactions containing ATPase-dead HLTF[DEAA] or ΔHIRAN[DEAA] proteins, or of a mock reaction with no protein are shown as negative controls. Quantification of data is shown (mean ± S.D.) from three independent measurements is shown in (**C**) and (**D**). The 40% weaker DNA signal in the ΔHIRAN[DEAA] lanes are the result of a contaminating nuclease activity in that protein preparation.

fully paired fork (Figure 3.7A and3.7C). These data suggest that HIRAN assists in the transition from 3-way fork to 4-way junction by displacing the nascent strands away from the template. Interestingly, complete removal of either nascent strand caused a modest (2-5-fold) reduction in the rate of regression by full-length HLTF as compared to the four-stranded forks (Figure 3.7B, 3.7D and Table 3.1), indicating that HLTF prefers a four-stranded fork and that the nascent strands are not required for fork regression.

Table 3.1 HLTF fork regression rates. Rates of fork regression (min^{-1}) derived from exponential fits to the data shown in Fig. 3.6. Values are mean ± S.D., n=3.

	WT	ΔHIRAN	mock
Fork	0.8 ± 0.1	0.01 ± 0.001	
Unpaired	0.9 ± 0.03	0.02 ± 0.01	
Lead gap	0.7 ± 0.1	0.06 ± 0.01	
Lag gap	0.5 ± 0.1	0.04 ± 0.001	
Lead flap	0.2 ± 0.03	0.04 ± 0.02	0.01 ± 0.006
Lag flap	0.4 ± 0.2	0.05 ± 0.01	0.01 ± 0.002

Consistent with the result obtained with the gapped substrates, the ΔHIRAN mutant showed only a very modest reduction in activity against these partial forks (Figure 3.7B and 3.7D), further supporting a model in which HIRAN's main function is to promote the transition from a 3-way to a 4-way junction. Removal of the nascent leading strand appeared to have a slightly greater effect than removal of the lagging strand, but this difference can be attributed to the inherent instability of these flap substrates and not to an HLTF-mediated effect since the same effect was observed in the rates of spontaneous fork regression (Figure 3.7B, 3.7D and Table 3.1).

In the cell, ssDNA present at a stalled fork would be bound by RPA, which interacts with other replication factors to promote restart. SMARCAL1 is recruited to forks through a direct interaction with RPA (Bansbach *et al*, 2009; Ciccia *et al*, 2009; Yuan *et al*, 2009; Yusufzai *et al*, 2009; Postow *et al*, 2009). RPA imposes a bias on SMARCAL1 toward forks containing a leading strand gap (Bétous *et al*, 2013; Bhat *et al*, 2015). On the same gapped substrates used in Figure 5C, SMARCAL1 activity was stimulated when RPA was bound to the leading strand gap and inhibited by RPA on the lagging strand. In contrast,

we found that RPA binding to these ssDNA regions modestly inhibited regression by HLTF, consistent with a previous study (Bhat *et al*, 2015). To test if the modest inhibition seen is because RPA is acting as a physical block to HLTF and not due to a direct interaction between HLTF and RPA, yeast RPA was also bound to these fork substrates and tested. Binding of yRPA to these ssDNA regions resulted in the same modest inhibition as seen with human RPA. These data show that RPA does not impose a specificity for a leading or lagging gapped fork and suggest that HLTF and SMARCAL1 use different mechanisms for fork recognition and remodeling.

HLTF catalyzes fork restoration

In addition to their ability to regress stalled forks, SMARCAL1 and ZRANB3 have been shown to catalyze fork restoration *in vitro*, whereby a partially regressed fork is pushed back toward the 3-way replication fork structure (Figure 3.8A) (Bétous *et al*, 2013). We tested whether HLTF also possesses this activity using the previously established assay with the same substrates used for SMARCAL1 and ZRANB3 (Bétous *et al*, 2013). In this assay, partially regressed fork substrates containing either nascent leading or lagging strand ssDNA extensions are used to resemble regressed forks with leading or lagging strand gaps, respectively (Figure 3.8A). We found that HLTF catalyzed ATP-dependent restoration of both forks, with significantly greater activity on the fork containing an extended leading strand (Figures 3.8B-D). Under the assumption that HLTF drives restoration by translocating 3′-5′ (Blastyák *et al*, 2010) toward the fork on either leading or lagging arm, we would predict restoration of the lagging gap substrate to be

74
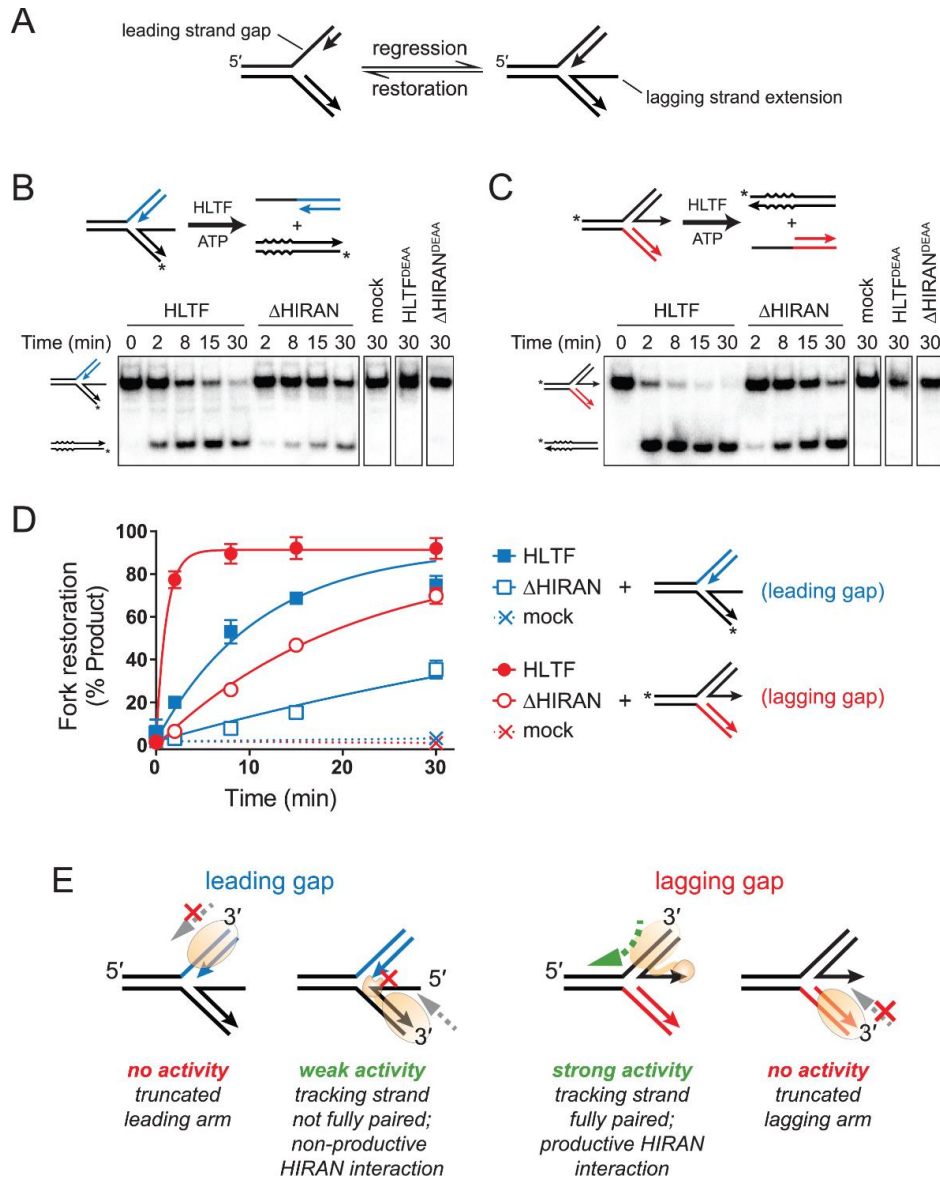
Figure 3.9 HLTF catalyzes fork restoration. A. Schematic of fork restoration. **B,C.** Restoration of partially regressed fork substrates containing ssDNA tails on the lagging (**B**) or leading (**C**) strands. Assay reaction schematics are shown above each representative gel. Restoration activity is defined by annealing the ssDNA tail to the complementary parental strand (which contain four base mismatches to prevent spontaneous annealing), subsequently denaturing the parental duplex. Leading and lagging arms are non-complementary to prevent regression. Asterisks denote ^{32}P-labels. Representative gels show time-dependent accumulation of products for HLTF and ΔHIRAN proteins, as well as 30-min endpoints of reactions containing no enzyme (mock) or ATPase-dead HLTF[DEAA] or ΔHIRAN[DEAA] proteins. **D.** Quantitation of data shown in panels **B** and **C** (n=3, mean ± S.D.). **E.** Schematic model for fork restoration by HLTF. Fork restoration is driven by 3′-5′ translocation HLTF on leading or lagging duplex arms toward the fork and can only be supported when the protein has a full duplex on which to translocate.

most productive since translocation along the fully base-paired leading arm would keep the protein engaged to dsDNA throughout the annealing reaction (Figure 6E). In contrast, the weaker activity of the leading gap substrate can be explained by the fact that a protein tracking 3′-5′ along the nascent lagging stand would eventually encounter the ssDNA tail, which would impede dsDNA translocation until it becomes annealed. We also found that the ΔHIRAN mutant reduced restoration of both substrates, indicating that HIRAN plays a role in this process (Figures 3.8B-D). The residual activity of the ΔHIRAN mutant further indicates that HIRAN is not essential to branch migration. Although the molecular basis for the effect of HIRAN on the restoration reaction is not entirely clear, we speculate that HIRAN interaction with the nascent leading 3′-end stimulates annealing of the leading arm and inhibits annealing of the lagging arm (Figure 3.8E). Consistent with this, occluding access to the nascent leading 3′-end by binding of RPA inhibited restoration of the lagging gap fork, whereas RPA binding to the 5′ ssDNA lagging strand extension had no effect on restoration of the leading gap fork (Figure 3.9C and 3.9D).

Discussion

This work provides experimental evidence for and clarifies some aspects of our previous model for fork binding and remodeling by HLTF (Kile *et al*, 2015). DNase footprinting reveals that HLTF engages all three arms of a stalled fork, both ahead of and behind the junction. The preferential DNase protection of the lagging template ahead of the fork is consistent with the dsDNA motor domain translocating in a 3′-5′ direction (Blastyák *et al*, 2010) toward the fork. Translocation into the junction would conceivably promote annealing of the template strands, which would be facilitated by prior or

concomitant unwinding of daughter duplexes. Indeed, our data show that the motor domain alone is capable of catalyzing regression of forks, but only when they contain ssDNA regions at the junction that lower the barrier to reannealing of template strands. We provide three pieces of evidence to support the conclusion that HIRAN facilitates 4-way junction formation by remodeling the nascent leading strand from its template. First, the motor domain is unable to regress a fully paired fork in the absence of the HIRAN-3′-end interaction. Second, deletion of the HIRAN domain reduced fork regression activity against a lagging gap fork more than against a leading gap fork, indicating that HIRAN is more important to regression when the end of the nascent leading strand is directly at the fork. Third, HLTF binding to a 3′-frayed fork causes a specific DNase protection and sensitivity pattern along the leading template that tracks with the position of the nascent 3′-end and that is dependent on the presence of the HIRAN domain, indicating that the leading arm at the junction is remodeled by the HIRAN-3′-interation.

Since we now know that the HIRAN domain requires an already frayed 3′-end to bind, how the nascent 3′-end is initially denatured is a remaining question. In our *in vitro* assays, we speculate that the activity of the HLTF ATPase motor pushing into the junction is sufficient to destabilize and eventually denature the 3'-end. In the cell, a stalled fork containing a partially unpaired or frayed leading 3′-end is likely to be formed by nucleotide misincorporation by DNA polymerase encountering a damaged template, and/or by dissociation of the DNA polymerase from the template (Fernandez-Leiro *et al*, 2017). Given that the HIRAN-3′-end interaction facilitates but is not required for fork regression, we cannot rule out the possibility that the HIRAN domain also protects a frayed leading end in addition to its role in initiating 4-way junction formation. *In vitro* binding data and co-crystal structures of HIRAN with different DNA sequences are consistent with

accommodation of any nucleobase in the 3′-binding pocket (Kile *et al*, 2015; Hishiki *et al*, 2015; Achar *et al*, 2015) and thus we do not expect HIRAN to have a sequence preference for the frayed 3′-end.

The precise roles of HLTF, SMARCAL1, and ZRANB3 and the need for three related remodelers in the cell remains to be determined. However, the use of the same assays and substrates here and in previous work (Bétous *et al*, 2013) enables a direct comparison of how the unique SRD of each protein imparts a specificity for a particular stalled fork or DNA intermediate generated during fork remodeling. Differences in footprinting on the parental duplex ahead of the fork suggest that HLTF and SMARCAL1 track with opposite polarity on the lagging and leading strands, respectively (Figure 3.10). This difference in strand tracking is consistent with activation of SMARCAL1, but not HLTF, by RPA on leading strand gaps (Bétous *et al*, 2013). It also may suggest that SMARCAL1 processes forks stalled on the leading strand, while HLTF processes lagging strand stalls. Because lesions on the discontinuously replicated lagging strand are less of a threat than impediments on the leading strand, this strand bias helps to explain the stronger cellular phenotypes observed from SMARCAL1 deficiency (Poole & Cortez, 2017). It is also intriguing to speculate that regression of forks stalled on the leading strand by SMARCAL1 creates a substrate recognized by HLTF by pushing the junction closer to the nascent leading strand. Such a sequential operation of each protein would be consistent with a specific recognition of non-duplex DNA structures by the SMARCAL1 HARP domain (Mason *et al*, 2014) and the specific interaction with the nascent 3′-end by HIRAN.
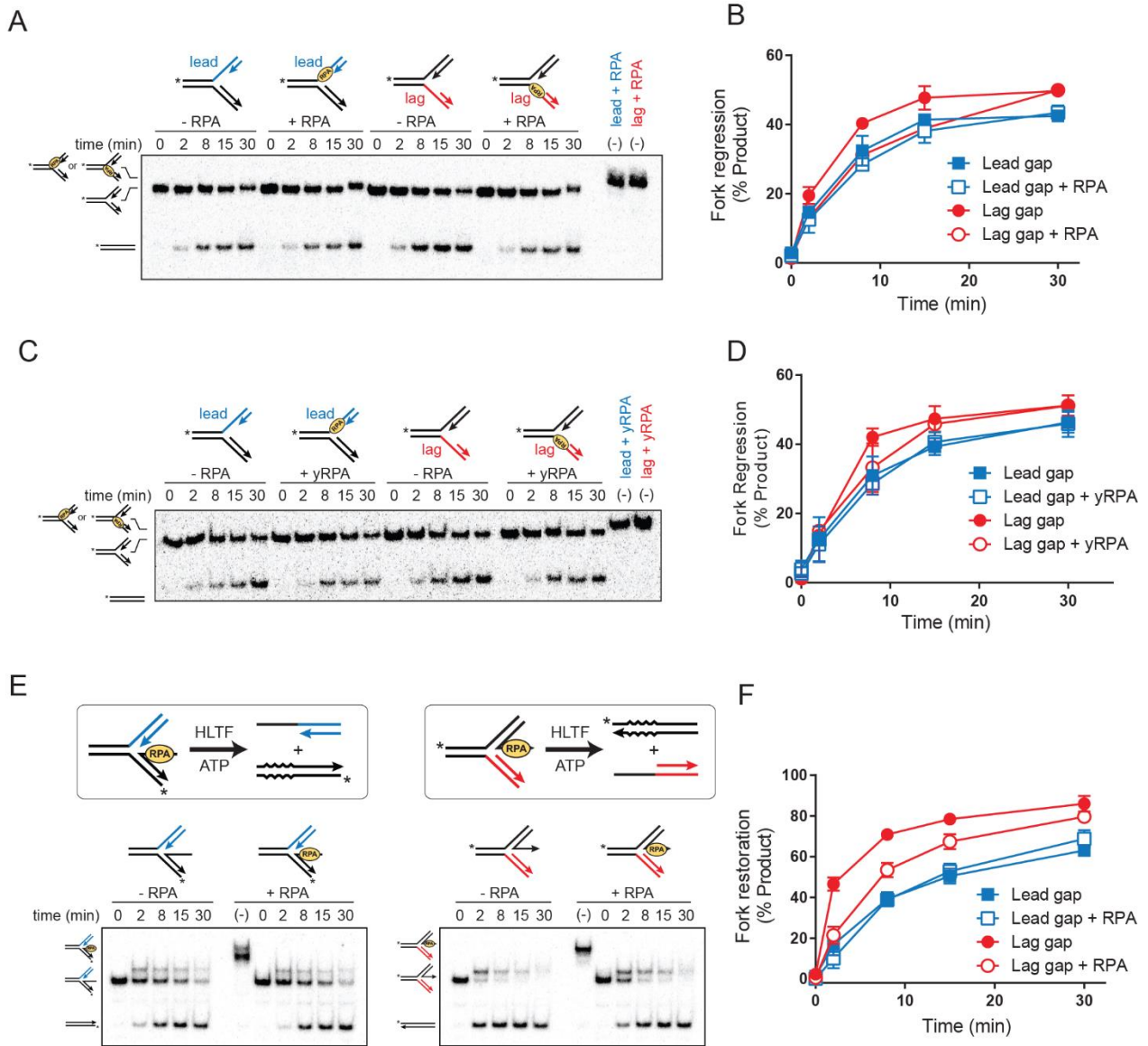
Figure 3.10 Effect of RPA on HLTF fork remodeling activities. Fork regression (**A, B, C, D**) and restoration (**E, F**) by HLTF. Fork substrates contain 30-nucleotide ssDNA regions on either the template (**A, B, C, D**) or nascent strands (**E, F**) that are bound by RPA (**A, B**) or yeast RPA, annotated as yRPA (**C, D**). Experimental details are as specified in Materials and Methods, except DNA was pre-incubated with 3 nM RPA for 30 min prior to addition of HLTF. Representative native gels show ratios of substrates and products over time. The samples lanes labeled "(-)" were not treated with proteinase K prior to loading to verify RPA was bound to the substrates under the experimental conditions. The plots show average data from three experiments (mean ± S.D.).
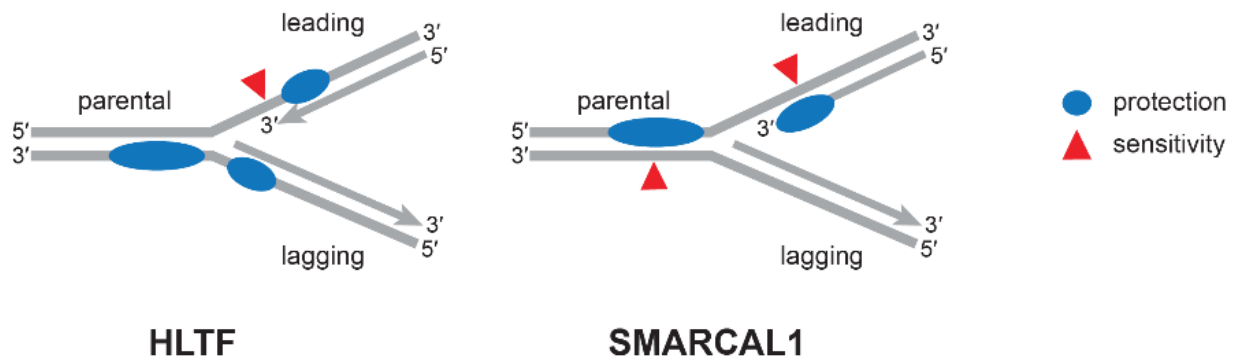
Figure 3.11 Comparison of HLTF and SMARCAL1 footprinintg. Diagram of the observed footprint of HLTF and SMARCAL1 when bound to a fork substrate. Areas of DNase protection are shown in blue and areas of sensitivity are shown in red.

In addition to the template strand bias, there is a marked difference in the intrinsic and RPA-dependent regression and restoration activities of HLTF, SMARCAL1 and ZRANB3 (Bétous *et al*, 2013) (Table 3.2). SMARCAL1 regression of forks is enhanced by RPA bound to the leading template and inhibited by RPA on the lagging template, despite a slight intrinsic preference to regress forks containing a ssDNA gap on the lagging arm. In contrast, neither HLTF nor ZRANB3 shows a preference for regression of forks containing leading and lagging gaps. As noted above, this difference is consistent with putative operation of SMARCAL1 on the leading strand template. Interestingly, whereas HLTF regression is modestly inhibited by RPA bound to the lagging template, ZRANB3 is inhibited by RPA on the leading strand, suggesting a preference for a particular type of stalled fork by these two proteins. Similar differences are evident in fork restoration activities, in which HLTF and ZRANB3 are similar to one another and markedly different from SMARCAL1.

Most strikingly, SMARCAL1 restores leading and lagging gapped forks equally unless bound by RPA, which dramatically stimulates restoration of a lagging gapped fork.

Table 3.2 Effect of RPA on fork remodeling proteins. Comparison of effects of RPA on the fork regression and restoration activities of HLTF, SMARCAL1, and ZRANB3. Preferences for either leading or lagging gap substrates are specified. Data for SMARCAL1 and ZRANB3 are summarized from data reported in Bétous et al (2013) Cell Rep, 3: 1958-69.

	Reversal		Restoration	
	- RPA	+RPA	- RPA	+RPA
HLTF	No preference	Inhibits lag gap	Lag gap	Inhibits lag gap
SMARCAL1	Lag gap	Lead gap	No preference	Lag gap
ZRANB3	No preference	Inhibits lead gap	Lag gap	Inhibits lag gap

In contrast, HLTF and ZRANB3 show an intrinsic preference to restore forks containing a gap on the lagging strand, and this preference is inhibited by RPA. Because RPA coating of ssDNA at a stalled fork is an early response to replication stress, the RPA stimulation of fork remodeling activities of SMARCAL1, but not HLTF and ZRANB3, further suggests that SMARCAL1 operates earlier in the replication stress response than HLTF and ZRANB3. Thus, a more comprehensive picture of the differences and similarities in remodeling activities of HLTF, SMARCAL1, and ZRANB3 is starting to emerge that will facilitate further studies into their non-redundant roles in the cell.

Materials and Methods

Protein purification.

The HIRAN-UvsW chimera was constructed by the overlap extension PCR method, in which the gene sequences corresponding to the HIRAN domain (HLTF residues 55-180) and the UvsW motor (residues 83- 503) were inserted into a pBG101 expression vector (Vanderbilt Center for Structural Biology). The chimera was expressed

as a His$_6$-GST tagged protein in *E. coli* BL-21 RIL cells at 16°C for 16 hr. Cells were lysed using a Avestin Emulsifier C3 homogenizer in 50 mM Tris-HCl pH 7.5, 500 mM NaCl, 20% glycerol, 0.01% NP-40, 0.1% Triton-X100, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), and an EDTA-free protease inhibitor tablet (Roche). The clarified lysate was passed over a Ni-NTA column, eluted with lysis buffer containing 500 mM imidazole, and the His$_6$-GST tag removed by treatment with Rhinovirus 3C (PreScission) protease at 4°C. The ionic strength of the sample was reduced to 300 mM NaCl by dilution with 50 mM Tris-HCl pH 7.5, 20% glycerol, 28 mM 2-mercaptoethanol (2ME) buffer, and then purified over a heparin-sepharose column, eluted by a 300—2000 mM NaCl gradient. The protein was then purified over a Superdex S200 size exclusion column (GE Healthcare) in 50 mM Tris-HCl pH 7.5, 20% glycerol, 500 mM NaCl, and 1 mM dithiothreitol (DTT), concentrated using an Amicon-Ultra 3-kDa concentrator, and stored at -80°C.

The HIRAN domain from human HLTF was purified as previously described (Kile *et al*, 2015). Full-length HLTF was cloned into a pFastBac 438-C vector (Addgene) and expressed as a His$_6$-MBP tagged protein from baculovirus-infected Hi5 insect cells. Cells were harvested 48 hours after infection and lysed using a dounce homogenizer in Buffer A (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 10% glycerol, 0.01% NP-40) supplemented with 1 µg/ml aproptin, 5 µg/ml leupeptin, and 2 µg/ml pepstatin A, and 20 mM imidazole. The clarified lysate was incubated with Ni-NTA resin and eluted with 300 mM imidazole in Buffer A. The protein sample was incubated with amylose resin in 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 % glycerol, and 0.01% NP-40 and eluted by on-column cleavage of the His$_6$-MBP tag by TEV protease at 4°C. Imidazole was added to a final concentration of 30 mM and repassed through a nickel affinity column. The sample was concentrated using an Amicon-Ultra 30 kDa concentrator, buffer exchanged into 50 mM Tris-HCl pH

7.5, 250 mM NaCl, 20% glycerol, 0.5 mM EDTA, and 1 mM tris (2-carboxyethyl)phosphine (TCEP), and stored at -80°C.

The HLTF^{DEAA} (D557A/E558A) construct was prepared by cloning human HLTF into pBG101, followed by mutation of the codons corresponding to D557 (GAT→GCT) and E558 (GAA→GCG) using the QuikChange Mutagenesis Kit (Agilent). Protein was overexpressed with an N-terminal His$_6$-GST tag in *E. coli* Rosetta cells at 16°C for 16 hrs. Cells were lysed using a Avestin Emulsifier C3 homogenizer in Buffer A containing 5 mM imidazole, 1 µg/ml aproptin, 5 µg/ml leupeptin, and 2 µg/ml pepstatin A and purified by Ni-NTA affinity chromatography using 500 mM imidazole in Buffer A. The protein sample was incubated with glutathione resin and eluted by an on-column cleavage by Rhinovirus 3C (PreScission) protease at 4°C. The sample was diluted to 250 mM NaCl by addition of a Tris-HCl pH 7.5, 20% glycerol, 1 mM TCEP buffer and purified from heparin-sepharose column using a 250-2000 mM NaCl gradient. Protein was concentrated and stored in the same manner as wild-type HLTF.

The ΔHIRAN mutant (residues 181-1009) gene was cloned into pFastBac-HTb vector (Invitrogen) and expressed as a His$_6$-tagged protein in bacolovirus-infected Hi5 insect cells. Cells were lysed 48-hours after infection using a dounce homogenizer in Buffer A and protein purified by Ni-NTA affinity chromatography, eluting with 300 mM imidazole in Buffer A. The sample was diluted to 250 mM NaCl by addition of a Tris-HCl pH 7.5, 20% glycerol, 1 mM TCEP buffer and purified by heparin-sepharose with a 200—1000 mM NaCl gradient elution. Protein was concentrated, buffer exchanged into 50 mM Tris-HCl pH 7.5, 250 mM NaCl, 20% glycerol, 0.5 mM EDTA, and 1 mM TCEP, and stored at -80°C. ΔHIRAN^{DEAA} mutant was constructed by introducing D557A and D558A mutations into the ΔHIRAN/pFastBac vector and purified the same as ΔHIRAN protein.

DNA substrate preparation.

Oligodeoxynucleotides used for DNA binding, footprinting, and enzymatic assays are shown in Table S1 in Supporting Information. Annealing reactions were carried out in 1X SSC buffer (15 mM sodium citrate pH 7.0 and 150 mM NaCl) and decreasing the temperature from 95° to 25°C at 1°/min using a thermal cycler. Duplexes containing 5′-overhangs used in HIRAN binding experiments were generated by annealing oligonucleotide *ov_F*, which contained a 5′-^{32}P and a 3′-6-carboxyfluorescein (FAM), to either *ov_0*, *ov_1*, *ov_2*, or *ov_3* in a 1:1 ratio. Heterologous forks used in EMSA and DNase footprinting studies were generated by annealing all four oligonucleotides using a 1.2-fold molar excess of each unlabeled strand. Fork regression substrates were generated by annealing leading (*A*/*B* or *48*/*50*) and lagging (*C*/*D* or *52*/*53*) arms of the fork separately. The two duplexed arms were then combined using a two-fold molar excess of the unlabeled arm and incubated at 37°C for 30 min. Fork restoration substrates were generated by annealing three arms of the fork (*R1*/*R2*/*R3* or *R1*/*R2*/*R5*), followed by addition of five-fold excess of the R5 or R6 oligo and incubation at 37°C for 30 min. With the exception of *48*/*50*/*52*/*53*, fork substrates were PAGE purified using a 6% 0.5X TBE DNA Retardation Gel (Invitrogen). Substrates were excised from the gel and electroeluted in 0.5X TBE, concentrated using a 10K Amicon-Ultra 0.5 ml centrifugal filter, and stored at -20°C.

DNA binding.

All DNA binding reactions were carried out for 20 min at 25°C using 0-50 nM protein and either 5 nM (5′-overhang duplex) or 25 nM (fork) DNA. HIRAN binding was performed in 20 mM Tris-HCl pH 8.3, 100 mM NaCl, 2 mM DTT and 0.1 mM EDTA, and

electrophoresed on a 5% 29:1 (acrylamide:bis-acrylamide) 0.5X TBE gel. HLTF (wild-type and ΔHIRAN) binding was performed in 20 mM Tris-HCl pH7.76, 3% glycerol, 0.5 mM CaCl$_2$, 5 mM MgCl$_2$, 0.2% NP-40 and 0.05 mM TCEP and electrophoresed on a 5% 79:1 (acrylamide:bis-acrylamide) 0.5X TBE 3% glycerol gel. Gels were run with 0.5X TBE at 200 V for 0.5-1 hr. Gels were either phosphorimaged (^{32}P-DNA) or directly fluorimaged at 532 nm excitation and 526 nm emission wavelengths on a Typhoon Trio variable mode imager. Band intensities were quantified with GelAnalyzer and data were plotted using GraphPad Prism 6.

Fork regression and restoration assays.

Fork remodeling reactions were carried out at 37°C and contained 10 nM HLTF, 1 nM ^{32}P-labeled fork substrate, 2 mM ATP, 40 mM Tris-HCl pH 7.8, 50 mM NaCl, 5 mM MgCl$_2$, and 1 mM TCEP. At various time points, 10 μl of the reaction were stopped by adding 320 mU Proteinase K (Sigma) and electrophoresed on a 6% 19:1 (acrylamide:bis-acrylamide) 1X TBE gel at 6 W for 1-1.5 hr. Gels were phosphorimaged and quantified as above.

DNase footprinting.
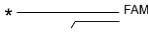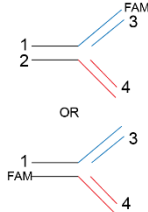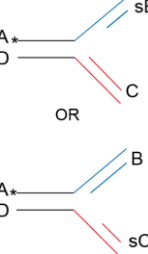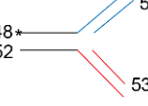
100 nM fork substrate, in which either the leading or lagging template strand contained a 3′-FAM label (Figures 3.5, 3.6, Table 3.3) was incubated for 20 min at 25°C with 0, 25, 50, 100, or 200 nM HLTF in 20 mM Tris-HCl pH 7.8, 3% glycerol, 0.5 mM CaCl$_2$, 5 mM MgCl$_2$, and 0.05 mM TCEP. DNA was digested with 80 mU Benzonase (Sigma) for 2 min at 25°C. The digestion reaction was stopped by addition of Proteinase K, denatured with loading buffer (80% formamide, 10 mM EDTA) at 70° for 5 minutes and

electrophoresed on a 15% 19:1 (acrylamide:bis-acrylamide) 1X TBE 7M urea denaturing gel for 3 hr. Gels were fluorimaged and quantified as above. The changes in band intensities between only the 0 and 200 nM HLTF samples were determined for each band and normalized to the band intensities from the 0 HLTF sample. The mean changes were calculated separately for positive and negative slopes. Bands with a normalized slope of 0.5 standard deviations from the mean were labeled as significant.

ATPase activity assays.

5 nM protein was incubated for 30 min. with 0, 5, or 30 nM fork DNA substrates for ΔHIRAN ATPase assays, and 5 nM protein was incubated for 0, 5, 15, 30, 60, or 120 min. with 100 nM DNA for assays with varying DNA substrates. Oligonucleotides used for these assays are listed in the table under ATPase assays, with the exception of *48* and *52* which were used to make the dsDNA60 substrate.10-μL reactions containing 40 mM Tris-HCl pH 7.8, 5 mM $MgCl_2$, 50 mM NaCl, 1 mM TCEP, and 1.67 uM ^{32}P-γATP and were incubated at 37°C. Reactions were stopped by adding EDTA to a final concentration of 25 mM. ATP and P_i were separated by thin layer chromatography, in which a 1 μL reaction aliquot was spotted on PEI cellulose F TLC plates (Milllipore) and resolved in 0.25 M LiCl/1 M formic acid for 30 min. Plates were phosphorimaged (20 min exposure) and quantified as above.

Table 3.3 Oligodeoxynucleotides used in Chapter III.			
Substrate	**Name**	**Length**	**Sequence (5′-3′) [a]**
EMSA & footprinting			
	ov_F	30	^{32}P–TTTTTTTTTTTTTTTTCCGCTACGCATGTCC–FAM
	ov_0	15	GGACATGCGTAGCGG
	ov_1	16	GGACATGCGTAGCGGT
	ov_2	17	GGACATGCGTAGCGGTT
	ov_3	18	GGACATGCGTAGCGGTTT
	1_F	55	GCATCCGACTCAGTTCGCTCGAGCTAGCCCTGATATCGATGGATCTAGAGCTACC–FAM
	1	55	GCATCCGACTCAGTTCGCTCGAGCTAGCCCTGATATCGATGGATCTAGAGCTACC
	2	55	GGCGAAGGGATCCGTAGGCACAGTTCCCCTAGCTCGAGCGAACTGAGTCGGATGC
	2_F	55	GGCGAAGGGATCCGTAGGCACAGTTCCCCTAGCTCGAGCGAACTGAGTCGGATGC–FAM
	3	28	GGTAGCTCTAGATCCATCGATATCAGGG
	3_TT	28	GGTAGCTCTAGATCCATCGATATCAGTT
	3_PO4	28	GGTAGCTCTAGATCCATCGATATCAGTT–PO$_4$
	4	28	GGGAACTGTGCCTACGGATCCCTTCGCC
	3gap_TT	19	GGTAGCTCTAGATCCATTT
	A	122	^{32}P–CGTGACTTGATGTTAACCCTAACCCTAAGATATCGCGTTATCAGAGTGTGAGGATACATGTAGGCAATTGCCACGTGTCTATCAGCTGAAGTTGTTCGCGACGTGCGATCGTCGCTGCGACG
	B	82	CGTCGCAGCGACGATCGCACGTCGCGAACAACTTCAGCTGATAGACACGTGGCAATTGCCTACATGTATCCTCACACTCTGA
	C	82	TCAGAGTGTGAGGATACATGTAGGCAATTGCCACGTGTCTATCAGCTGAAGTTGTTCGCGACGTGCGATCGTCGCTGCGACG
	D	122	CGTCGCAGCGACGATCGCACGTCGCGAACAACTTCAGCTGATAGACACGTGGCAATTGCCTACATGTATCCTCACACTCTGAATACGCGATATCTTAGGGTTAGGGTTAACATCAAGTCACG
	sB	52	CGTCGCAGCGACGATCGCACGTCGCGAACAACTTCAGCTGATAGACACGTGG
	sC	52	CCACGTGTCTATCAGCTGAAGTTGTTCGCGACGTGCGATCGTCGCTGCGACG
	48	60	^{32}P–ACGCTGCCGAATTCTACCAGTGCCTTGCTAGGACATCTTTGCCCACCTGCAGGTTCACCC
	50	30	GGGTGAACCTGCAGGTGGGCAAAGATGTCC
	50-GG	30	GGGTGAACCTGCAGGTGGGCAAAGATGTGG
	52	60	GGGTGAACCTGCAGGTGGGCAAAGATGTCCCAGCAAGGCACTGGTAGAATTCGGCAGCGTC
	53	30	GGACATCTTTGCCCACCTGCAGGTTCACCC

87

Fork restoration

	R1	62	^{32}P– TAGGCAATTGCCACGTGTCTATCAGCTGAAGTACAAGCGCTGCACCCTAGGTCCGACGCTGC
	R2	62	^{32}P– CGTCGCAGCCTGGATCCCACGTCGCGAACAACTTCAGCTGATAGACACGTGGCAATTGCCTA
	R3	62	GCAGCGTCGGACCTAGGGTGCAGCGCTTGTTGTTCAGGTGATACACACGCGGCAAATGCCTA
	R4	30	TGTTCGCGACGTGGGATCCAGGCTGCGACG
	R5	30	GCAGCGTCGGACCTAGGGTGCAGCGCTTGT
	R6	62	TAGGCATTTGCCGCGTGTGTATCACCTGAACATGTTCGCGACGTGGGATCCAGGCTGCGACG
	40	40	CTCAGGACTCAGTTCGTCAGCCCTTGACAGCGATGGAAGC
	Lead	20	GCTTCCATCGCTGTCAAGGG
	Lead com	20	CCCTTGACAGCGATGGAAGC
	Lag	20	GGGAACTGTCGCTACCTTCG
	F20.40	40	CGAAGGTAGCGACAGTTCCCCTGACGAACTGAGTCCTGAG

[a] Underlined nucleotides form mismatched base pairs
[b] Sequences for the *ABCD*-type fork regression and fork restoration substrates were taken from Bétous *et al* (2013) *Cell Reports*, 3: *1958-69.*

CHAPTER IV

DISCUSSION AND FUTURE WORK

Summary of work

Understanding the mechanisms that are used to stabilize stalled forks is essential for understanding how DNA is replicated in a complete and accurate manner. Consequently, much work has been done to characterize fork remodeling proteins by determining the different specificities and activities that the domains within these proteins have. Characterization of the domains and their activities will lead to a better understanding of the mechanisms these proteins use to stabilize forks during replication stress, which may ultimately be used to develop treatments for cancer and other diseases caused by defects in DNA replication.

The work presented in this thesis expands our knowledge of HLTF. Prior to this work, only two of HLTF's domains, the RING and SNF2 motor, had been characterized. These two domains allow HLTF to act as a ubiquitin ligase to polyubiquitinate PCNA to promote the template switching pathway of DDT and allow HLTF to translocate on DNA to catalyze fork regression. In contrast, the activity of HLTF's HIRAN domain had remained elusive. The focus of this work was to characterize HLTF's HIRAN domain and understand how HIRAN promotes HLTF fork regression. In summary, I have established that HIRAN is a substrate recognition domain. Using biochemical and structural methods, I showed that HIRAN specifically recognizes and binds to the 3´OH of ssDNA. Structural analysis of HIRAN showed that it has a unique binding pocket where several residues

interact with ssDNA to form the basis for this specificity. I also demonstrated that HIRAN can bind to a ssDNA-dsDNA junction, and a fork substrate, as long as there are least two unpaired nucleotides on the 3'-end. HIRAN binding to DNA is also essential for HLTF fork regression *in vitro.* Furthermore, I have also demonstrated that HIRAN binding to a fork helps position HLTF at a fork, and enables HLTF fork regression by facilitating the transition from a three-way to a four-way junction. Additional work by our collaborators also demonstrated that HIRAN is essential for HLTF function *in vivo,* as impairing HIRAN DNA binding resulted in altered fork progression in cells under stress.

Remaining questions about HLTF function

HLTF fork remodeling

Our work has led to a model of HLTF fork regression where the SNF2 motor binds the duplex DNA ahead of the fork and places HIRAN in close proximity to the 3'-end of the nascent leading strand. In this model, HIRAN binding to the nascent 3'-end helps facilitate the formation of a four-way junction to drive fork regression. Even though this model helps us to visualize the cooperation between HLTF's motor and the HIRAN domain, it needs to be further tested and leaves unanswered questions. Due to the nature of the fork substrates, which contain a mismatch at the junction to prevent spontaneous regression, the fork regression assays used in these studies only measure the ability of HLTF to translocate past the mismatch at the fork junction. It is possible that the regression of the fork arms past the mismatch junction in these relatively short model fork substrates is simply a spontaneous reannealing and not enzyme catalyzed. Therefore, it

is still unknown how HLTF actively promotes strand reannealing at forks on longer substrates. Single-molecule studies that can measure continuous fork reversal will be useful to characterize HLTF fork remodeling.

Another outstanding question is, if HIRAN binds an already frayed 3´-end, how does this occur? Perhaps the motor, by pushing on the junction, causes the 3´-end to be frayed allowing HIRAN to bind. Recent work using single molecule FRET assays have shown that ATP binding induces Rad5 to unwind the leading strand of a fork to catalyze fork regression rad5 (Shin *et al*, 2018). It is possible that like its yeast homolog, upon binding ATP HLTF unwinds the leading strand allowing HIRAN to bind. Once HIRAN is bound, the motor hydrolyses the bound ATP and in conjunction with HIRAN catalyzes the formation of a four-way junction to start fork regression. Another question that remains unanswered is how long HIRAN binds to the 3´-end during fork regression. One possibility is that HIRAN's interaction with the 3´-end is transient, and HIRAN disengages from the nascent end almost as soon as it binds to it. Another possibility is that HIRAN binds to the nascent 3´-end and is bound to the regressed arm until HLTF is done reversing the fork. This would also keep the ends of the regressed arm a certain distance from the motor, which may be important for adequate HLTF function. Perhaps HIRAN's interaction with the 3'-end is controlled by a change in the DNA structure as the four-way junction is formed, or by a conformational change in the motor itself induced by ATP hydrolysis. Single molecule FRET fork regression assays in which the HIRAN domain and the 3´-end of the nascent leading strand at a fork are labeled would help determine how long HIRAN is bound to the fork during HLTF fork regression.

Understanding how the different domains of HLTF engage a fork substrate and how they are spatially placed in relation to each other will help to answer the remaining questions about HLTF fork remodeling mechanisms. A structure of full-length HLTF bound to a fork substrate is needed for this. Currently, there are no structures of full-length SNF2 fork remodeling enzymes as conducting structural studies on these enzymes has proved to be a challenge. These proteins are large, usually over 1000 amino acids,
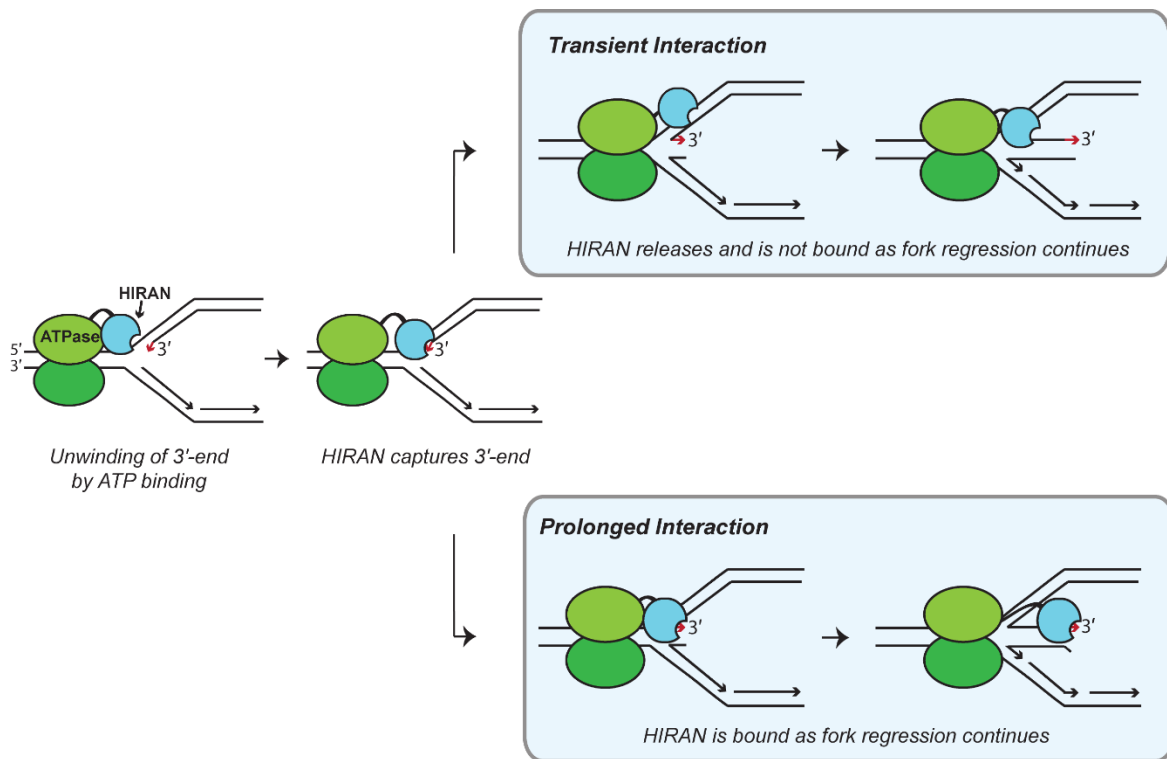
Figure 4.1 Models for HIRAN interaction with 3'-end. Models for interaction of the HIRAN domain with the nascent 3'-end during fork regression. HIRAN could have a transient (top) or prolonged (bottom) interaction with the nascent 3'-end (red) during fork regression. The two ATPase lobes of HLTF are shown in green and HIRAN in blue.

and can be difficult to express and purify. It is also common for them to have flexible regions that make them even more challenging to work with. In addition to these issues

common to SNF2 enzymes, HLTF has a predicted disordered region within the motor, between motifs *I* and *Ia*, that make HLTF a poor crystallographic target.

To bypass these difficulties, we have opted to pursue crystallographic studies with HLTF orthologs that lack this disordered region within their motor. We have identified several HLTF proteins that lack the insertion and that have the same domain structure as human HLTF. These identified proteins, which belong to plant or fungal species, are also smaller than human HLTF (less than 870 residues) that should aid with expression and purification. With the help of the EMB core of the Structural Biology of DNA Repair program, we have expressed all of these homologs as a His$_6$-MBP tagged protein from baculovirus-infected Hi5 insect cells and purified using Ni-NTA, amylose and size-exclusion chromatography (Table 4.1). I have begun testing these orthologs for DNA binding, ATPase, and fork regression activity to ensure that these proteins have similar functions to human HLTF and to identify the best candidate for future crystallographic studies. Based on these results and protein purity, I have identified the *Neurospora crassa*

Figure 4.2 HLTF homolog domain architecture. Domain structure of the different HLTF homologs in comparison to human HLTF. The corresponding organism name and length is listed in Table 4.1 (below).
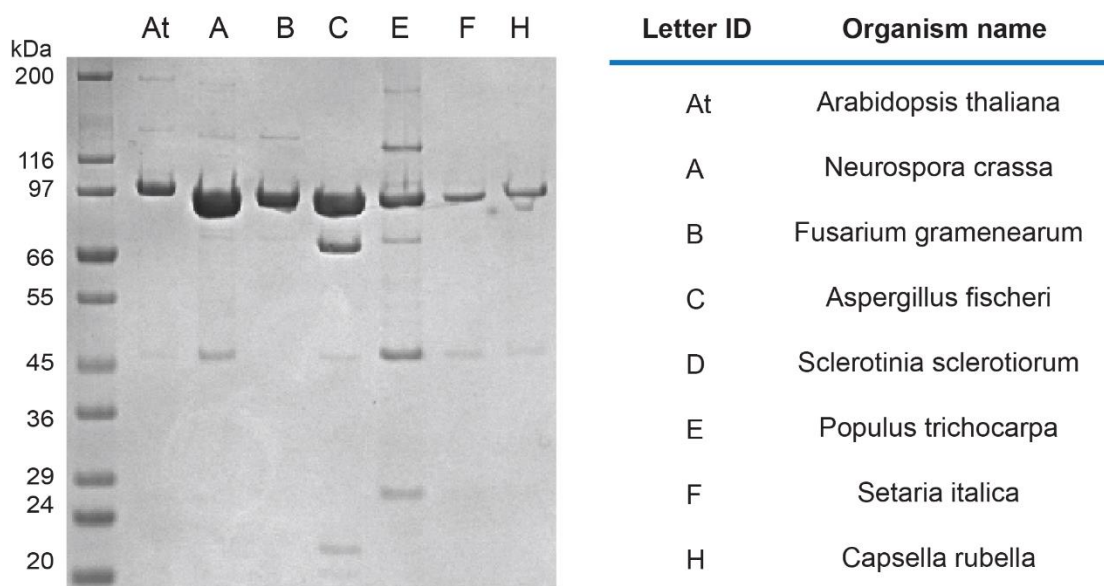
Figure 4.3 HLTF homolog purity. SDS-PAGE of the final purity of the different homologs. D was not purified due to lack of solubility.

Table 4.1 Homolog purification notes. pH of the lysis buffer used (50 mM Tris-HCl pH 7.3, or 7.8, 500 mM NaCl, 15% glycerol, 0.01% NP-40) for the different homologs. Also listed are the different isoelectric points, and the size of each (kDa) in the presence or absence of the 6XHis-MBP tag.

	Buffer pH	pI with tag	Size with tag	pI without tag	Size without tag
At	7.3	7.8	142	8.6	96
A	7.8	6.8	136	7.9	92
B	7.8	7.2	137	8.3	92
C	7.8	6.8	140	7.9	95
D	7.3	8.2	138	8.7	93
E	7.3	6.9	133	8.2	88
F	7.3	6.8	138	7.9	93
H	7.8	7.1	140	8.4	95
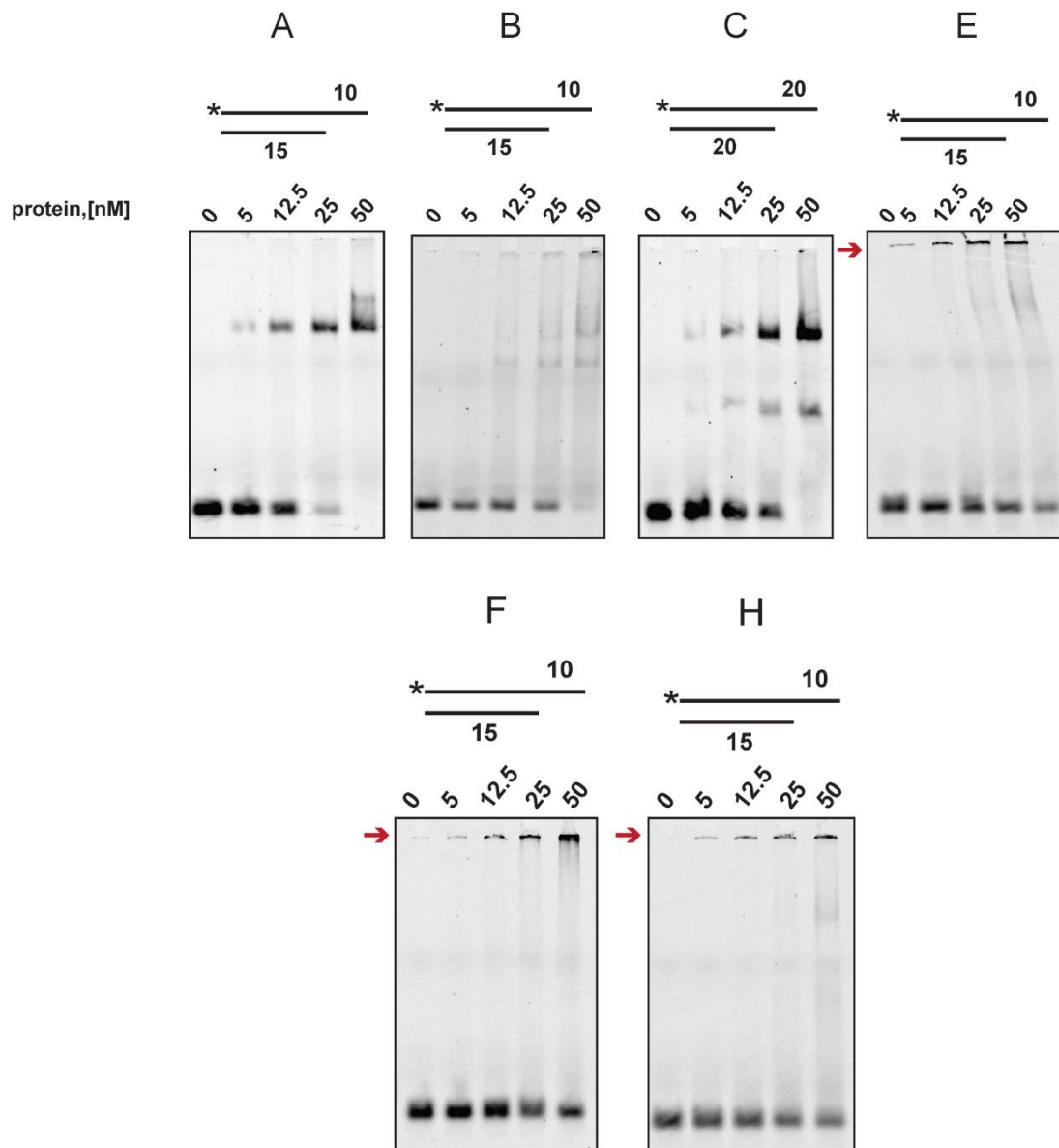
Figure 4.4 HLTF homolog EMSAs. Binding of HLTF homologs to 3' overhang fluorescein (FAM)-labeled substrates. Binding was observed with A and C homologs. Binding results of E, F, and H were not conclusive because the DNA did not migrate into the gel (red arrows). Position of the FAM is indicated by a star and the numbers on the DNA schematics indicate the length of DNA.

Table 4.2 HLTF homolog characterization summary. Summary of work done to purify and characterize the stability, DNA binding and ATPase activity of the HLTF homologs.
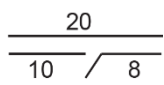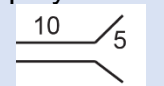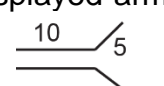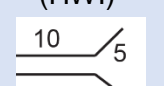
	Length (AA)	Expression/Solubility	Stability	DNA Binding	Activity
At	862	Good/High	Stable	Binds[H]	Low activity[A]
A	813	High/High	Stable	Binds[E]	High activity[A,F]
B	819	High/High	Stable	No binding[E]	ND
C	845	Good/High	Degrades	Binds[E]	ND
D	837	High/None	Stable	ND	ND
E	792	Good/Poor	Degrades	ND[E]	ND
F	832	Good/High	Stable	ND[E], Binds[H]	ND
H	846	Good/Poor	Stable	ND[E]	ND

Binds[H] -tested by binding to Heparin column; Binds[E]/ No binding[E] -tested by EMSA
ND- not determined; ND[E] – not determined although tested by EMSA;
Low activity[A] – activity tested by ATPase assay; High activity[A,F]- activity tested by both ATPase assay and fork regression

Table 4.3 Homolog crystallographic trials. Crystallography screening conditions used with homologs. A schematic of some of the DNA substrates are included for clarity.

	Condition	Protein: DNA ratio	Protein concentration	Number of conditions
At	alone	NA	2 mg/ml	192
	10 dsDNA	1:1.2	2 mg/ml	864
A	alone (w/ 5% glycerol)	NA	4 mg/ml	1056
	alone (HWI)[*]	NA	4 mg/ml	1500
	ADP MgCl$_2$	NA	4 mg/ml	672
	AMP-PNP	NA	4.6 mg/ml	1056
	18 dsDNA	1:1.2	4 mg/ml	864
	flap DNA	1:1.2	9 mg/ml	672
	splayed-arm	1:1.2	8 mg/ml	1248
	splayed-arm	1:1.8	12 mg/ml	192
	splayed-arm (HWI)[*]	1:2	9.5 mg/ml	1500
B	alone	NA	5 mg/ml	576
C	alone	NA	3.4 mg/ml	576

HWI[*]- conditions set-up at Hauptman-Woodward Institute

HLTF ortholog (also referred to as HLTF-A) as the best candidate for future structural studies, but have started crystallographic trials with At, A, B, and C (Table 4.2). Other structural techniques, such as small-angle x-ray scattering (SAXS) and electron microscopy (EM) can be used as alternative approaches to obtain a low-resolution structure if crystallography is not successful.

The connection between HLTF E3 ligase and fork remodeling activities

Several aspects of HLTF function remain unclear. Perhaps most importantly, it is still not well understood whether HLTF acts primarily as an E3 ligase or a fork remodeler, or if these activities are connected. HLTF is part of the RING E3 ligases, which constitutes the majority of E3 ligases in cells (Metzger *et al*, 2014). RING E3 ligases use a RING domain, which coordinates two zinc ions in a cross-braced arrangement, to create a platform for ubiquitin-conjugated E2 enzymes to bind. E3s catalyze the transfer of ubiquitin from E2 to a substrate and are essential to determining the specificity of the ubiquitin modification reaction (Metzger *et al*, 2014).

Interestingly, HLTF's RING domain is located between the two ATPase lobes of its SNF2 motor. It is unclear if this domain architecture is functionally relevant. It is possible that the close placement of these two domains allow them to regulate each other. Perhaps when the RING domain is engaged to its substrate or actively catalyzing ubiquitination of PCNA, it hinders the two ATPase lobes from orienting themselves in a catalytic active conformation. Contrarily, when the ATPase lobes are engaged in a correct orientation to bind and hydrolyze ATP, the RING domain cannot catalyze ubiquitination

because it is not able to engage with the ubiquitinated E2 or its substrate. Another question is if in addition to regulating each other, the activities of the two are connected. It is not clear if HLTF only regress a fork after it polyubiquitinates PCNA, or if its fork remodeling activity is independent from its E3 ligase activity. Alternatively, HLTF first uses its fork remodeling activity to try to regress a stalled fork but if it is unsuccessful, then it polyubiquitinates PCNA to recruit ZRANB3.

In order to investigate these questions, biochemical and functional studies using HLTF mutant proteins need to be pursued. To determine if the placement of the RING domain is functionally relevant, the ATPase and E3 ligase activities of a HLTF protein where the RING domain is moved from within the motor should be compared to wild-type HLTF. Additional studies using HLTF proteins with separation of function mutations that inhibit either ATP hydrolysis or inhibit E3 ligase activity should be pursued to determine if the two activities are independent. New methods will need to be developed to characterize the relevance of HLTF's E3 ligase and fork remodeler functions *in vivo* for maintaining genomic stability.

Other remaining questions about HLTF function

There are other unclear aspects about HLTF function. For example, what is the relevance of HLTF's fork restoration activity; does HLTF catalyze fork restoration *in vivo* and how is this function regulated? Furthermore, if HLTF does travel with a fork even under normal replication, how is it recruited? Is HLTF recruited through its interaction with PCNA, or does it bind other proteins at a fork? Could it also be that HLTF associates with

a fork by translocating on the DNA behind a moving fork? To begin testing this, HLTF proteins with mutations that disrupt its interaction with PCNA or its translocation activity could be expressed in cells and tested with iPOND to determine if HLTF is still recruited to a fork.

Although many studies have investigated DDT pathway choice, it is still unknown how it is regulated and if HLTF is somehow involved. One possibility is that HLTF is crucial to pathway choice. After stress is encountered and PCNA is monoubiquitinated by Rad6 and Rad18, HLTF binds Rev1, just as Rad5 does, to help mediate the interaction between Rev1, TLS polymerases, and PCNA to promote the TLS pathway (Xu *et al*, 2016). Instead of binding to Rev1, HLTF could bind to the Ubc13-Mms2 complex to polyubiquitinate PCNA and to promote template switching. In these scenarios, HLTF is essential for the stability of these complexes and as such is the key mediator in promoting either TLS or template switching.

It is also not clear what ubiquitination mechanism HLTF uses to modify PCNA, as E3 RING ligases can use different mechanisms to polyubiquitinate a substrate (Hochstrasser, 2006). One of these mechanisms is the sequential mechanism, where ubiquitin is added on to the substrate successively to form a chain. Alternatively, in the *en-bloc* mechanism a ubiquitin chain is added after being pre-assembled on the E2. *En-bloc* mechanisms can be sequential, where the chain is pre-assembled on the E2 before its transferred to the substrate or can be 'see-saw' where the growing chain is reciprocally transferred between the E2 and the E3 before its transferred to the substrate. It is not yet clear if the see-saw mechanism that HLTF uses *in vitro* to ubiquitinate PCNA is relevant *in vivo*. In fact, in the same study where HLTF's see-saw mechanism was determined,
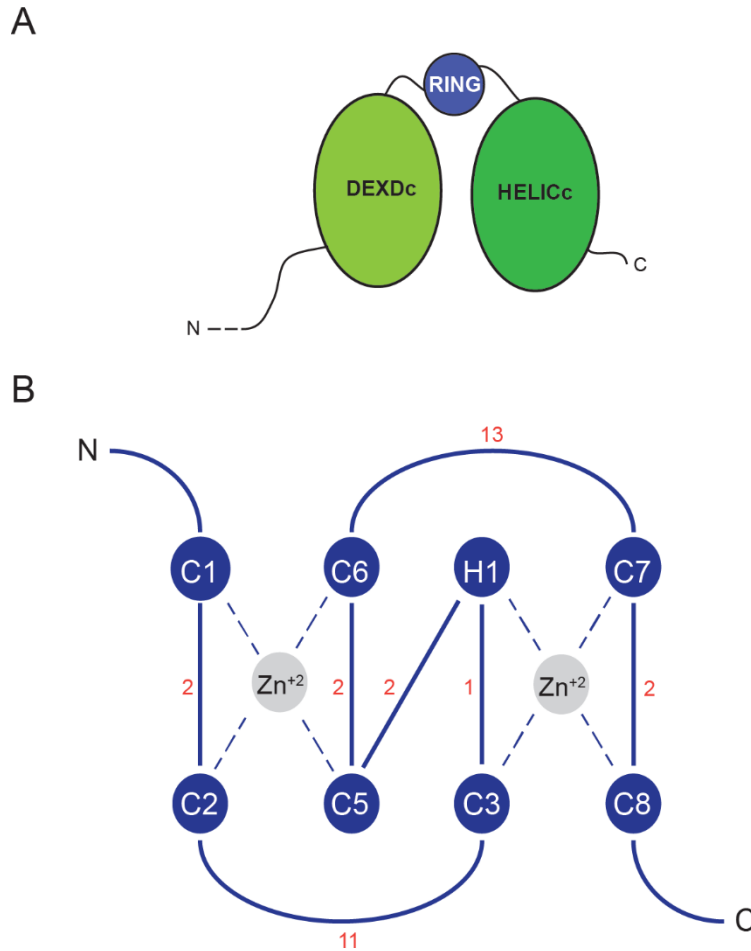
Figure 4.5 HLTF RING domain. A. Schematic of the location of the RING domain in between the two ATPase lobes. **B.** Depiction of the arrangement of the different residues coordinating the zinc ions in the HLTF RING domain. The numbers in red indicate the number of residues in between the coordinating residues.

HLTF was also observed to use a sequential mechanism, although less ubiquitination was seen with this method (Masuda *et al*, 2012).

Another unclear aspect of HLTF function is how it is regulated so that its E3 ligase and fork remodeling activities are inhibited until stress in encountered. One possible answer is that HLTF, like many other proteins, is regulated by post-translational modifications. E3 ligases commonly regulate themselves through self-ubiquitination in a

non-proteolytic manner. Self-ubiquitination can either enhance or inhibit activity or can serve to recruit protein partners (de Bie & Ciechanover, 2011). It could be that HLTF is ubiquitinated under normal conditions and is de-ubiquitinated once replication stress is encountered. Self-ubiquitination may also lead HLTF to recruit protein partners that can post-translationally modify and further regulate its activity. It is possible that phosphorylation of HLTF in a similar manner to SMARCAL1 also regulates its fork remodeling and E3 ligase activities. SMARCAL1 phosphorylation on S652, found in between the two ATPase lobes, diminishes its fork remodeling and upregulates its ATPase activity (Couch *et al*, 2013). Although there have not been any studies done to characterize HLTF post-translational modifications, there are many residues within HLTF's motor that have a high probability of being phosphorylated. In addition to its predicted phosphorylation sites, HLTF also has predicted ubiquitination, SUMOylation, and SUMO-interaction sites. Mass-spectrometry to map and characterize HLTF post-translational modifications in conjunction with additional studies that look at the effect of these modifications on HLTF activity both in vitro and in vivo will greatly aid our understanding of HLTF function.

HLTF could have additional functions to promote genomic stability during replication. Perhaps HIRAN's unique binding specificity might be important for recognizing DNA structures outside of a fork. In fact, recent data on Rad5 suggests that the Rad5 HIRAN domain recognizes the 3'OH generated by apurinic/apyrimidic endonucleases in the BER pathway during S phase. The authors of this study propose Rad5 binding to these sites recruit TLS polymerases to bypass these abasic sites (AP sites) to alleviate replication stress (Xu *et al*, 2016). Like its yeast homolog, it is possible

that HLTF also recognizes and binds to specific DNA structures that have an exposed 3'OH group. These structures could be unprocessed intermediates generated by other pathways, such as BER or NER. Recognition of these structures by HLTF could lead to recruitment of DNA repair factors during S phase.
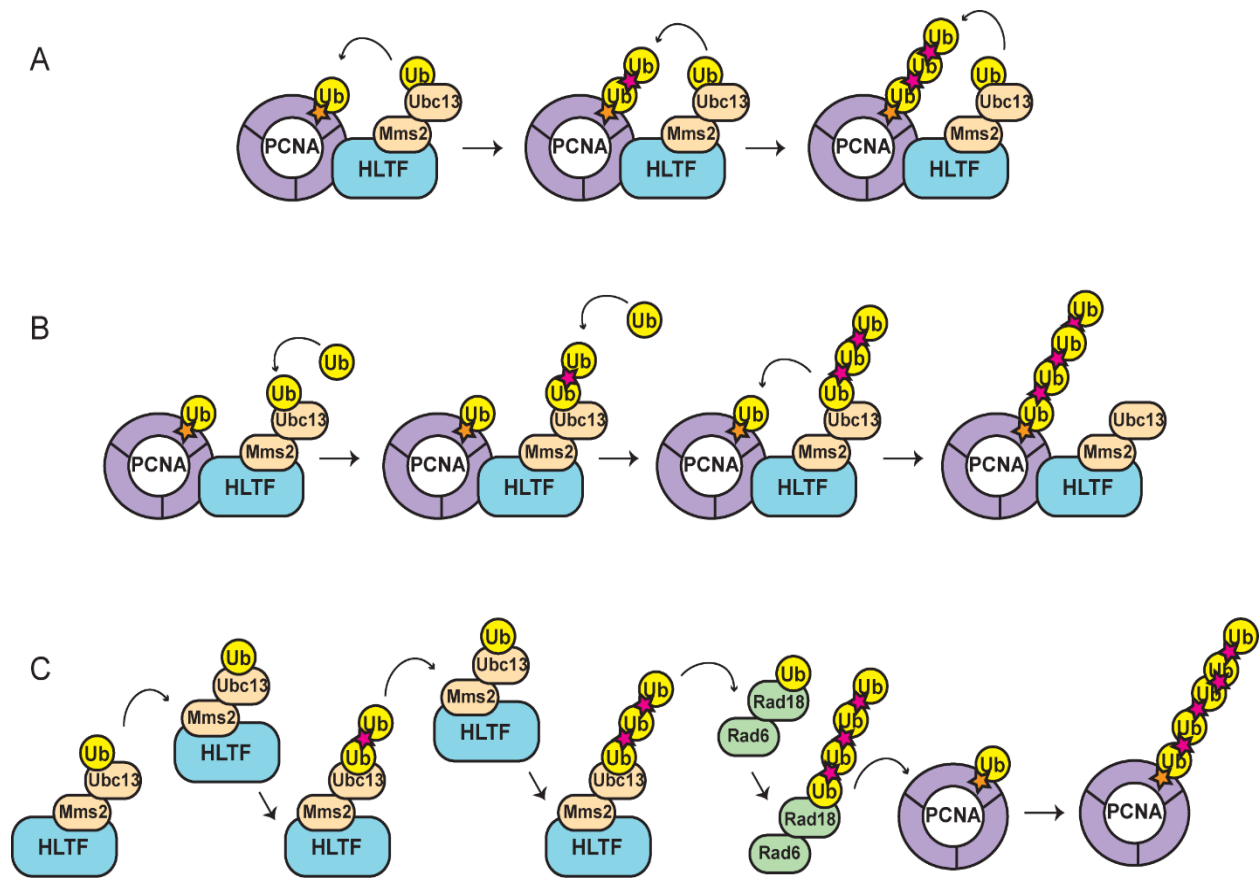
Figure 4.6 HLTF ubiquitination mechanisms. Schematic of PCNA polyubiquitination reactions using a sequential (**A**), *en-bloc* (**B**), or *en-bloc* see-saw (**C**) mechanism. Rad6 and Rad18 are shown to mediate the transfer of the ubiquitin chain to PCNA (**C**) as previously characterized in Masuda *et al*, 2012.

Why so many fork remodelers?

A crucial aspect to understand fork stability mechanisms, is to determine and characterize why cells possess several SNF2 fork remodeling enzymes that catalyze similar reactions in a cell. One clue might be found in the different SRDs that these fork remodeling proteins have. In comparison to other characterized SRDs, HIRAN's 3´-end binding specificity is quite unique. Unlike HIRAN, the HARP and HARP-like SRDs found in SMARCAL1 and ZRANB3 recognize ss-dsDNA junctions (Badu-Nkansah *et al*, 2016; Bétous *et al*, 2012; Mason *et al*, 2014). These differences in specificities, allow us to hypothesize that these fork remodelers recognize different fork structures *in vivo.* In fact, SMARCAL1 has a slight intrinsic preference to regress forks containing a ssDNA gap on the lagging arm (Bétous *et al*, 2013). In contrast, neither HLTF nor ZRANB3 shows a preference for regression of forks containing leading and lagging gaps (Bétous *et al*, 2013; Chavez *et al*, 2018). It could also be that HLTF binds to a stalled fork when there is a frayed 3´-end near the junction, whereas SMARCAL1 and ZRANB3 bind to fork structures that do not have a frayed 3´-end.

In addition to the SRDs, these proteins have additional domains that allow them to interact with specific protein partners. SMARCAL1 has an RPA binding domain that is required for its localization to sites of replication stress (Bansbach *et al*, 2009; Ciccia *et al*, 2009; Postow *et al*, 2009; Yuan *et al*, 2009; Yusufzai *et al*, 2009). In addition, this interaction with RPA regulates SMARCAL1 fork remodeling activity (Bétous *et al*, 2013; Bhat & Cortez, 2018). ZRANB3, has a PCNA interacting protein (PIP)-box and an AlkBH2-PCNA interacting motif (APIM) sequence that allow it to bind to PCNA and a Npl4 zinc finger (NZF) that binds ubiquitin. Together, these domains allow ZRANB3 to bind

polyubiquitinated PCNA and be recruited to sites of replication stress (Ciccia *et al*, 2012; Weston *et al*, 2012; Yuan *et al*, 2012). It is possible that these different protein interactions might bestow different specificities for fork substrates. Some lesions may generate stretches of ssDNA, caused by the uncoupling of the helicase and polymerase that are bound by RPA, and then recruit SMARCAL1 (Cimprich & Cortez, 2008; Poole & Cortez, 2017). Other lesions, such as interstrand cross-links, cannot generate ssDNA, and might lead to the recruitment of other fork remodelers such as HLTF and ZRANB3. It is reasonable to consider that these protein interactions also regulate the timing of recruitment to a fork. RPA binding to stretches of ssDNA is an early response to damage that could recruit SMARCAL1 before HLTF and ZRANB3. Additionally, PCNA post-translational modifications could also regulate protein recruitment to a fork. Since PCNA polyubiquitination is the signal for recruitment of ZRANB3 to a fork, it is possible that other modifications such as SUMOylation, acetylation, or methylation help to recruit and regulate the activity of fork remodelers. Once sites of post-translational modifications are identified, HLTF, SMARCAL1, and ZRANB3 proteins with mutations that inhibit these post-translational modifications could be tested for DNA damage response defects *in vivo*.

Protein interactions at a fork could also regulate the enzymatic function of these proteins. SMARCAL1 fork regression is regulated by RPA. SMARCAL1 activity is stimulated when RPA is bound to the leading template and inhibited when bound to the lagging template (Bétous *et al*, 2013). Interestingly, whereas HLTF regression is modestly inhibited by RPA bound to the lagging template, ZRANB3 is inhibited by RPA

on the leading strand (Chavez *et al*, 2018; Bétous *et al*, 2013) . This data suggests that RPA can regulate the preference of these proteins for a particular type of stalled fork.

It is also possible that these fork remodelers are directed to forks based on their specific functions. ZRANB3 has an HNH endonuclease domain that binds DNA and cleaves different fork-like structures to generate a 3'OH on the leading template strand (Badu-Nkansah *et al*, 2016; Sebesta *et al*, 2017; Weston *et al*, 2012). This activity might be required to stabilize specific fork structures, and as a result, ZRANB3 might be preferentially recruited to these forks. Certain stalled forks might call for HLTF to polyubiquitinate PCNA to promote template switching pathways. Once recruited, it is possible that HLTF catalyzes fork reversal or D-loop formation.

Yet another hypothesis is that these proteins work together to stabilize forks, where the remodeling reaction of one generates the substrate for the next. One can speculate that since ssDNA can be generated as an early response to stress, SMARCAL1 first binds to RPA and remodels a fork to generate a substrate that has a 3'-end in close proximity to the fork junction that would be recognized by HLTF. Once present, HLTF could then generate the polyubiquitinated PCNA that recruits ZRANB3 to a fork.

Concluding remarks

My work has led to a model of how HLTF uses its HIRAN and ATPase motor domains to recognize and catalyze fork regression. Furthermore, utilizing the combined knowledge gained from this work on HLTF and previous work on SMARCAL1 and ZRANB3 fork remodeling mechanisms, we have made predictions for how these enzymes

work *in vivo* to promote genomic stability. Studies to test our hypotheses and to fully understand these fork remodeling proteins will require the development of new methods that can identify different fork structures in a cell, determine the timing of recruitment of these enzymes to a fork, and assay whether these fork remodelers are recruited to the same fork within a cell. Comprehending the mechanisms used by these enzymes will be invaluable for understanding the diseases associated with defects in fork remodeling enzymes and fork repair.

REFERENCES

Achar YJ, Balogh D & Haracska L (2011) Coordinated protein and DNA remodeling by human HLTF on stalled replication fork. *Proc. Natl. Acad. Sci. U.S.A.* **108:** 14073–8

Achar YJ, Balogh D, Neculai D, Juhasz S, Morocz M, Gali H, Dhe-Paganon S, Venclovas Č & Haracska L (2015) Human HLTF mediates postreplication repair by its HIRAN domain-dependent replication fork remodelling. *Nucleic Acids Res.* **43:** 10277–91

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ & Grosse-Kunstleve RW (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66:** 213–221

Aguilera A & García-Muse T (2012) R loops: from transcription byproducts to threats to genome stability. *Mol. Cell* **46:** 115–24

Ahuja A, Jodkowska K, Teloni F, Bizard A, Zellweger R, Herrador R, Ortega S, Hickson I, Altmeyer M, Mendez J & Lopes M (2016) A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. *Nature communications* **7:** 10660

Alabert C, Bukowski-Wills J & Lee S (2014) Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nature Cell Biology* **16:** 281–93

Badu-Nkansah A, Mason AC, Eichman BF & Cortez D (2016) Identification of a Substrate Recognition Domain in the Replication Stress Response Protein Zinc Finger Ran-binding Domain-containing Protein 3 (ZRANB3). *J. Biol. Chem.* **291:** 8251–7

Bansbach C, Bétous R, Lovejoy C, Glick G & Cortez D (2009) The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes & development* **23:** 2405–14

Bartholomew B (2014) ISWI chromatin remodeling: one primary actor or a coordinated effort? *Current opinion in structural biology* **24:**150-155

Berti M & Vindigni A (2016) Replication stress: getting back on track. *Nat. Struct. Mol. Biol.* **23:** 103–9

Bessho Y, Shibata R, Sekine S, Murayama K, Higashijima K, Hori-Takemoto C, Shirouzu M, Kuramitsu S & Yokoyama S (2007) Structural basis for functional mimicry of long-variable-arm tRNA by transfer-messenger RNA. *Proceedings of the National Academy of Sciences* **104:** 8293–8298

Bétous R, Couch FB, Mason AC, Eichman BF, Manosas M & Cortez D (2013) Substrate-selective repair and restart of replication forks by DNA translocases. *Cell Rep* **3:** 1958–69

Bétous R, Mason A, Rambo R, Bansbach C, Badu-Nkansah A, Sirbu B, Eichman B & Cortez D (2012) SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication. *Genes & development* **26:** 151–62

Bhat K, Bétous R & Cortez D (2015) High-affinity DNA-binding Domains of Replication Protein A (RPA) Direct SMARCAL1-dependent Replication Fork Remodeling. *J Biol Chem* **290:** 4110–4117

Blastyák A, Hajdú I, Unk I & Haracska L (2010) Role of Double-Stranded DNA Translocase Activity of Human HLTF in Replication of Damaged DNA. *Mol Cell Biol* **30:** 684–693

Bochman ML, Paeschke K & Zakian VA (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* **13:** 770–80

Branzei D & Foiani M (2010) Maintaining genome stability at the replication fork. *Nature Reviews Molecular Cell Biology* **11:** 208–219

Branzei D & Psakhye I (2016) DNA damage tolerance. *Curr. Opin. Cell Biol.* **40:** 137–144

Brooks PJ & Theruvathu JA (2005) DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. *Alcohol* **35:** 187–93

Burgers P & Kunkel T (2017) Eukaryotic DNA Replication Fork. *Annual review of biochemistry* **86:** 417–438

Burkovics P, Sebesta M, Balogh D, Haracska L & Krejci L (2014) Strand invasion by HLTF as a mechanism for template switch in fork rescue. *Nucleic Acids Res.* **42:** 1711–20

Byun T, Pacek M, Yee M, Walter J & Cimprich K (2005) Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes & Development* **19:** 1040–1052

Cadet J & Wagner JR (2013) DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harb Perspect Biol* **5:**

Chang D & Cimprich K (2009) DNA damage tolerance: when it's OK to make mistakes. *Nature chemical biology* **5:** 82–90

Chaudhuri A, Hashimoto Y, Herrador R, Neelsen K, Fachinetti D, Bermejo R, Cocito A, Costanzo V & Lopes M (2012) Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nature structural & molecular biology* **19:** 417–23

Chaudhuri A, Callen E, Ding X, Gogola E, Duarte A, Lee J-E, Wong N, Lafarga V, Calvo J, Panzarino N, John S, Day A, Crespo A, Shen B, Starnes L, de Ruiter J, Daniel J, Konstantinopoulos P, Cortez D, Cantor S, et al (2016) Replication fork stability confers chemoresistance in BRCA-deficient cells. *Nature* **535:** 382–7

Chavez DA, Greer BH & Eichman BF (2018) The HIRAN domain of helicase-like transcription factor positions the DNA translocase motor to drive efficient DNA fork regression. *J. Biol. Chem.* **22:** 8484-8494

Choe K & Moldovan G-L (2017) Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Molecular cell* **65:** 380–392

Choi K, Batke S, Szakal B, Lowther J, Hao F, Sarangi P, Branzei D, Ulrich H & Zhao X (2015) Concerted and differential actions of two enzymatic domains underlie Rad5 contributions to DNA damage tolerance. *Nucleic acids research* **43:** 2666–77

Ciccia A, Bredemeyer AL, Sowa ME, Terret M-EE, Jallepalli PV, Harper JW & Elledge SJ (2009) The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. *Genes Dev.* **23:** 2415–25

Ciccia A & Elledge SJ (2010) The DNA damage response: making it safe to play with knives. *Mol. Cell* **40:** 179–204

Ciccia A, Nimonkar AV, Hu Y, Hajdu I, Achar YJ, Izhar L, Petit SA, Adamson B, Yoon JC, Kowalczykowski SC, Livingston DM, Haracska L & Elledge SJ (2012) Polyubiquitinated PCNA recruits the ZRANB3 translocase to maintain genomic integrity after replication stress. *Mol. Cell* **47:** 396–409

Cimprich K & Cortez D (2008) ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Bio* **9:** 616–627

Cortez D (2015) Preventing replication fork collapse to maintain genome integrity. *DNA Repair (Amst.)* **32:** 149–57

Couch FB, Bansbach CE, Driscoll R, Luzwick JW, Glick GG, Bétous R, Carroll CM, Jung SY, Qin J, Cimprich KA & Cortez D (2013) ATR phosphorylates SMARCAL1 to prevent replication fork collapse. *Genes Dev.* **27:** 1610–23

De Bie P & Ciechanover A (2011) Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. *Cell Death Differ.* **18:** 1393–402

De Septenville AL, Duigou S, Boubakri H & Michel B (2012) Replication fork reversal after replication-transcription collision. *PLoS Genet.* **8:** e1002622

Dhont L, Mascaux C & Belayew A (2016) The helicase-like transcription factor (HLTF) in cancer: loss of function or oncomorphic conversion of a tumor suppressor? *Cell. Mol. Life Sci.* **73:** 129–47

Ding L & Forsburg SL (2014) Essential domains of Schizosaccharomyces pombe Rad8 required for DNA damage response. *G3 (Bethesda)* **4:** 1373–84

Dong G, Nowakowski J & journal H-D (2002) Structure of small protein B: the protein component of the tmRNA–SmpB system for ribosome rescue. *The EMBO journal* **7:** 1845-1854

Dürr H, Flaus A, Owen-Hughes T & Hopfner K-PP (2006) Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. *Nucleic Acids Res.* **34:** 4160–7

Farnung L, Vos S, Wigge C & Cramer P (2017) Nucleosome–Chd1 structure and implications for chromatin remodelling. *Nature* **550:** 539

Fernandez-Leiro R, Conrad J, Yang J-CC, Freund SM, Scheres SH & Lamers MH (2017) Self-correcting mismatches during high-fidelity DNA replication. *Nat. Struct. Mol. Biol.* **24:** 140–143

Flaus A, Martin D, Barton G & Owen-Hughes T (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic acids research* **34:** 2887–905

Flaus A & Owen-Hughes T (2011) Mechanisms for ATP-dependent chromatin remodelling: the means to the end. *The FEBS journal* **278:** 3579–95

Follonier C, Oehler J, Herrador R & Lopes M (2013) Friedreich's ataxia–associated GAA repeats induce replication-fork reversal and unusual molecular junctions. *Nature Structural and Molecular Biology* **20:** nsmb.2520

Friedberg E (2005) Suffering in silence: the tolerance of DNA damage. *Nat Rev Mol Cell Bio* **6:** 943–953

Gabbai C & Marians K (2010) Recruitment to stalled replication forks of the PriA DNA helicase and replisome-loading activities is essential for survival. *DNA repair* **9:** 202–209

Gangavarapu V, Haracska L, Unk I, Johnson R, Prakash S & Prakash L (2006) Mms2-Ubc13-dependent and -independent roles of Rad5 ubiquitin ligase in postreplication repair and translesion DNA synthesis in Saccharomyces cerevisiae. *Molecular and cellular biology* **26:** 7783–90

Hauk G, McKnight JN, Nodelman IM & Bowman GD (2010) The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Molecular cell* **39:** 711–723

Hayward-Lester A, Hewetson A, Beale EG, Oefner PJ, Doris PA & Chilton BS (1996) Cloning, characterization, and steroid-dependent posttranscriptional processing of RUSH-1 alpha and beta, two uteroglobin promoter-binding proteins. *Mol. Endocrinol.* **10:** 1335–49

Helmer RA, Martínez-Zaguilán R, Dertien JS, Fulford C, Foreman O, Peiris V & Chilton BS (2013) Helicase-like transcription factor (Hltf) regulates G2/M transition, Wt1/Gata4/Hif-1a cardiac transcription networks, and collagen biogenesis. *PLoS ONE* **8:** e80461

Higgins, Kato & Strauss (1976) A model for replication repair in mammalian cells. *Journal of molecular biology* **101:** 417–25

Hishiki A, Hara K, Ikegaya Y, Yokoyama H, Shimizu T, Sato M & Hashimoto H (2015) Structure of a Novel DNA-binding Domain of Helicase-like Transcription Factor (HLTF) and Its Functional Implication in DNA Damage Tolerance. *J. Biol. Chem.* **290:** 13215–23

Hochstrasser M (2006) Lingering mysteries of ubiquitin-chain assembly. *Cell* **124:** 27–34

Holm L & Sander C (1993) Protein structure comparison by alignment of distance matrices. *Journal of molecular biology* **233:** 123–138

Hopfner K-P, Gerhold C-B, Lakomek K & Wollmann P (2012) Swi2/Snf2 remodelers: hybrid views on hybrid molecular machines. *Current Opinion in Structural Biology* **22:** 225–233

Hotchkiss RD (1974) Models of genetic recombination. *Annu. Rev. Microbiol.* **28:** 445–68

Iyer L, Babu M & Aravind (2006) The HIRAN domain and recruitment of chromatin remodeling and repair activities to damaged DNA. *Cell Cycle Georget Tex* **5:** 775–82

Jackson D & Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S …. *The Journal of Cell Biology* **140:** 1285–1295

Jones J & Nakai N- (1999) Duplex opening by primosome protein PriA for replisome assembly on a recombination intermediate1. *Journal of molecular biology* **289:** 503–516

Kanao R & Masutani C (2017) Regulation of DNA damage tolerance in mammalian cells by post-translational modifications of PCNA. *Mutation research*

Kile AC, Chavez DA, Bacal J, Eldirany S, Korzhnev DM, Bezsonova I, Eichman BF & Cimprich KA (2015) HLTF's Ancient HIRAN Domain Binds 3' DNA Ends to Drive Replication Fork Reversal. *Mol. Cell* **58:** 1090–100

Kobbe D, Kahles A, Walter M, Klemm T, Mannuss A, Knoll A, Focke M & Puchta H (2016) AtRAD5A is a DNA translocase harboring a HIRAN domain which confers binding to branched DNA structures and is required for DNA repair in vivo. *Plant J.* **88:** 521–530

Krijger P, Lee K-Y, Wit N, van den Berk P, Wu X, Roest H, Maas A, Ding H, Hoeijmakers J, Myung K & Jacobs H (2011b) HLTF and SHPRH are not essential for PCNA polyubiquitination, survival and somatic hypermutation: Existence of an alternative E3 ligase. *Dna Repair* **10:** 438–444

Lin J-RR, Zeman MK, Chen J-YY, Yee M-CC & Cimprich KA (2011) SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis. *Mol. Cell* **42:** 237–49

Liu X, Li M, Xia X, Li X & Chen Z (2017) Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544:** 440

Manosas M, Perumal S & Croquette V (2012) Direct observation of stalled fork restart via fork regression in the T4 replication system. *Science* **338:** 1217–1220

Manosas M, Perumal S, Bianco P, Bianco P, Ritort F, Benkovic S & Croquette V (2013) RecG and UvsW catalyse robust DNA rewinding critical for stalled DNA replication fork rescue. *Nat Commun* **4:** 2368

Mason A, Rambo R, Greer B, Pritchett M, Tainer J, Cortez D & Eichman B (2014) A structure-specific nucleic acid-binding domain conserved among DNA repair proteins. *Proc National Acad Sci* **111:** 7618–7623

Masuda Y, Suzuki M, Kawai H, Hishiki A, Hashimoto H, Masutani C, Hishida T, Suzuki F & Kamiya K (2012) En bloc transfer of polyubiquitin chains to PCNA in vitro is mediated by two different human E2-E3 pairs. *Nucleic Acids Res.* **40:** 10394–407

McMurray C (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet* **11:** 786–799

Metzger MB, Pruneda JN, Klevit RE & Weissman AM (2014) RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim. Biophys. Acta* **1843:** 47–60

Minca E & Kowalski D (2010) Multiple Rad5 activities mediate sister chromatid recombination to bypass DNA damage at stalled replication forks. *Molecular cell* **38:** 649–61

Motegi A, Liaw H-J, Lee K-Y, Roest H, Maas A, Wu X, Moinova H, Markowitz S, Ding H, Hoeijmakers J & 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* **105:** 12411–12416

Neelsen K, Zanini I, Herrador R & Lopes M (2013) Oncogenes induce genotoxic stress by mitotic processing of unusual replication intermediates. *J Cell Biol* **200:** 699–708

Neelsen K & Lopes M (2015) Replication fork reversal in eukaryotes: from dead end to dynamic response. *Nature Reviews Molecular Cell Biology* **16:** 207–220

Nelson S & Benkovic S (2010) Response of the Bacteriophage T4 Replisome to Noncoding Lesions and Regression of a Stalled Replication Fork. *Journal of Molecular Biology* **401:** 743–756

Ortiz-Bazán MÁÁ, Gallo-Fernández M, Saugar I, Jiménez-Martín A, Vázquez MVV & Tercero JAA (2014) Rad5 plays a major role in the cellular response to DNA damage during chromosome replication. *Cell Rep* **9:** 460–8

Pierce B & Wendland B (2009) Sequence of the yeast protein expression plasmid pEG (KT). *Yeast* **26:** 349–353

Poole LA & Cortez D (2017) Functions of SMARCAL1, ZRANB3, and HLTF in maintaining genome stability. *Crit. Rev. Biochem. Mol. Biol.* **52:** 696–714

Postow L, Woo E, Chait B & Funabiki H (2009) Identification of SMARCAL1 as a component of the DNA damage response. *The Journal of biological chemistry* **284:** 35951–61

Potenski CJ & Klein HL (2014) How the misincorporation of ribonucleotides into genomic DNA can be both harmful and helpful to cells. *Nucleic Acids Res.* **42:** 10226–34

Ran F, Hsu P, Wright J, Agarwala V & Scott D (2013) Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8:** 2281–2308

Saugar I, Ortiz-Bazán M & Tercero J (2014) Tolerating DNA damage during eukaryotic chromosome replication. *Experimental cell research* **329:** 170–7

Sebesta M, Cooper CDOD, Ariza A, Carnie CJ & Ahel D (2017) Structural insights into the function of ZRANB3 in replication stress response. *Nat Commun* **8:** 15847

Sheridan PL, Schorpp M, Voz ML & Jones KA (1995) Cloning of an SNF2/SWI2-related protein that binds specifically to the SPH motifs of the SV40 enhancer and to the HIV-1 promoter. *J. Biol. Chem.* **270:** 4575–87

Shin S, Hyun K, Kim J & reports H-S (2018) ATP Binding to Rad5 Initiates Replication Fork Reversal by Inducing the Unwinding of the Leading Arm and the Formation of the Holliday Junction. *Cell reports* **6:** 1831-1839

Singleton MR, Dillingham MS & Wigley DB (2007b) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* **76:** 23–50

Stingele J, Bellelli R & Boulton S (2017) Mechanisms of DNA-protein crosslink repair. *Nature reviews. Molecular cell biology* **18:** 563–573

Theobald DL, Mitton-Fry RM & Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32:** 115–33

Tsutakawa SE, Classen S, Chapados BR, Arvai AS, Finger LD, Guenther G, Tomlinson CG, Thompson P, Sarker AH, Shen B, Cooper PK, Grasby JA & Tainer JA (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145:** 198–211

Unk I, Hajdú I, Fátyol K, Hurwitz J, Yoon J-H, Prakash L, Prakash S & Haracska L (2008) Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination. *Proc National Acad Sci* **105:** 3768–3773

Weston R, Peeters H & Ahel D (2012) ZRANB3 is a structure-specific ATP-dependent endonuclease involved in replication stress response. *Genes Dev.* **26:** 1558–72

Xu X, Lin A, Zhou C, Blackwell S, Zhang Y, Wang Z, Feng Q, Guan R, Hanna M, Chen Z & Xiao W (2016) Involvement of budding yeast Rad5 in translesion DNA synthesis through physical interaction with Rev1. *Nucleic Acids Research* **44:** 5231–5245

Yuan J, Ghosal G & Chen J (2009) The annealing helicase HARP protects stalled replication forks. *Genes & development* **23:** 2394–9

Yuan J, Ghosal G & Chen J (2012) The HARP-like domain-containing protein AH2/ZRANB3 binds to PCNA and participates in cellular response to replication stress. *Mol. Cell* **47:** 410–21

Yusufzai T & Kadonaga J (2008) HARP Is an ATP-Driven Annealing Helicase. *Science* **322:** 748–750

Yusufzai T & Kadonaga JT (2010) Annealing helicase 2 (AH2), a DNA-rewinding motor with an HNH motif. *Proc. Natl. Acad. Sci. U.S.A.* **107:** 20970–3

Yusufzai T, Kong X, Yokomori K & Kadonaga J (2009) The annealing helicase HARP is recruited to DNA repair sites via an interaction with RPA. *Genes & development* **23:** 2400–4

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

Zeman M & Cimprich K (2014) Causes and consequences of replication stress. *Nature Cell Biology* **16:** 2–9