

Summer 8-15-2017

DNA Replication Challenges: Telomeres and R loops

Shankar Parajuli

Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds



Part of the [Biochemistry Commons](#)

Recommended Citation

Parajuli, Shankar, "DNA Replication Challenges: Telomeres and R loops" (2017). *Arts & Sciences Electronic Theses and Dissertations*. 1209.

https://openscholarship.wustl.edu/art_sci_etds/1209

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS
Division of Biology and Biomedical Sciences
Biochemistry

Dissertation Examination Committee:
Sheila A. Stewart, Chair
Peter M. Burgers
Nima Mosammaparast
Alessandro Vindigni
Zhongsheng You

DNA Replication Challenges: Telomeres and R loops

by
Shankar Parajuli

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2017
St. Louis, Missouri

© 2017, Shankar Parajuli

Table of Contents

List of figures.....	iv
Acknowledgment.....	v
Abstract.....	viii
Chapter 1: Introduction to DNA replication, repair, and challenges.....	1
Introduction.....	2
1. Eukaryotic DNA replication and DNA damage repair.....	4
1.1 DNA replication.....	4
1.2 DNA repair.....	9
2. Telomere biology.....	14
2.1 Telomere structure and proteins.....	14
2.2 Telomere replication and maintenance.....	17
2.3 Telomere physiology.....	22
3. R loops: Structure, proteins, and physiology.....	27
3.1 Structure.....	27
3.2 R loop proteins.....	29
3.3 R loop physiology.....	30
References.....	35
Chapter 2: Flap endonuclease 1 limits telomere fragility on the leading strand.....	50
Introduction.....	51
Experimental Procedures.....	55
Results.....	63
FEN1 depletion and transcription inhibition induce replication stress and a DNA.....	63
Inhibition of transcription exacerbates the telomere fragility observed upon FEN1 depletion.....	66
Leading strand-specific telomere fragility is caused by RNA:DNA hybrids.....	67
FEN1 flap endonuclease activity is required for limiting telomere fragility.....	70
Discussion.....	72
Acknowledgments.....	76
References.....	89
Chapter 3: Human Ribonuclease H1 resolves R loops and thereby enables progression of the DNA replication fork.....	97
Introduction.....	98
Experimental Procedures.....	100
Results.....	106

RNH1 contributes to genome stability and preserves telomere integrity	106
Nuclear RNA:DNA hybrid levels increase upon RNH1 depletion	107
RNH1 depletion results in replication fork slowing and increased termination and stalling	108
Nuclease activity of RNH1 is required for efficient replication fork movement	110
Discussion.....	112
Acknowledgments.....	114
References.....	123
Chapter 4: Conclusions and future directions.....	128
Conclusions.....	129
FEN1 limits telomere fragility on the leading strand	129
Human RNH1 limits R loops and facilitates DNA replication	130
Future Directions.....	131
Structure and physiology of telomere fragility	131
Regulation of RNH1 during DNA replication	134
References.....	136

List of figures

Chapter 2: Flap endonuclease (FEN1) limits telomere fragility on the leading strand

Figure 2.1: α -amanitin treatment abrogates expression of mRNAs with short half lives but does not alter steady-state TERRA levels	77
Figure 2.2: FEN1 depletion and transcription inhibition induce replication stress, a DNA damage response, and telomere fragility	80
Figure 2.3: RNA:DNA hybrids are responsible for FEN1 depletion-induced leading strand-specific telomere fragility	81
Figure 2.4: RNA:DNA hybrids are responsible for α -amanitin-induced telomere fragility	83
Figure 2.5: FEN1 flap endonuclease activity is required to limit leading strand-specific telomere fragility	86
Figure 2.6: A model of FEN1's role following co-directional replisome–RNAP collisions	87

Chapter 3: Human ribonuclease H1 limits R loops and facilitates DNA replication

Figure 3.1. RNH1 contributes to genome stability and preserves telomere integrity	115
Figure 3.2. Nuclear RNA:DNA hybrid levels increase upon RNH1 depletion	117
Figure 3.3. RNH1 depletion results in replication fork slowing and increased termination and stalling	119
Figure 3.4. Nuclease activity of RNH1 is required for efficient replication fork movement	121

Acknowledgments

First and foremost, I thank my mentor, Dr. Sheila Stewart, for her guidance and unwavering support throughout my Ph.D. journey. Since joining her laboratory in January 2012, she has served as a valuable mentor and helped me become the scientist that I am today. I always felt welcomed by her open-door policy and that encouraged me to stop by her office to ask any question without having to schedule a meeting. She was always available when I needed to consult or verify my reasoning as I planned my experiments. In the meantime, she granted me the freedom and resources to test my ideas and to become independent. I very much appreciated her balanced approach of mentorship that has helped me grow and develop both as a scientist and as an adult. One of the most valuable lessons that Sheila taught was to work hard and learn from my mistakes and failures; I will treasure this lesson and continue to apply it in my personal and professional life. I will forever be thankful for her training.

I thank the members of my thesis committee- Dr. Peter M. Burgers, Dr. Nima Mosammaparast, Dr. Alessandro Vindigni and Dr. Zhongsheng You. Each brought a unique perspective to my individual and group meetings and pushed me to think critically and ask tough questions. I offer special thanks to the members of Dr. Vindigni's group at Saint Louis University including Jessica Jackson and Dr. Matteredo Berti for both teaching and sharing their reagents.

I am also grateful to the two WashU departments – Department of Biochemistry and Molecular Biophysics and the Department of Cell biology and Physiology –for providing a

resourceful environment to carry out my Ph.D. work. I am thankful to the biochemistry program coordinators, Melissa Torres, Andrea Krussel, and Stacia Burd, who worked tirelessly to ensure that I met all the Ph.D. deadlines and filed my paperwork on time. I am also thankful to the ICCE institute and all its members for providing both a collegial and a fun environment to work.

My time in graduate school would not have been as rewarding without the support of the members of the Stewart laboratory. I will forever be grateful to all the lab members, both past and present. Megan Ruhland, Elise Alspach, and Hayley Moore were thoughtful scientists and I learned a great deal from interacting with them on a daily basis. I am especially thankful to Daniel Teasley; he has been a tremendous teacher from my first day the laboratory. Even after his graduation, Daniel continues to advise me in my post Ph.D. career plan. I also want to recognize and appreciate Kevin Flanagan who became a close friend in and outside of the laboratory. I have also found friendship in Bhavna Murali, Yujie Fu, Xianmin Luo, and Qiaho Ren; they have been responsible for creating a caring lab environment.

I want to thank all my friends who stood by me and supported as I went through ups and downs of graduate school. In St. Louis, I was graced with the company of outstanding individuals including Biva Rajbhandari, Sharad Paudyal, Joo Young Park, Katherine Mann, Catherine Kuzmicki, Shannon Ohlemacher, John Jimah, Hirak Biswas, Samarth Hegde, Andrew Chang, Nick Dietrich, and Melissa Li. Each made my living in Saint Louis enjoyable and full of good memories. I also relished the continued support of my old

friends— Pravin Paudel, Prabin Chandra Subedi, Rashik Adhikari, Ajaya Sharma, Jagat Adhikari, Rajat Thapa, Dipendra Rokaha, Pradeep Subedi, and Amrit Godar. These friends have provided unwavering support throughout my life's successes and challenges.

Lastly, I want to thank my immediate family in Nepal for their unconditional love, support, and care without which I would not be where I am today. I thank my mom (mommy) and dad (baba) for instilling good values on me from an early age and for always believing in me. Their sacrifice is the only reason that I was able to fly to the other side of the world to obtain the best education. Their teaching that nobody is perfect and that there is always room to be a better person is what keeps me grounded and motivated every day. Therefore, I owe all my success to them. I am also forever grateful to my two loving siblings, my older brother and younger sister. They are my closest friends, biggest admirers, and greatest critics. Even from afar, they are a constant source of my joy and happiness.

I am also grateful to the Predoctoral Cancer Biology Pathway Grant from the Siteman Cancer Center and Barnes Jewish Hospital in Saint Louis for funding portions of my Ph.D. work.

Shankar Parajuli

Washington University in St. Louis

August 2017

ABSTRACT OF THE DISSERTATION

DNA Replication Challenges: Telomeres and R loops

by

Shankar Parajuli

Doctor of Philosophy in Biology and Biomedical Sciences

Biochemistry

Washington University in St. Louis, 2017

Professor Sheila A. Stewart, Chair

Faithful DNA replication and repair are essential for maintaining genome stability and preventing various diseases including cancer. Both processes are executed by numerous redundant mechanisms to ensure that these processes are uninterrupted even when a primary mechanism fails. Despite this, they are not immune to challenges and failures leading to DNA damage and genome instability. These problems are more evident at the difficult-to-replicate regions of the genome such as the telomeres that cap and protect linear chromosome ends. Additionally, topological structures such as RNA:DNA hybrids, commonly referred to as R loops, can also present severe challenges to the DNA replication and repair machineries. Herein we report the functions of two distinct DNA replication and repair proteins—flap endonuclease 1 (FEN1) and ribonuclease H1 (RNH1)—that preserve genome stability. First, we show that FEN1 limits telomere fragility in leading strand replicated telomeres. This is mediated by its flap endonuclease activity

independent of its gap endonuclease activity and C-terminal interactions. We show that the fragility phenotype is increased by RNA polymerase II inhibition and rescued by ectopic RNH1 expression. Because the telomere is transcribed and can form hybrids, these data suggest that the FEN1-mediated telomere fragility depends on RNA:DNA hybrids that accumulate from co-directional replisome-RNAP collision at the leading strand replicated telomere. These findings are the first to assign a leading strand specific function of FEN1, which is a canonical lagging strand protein. Second, we uncover a novel role for human RNH1 in DNA replication in the nucleus. We show that RNH1 depletion results in a global DNA damage response as well as telomere loss phenotype. Because RNH1 resolves RNA:DNA hybrids, we measured those hybrid levels and found that they increase upon RNH1 depletion. Given these hybrids could pose barriers to a moving replication machinery, we interrogated replication efficiency and discovered that RNH1 facilitates the replication fork movement, possibly by clearing hybrids. These data shed light onto the role of RNH1 in global DNA replication. Together, our work underscore the complexity of DNA replication and repair processes and highlight the varied roles that FEN1 and RNH1 play to maintain genome stability.

Chapter 1:

Introduction to DNA replication, repair, and challenges

Shankar Parajuli and Sheila A. Stewart

Introduction

Faithful DNA replication is crucial for maintaining genome stability. This is ensured by numerous complex and redundant mechanisms aimed at fulfilling various functions during replication. Redundancy is required to compensate for a function even when a primary mechanism fails to carry out its designated role. Despite this, replication is not immune to challenges and failure. For example, variations in genome landscape and architecture such as regions of high GC content, repetitive sequences, or secondary structures produce difficult-to-replicate templates increasing the susceptibility to a replication dysfunction (1), (2). Similarly, DNA replication is also affected by genotoxic and cytotoxic agents in various forms from use of radiation and chemotherapy to exposure to environmental insults and food habits. These dysfunctions in DNA replication- either partial or complete - can lead to genome instability. Such problems are evident in many diseases including cancer, neurological disorders, and aging syndromes. In fact, these are the underlying causes of these diseases in many cases. Therefore, understanding the molecular mechanism of DNA replication, its dysfunction, and its repair is vital to developing therapies and advancing cures to multiple ailments.

The structure of the genome yields valuable insights into DNA replication process. The genome in higher eukaryotes is packaged into linear chromosomes to maximize both the organizational and functional efficiency. Proteins of the histone family play an important role in this endeavor (3). Furthermore, the linear chromosomes are capped at their termini by nucleoprotein structures known as telomeres. Telomeres serve multiple functions to maintain genome stability. First, telomeres prevent chromosome ends from being

recognized as double strand breaks. Second, they minimize the physical loss of DNA from end-replication problem that arises from the inability of the lagging strand machinery to completely replicate the terminal end of a linear chromosome. Third, telomeres regulate access to replication and repair proteins. Despite a critical role that telomeres play to protect chromosomal integrity, they pose a significant challenge to the replication machinery. This is largely due to the G- rich repetitive nature of the sequences that can form G-quadruplex secondary structures (4). This problem is exacerbated by the absence of a compensating origin-firing from the opposite direction in the event of a centromere-distal fork collapse leading to a loss of telomere (5). Therefore, several DNA replication and repair proteins along with the telomere specific proteins play an important role in minimizing these problems associated with the innately difficult-to-replicate telomere template.

In addition to a sequence-specific challenges at the telomere, replication machinery also encounters several topological obstacles throughout the genome. One such challenge is RNA: DNA hybrids, commonly referred to as R-loops. An R-loop is formed when a nascent RNA emanating from an RNA polymerase hybridizes with a template DNA thereby producing a three-strand structure consisting of an RNA bound to a template DNA and a flanking non-template DNA (6). These structures are formed both at the telomeres as well as at other regions across the genome at a frequency much higher than previously appreciated. In recent years, these structures have received renewed interest due to the diverse and at times opposing cellular roles that they exhibit. While R loops play important physiological roles in diverse functions including during class switch recombination (CSR)

and RNA primer generation, they also pose significant threats to genome integrity (7), (8). Several studies have shown that RNA: DNA hybrids increase the rate of DNA mutation, DNA recombination, and impairment of replication and transcription (9). Therefore, cells have developed multiple mechanisms to either prevent or resolve such structures. Together, the findings that replication machinery faces both a difficult to replicate template such as a telomere as well as a topological challenge such as an R loop underscore the need for further investigation into replication process and its efficiency under various cellular stress to preserve genome stability.

1. Eukaryotic DNA Replication and DNA Damage Repair

1.1 DNA Replication

DNA replication is a fundamental mechanism by which the genome is duplicated during cell division. The seminal discovery by Watson and Crick in 1953 revealed that the DNA structure was a double helix with two strands running antiparallel to each other (10). These findings laid the groundwork for various molecular insights into how DNA replication occurs. A few years later, an experiment performed by Meselson- Stahl proved the semi conservative nature of replication whereby a parent DNA strand acts as a template to produce a daughter strand resulting in a double strand daughter DNA composed of one old (parent) and one new (daughter) strand (11). In the last 60 years since, advancement in genetic tools, *in-vitro* biochemistry, and sequencing has identified numerous proteins and illuminated a detailed mechanism of DNA replication. Furthermore, these tools combined with a high-resolution microscopy have offered a

spatial and temporal insight into DNA replication. In addition to the detailed understanding of this process, these scientific advancements have allowed discovery and development of therapies to combat various diseases.

The bulk of the DNA replication occurs during S-phase of the cell cycle and is carried out in 3 distinct phases: initiation, elongation, and termination. Each of these steps is a concerted effort of many proteins and is tightly regulated.

Initiation

Initiation of DNA replication begins in G1 and ends in early S-phase of the cell cycle. The first step of this process is the demarcation of the sites of replication initiation, commonly referred to as origins of replication. This is initiated by the binding of several proteins collectively referred to as Origin Recognition Complex (ORC). ORC binding serves as a platform for recruitment and assembly of Cdc6 (helicase loader), Cdt1 (chaperone), and the MCM2-7 (replicative helicase) complex, collectively referred to as pre-replication complex (pre-RC) (1), (12). This facilitates the loading of the helicase activators Cdc45 and GINS to form CMG complex followed by the formation of pre-initiation complex, initiation of origin DNA unwinding by Mcm10, and subsequent activation of bidirectional replication forks. In lower eukaryotes such as budding yeast, ORC displays a degree of sequence specificity but in metazoans such as humans, binding is promiscuous (13), (14). Because DNA replication can only take place once per cell division, the replication machinery is under immense pressure to complete replication of the entire genome. This necessitates many origins to prime replication. As a result, for example, a budding yeast

genome of about 12 mega base pairs contains about 400 origins whereas a human genome of about 3 giga base pairs contains between 30,000 and 50,000 origins to replicate a vast amount of DNA (15).

Elongation

Elongation of DNA replication takes place in S phase of the cell cycle and involves the addition of deoxyribonucleotides to produce two daughter DNA strands. This process involves a careful co-ordination between replication initiation proteins and replisome complex consisting of core proteins including 11- subunit CMG helicase, topoisomerase, helicase, primase, DNA polymerase α -primase, leading strand polymerase ϵ , lagging strand polymerase δ , RFC clamp loader, Proliferating Cell Nuclear Antigen (PCNA) clamp, and single strand binding protein RPA (16), (17). Cdc45 plays an important role at the interface and transition of replication initiation into elongation. As part of the helicase activating complex, CMG, Cdc45 physically interacts with DNA template and is expected to play a role in recruitment of the DNA polymerases (18). The high resolution cryo EM studies of DNA-bound fly CMG and yeast CMG reveal Cdc45 complex structure is similar to that of Cdc45 crystal structure indicating that the groove inside the CMG complex is normally blocked and therefore would require further interactions for its opening and fork movement (19), (20). In fact, in vitro studies have shown that CMG directly interacts with both leading strand polymerase ϵ and lagging strand polymerase δ to maximize their replication efficiency (21). These findings also highlight an intricate interaction that must occur with diverse proteins to initiate elongation stage of DNA synthesis. A simpler model of replication exists in bacteria whereby a duplex DNA is unwound by DNA helicase at

the leading edge of the replication fork, with replicative DNA polymerase trailing (22). In eukaryotes, however, this process is complex with a leading strand polymerase anchored directly on the front of the helicase while the lagging strand polymerase and primase complex is linked at the back of the helicase (23).

Each DNA double helix consists of two strands- Leading and lagging- that run antiparallel to each other. Their replications differ on two key fronts. First, because the synthesis of a new DNA strand can only occur in a 5' to 3' direction, DNA replication of a leading strand (3'-5' strand) occurs in the direction of replication fork movement and is continuous whereas that of a lagging strand (5'-3' strand) occurs in the opposite direction of the moving fork and is therefore, discontinuous (24). As a result, lagging strand replication produces DNA fragments (180 and 200 bp in eukaryotes) called Okazaki fragments that are subsequently processed and joined together by several enzymes including DNA ligase and Flap endonuclease 1 (FEN1) to generate a new strand. Another significant difference between the replication of leading and lagging strand involves the utilization of different DNA polymerases. DNA polymerase ϵ (Pol ϵ) is responsible for bulk of the leading strand replication whereas DNA polymerases α (Pol α) and δ (Pol δ) are responsible for initiation of lagging strand replication and its elongation and maturation respectively.

DNA replication on both strands is initiated by a 7-10 nucleotide long RNA primer that is synthesized by DNA polymerase α - primase. Pol α extends these by about 15 deoxynucleotides and hands the replication task over to Pol ϵ and Pol δ for replication of

leading and lagging strands respectively (25). Replication fidelity is key to the massive task of completing replication of a vast genome. All eukaryotic DNA polymerases can add between 1000 to 2000 nucleotides of DNA per minute. However, pol ϵ and Pol δ have a much higher fidelity (1 error per 10^6 - 10^7 nucleotide synthesis) compared to Pol α (1 error per 10^4 - 10^5 nucleotide synthesis). It is for this reason that the bulk of the replication is carried out by Pol ϵ and Pol δ after initiation by Pol α . The main reason for a lower fidelity of Pol α is because it lacks the proofreading exonuclease activity that the other two polymerases ϵ and δ possess. DNA replication efficiency of these polymerases is also enhanced by a clamp loader protein PCNA by increasing their processivities (26). Replication is tightly regulated to ensure that replication takes place only once during a cell division. This is ensured by several cell cycle regulatory proteins such as Cdks to prevent secondary loading at the origin of replication sites.

Termination

DNA replication is terminated by a concerted effort of several proteins. It involves completion of DNA synthesis, decatenation of daughter DNA strands, and dissociation of replisome complex. While previous reports suggested that replication termination in eukaryotes was a result of accumulation of replication forks at the replication pause sites, it is, however, unclear if forks would stall and eventually terminate at those sites (27). Recent work suggests that the replicative helicase CMG plays an important role in termination whereby it is removed from DNA by the ATPase p97 following ubiquitination of MCM7 (28). This view is also contested because another study has implied that DNA replication can terminate even in the absence of CMG removal from DNA. To address

these competing hypotheses of replication termination, Walter's group utilized *in vitro* Xenopus extract system and showed that DNA synthesis does not pause or slow down as two replication forks converge during termination (29). Instead, leading strands pass each other unaffected and position next to lagging strands before undergoing ligation. Moreover, they reveal that CMG helicase remains associated and is removed only after the ligation of the leading strand of one fork with the lagging strand of the opposite fork thereby negating a previously proposed model of unloading of CMG as a first step of termination. In addition, the new model suggests that decatenation of daughter strands occurs concurrently with ligation before CMG unloading. According to this model, even if one fork stalls, it would remain stable until a converging fork arrives to ensure complete DNA synthesis by minimizing the possibility of premature replisome disassembly.

1.2 DNA Repair

Despite a high fidelity and robust proofreading capabilities of DNA polymerases and a tight regulation, DNA is not immune to damage that if not repaired can lead to genome instability. This can result from both internal and external sources such as failure in the part of replication proteins, natural replication challenges, and exogenous DNA damaging agents. Therefore, cells have evolved various repair mechanisms to sense such damage and repair it to ensure that the damaged DNA is not propagated through DNA replication and cell division. Various DNA repair systems work in tandem with the cell cycle regulatory machinery to that endeavor. Therefore, when DNA is damaged, DNA replication is halted and the DNA repair machinery is recruited to the site of the damage. If the damage is fixed, replication resumes (30). However, if DNA damage is unresolved,

cells can activate a cell cycle arrest to recruit additional factors to resolve it. In the event of severe damage or if the repair machinery fails to fix the lesion(s), a cell will either enter senescence or a programmed cell death to stop the damage propagation.

DNA Damage Response (DDR)

A coordinated cellular response to resolving damaged DNA is called a DNA damage response (DDR). DDR includes diverse repair systems, each with a task of repairing a distinct DNA lesion. DNA double strand breaks (DSBs) are the most severe forms of DNA damage because of their high mutagenic potential (31). Such DNA lesions are repaired mainly through one of two pathways- homologous recombination (HR) or non-homologous end joining (NHEJ). HR involves utilization of a homologous template to recover genetic information that is lost during the initial breakage step or subsequent processing of DNA. HR initiates with resection of the broken ends from the 5' to 3' direction to produce a 3' single stranded DNA (ssDNA) (32). Initial resection is carried out by the action of CtIP nuclease in concert with the Mre11, Rad50, and Nbs1 (MRN) proteins. This is followed by extensive resection by enzymes such as DNA2, which exposes a long stretch of ssDNA that is then bound by RPA. The function of RPA is to protect ssDNA from further degradation and recruit Rad51 recombinase to form a nucleoprotein filament. Once formed, nucleoprotein filaments containing Rad51 along with several other proteins such as Rad52, BRCA1 and BRCA2 are involved in homology search, strand invasion, displacement loop (D loop) formation and DNA synthesis to form a branched nucleic acid structure called a Holliday junction. These structures are then resolved through either double-strand break repair (DSBR) pathway producing crossover

products or synthesis-dependent strand annealing (SDSA) producing non-crossover products. Generally, HR results in non-crossover products to restore parent DNA sequence.

In contrast to HR, NHEJ is a rapid but an error prone pathway for DSB pathway. It involves binding of Ku proteins (Ku70/80) to the broken ends followed by end processing and ligation by DNA ligase 4 (33). DNA protein kinase (DNAPK) plays an important role in signaling and recruitment of accessory proteins during NHEJ. In cells where NHEJ is blocked or inhibited, DNA double strand breaks can also be repaired by alternative NHEJ called micro-homology mediated end joining (MMEJ) (34). MMEJ involves resection or DNA unwinding to expose short, single strand microhomologies. Such microhomology sequences are then annealed and DNA repair is completed by cleavage of overhanging of 3' flaps, DNA synthesis, and ligation. MMEJ utilizes ligase 1 and ligase 3 instead of ligase 4. Notably, MMEJ is mutagenic and therefore results in DNA deletions and loss of genetic information.

In addition to DNA double strand breaks (DSBs), cells encounter several other types of DNA damage that are processed by distinct repair pathways. For example, ultraviolet light- induced DNA lesions and bulky DNA adducts are repaired by nucleotide excision repair (NER) (35). Nucleotide base lesions (single or multiple) are repaired by base excision repair (BER) whereas base mismatches are fixed by mismatch repair (MMR). Similarly, DNA crosslinks are repaired by Fanconi anemia (FA) pathway.

Regulation of the DDR

DDR pathways are subject to tight regulation to ensure repair processes are only active at the right place and the right time. The regulatory network consists of DNA damage sensors, signal mediators, signal transducers and signal effectors, all working in tandem (36). The factors involved in post-translational modifications (PTMs) are a major part of this network. The most common forms of such modifications are phosphorylation, ubiquitination, acetylation, methylation, sumoylation and neddylation. In many cases, these modifications are reversible and therefore are a function of two opposing enzymes.

Among a list of many proteins, two phosphatidylinositol 3-kinase related kinases (PIKKs)- ATM (ataxia telangiectasia mutated) and ATR (ATM and Rad3 related)- play a central role in DDR regulation. Upon DNA damage, the recruitment of the MRN complex to the DSBs triggers the auto-phosphorylation of ATM at S1981 turning it from an inactive dimer to an active monomer (37). ATR, on the other hand, is activated in response to replication stress and replication fork collapse and recruited to the DNA damage site upon binding of Rad9-Hus1-Rad1 (9-1-1) clamp complex (38). The activation of both ATM and ATR can lead to the phosphorylation of over 700 downstream substrates that are involved in various DNA damage responses including DNA repair, cell cycle progression and apoptosis (39). One of the well-studied substrates of ATM/ATR is histone H2A variant H2AX. It is an important sensor of DDR and when phosphorylated on S139, referred to as γ -H2AX, it recruits plethora of DDR factors onto a damaged DNA to enact an appropriate response (40). The regulation of DDR via ATM/ATR is also closely coordinated with the cell cycle checkpoint kinases, Chk1 and Chk2, which transduce and

amplify the DDR signal. In fact, ATM directly phosphorylates and activates Chk2 in response to DSB while ATR phosphorylates Chk1 in response to several other DNA damages including single strand breaks (41). Additionally, ATM and ATR phosphorylate master regulator and central tumor suppressor protein, p53, thereby activating the transcription of several genes that regulate cell cycle arrest, DNA repair, senescence, and apoptosis (42). The interplay between master regulators of DDR including ATM, ATR, Chk1, Chk2, and p53 highlight the complexity and redundancy involved in DDR regulation.

In addition to kinases, several other PTM factors play crucial roles in DDR regulations. Phosphatases such as WiP1 and PP2A directly regulate the activity of ATM, ATR, DNA-PKcs, Chk1, Chk2, and p53 to antagonize the kinase signaling cascades (43). Similarly, ubiquitin E3 ligases such as RNF8 and RNF168 and deubiquitination (DUBs) enzymes such as USP, UCH, OUT, and MJD control ubiquitination of H2A histones to regulate DDR response (44), (45). In addition to ubiquitination, sumoylation is also emerging as a new mode of regulating DDR via PTM (46).

Aside from PTM factors, non-coding RNAs (ncRNAs) are also evolving as key regulators of the DDR. This group of ncRNAs includes both the small ncRNAs such as microRNAs (miRNAs) and small interfering RNAs (siRNAs) as well as long ncRNAs such as XIST and HOTAIR. miRNAs, for example, are known to directly bind to the 3'UTR of p53 to disrupt its interaction with Mdm2 following DNA damage (47). Similarly, lncRNAs, for

example, interact with chromatin remodeling complexes to allow necessary response to DNA damage.

2. Telomere biology

2.1 Telomere structure and proteins

Telomeres are the nucleoprotein structures that cap the termini of linear chromosomes in eukaryotes. They consist of telomeric DNA and proteins. Telomeric DNA length varies between organisms and can range from several hundred base pairs in yeast to tens of kilobases in mammals.(5) In humans, telomere length is between 5-15 kb whereas that in lab mouse, it is much longer in the order of 40-50 kb. In mammals, telomeric DNA is composed of G-rich –TTAGGG- hexanucleotide repeats. Although bulk of the telomeric DNA is double stranded, mammalian telomeres end with a distinct protrusion of a 50-300 nucleotide single stranded repeats at the 3'end. These 3' overhangs are called G-tails or G overhangs (48). By forming a secondary structure commonly referred to as t-loop (telomeric loop), these 3' overhangs prevent telomeres from being identified as double strand breaks.

In addition to the DNA, telomeres consist of many proteins. In mammals, telomeres are characterized by a basic complex of six proteins, referred together as shelterin complex (49). The Shelterin complex recognizes the dsDNA and the G-overhang and are responsible for sheltering telomeres from unwanted activities including the repair machinery that targets double strand breaks. Furthermore, the shelterin complex is

involved in facilitating t loop formation at the telomere. The proteins comprising the shelterin complex are telomeric repeat-binding factor 1 (TRF1), telomeric repeat-binding factor 2 (TRF2), repressor and activator protein 1 (RAP1), TRF1-interacting nuclear protein 2 (TIN2), protection of telomeres 1 (POT1), and POT1-and TIN2-interacting protein (TPP1). TRF1 and TRF2 bind the dsDNA while POT1 binds the single stranded 3' overhang. TIN2 and TPP1 are linker proteins that hold the complex together. Several interactions exist amongst these shelterin proteins to form a stable complex to maintain telomere integrity. For example, TIN2 interacts with TRF1, TRF2 and TPP1 and works as a lynchpin. TPP1 interacts with TIN2 and POT1 and forms a bridge to hold dsDNA and single stranded 3' overhang together. RAP1 interacts with TRF2 to stabilize the complex. Importantly, all 6 proteins work in a coordinated fashion to ensure the proper replication and protection of the telomere.

The physiological functions of the shelterin proteins are evident from various *in-vitro* and *in-vivo* studies (5), (50). TRF1 and TRF2 are both essential in mice. In addition, conditional deletion of either TRF1 or TRF2 results in destabilization of shelterin complex and telomere deprotection resulting in a telomeric DNA damage response. While they work as a unit, they also have specific functions during telomere replication. POT1 and TPP1 work together as a heterodimer, while TPP1 does not bind directly to the telomere, it is required to localize POT1 to the telomere as POT1 lacks its own nuclear localization signal. Additionally, TPP1 enhances the telomere-binding efficiency of POT1. POT1/TPP1 heterodimers execute two major functions at the telomere- protection of 3' G-overhang and control of telomerase activity. First, by binding to the G-overhang, the

POT1/TPP1 heterodimer prevents RPA-binding at the G overhang and protect from activation of an unwanted ATR- mediated DNA damage response. Second, these two proteins have opposing functions in regulating telomerase activity; POT1 can block telomerase access to 3'overhang while TPP1 can recruit telomerase and stimulate its activity. However, how these two opposing activities are coordinated remains unclear. Another shelterin protein, TIN2, is involved in stabilizing the complex by binding to both TRF1 and TRF2. It also increases the DNA binding efficiency of both proteins. TIN2-interaction with TPP1 is also responsible for facilitating the POT1-TPP1 interaction and POT1's binding to the 3' G-overhang. TIN2 also facilitates the interaction and recruitment of non-shelterin proteins to the telomere. Notably, TIN2 mutations are found in patients with short telomeres and those with dyskeratosis congenital. The final shelterin protein, RAP1, is unique in that it does not directly bind to the telomere DNA but with only TRF2. Furthermore, it is the only shelterin protein that is not essential in mice. However, its deletion results in increased recombination without affecting NHEJ or chromosomal fusions. Its function in humans is predicted to be slightly different from that in mice; it is expected to function with TRF2 to prevent NHEJ and chromosome fusion. RAP1, like TIN2, can also recruit or function with non-shelterin proteins, most notably in NFkB activation.

In addition to the six primary proteins of the Shelterin complex, a growing list of other telomere-associated proteins have also been identified in mammalian cells. These non-shelterin proteins either aid shelterin proteins or have their own independent functions during telomere replication, repair, and maintenance. For example, the CST protein

complex comprising of CTC1, STN1 and TEN1 in mammalian cells and Cdc13, Stn1, and Ten1 in *S. cerevisiae* bind to the 3' overhang and function in telomere protection and replication (51), (52). Similarly, proteins of the RecQ helicases, WRN and BLM, interact with members of the shelterin complex and play a role in lagging strand telomere replication (53), (54). Additionally, our laboratory's findings demonstrating the role of flap endonuclease 1 (FEN1) in telomere replication and maintenance highlight the diversity of non-telomere proteins that are recruited to telomeres (55), (56), (57). In addition to telomere DNA and proteins, telomeric RNA, commonly referred to as TERRAs, play an important role in telomere homeostasis either as free or bound to the telomere (58).

2.2 Telomere replication and maintenance

Telomeres must be replicated during S-phase to ensure cell division continuity. Telomeres are replicated primarily by one of two mechanisms: semi conservative DNA replication and by the action of the enzyme telomerase. In addition, there are two specialized mechanisms—recombination-based and alternative lengthening of telomeres (ALT) –which are activated under certain conditions in distinct cell types to maintain telomere lengths.

Semi conservative telomere replication and challenges

The most common mode of telomere replication in somatic cells is semi conservative DNA replication during which telomeres are replicated along with the rest of the genome. It is a multistep process that involves the movement of replication forks along the telomeric DNA on both leading and lagging strands followed by the processing of the DNA

ends to produce a terminal 3' overhang structure. Replication machinery, however, faces two major challenges at the telomere compared to other loci in the genome (59). First, telomere replication is difficult due to its terminal location on the chromosome. While replication of any other genomic region is carried out by more than one replication forks originating from different origins, telomeres must be replicated by one centromere-distal replication fork. This can result in telomeres loss due to replication failure in the event of a fork collapse or damage. This problem is absent at other regions of the genome because a compensating replication fork traveling in the opposite direction can at least theoretically complete replication even when one fork collapses or is no longer functional. Furthermore, because telomere replication is thought to be initiated from origins located in the sub-telomeric region, replication machinery is under immense pressure to ensure that the last fired-origin replicate correctly (60). Failure to do so would result in incomplete telomere replication. Another major challenge in telomere replication emanates from its DNA sequence. Telomeres contain G-rich hexanucleotide repeats that form secondary structures such as G-quadruplexes and T-loops. Such secondary structures need to be resolved into a linear DNA molecule for replication to occur. While there are mechanisms for resolution of such structures, this makes DNA replication vulnerable to dysfunction resulting in incomplete replication or complete loss of telomeres.

Telomere extension by telomerase

Gradual telomere shortening is an inevitable consequence of semi conservative DNA replication and end-processing. Therefore, cells have evolved a mechanism to extend

telomeres to restore their lengths and facilitate cell division. This mechanism is a function of a specialized enzyme called telomerase.

Telomerase is a ribonucleoprotein (RNP) complex comprised of two core subunits: telomerase RNA (TER) and telomerase reverse transcriptase (TERT) (61). TERT utilizes TER as a template to synthesize telomeric repeat sequences. While TERs are divergent in sequence and secondary structures amongst organisms, all TERS contain two common motifs- template/pseudoknot (PK) and TERT binding CR 4/5 domain. On the other hand, TERTs are evolutionarily conserved with four domains: the TERT N-terminal (TEN) domain, the TERT RNA binding domain (TRBD), the Reverse Transcriptase (RT) domain, and the C-terminal extension (CTE). While the RT and the CTE domains are conserved between TERT and other reverse transcriptases, TEN and TRBD are unique to TERT.

Telomerase, like other polymerases, catalyzes the addition of nucleotide to a primer 3' hydroxyl group. However, telomerase is unique in its ability to interact with both its RNA template as well as the telomere DNA template. One notable difference is the ability of telomerase to extend without product dissociation from the enzyme, a process referred to as repeat addition processivity (RAP) (62). Furthermore, telomerase can copy the 3' end of the template DNA with less base pairing compared to the 5' end thereby increasing its efficiency. Furthermore, telomeric repeats containing RNA, commonly referred to as TERRA, is believed to play an important role in recruitment of TERT to short telomeres to facilitate their extension (63).

Telomere extension through the action of telomerase is vital during development in organisms. Therefore, telomerase activity is robust in embryonic stem cells as well as in somatic cells during development. However, this activity is significantly reduced or absent in adult human somatic cells (64). This results in progressive shortening of telomeres with each division in adults. The absence of telomerase in adult somatic cells has both beneficial as well as deleterious effects. For example, several human diseases including syndromes of bone marrow failure and pulmonary fibrosis arise from insufficient telomerase activity (65). In contrast, the absence of telomerase activity is a safeguard mechanism to preventing cancer as aberrant telomerase reactivation is directly linked to tumorigenesis (66). Therefore, the function of telomerase in telomere elongation is very consequential and underscores the importance of understanding its regulation in multitude of biological processes.

Telomere recombination

In addition to telomere elongation by semi conservative replication and by the action of telomerase, recombination is a pathway utilized by both yeast and mammals to maintain telomere length and for survival. Recombination-based telomere length was reported following an observation that some yeast cells lacking telomerase (*est1* Δ) had survived senescence and formed viable colonies (67). This was corroborated by the fact that deletion of recombination incompetent cells (*rad52* Δ) in *est1* Δ delta cells led to loss of all survivors. In addition to RAD52, all survivor cells require replication protein *Pol32p* suggesting that replication is a part of the recombination based telomere maintenance

(68). These survivor cells are broadly categorized into Type I and Type II (59). Type I survivors grow slowly and contain multiple Y' sub telomere repeats and short terminal telomere repeats. In addition to *RAD52* and *POL32*, type I survivors need *RAD51*, *RAD54*, *RAD57*, and most likely *RAD55* for survival. Type II survivors, on the other hand, have few sub-telomere repeats but high degree of telomere repeat amplification. Type II survivors are also a result of type I transformations and require MRX complex, *RAD59*, and *SGS1* helicase for survival. More recently, recombination based cellular survival in telomerase-negative cells is proposed to be regulated by *TERRA* and telomere RNA: DNA hybrids thereby highlighting the roles of *TERRA* in telomere homeostasis (69).

Alternative lengthening of telomeres

While 85% of all human cancers have active telomerase for maintaining telomere length and cell division, the remaining 15% utilize one or more mechanisms referred to as alternative lengthening of telomeres (ALT) (70). ALT is commonly found in tumors of mesenchymal origin such as glioblastoma multiforme, osteosarcomas, and some soft tissue sarcomas (71), (72). ALT cells are identified and characterized based on their several unique characteristics. One such prominent feature is the presence of telomeric DNA sequences, as independent DNAs separate from chromosomes. These are found as double stranded telomeric circles (t-circles), partially single stranded C-rich strands (C-circles), linear double stranded DNA and t-complex structures with highly branched DNA. Another distinguishing feature of ALT cells is the presence of telomeric DNA in a subset of promyelocytic leukaemia nuclear bodies (PML). Such PML bodies are therefore, referred to as ALT-associated PML bodies and these are believed to play an important role in

several processes including senescence and DNA damage response. ALT cells are also marked by high level of telomere recombination and therefore telomere sister chromatid exchanges (T-SCEs) are abundant in ALT cells. Although still unclear, two current models exist to explain recombination-dependent telomere elongation in ALT cells. According to the first model called unequal T-SCE model, unequal T-SCEs occurs between a daughter cell with a longer telomere and another cell with shorter telomere thereby providing prolonged proliferative advantage to the cell population. The second model called homologous recombination-dependent DNA replication, which is a more favorable one proposes that ALT is a result of shorter telomeres undergoing homologous recombination-based synthesis. The substrate for telomere elongation, according to this model, can be a telomere from another chromosome, the same telomere or any linear extrachromosomal telomeric DNA pieces. Regardless of the model, ALT cells depend on several DNA replication and repair proteins including the MRN complex and shelterin complex proteins. Interestingly, a growing body of evidence suggests that telomere RNA (TERRA) and telomere hybrids play an important role in telomere maintenance in ALT cells (73).

2.3 Telomere Physiology

Telomeres are indispensable modules for maintaining genome stability. Despite robust cellular mechanisms aimed at replicating and protecting telomeres, they are not immune to challenges. Furthermore, gradual telomere attrition due to the “end replication problem” can produce several consequences in the cell. Some of the most significant

manifestations are cellular senescence, cell death, and genome instability leading to several pathologies including cancer and aging disorders.

Cellular senescence and apoptosis

The most prominent consequence of telomere shortening and telomere damage is cellular senescence. It is a state of irreversible cell growth arrest along with several physiological changes. Telomere length is a primary determinant of cell division. When a telomere length reaches a critical length, primarily due to the end replication problem, commonly referred to as the Hayflick limit, a cell ceases to grow and undergoes growth arrest (74). Cells also become senescent when they sustain severe damage to the telomere or any other region of the chromosomes. Cellular senescence has both desirable as well as deleterious effects depending on the context. While it is a mechanism to suppress tumorigenesis at times of challenges from DNA damage as well as oncogenic stimuli, it can also promote pro-tumorigenic effects. Therefore, to prevent cellular senescence that can result from telomere shortening, cells utilize telomerase. This is a feature utilized by most cancer cells to achieve immortality. Ectopic expression of telomerase is also commonly used to immortalize cells for research purpose. However, telomerase cannot prevent cellular senescence that results in response to oncogenes such as RAS or DNA damaging agents. Furthermore, inactivation of the tumor suppressor genes such as p53 and pRb also prevents both human and mice cells from undergoing senescence in response to DNA damage and oncogenic RAS (75).

Apoptosis is another consequence of telomere dysfunction. It is a controlled cell death program activated when a cell sustains irreparable damage. Two tumor suppressors p53 and pRb are critical components controlling the fate of a cell post DNA damage including post-telomere damage (76). Apoptosis, in many cases, is an alternative to senescence. For example, when both p53 and pRb are inactive, cells continue dividing until telomeres are extremely short. Cells at that state enter a phase known as crisis where they sustain chromosomal instabilities along with telomere attrition, resulting in apoptosis. However, some cells may still maintain their telomere lengths and hence escape apoptosis; these cells can proliferate indefinitely but are at a high risk for malignant transformation.

Cancer and genome instability

The importance of telomere maintenance in cancer is underscored by the fact that about 85% of all cancers have activated telomerase and the remaining 15% have the ALT pathway of telomere elongation (75). Because telomere length can control cell division, a cancer cell must maintain its telomere length to enable its uncontrolled division to successfully establish a tumor and later metastasize. Telomerase expression in somatic cells has provided insights into direct relationship between telomere and cancer. Because telomerase expression is much higher in mice tissues compared to that in humans, mice are more cancer-prone than humans. Furthermore, telomerase co-operates with oncogenic stimuli to promote tumorigenesis as demonstrated by the requirement of telomerase in immunocompromised mice for tumor formation. Similar observations were also observed in transgenic mice whereby constitutive expression of telomerase in the skin promoted skin carcinogenesis (77). In addition, mutations in shelterin proteins are

well documented in numerous tumors (78). Similarly, numerous DNA replication and repair proteins including Ku 70/86, and Rad 51 have altered expression changes in cancers. These direct and indirect roles of telomere proteins in cancer underscore the significance of telomere protection in preventing tumorigenesis.

Telomere dysfunction is also closely tied to genome instability, which is a hallmark feature of cancer. In fact, genome instability is a primary cause of many cancers (79). Specifically, loss of function of DNA replication and repair proteins as well as shelterin proteins are well-documented in several cases of genome instability and cancer. Telomeres play an important role in protecting chromosome ends and maintaining genome stability. When this protection is compromised due to internal and external challenges, chromosomes become unstable. When severe, it results in loss of genetic material due to chromosome breakage, fusions, and recombination. This can have deleterious effects including cancer development and cell death. Telomere dysfunction- caused by telomere shortening, telomere damage and loss of functions of telomere associated proteins- produces one of three outcomes for a cell: cellular senescence, cell death, or genome instability. Cellular senescence is a preferred state adopted by a cell in many cases to avoid genome instability. This can prevent division of cells with dysfunctional telomeres thereby guarding against tumor development. However, somatic mutations that accumulate with age can prevent activation of senescence. This is illustrated by studies showing loss of heterozygosity and mutations in the *p53* tumor suppressor as well as *RAS* oncogene even in young and normal tissues (80), (81), (82). These cells fail to activate senescence or even apoptosis and in turn, result in genome instability that presents an enormous risk

for neoplastic transformation. A simple mutation or epigenetic change in these cells can result in transformation.

Aging and non-cancer disorders

Telomere dysfunction is also associated with aging and several other non-cancer pathologies. The effect of telomere shortening on senescence extends beyond a cell growth arrest. A senescent cell secretes a whole host of factors that are collectively referred to as the senescence-associated secretory phenotype (SASP) (83). The SASP can exert both cell-autonomous and non-autonomous effects and can change tissue homeostasis by altering cell integrity and function. The effects of SASP factors are studied in detail in fibroblasts, which are the major component of stroma. Senescent fibroblasts secrete many factors including enzymes that remodel extracellular matrix, cytokines that change the immune environment as well as growth factors that dictate cellular proliferation. These functions of senescent fibroblasts can have either beneficial roles in tissue homeostasis or detrimental consequences in various pathologies, including cancer.

The importance of telomere integrity is also reflected by numerous non-cancer diseases. These diseases arise because of either mutations in specific telomere-related genes or due to changes in telomere length. Many premature aging syndromes manifest with mutations in one or more telomere genes. One such example is dyskeratosis congenital (DC) (84). DC Patients harbor mutations in the components of the telomerase complex that results in reduced telomerase stability and therefore shorter telomeres. DC is

classified as either autosomal dominant if mutations affect the *TER* gene or X-linked if mutation is on the dyskeratosis congenital 1, dyskerin gene (*DKC1*). DC patients display elevated levels of chromosomal instability and develop several pathologies including abnormal skin pigmentation, short stature, bone marrow failure, hypogonadism and infertility and premature death. These patients are also predisposed to spontaneous cancers. Other common premature aging syndromes include ataxia telangiectasia (mutations in *ATM* gene), Werner syndromes (mutations in *WRN* gene) and Bloom syndromes (mutations in *BLM* gene) with common features of shorter telomeres and genome instability. An important point is that patients with some of these premature syndromes have strikingly different phenotypes compared to the mice models of these diseases. This presents a challenge in interrogating molecular mechanism of these pathologies. In addition to diseases resulting from telomere related gene mutations, many age-related diseases have been identified due to their direct or indirect correlation with telomere lengths. Some of these disorders include heart failure, digestive tract atrophies, infertility, reduced angiogenic potential, reduced wound healing, and loss of body mass (85). Overall, manifestations of telomere dysfunctions through mutations of genes involved in telomere replication, telomere repair, and telomere maintenance highlight a diversity of roles telomere plays in genome stability and tissue homeostasis.

3. R loops: Structure, proteins, and physiology

3.1 Structure

RNA:DNA hybrids, commonly referred to as R-loops, are a three-stranded nucleic acid structure composed of an RNA-DNA hybrid and a displaced single stranded DNA. While R loops arise during several biological processes, the most accepted mechanism of R loop formation is transcription during which a nascent RNA emanating from an RNA polymerase hybridizes with a template DNA leaving a non-template DNA flanking. In vivo, R loops are formed as natural intermediates during the initiation of DNA replication in mitochondria, bacterial plasmids, the bacteriophages ColE1 and T4, and in immunoglobulin (Ig) class-switch recombination (CRV) (6). Depending on its origin, R loops can range from 8-10 base pairs to several thousand base pairs in length.

R loop formation depends on three main features: G:C content, DNA supercoiling, and DNA cleavage (86). Using plasmid constructs in bacteria, it is well-established that clusters at the non-transcribed strand are required for initiation of R loop formation whereas the high G density region is required for stabilization and elongation of such structures. Similarly, negative supercoiling is considered an important factor for R loop formation *in vivo*. In fact, high negative superhelicity can reduce the G dependency. A third element, DNA cleavage or a DNA nick, is an important contributor of R loop formation based on findings that DNA nicks when placed downstream of the T7 promoter driving transcription enhanced R loop formation. Because an RNA strand would need to intertwine with the complementary DNA strand, negative supercoiling and high G content would make that event favorable by initiation of the bubble between DNA strands. Once formed, R loop is also favored by the fact that a RNA:DNA hybrid is thermodynamically more stable than a DNA duplex of identical length.

3.2 R loop proteins

While naturally occurring R-loops have beneficial roles, undesirable R loop formation and stabilization have deleterious effects in the cells. Because replication and transcription occur simultaneously in higher eukaryotes, R loops by virtue of their large topological size can become a barrier to a moving replication fork as well as transcription apparatus. Similarly, R loops can induce DNA mutations on non-template DNA as well as increase the frequency of unwanted recombination (87), (88). Therefore, cells have evolved numerous mechanisms to either prevent the formation of undesirable R loops or to process them after their formation. To that end, several nucleases and helicases are tasked with the removal of R loops and hence prevent their accumulation. Ribonuclease H1 (RNH1) and ribonuclease H2 (RNH2) are two major nucleases implicated in the degradation of the RNA component of the hybrid (89). In yeast, both ribonucleases can resolve R loops and are therefore considered redundant in R-loop resolution. However, in humans, RNH1 is considered the primary enzyme responsible for R loop resolution while RNH2 is primarily tasked with the removal of single mis-incorporated ribonucleotide from DNA during DNA replication. Aside from nucleases, several helicases have also been identified. In humans, DNA helicase Pif1, RNA helicase DHX9, DNA/RNA helicase SETX, and RNA helicase AQR have all been shown to resolve R loops; their loss is associated with accumulation of R loops either in cells or mice models (90), (91), (92), (93). Similarly, accumulation of negative supercoiling is also associated with R loop formation. Therefore, enzymes that resolve negative supercoiling are deemed important. In yeast, cells lacking both topoisomerases Top1 and Top2 accumulate R loops in the

ribosomal DNA (rDNA) locus in addition to the stalling of RNA Polymerase and defects in pre-rRNA synthesis (94). In humans, cells deficient in TOP1 display DNA breaks and replication defects and these phenotypes are rescued by ectopic expression of RNH1 thereby suggesting that TOP1 prevents R loop accumulation (95). Another mechanism involved in the prevention of R loop formation is a messenger ribonucleoprotein particle (mRNP) assembly complex. This involves numerous proteins involved in nascent RNA synthesis as well as mRNA export machinery. By directly or indirectly binding and protecting a nascent RNA, mRNP biogenesis proteins prevent its hybridization with the transcribed DNA strand. The first evidence of the role of mRNP biogenesis in prevention of R loop formation was provided by yeast studies which showed that yeast cells with a mutant THO complex, a part of the mRNP biogenesis, led to the accumulation of R loops (96). Similarly, another mRNP factor, SRSF1, is important for R loop prevention as evidenced by findings that showed that chicken DT40 and human HeLa cells depleted of SRSF1 showed elevated levels of rearrangements that were suppressed by RNH1 ectopic expression (97). In addition to mRNP factors, the mRNA surveillance system also prevents R loop accumulation as evidenced by the increased levels of R loops in yeast cells depleted of Trf4 as well as in mouse embryonic stem cells depleted of the exoribonucleases exosome components EXOSC3 and EXOSC10 (98), (99).

3.3 R loops physiology

Recent genome wide analyses have revealed that R loops are present throughout the genome at a frequency much higher than previously appreciated. Furthermore, it is becoming increasingly clear that R loops have a wide range of functions in transcription,

genome dynamics, and chromatin changes. Similarly, they are also implicated in numerous human diseases including cancer and neurodegenerative diseases.

R loops in transcription

The role of R loops in transcription is evident from genome wide mapping studies. R loops are present at both the promoter and termination regions of several genes. A genome-wide capture of RNA:DNA hybrids followed by sequencing (DRIP-seq) revealed that R loops are enriched in CpG islands showing a strong GC skew (100), (101). Furthermore, R loops localize immediately after the transcription start site. Because these sites are hot spots for the action of gene-silencing DNA methyltransferase 3 B1 (DNMT3B1), R loops are believed to promote transcription activation of these genes. This hypothesis is also supported by the fact that regions of high GC skew are bound by methylated histones such as H3K4me3, H4K20me1 and H3K79me2 linked to transcription initiation and elongation.

In addition to their roles in transcription activation, R loops are also associated with transcription termination. High GC skews are present at the 3' end of some genes and the incidence of R loops in those regions raises a possibility that R loops are involved in protection from DNMTs just like in the promoter regions. In doing so, R loops may help avoid the transcriptional read-through and contribute to termination. In addition, roles of R loops in transcriptional termination are elucidated from their function in RNA pol II driven genes. The role of human SETX in XRN2-dependent termination is evidenced by accumulation of R loops at G rich termination pause sites upon depletion of SETX (92).

These data suggest that R loops at termination sites are required for RNA pol II pausing but subsequent removal would allow for release of RNA molecule and efficient transcription termination.

R loops in genome dynamics

Despite being key intermediates of various processes, R loops can modulate genome structure by inducing DNA damage and replication stress and by altering telomere dynamics. In doing so, R loops play an important role in genome instability. A non-template DNA strand of an R loop is susceptible to mutations because it is more accessible to the action of mutagenic enzymes such as activation induced cytidine deaminase (AID), which can convert cytosine into uracil (102). Furthermore, R loop mediated DNA breaks, recombination, and chromosome rearrangements have been demonstrated in yeast and human cells. Because R loops are large topological structure, most of these DNA damage effects could be a result of the capacity of R loops to stall replication fork progression. R loops also interfere with the replication process in a transcription-dependent manner due to the collision of replication and transcription machineries in either co-directional or head-on directions. Additionally, recent findings that tumor suppressor proteins BRCA1 and BRCA2 are associated with R loop homeostasis suggest an expanded role of R loops in DNA repair and replication dynamics (103).

R loops also play prominent roles at the telomere to protect chromosomal stability. Telomeric RNA, commonly referred to as TERRAs, hybridize with telomeric DNA to form

telomeric R loops. These R loops are equally important in both yeast as well as human cells. These R loops accumulate in yeast strains lacking RNase H or proteins of mRNP complex such as THO (104). Similarly, R loops play a critical role in maintaining telomeres in telomerase negative yeast cells whereby R loops activate recombination events that are necessary for telomere elongation and senescence delay (105). In human cells, telomeric RNA:DNA hybrids are important in cells that maintain telomere length via ALT mechanism, which requires R loops to initiate recombination (73). This highlights the varied roles that telomeric RNA:DNA hybrids play to maintain telomere integrity and prevent genome instability in yeast and human cells.

R loops in human disease

R loops are associated with numerous human diseases. The neurodegenerative disorders are one such group that has received a renewed interest. Multiple neurological disorders are caused by the expansion of trinucleotide repeats. Transcription of CTG repeats is shown to form R loop both *in vitro* and *in vivo* and this is linked to repeat instability in human cells (106). Similarly, FRDA is another neurological disease caused by expansion of GAA repeat in the first intron of *FXN* gene (107). *In vitro* studies and those in bacteria have revealed that RNA:DNA hybrids form on GAA repeat sequences leading to RNA polymerase arrest and reduced transcription of *FXN* expression. Similar R loop associated inhibition of transcription are also observed in FXS and fragile X associated tremor/ataxia syndrome (FXTAS) (108). In addition to direct role of R loops, R-loop removing enzymes are also implicated in various neurodegenerative disorders. For example, mutations of helicase SETX is associated with ataxia-ocular apraxia type 2

(AOA2) and with ALS type 4 (ALS4) (109), (110).

In addition to neurodegenerative disorders, R loops are also important in cancer. This is largely because R loops are a source of genome instability and replication stress, both of which are hallmark features of cancer (111). An example of this is the role of tumor suppressors BRCA1 and BRCA2 in the prevention of R loop accumulation (103). Human cells depleted of BRCA1 or BRCA2 show elevated levels of R loops as well as double strand breaks, which are both rescued by ectopic expression of RNH1. In addition, there are various cancers with mutated R loop related genes whose functions remain unclear. One such example is a *FIP1L1* gene that encodes a cleavage and poly (A) factor; it is affected by translocations between its amino-terminal domain and platelet-derived growth factor receptor α (PDGFR α) in 10-20% of eosinophilic leukemia cases (112). Similarly, another R loop linked protein, splicing factor SRSF1, is also overexpressed in several cancer types (113). Similar inferences can be made of results from cells infected with the cancer-causing Kaposi's sarcoma associated herpesvirus (KSHV) (114). These cells have elevated DNA damage and increased R loops due to sequestration of RNA export factor TREX by a viral protein. It is unclear whether these R loop related genes have a causative link. Regardless, the role of R loops in genome instability and multiple cancers is undisputed and deserves further investigation.

References

1. Bell S, Dutta A. DNA replication in eukaryotic cells. *Annu Rev Biochem* [Internet]. 2002;71:333–74.
2. Postow L, Crisona NJ, Peter BJ, Hardy CD, Cozzarelli NR. Topological challenges to DNA replication: conformations at the fork. *Proc Natl Acad Sci U S A* [Internet]. 2001;98(15):8219–26.
3. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 2000;403(6765):41–5.
4. Neidle S, Parkinson G. Telomere maintenance as a target for anticancer drug discovery. *Nat Rev Drug Discov* [Internet]. 2002;1(5):383–93.
5. Stewart JA, Chaiken MF, Wang F, Price CM. Maintaining the end: Roles of telomere proteins in end-protection, telomere replication and length regulation. *Mutat Res - Fundam Mol Mech Mutagen*. Elsevier B.V.; 2012;730(1–2):12–9.
6. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. *Nat Rev Genet* [Internet]. Nature Publishing Group; 2015;16(10):583–97.
7. Stavnezer J, Guikema JEJ, Schrader CE. Mechanism and Regulation of Class Switch Recombination. *Annu Rev Immunol*. 2008;26:261–92.
8. Lee DY, Clayton DA. Initiation of mitochondrial DNA replication by transcription and R-loop processing. *J Biol Chem*. 1998;273(46):30614–21.

9. Aguilera A, García-Muse T. R loops: from transcription byproducts to threats to genome stability. *Mol Cell* [Internet]. 2012 Apr 27 [cited 2014 Sep 2];46(2):115–24.
10. Crick FHC, For IE, Fibrous THE, Of N. D.n.a.
11. Meselson M, Stahl FW. The replication of DNA in. *Proc Natl Acad Sci*. 1958;44:671–82.
12. Parker MW, Botchan MR, Berger JM. Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit Rev Biochem Mol Biol* [Internet]. 2017;9238(March):1–41.
13. Vashee S, Cvetic C, Lu W, Simancek P, Kelly TJ, Walter JC. Sequence-independent DNA binding and replication initiation by the human origin recognition complex. 2003;1894–908.
14. Bell SP, Stillman B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* [Internet]. 1992;357(6374):128–34.
15. Leonard AC, Méchali M. DNA replication origins. *Cold Spring Harb Perspect Biol* [Internet]. 2013;5(10):a010116–a010116.
16. Spring C. Organization of DNA Replication. 2010;1–14.
17. Zhang D, O'Donnell M. The Eukaryotic Replication Machine [Internet]. 1st ed. Vol. 39, *Enzymes*. Elsevier Inc.; 2016. 191-229 p.
18. Pellegrini L, Costa A. New Insights into the Mechanism of DNA Duplication by the Eukaryotic Replisome. *Trends Biochem Sci* [Internet]. Elsevier Ltd;

- 2016;41(10):859–71.
19. Abid Ali F, Renault L, Gannon J, Gahlon HL, Kotecha A, Zhou JC, et al. Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat Commun* [Internet]. Nature Publishing Group; 2016;7:10708.
 20. Yuan Z, Bai L, Sun J, Georgescu R, Liu J, O'Donnell ME, et al. Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat Struct Mol Biol* [Internet]. Nature Publishing Group; 2016;23(3):217–24.
 21. Georgescu RE, Schauer GD, Yao NY, Langston LD, Yurieva O, Zhang D, et al. Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *Elife*. 2015;2015(4):1–20.
 22. Yao NY, O'Donnell M. The Replisome. *Cell* [Internet]. 2010;141(6):1088–1088.e1. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0092867410006173>
 23. Simon AC, Zhou JC, Perera RL, van Deursen F, Evrin C, Ivanova ME, et al. A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. *Nature* [Internet]. 2014;510(7504):293–7.
 24. Bambara RA, Murante RS, Henricksen LA. Eukaryotic DNA Replication Fork *. *Biochemistry*. 2017;(41):3–6.
 25. Burgers PMJ. Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem*. 2009;284(7):4041–5.
 26. Celis JE, Petersen G. Progressin : a novel proliferation-sensitive and cell cycle-

- regulated human protein whose rate of synthesis increases at or near the G1/S transition border of the cell cycle. 1987;223(2):0–5.
27. Fachinetti D, Bermejo R, Cocito A, Minardi S, Katou Y, Kanoh Y, et al. Replication Termination at Eukaryotic Chromosomes Is Mediated by Top2 and Occurs at Genomic Loci Containing Pausing Elements. *Mol Cell* [Internet]. Elsevier; 2010;39(4):595–605.
 28. Maric M, Maculins T, De Piccoli G, Labib K. Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. *Science* [Internet]. 2014;346(6208):1253596.
 29. Dewar JM, Budzowska M, Walter JC. The mechanism of DNA replication termination in vertebrates. *Nature* [Internet]. 2015;525(7569):345–50. Available from: <http://dx.doi.org/10.1038/nature14887>
 30. Campisi J, d'Adda di Fagagna F. Cellular senescence: when bad things happen to good cells. *Nat Rev Mol Cell Biol*. 2007;8(9):729–40.
 31. Rothkamm K, Krüger I, Thompson LH, Kru I, Lo M. Pathways of DNA Double-Strand Break Repair during the Mammalian Cell Cycle Pathways of DNA Double-Strand Break Repair during the Mammalian Cell Cycle. *Mol Cell Biol*. 2003;23(16):5706–15.
 32. Rodgers K, Mcvey M. Error-Prone Repair of DNA Double-Strand Breaks. *J Cell Physiol*. 2016;231(1):15–24.
 33. Burma S, Chen BPC, Chen DJ. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA Repair (Amst)*. 2006;5(9–10):1042–8.

34. Bennardo N, Cheng A, Huang N, Stark JM. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet.* 2008;4(6).
35. Toueille M, Hübscher U. Regulation of the DNA replication fork: A way to fight genomic instability. *Chromosoma.* 2004;113(3):113–25.
36. Liu Y, Li Y, Lu X. Regulators in the DNA damage response. *Arch Biochem Biophys* [Internet]. Elsevier Inc; 2016;594:18–25.
37. Kozlov S V., Graham ME, Jakob B, Tobias F, Kijas AW, Tanuji M, et al. Autophosphorylation and ATM activation: Additional sites add to the complexity. *J Biol Chem.* 2011;286(11):9107–19.
38. Delacroix S, Wagner JM, Kobayashi M, Yamamoto KI, Karnitz LM. The Rad9-Hus1-Rad1 (9-1-1) clamp activates checkpoint signaling via TopBP1. *Genes Dev.* 2007;21(12):1472–7.
39. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER, Hurov KE, Luo J, et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science (80-)* [Internet]. 2007;316(5828):1160–6.
40. Burma S, Chen BP, Murphy M, Kurimasa A, Chen DJ. ATM Phosphorylates Histone H2AX in Response to DNA Double-strand Breaks. *J Biol Chem.* 2001;276(45):42462–7.
41. Smith J, Mun Tho L, Xu N, A. Gillespie D. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer [Internet]. 1st ed. Vol. 108, *Advances in Cancer Research.* Elsevier Inc.; 2010. 73-112 p.

42. Cheng Q, Chen J. Mechanism of p53 stabilization by ATM after DNA damage. *Cell Cycle*. 2010;9(3):472–8.
43. Lee DH, Chowdhury D. What goes on must come off: Phosphatases gate-crash the DNA damage response. *Trends Biochem Sci* [Internet]. Elsevier Ltd; 2011;36(11):569–77.
44. Huen MSY, Grant R, Manke I, Minn K, Yu X, Yaffe MB, et al. RNF8 Transduces the DNA-Damage Signal via Histone Ubiquitylation and Checkpoint Protein Assembly. *Cell*. 2007;131(5):901–14.
45. Nijman SMB, Luna-Vargas MPA, Velds A, Brummelkamp TR, Dirac AMG, Sixma TK, et al. A genomic and functional inventory of deubiquitinating enzymes. *Cell*. 2005;123(5):773–86.
46. Saitoh H, Hinchey J. Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J Biol Chem*. 2000;275(9):6252–8.
47. Liu Y, Lu X. Non-coding RNAs in DNA damage response. *Am J Cancer Res* [Internet]. 2012;2(6):658–75.
48. O'Sullivan RJ, Karlseder J. Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol* [Internet]. Nature Publishing Group; 2010;11(3):171–81.
49. Palm W, de Lange T. How Shelterin Protects Mammalian Telomeres. *Annu Rev Genet* [Internet]. 2008;42(1):301–34.
50. Linger BR, Price CM. Conservation of telomere protein complexes: shuffling

- through evolution. *Crit Rev Biochem Mol Biol* [Internet]. 2009;44(6):434–46.
51. Surovtseva Y V., Churikov D, Boltz KA, Song X, Lamb JC, Warrington R, et al. Conserved Telomere Maintenance Component 1 Interacts with STN1 and Maintains Chromosome Ends in Higher Eukaryotes. *Mol Cell* [Internet]. Elsevier; 2009;36(2):207–18.
 52. Miyake Y, Nakamura M, Nabetani A, Shimamura S, Tamura M, Yonehara S, et al. RPA-like Mammalian Ctc1-Stn1-Ten1 Complex Binds to Single-Stranded DNA and Protects Telomeres Independently of the Pot1 Pathway. *Mol Cell* [Internet]. Elsevier; 2009;36(2):193–206.
 53. Crabbe L, Verdun RE, Haggblom CI, Karlseder J. Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science* [Internet]. 2004;306(5703):1951–3.
 54. Barefield C, Karlseder J. The BLM helicase contributes to telomere maintenance through processing of late-replicating intermediate structures. *Nucleic Acids Res.* 2012;40(15):7358–67.
 55. Saharia A, Guittat L, Crocker S, Lim A, Steffen M, Kulkarni S, et al. Flap endonuclease 1 contributes to telomere stability. *Curr Biol* [Internet]. 2008 Apr 8 [cited 2013 Nov 25];18(7):496–500.
 56. Saharia A, Teasley DC, Duxin JP, Dao B, Chiappinelli KB, Stewart S a. FEN1 ensures telomere stability by facilitating replication fork re-initiation. *J Biol Chem* [Internet]. 2010 Aug 27 [cited 2013 Nov 25];285(35):27057–66.
 57. Teasley DC, Parajuli S, Nguyen M, Moore HR, Alspach E, Lock YJ, et al. Flap

- endonuclease 1 limits telomere fragility on the leading strand. *J Biol Chem* [Internet]. 2015;jbc.M115.647388.
58. Xu Y, Komiyama M. Structure, function and targeting of human telomere RNA. *Methods* [Internet]. Elsevier Inc.; 2012;57(1):100–5.
 59. Wellinger RJ, Zakian VA. Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: Beginning to end. *Genetics*. 2012;191(4):1073–105.
 60. Wellinger RJ, Wolf AJ, Zakian VA. Origin Activation and Formation of Single-Strand TG1.3 Tails Occur Sequentially in Late S Phase on a Yeast Linear Plasmid. *Mol Cell Biol*. 1993;13(7):4057–65.
 61. Wu RA, Upton HE, Vogan JM, Collins K, Collins K, Hall B. Telomerase Mechanism of Telomere Synthesis. 2017;(January):1–22.
 62. Collins K. Single-stranded DNA repeat synthesis by telomerase. *Curr Opin Chem Biol* [Internet]. Elsevier Ltd; 2011;15(5):643–8.
 63. Cusanelli E, Romero C, Chartrand P. Telomeric Noncoding RNA TERRA Is Induced by Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres. *Mol Cell*. 2013;51(6):780–91.
 64. Wang LQ, Liu DW. Telomere, telomerase and stem cells. *J Clin Rehabil Tissue Eng Res*. 2009;13(10):1985–8.
 65. Armanios M, Blackburn EH. The telomere syndromes. *Nat Rev Genet* [Internet]. Nature Publishing Group; 2012;13(10):693–704.

66. Shay JW. Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Discov* [Internet]. 2016;(June):584–94.
67. Lundblad V, Blackburn EH. An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell*. 1993;73(2):347–60.
68. Lydeard JR, Jain S, Yamaguchi M, Haber JE. Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature*. 2007;448(7155):820–3.
69. Balk B, Maicher A, Dees M, Klermund J, Luke-Glaser S, Bender K, et al. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* [Internet]. 2013;20(10):1199–205.
70. Cesare AJ, Reddel RR. Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* [Internet]. Nature Publishing Group; 2010;11(5):319–30.
71. Bryan TM, Englezou A, Dalla-Pozza L, Dunham MA, Reddel RR. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat Med* [Internet]. 1997;3(11):1271–4.
72. Henson JD. A robust assay for alternative lengthening of telomeres (ALT) in tumors demonstrates the significance of ALT in sarcomas and astrocytomas. *Clin Cancer Res*. 2005;11:217–25.
73. Arora R, Lee Y, Wischnewski H, Brun CM, Schwarz T, Azzalin CM. RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* [Internet]. Nature Publishing Group; 2014 Jan [cited

2014 Oct 22];5:5220.

74. Ruben AJ, Biology E. Hayflick , his limit , and cellular ageing. 2000;1(October):72–6.
75. Campisi J, Kim S ho, Lim CS, Rubio M. Cellular senescence, cancer and aging: The telomere connection. *Exp Gerontol.* 2001;36(10):1619–37.
76. Sherr CJ, McCormick F. The RB and p53 pathways in cancer. *Cancer Cell.* 2002;2(2):103–12.
77. González-Suárez E, Samper E, Ramírez A, Flores JM, Martín-Caballero J, Jorcano JL, et al. Increased epidermal tumors and increased skin wound healing in transgenic mice overexpressing the catalytic subunit of telomerase, mTERT, in basal keratinocytes. *EMBO J.* 2001;20(11):2619–30.
78. Blasco MA. Telomeres and human disease: ageing, cancer and beyond. *Nat Rev Genet [Internet].* 2005;6(8):611–22.
79. Gray JW, Collins C. Genome changes and gene expression in human solid tumors. *Carcinogenesis [Internet].* 2000;21(3):443–52.
80. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science.* 1996;274(5295):2057–9.
81. Jonason AS, Kunala S, Price GJ, Restifo RJ, Spinelli HM, Persing JA, et al. Frequent clones of p53-mutated keratinocytes in normal human skin. *Proc Natl Acad Sci U S A [Internet].* 1996;93(24):14025–9.
82. Cha RS, Thilly WG, Zarbl H. N-nitroso-N-methylurea-induced rat mammary

- tumors arise from cells with preexisting oncogenic Hras1 gene mutations. *Proc Natl Acad Sci U S A*. 1994;91(9):3749–53.
83. Campisi J. Replicative senescence: an old wives tale. *Cell*. 1996;84:497–500.
 84. Savage SA, Bertuch AA. The genetics and clinical manifestations of telomere biology disorders. *Genet Med [Internet]*. 2010;12(12):753–64.
 85. Khan S, Chaturgoon a a, Naidoo DP. Telomeres and atherosclerosis. *Cardiovasc J Afr [Internet]*. 2012;23(10):563–71.
 86. Aguilera A, García-Muse T. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol Cell*. 2012;46(2):115–24.
 87. Polak P, Arndt PF. Transcription induces strand-specific mutations at the 5' end of human genes. *Genome Res*. 2008;18(8):1216–23.
 88. Aguilera A, García-Muse T. Causes of genome instability. *Annu Rev Genet [Internet]*. 2013;47(July):1–32.
 89. Cerritelli SM, Crouch RJ. Ribonuclease H: The enzymes in eukaryotes. *FEBS J*. 2009;276(6):1494–505.
 90. Boué JB, Zakian VA. The yeast Pif1p DNA helicase preferentially unwinds RNA-DNA substrates. *Nucleic Acids Res*. 2007;35(17):5809–18.
 91. Chakraborty P, Grosse F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst) [Internet]*. Elsevier B.V.; 2011;10(6):654–65.
 92. Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human senataxin resolves

- RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol Cell* [Internet]. 2011 Jun 24 [cited 2014 Jul 10];42(6):794–805.
93. Sollier J, Stork CT, García-Rubio ML, Paulsen RD, Aguilera A, Cimprich KA. Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Mol Cell* [Internet]. 2014 Nov [cited 2014 Nov 27];777–85.
 94. El Hage A, French SL, Beyer AL, Tollervey D. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev*. 2010;24(14):1546–58.
 95. Tuduri S, Crabbé L, Conti C, Tourrière H, Holtgreve-Grez H, Jauch A, et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat Cell Biol* [Internet]. Nature Publishing Group; 2009;11(11):1315–24.
 96. Domínguez-Sánchez MS, Barroso S, Gómez-González B, Luna R, Aguilera A. Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet*. 2011;7(12):19–22.
 97. Li X, Manley JL. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*. 2005;122(3):365–78.
 98. Gavaldà S, Gallardo M, Luna R, Aguilera A. R-Loop Mediated Transcription-Associated Recombination in *trf4* Mutants Reveals New Links between RNA Surveillance and Genome Integrity. *PLoS One*. 2013;8(6):1–8.

99. Pefanis E, Wang J, Rothschild G, Lim J, Kazadi D, Sun J, et al. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* [Internet]. Elsevier Inc.; 2015;161(4):774–89. Available from: <http://dx.doi.org/10.1016/j.cell.2015.04.034>
100. Ginno PA, Lim YW, Lott PL, Korf I, Chédin F. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *TL - 23. Genome Res* [Internet]. 2013;23 VN-r(10):1590–600.
101. Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Mol Cell*. 2012;45(6):814–25.
102. Tomas L. Instability and decay of the primary structure of DNA.
103. Bhatia V, Barroso SI, García-Rubio ML, Tumini E, Herrera-Moyano E, Aguilera A. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* [Internet]. 2014 Jun 1 [cited 2014 Jul 9];
104. Balk B, Maicher A, Dees M, Klermund J, Luke-Glaser S, Bender K, et al. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* [Internet]. Nature Publishing Group; 2013;20(10):1199–205.
105. Yu T-Y, Kao Y-W, Lin J-J. Telomeric transcripts stimulate telomere recombination to suppress senescence in cells lacking telomerase. *Proc Natl Acad Sci U S A* [Internet]. 2014;111(9):3377–82.
106. Lin Y, Dent SYR, Wilson JH, Wells RD, Napierala M. R loops stimulate genetic instability of CTG.CAG repeats. *Proc Natl Acad Sci U S A*. 2010;107(2):692–7.

107. Grabczyk E, Mancuso M, Sammarco MC. A persistent RNA:DNA hybrid formed by transcription of the Friedreich ataxia triplet repeat in live bacteria, and by T7 RNAP in vitro. *Nucleic Acids Res.* 2007;35(16):5351–9.
108. Loomis EW, Sanz LA, Chédin F, Hagerman PJ. Transcription-Associated R-Loop Formation across the Human FMR1 CGG-Repeat Region. *PLoS Genet.* 2014;10(4).
109. Moreira M-C, Klur S, Watanabe M, Németh AH, Le Ber I, Moniz J-C, et al. Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat Genet.* 2004;36(3):225–7.
110. Chen Y-Z, Bennett CL, Huynh HM, Blair IP, Puls I, Irobi J, et al. DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet [Internet].* 2004;74:1128–35.
111. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol [Internet].* 2010;11(3):220–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20177397>
112. Stirling PC, Chan YA, Minaker SW, Aristizabal MJ, Barrett I, Sipahimalani P, et al. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes Dev.* 2012;26(2):163–75.
113. Li X, Niu T, Manley JL. The RNA binding protein RNPS1 alleviates ASF / SF2 depletion-induced genomic instability The RNA binding protein RNPS1 alleviates ASF / SF2 depletion-induced genomic instability. 2007;2108–15.
114. Jackson BR, Noerenberg M, Whitehouse A. A Novel Mechanism Inducing

Genome Instability in Kaposi's Sarcoma-Associated Herpesvirus Infected Cells.

PLoS Pathog. 2014;10(5).

Chapter 2:
Flap endonuclease 1 limits telomere fragility
on the leading strand

Daniel C. Teasley, Shankar Parajuli, Mai Nguyen, Hayley R. Moore,
Elise Alspach, Ying Jie Lock, Yuchi Honaker, Abhishek Saharia,
Helen Piwnica-Worms, and Sheila A. Stewart

Shankar Parajuli was a contributing author to this work.

This chapter was originally published in *The Journal of Biological Chemistry*, 2015; 290: 15133–15145. © the American Society for Biochemistry and Molecular Biology

Introduction

DNA replication and repair are high-fidelity processes that maintain genome stability. Due to the importance of these processes, robust mechanisms have evolved to ensure they are completed even when components of the replication and repair pathways are compromised or absent due to mutation. In some instances, this compensation is inadequate. Indeed, mutations in specific replication or repair proteins give rise to genetic disorders such as ataxia telangiectasia, Bloom syndrome, and Fanconi anemia. Cells from these patients reveal that while gross DNA metabolism continues largely unabated, mild replication defects and sensitivity to DNA damaging agents or ionizing radiation contribute to genomic instability and increased cancer incidence (1,2).

While the redundancy of replication and repair mechanisms ensures faithful replication of the bulk genome, regions with repetitive sequence or an ability to form secondary structures are problematic and thus particularly sensitive to mutations in DNA replication and repair proteins (3). This is best illustrated at common fragile sites, where replication stressors lead to replication defects and genomic instability. Why particular regions of the genome manifest as fragile sites remains obscure, but insufficient replication origins, repetitive sequences, and replication–transcription interference have all been implicated (4–6).

Recently, telomeres have also been described as fragile sites because treatment with aphidicolin, a potent inducer of replication stress, results in reduced replication fork progression and abnormal telomere structures (7,8). In checkpoint-competent backgrounds, aphidicolin treatment increases telomere fragility by 1.5 to 4.5-fold (7–9),

while suppression of the ataxia and telangiectasia and Rad3-related (ATR) kinase is sufficient to induce a 1.7-fold increase in telomere fragility in murine Seckel cells (9). Telomere fragility is also induced in the absence of telomere-binding proteins that participate in telomere replication. Indeed, knockout of the Shelterin complex member TRF1, which is required for replication fork progression through the telomere, increases the rate of telomere fragility in murine cells by 3.0-to 4.5-fold (7,8,10); similarly, depletion of the CST complex members CTC1 or STN1, which are important for replication fork restart at the telomere, causes between a 2.0- and 3.0-fold increase in telomere fragility in human cells (11).

DNA replication and repair proteins are also important in maintaining telomere stability by preventing or suppressing telomere fragility. We previously reported that depletion of flap endonuclease 1 (FEN1) results in a 2.0-fold increase in telomere fragility (12). Loss of the DNA glycosylase Nth1, which participates in the repair of oxidative stress-induced lesions, causes a 1.8-fold increase in telomere fragility (13). Helicases and topoisomerases also play roles in reducing telomere fragility. Depletion of Topolla causes up to an approximately 7-fold increase in telomere fragility, and depletion of the RecQ helicase BLM induces a 1.9-fold increase in telomere fragility (7,14). Similarly, RTEL1 depletion or deletion induces 2.3-fold and 4.0-fold increases in telomere fragility, respectively (7,10). These studies demonstrate the wide range of genetic manipulations that can induce telomere fragility with varying levels of severity.

The mechanism(s) by which telomere fragility occurs is not clear, but the large number of proteins implicated in the phenotype suggests that multiple mechanisms exist. G-

quadruplexes may play a role, as telomere fragility induced by RTEL1 deletion is exacerbated by treatment with the G-quadruplex-stabilizing drug TMPyP4 (10). Given these data, if the molecular event inducing telomere fragility occurs after the replication fork has passed, RTEL1-induced telomere fragility would be expected to exhibit lagging strand specificity; however, few studies have examined strand-specific telomere fragility. Sfeir et al. examined TRF1 knockout mouse cells using chromosome-orientation fluorescent in situ hybridization (CO-FISH), which is capable of distinguishing telomeres replicated by the leading versus lagging strand DNA replication machinery; they found that telomere fragility induced by loss of TRF1 did not exhibit strand specificity (7). Similarly, Chawla et al. identified UPF1, an ATPase and helicase associated with cytoplasmic RNA quality control, as a telomere binding protein; in UPF1-depleted cells, telomere fragility increased at both the leading and lagging strands, with a slightly larger increase observed at the leading strand (15). Most recently, Arora et al. demonstrated that ectopic expression of ribonuclease H1 (RNase H1) reduced fragile telomere formation on the leading strand in alternative lengthening of telomeres (ALT)-positive cells (16).

Among the stressors the replisome encounters, transcription has a significant impact on DNA replication. Indeed, head-on collisions between the replisome and RNA polymerase (RNAP) are extremely damaging to the replication process (17). In contrast to head-on collisions, co-directional replisome–RNAP collisions in bacteria are more common and better tolerated by the cell (18,19). This may be due to a mechanism recently elucidated in viral and prokaryotic polymerases: following a co-directional collision with RNAP on the leading strand-replicated DNA, DNA polymerase III is

removed from the template, moves forward to the 3' end of the nascent transcript, displaces RNAP, and restarts DNA synthesis using the transcript as a primer (20). Despite this mechanism, which would seem to permit damage-free replication across a region being transcribed, co-directional collisions between the replisome and RNAP can lead to unresolved RNA:DNA hybrids. If such collisions occur in mammalian cells, the RNA:DNA hybrids left behind would likely lead to DNA double strand breaks, an ataxia telangiectasia mutated (ATM)-mediated DNA damage response (DDR), and genomic instability (21,22). Thus, robust mechanisms would need to evolve to remove the RNA:DNA hybrids produced by a collision event.

The known role of FEN1 in limiting telomere fragility (12), as well as the idea that telomere fragility might be the result of replication stress or interference with transcription, led us to explore the mechanism by which FEN1 reduces telomere fragility. We show that treatment with α -amanitin, which reduces the rate of RNAP elongation and thus may increase the rate of stochastic co-directional replisome–RNAP collisions, exacerbates the telomere fragility induced upon FEN1 depletion. Additionally, we find that the telomere fragility phenotype induced by FEN1 depletion and collision induction is RNA:DNA hybrid-dependent by rescuing telomere fragility with ectopic expression of RNase H1. FEN1's role in limiting telomere fragility is distinct from its role in limiting sister telomere loss, as FEN1 depletion-induced telomere fragility is restricted to the leading strand. Neither FEN1's classical replication role as mediated by its interaction with proliferating cell nuclear antigen (PCNA), nor FEN1's DNA repair function mediated by its C-terminal interactions with numerous repair proteins are required for its activity in limiting telomere fragility. We find that FEN1's gap

endonuclease and exonuclease activities are also dispensable for limiting telomere fragility, but that FEN1's flap endonuclease activity is required. Our data support a model in which co-directional replisome–RNAP collisions on the leading strand-replicated telomere produce RNA:DNA hybrid/flap structures that accumulate in the absence of FEN1. We propose that FEN1, a classical lagging strand replication protein, acts on the leading strand during telomere replication to resolve RNA:DNA hybrid/flap structures resembling Okazaki fragment substrates; in the absence of this activity, the subsequent replication stress and DNA damage manifests as telomere fragility. We believe this to be the first report placing an Okazaki fragment-processing protein explicitly on the leading strand during DNA replication.

Experimental Procedures

Cell culture

Cells were cultured at 37 °C in 5% carbon dioxide and atmospheric oxygen, as reported previously (12,23,24). 293T cells and HEK 293 cells were cultured in high-glucose Dulbecco's modified Eagle's medium containing 10% heat-inactivated fetal bovine serum (Δ FBS) and 1% penicillin/streptomycin (P/S) (Sigma-Aldrich, St. Louis, MO). BJ fibroblasts and BJ fibroblasts expressing Large T Antigen (BJL) were cultured in high-glucose Dulbecco's modified Eagle's medium containing 15% Medium 199 (HEPES modification), 15% Δ FBS, and 1% P/S. RPE1 cells were cultured in Dulbecco's modified Eagle's medium (F12 modification) containing 7.5% Δ FBS and 1% P/S (Sigma-Aldrich, St. Louis, MO). Treatment with α -amanitin (Sigma-Aldrich, St. Louis, MO) was performed at 10 μ g/mL for 12 hours prior to collection. All cell cultures were verified free of *Mycoplasma* contamination by PCR analysis. RPE1 cells were obtained from ATCC;

all other cells were obtained from Dr. Robert Weinberg (Massachusetts Institute of Technology).

Virus production and infections

Lentiviral production and transductions were carried out as reported previously (25). Briefly, 293T cells were transfected with an 8:1 ratio of pHR'-CMV-8.2ΔR packaging plasmid and pCMV-VSV-G, and a pLKO.1-puro plasmid carrying an shRNA using TransIT-LT1 (Mirus Bio, Madison, WI). Supernatant-containing virus was collected 48 hours post-transfection and 72 hours post-transfection and filtered through a 0.45-μm PVDF membrane. Target cells were infected for four hours each on two consecutive days in the presence of 8 μg/mL protamine sulfate (Sigma-Aldrich, St. Louis, MO). Following infection, transduced BJ and BJL cells were selected with 1 μg/mL puromycin sulfate (Sigma-Aldrich, St. Louis, MO); transduced RPE1 cells were selected with 15 μg/mL puromycin sulfate.

Production of recombinant adenovirus type 5 was carried out using the AdEasy adenoviral vector system (Agilent Technologies, La Jolla, CA) according to the manufacturer's protocol. Following collection of primary adenoviral stock, secondary and tertiary viral stocks were prepared by sequential infection of HEK 293 cells and purification from a cesium gradient. Briefly, infected cells were lysed in 0.5% Nonidet P-40 and cell debris was cleared by centrifugation. Viral particles were precipitated from the lysate with 6.7% PEG 8000, 0.83 M sodium chloride, collected by centrifugation, and washed in PBS. Viral particles were suspended in 1.32 g/mL cesium chloride and centrifuged at 33,000 rpm for 18 hours at 4 °C in a swinging-bucket rotor. Intact viral

particles were collected from the cesium gradient, dialyzed in PBS, suspended in 33% glycerol, and frozen. Viral stocks were quantified using the AdEasy viral titer kit (Agilent Technologies, La Jolla, CA) according to the manufacturer's instructions.

Adenoviral transduction was carried out following lentiviral transduction. Cells were lifted, combined with concentrated adenovirus, and re-plated in media containing puromycin to select for lentiviral integration. Adenovirus was used at a multiplicity of infection of 20 on RPE1 cells. Following 48 hours of simultaneous selection and adenoviral infection, the media was replaced.

Western blot analysis

Western blots were conducted as described previously (26). Briefly, cells were washed with PBS and lysed in mammalian cell lysis buffer (100 mM sodium chloride, 50 mM tris-HCl pH 8, 5 mM EDTA, 0.5% Nonidet P40) supplemented with 2 mM dithiothreitol, 1 mM Microcystin-LR, 2 mM phenylmethylsulfonyl fluoride, 1 mM sodium fluoride, 1 mM sodium orthovanadate, protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO), and phosphatase inhibitor cocktail set I (EMD Millipore, Billerica, MA). Following centrifugation, clarified lysate was quantified using the Bio-Rad Protein Assay (Bio-Rad, Hercules, CA). Lysates were resolved by SDS-PAGE and transferred to PVDF membranes for western blotting. The following antibodies were used: mouse monoclonal anti-Chk1 (sc8408, Santa Cruz Biotechnology, Santa Cruz, CA); rabbit monoclonal anti-Chk1, phospho-S345 (2348, Cell Signaling Technology, Danvers, MA); rabbit polyclonal anti-FEN1 (A300-255A, Bethyl Laboratories, Montgomery, TX); mouse monoclonal anti-RNase H1 (H00246243-M01, Novus Biologicals, Littleton, CO); rat

monoclonal anti- α -tubulin (ab6160, Abcam, Cambridge, UK); mouse monoclonal anti- β -catenin (610154, BD Biosciences, San Jose, CA); rabbit polyclonal anti- γ H2AX (07-164, Millipore, Billerica, MA).

Metaphase chromosome preparation

Metaphase chromosome spreads were prepared as described previously (27). Briefly, BJ and BJL fibroblasts were cultured in the presence of 0.1 μ g/mL colcemid (Sigma-Aldrich, St. Louis, MO) for five hours; RPE1 cells were cultured in the presence of 0.3 μ g/mL colcemid for four hours. Following arrest, metaphase cells were collected by mitotic shake-off, swollen in 75 mM potassium chloride, and fixed in 3:1 methanol:acetic acid. Chromosomes were spread by dropping onto glass slides and aged for 18 hours at 65 °C. When metaphases were to be analyzed by CO-FISH, 0.3 μ g/mL of 5-bromo-2'-deoxyuridine (Sigma-Aldrich, St. Louis, MO) and 0.1 μ g/mL of 5-bromo-2'-deoxycytidine (MP Biomedicals, Santa Ana, CA) were added to the culture media 18 hours prior to collection of the cells.

Fluorescent in situ hybridization (FISH)

FISH was performed as described previously (27). Metaphase chromosomes were probed with a Cy3-(CCCTAA)₃ (telomere) peptide nucleic acid (PNA) probe at 0.03 μ g/mL and a FAM-CENPB (centromere) PNA probe at 0.03 μ g/mL (PNA Bio, Thousand Oaks, CA) and mounted using ProLong Gold (Life Technologies, Grand Island, NY) with 125 ng/mL DAPI (Sigma-Aldrich, St. Louis, MO).

Chromosome-orientation FISH (CO-FISH)

CO-FISH was conducted as described previously (28) with modifications. Briefly, metaphase chromosomes were rehydrated and treated with 100 µg/mL RNase for 10 minutes at 37 °C, rinsed, and re-fixed in 4% paraformaldehyde for 10 minutes at room temperature. Chromosomes were UV sensitized in 0.5 µg/mL Hoechst 33258 (Sigma-Aldrich, St. Louis, MO) in 2x SSC for 15 minutes and exposed to 365 nm UV light for 60 minutes using a UV crosslinker (Vilber-Lourmat, Marne-la-Vallée, France). Chromosomes were then digested with 3 U/µL exonuclease III (Promega, Madison, WI) for 15 minutes at room temperature, denatured in 70% formamide in 2x SSC at 72 °C for 90 seconds, and immediately dehydrated in cold ethanol before hybridization. Metaphase chromosomes were probed first with a FAM-(TTAGGG)₃ (leading strand telomere) PNA probe at 0.03 µg/mL, then probed with a Cy3-(CCCTAA)₃ (lagging strand telomere) PNA probe at 0.03 µg/mL (PNA Bio, Thousand Oaks, CA) and mounted as described for FISH.

Immunofluorescence (IF) and IF-FISH

IF was carried out as described (29). For IF-FISH, following the completion of IF, the cells were probed as described for chromosomes above using a Cy3-(CCCTAA)₃ (telomere) PNA probe at 0.03 µg/mL (PNA Bio, Thousand Oaks, CA). Antibodies used were: rabbit polyclonal anti-γH2AX (07-164, Millipore, Billerica, MA) and goat anti-rabbit IgG–Alexa Fluor 488 (Life Technologies, Grand Island, NY).

Fluorescence imaging

Chromosomes were imaged on a Nikon 90i epifluorescence microscope using a 100x 1.40 NA Plan Apo VC objective (Nikon Instruments, Melville, NY) with Cargille Type FF or Cargille Type LDF immersion oil (Cargille-Sacher Laboratories, Cedar Grove, NJ) at room temperature. Cells were imaged using a 40x 1.0 NA Plan Apo objective (Nikon Instruments, Melville, NY) under the same conditions as those for chromosomes. Filter cube sets used were: DAPI-1160B-000-ZERO, FITC-2024B-000-ZERO, and CY3-4040C-000-ZERO (Semrock, Inc., Rochester, NY). Images were captured using a CoolSnap HQ2 CCD camera (Photometrics, Tucson, AZ). Individual channel lookup tables were auto-adjusted non-destructively and linearly, and images were deconvolved with a blind algorithm using NISElements AR (Nikon Instruments, Melville, NY) prior to quantification.

RNA preparation and northern hybridization

RNA was prepared using TRI Reagent (Life Technologies, Grand Island, NY). RNA was serially diluted, denatured as previously described (30), and spotted onto a Hybond-XL charged nylon membrane (GE Healthcare, Waukesha, WI) using a Bio-Dot Microfiltration apparatus (Bio-Rad, Hercules, CA) according to the manufacturers' instructions. Samples were also treated with ribonuclease A (Roche Applied Science, Penzberg, Germany) and spotted to identify any DNA contamination in the RNA preparation. Following UV crosslinking, the membrane was prehybridized in northern hybridization buffer (15% formamide, 1% BSA, 100 mM sodium phosphate pH 7.7, 1 mM EDTA, 7% SDS) for one hour at 65 °C. A purified 1.6 kb fragment consisting exclusively of vertebrate telomere repeats was random prime-labeled with [α -³²P]dCTP (3000 Ci/mmol) using the High Prime DNA Labeling Kit (Roche Applied Science,

Penzberg, Germany) according to the manufacturer's instructions to produce a telomere-specific DNA probe. Similarly, a purified cDNA of the human 5S ribosomal RNA was random prime-labeled to produce a 5S rRNA-specific DNA probe. Probes were purified using Illustra ProbeQuant G-50 Micro Columns (GE Healthcare, Waukesha, WI) and diluted to 1.2×10^6 dpm/mL in 10 mL of northern hybridization buffer. Probes were hybridized to the membrane overnight at 65 °C, after which the membrane was washed and imaged using either autoradiography or a storage phosphor screen and imager. Quantitation was performed in Fiji by first background subtracting the image and then computing the integrated density for each spot.

Quantitative reverse transcription PCR (qRT-PCR)

For qRT-PCR, cDNA was synthesized using Superscript III reverse transcriptase (Life Technologies, Grand Island, NY) according to the manufacturer's instructions. qRT-PCR was conducted using TaqMan Gene Expression Assays (Life Technologies, Grand Island, NY) according to the manufacturer protocol. Target genes used for verification of α -amanitin efficacy were MYC (Hs00153408_m1) and SIAH1 (Hs02339360_m1).

Statistical analysis

Telomere fragility events were defined as chromatid arms with telomere FISH signal observed as either multiple telomere signals or elongated smears as previously described (7). Fragility was counted in metaphase chromosome spreads; for each experimental condition, a minimum total of 600 chromosomes was counted. The minimum sample size was chosen based on its ability to consistently detect aphidicolin-induced and FEN1 depletion-induced telomere fragility. Chromosomes completely lacking telomere FISH signal or exhibiting no strand specificity in CO-FISH (indicating

the technical issue of incomplete brominated strand digestion) were excluded and not quantified. Image groups were blinded prior to quantification. Two or more independent biological replicates were carried out for each experiment.

Where data are shown as representative, the telomere fragility rate was computed for each metaphase chromosome spread (% fragile telomeres), and each experiment was statistically analyzed. Where data are shown as combined, telomere fragility rates were computed for each metaphase chromosome spread, and a normalized value was computed for each metaphase chromosome spread by dividing the raw value by the mean of the control values. The mean of the normalized values from each sample in two independent experiments was computed and graphed with error bars representing the standard error of the mean. For statistical analysis, raw values were centered by computing a *t*-statistic for each data point: the centered value for each chromosome spread was calculated by dividing the residual of each raw value relative to the control sample's mean by the median absolute deviation of the control values. Centered values from two independent experiments were then combined for statistical analysis. Data are represented either by scatter plots with mean and standard error of the mean marked by a line and error bars, or by a bar graph with bars indicating the mean and error bars indicating standard error of the mean marked.

For IF, γ H2AX foci were counted in each nucleus. A minimum of 30 nuclei was counted for each condition in an experiment, and two independent biological replicates were combined for data quantification. Data are represented by a box and whiskers plot with the box marking 25th and 75th percentiles, line marking the median, whiskers marking

the 5th and 95th percentiles, and dots marking data points outside the 5–95 percentile range.

For all data, *p*-values were computed using a two-tailed Mann-Whitney *U* test with a 95% confidence interval in Prism 5 (GraphPad Software, La Jolla, CA). The Mann-Whitney *U* test was chosen because not all samples exhibited normal distributions. All figures except the box and whiskers plot include standard error of the mean as an indicator of variance, and in all cases the variance within samples was similar.

Results

FEN1 depletion and transcription inhibition induce replication stress and a DNA damage response

Because telomeres are transcribed to produce telomeric repeat-containing RNA (TERRA) (31,32), and because interference between replication and transcription is a known cause of genomic instability (5,33,34), we turned our attention to the impact that putative collisions between the replication and transcription machinery would have on telomere stability. Unlike in *Schizosaccharomyces pombe*, where transcription of telomeres and subtelomeres occurs using both strands as templates (35), mammalian telomeres are transcribed exclusively using the C-rich leading strand as a template (31,32); as such, co-directional collisions are the only type that can occur at the telomere. In bacteria, co-directional collisions are resolved by a mechanism that leaves behind an RNA:DNA hybrid/flap structure (20) which would need to be resolved in a eukaryotic cell to avoid a DDR and genomic instability (21,22). FEN1 has been previously shown to reduce telomere fragility (12), and the yeast FEN1 homolog

Rad27p can hydrolyze RNA flaps (36). We hypothesized that co-directional collisions are a molecular origin of telomere fragility, and that FEN1 can prevent post-collision RNA:DNA hybrid/flap structures from accumulating, causing damage, and ultimately leading to fragile telomere formation.

To address this hypothesis, we first examined whether increasing the rate of stochastic collisions between the replisome and RNAP would increase replication stress and trigger a DDR in the context of FEN1 depletion. We treated BJ fibroblasts with the RNA polymerase II (Pol II) elongation inhibitor α -amanitin, a cyclic peptide toxin that reduces the rate of Pol II transcription approximately 100-fold, allowing chain elongation to continue without triggering transcript cleavage (37,38). We expected α -amanitin treatment to increase the frequency of stochastic collisions between the replisome and RNAP and thus increase replication stress and telomere fragility. Following transduction with a validated shRNA targeting the 3'-untranslated region of the FEN1 mRNA (shFEN1) (24) or a control hairpin (shCtrl), we treated BJ fibroblasts with either vehicle or 10 μ g/mL α -amanitin for 12 hours and collected both RNA and protein lysates from the cells. qRT-PCR analysis confirmed the efficacy of α -amanitin treatment by quantitation of two short-lived transcripts, c-Myc and SIAH1. α -amanitin-treated control cells retained 2% and 6% of the c-Myc and SIAH1 mRNAs, respectively, compared to the levels observed in vehicle-treated control cells. Similarly, in FEN1-depleted cells, α -amanitin treatment resulted in 4% and 10% of the levels of c-Myc and SIAH1 mRNAs, respectively, compared to vehicle-treated cells (Fig. 2.1A). Since transcription inhibition by α -amanitin might reduce steady-state TERRA levels and produce telomere phenotypes as a result of decreased TERRA, we carried out a northern blot analysis of

total RNA to detect TERRA. Because TERRA are expressed at low levels in BJ fibroblasts, we utilized a dot blot rather than a gel to maximize signal intensity and hybridized the membrane to a telomere repeat-specific probe; treatment with ribonuclease A was used to show the absence of contaminating DNA, and a 5S rRNA-specific probe was used as a loading control. Northern analysis of vehicle- and α -amanitin-treated cells demonstrated that the α -amanitin treatment conditions subsequently used for western and metaphase analysis did not impact steady state levels of TERRA in our system, demonstrating that the phenotypes resulting from the treatment were not due to a loss of TERRA, which are known to impact telomere stability (39,40) (Fig. 2.1B,C).

To determine if Pol II inhibition induces replication stress and a DDR in the context of FEN1 depletion, we performed western blot analysis to examine phosphorylation of Chk1 at S345 and phosphorylation of histone H2AX at S139 (γ H2AX), classical markers for the replication stress response and DDR, respectively. BJ fibroblasts transduced with the control hairpin and treated with vehicle displayed neither Chk1 phosphorylation nor H2AX phosphorylation (Fig. 2.2A). Treatment with α -amanitin induced a small but detectable increase in Chk1 phosphorylation, but did not induce γ H2AX, indicating that α -amanitin treatment can induce limited replication stress, but is not sufficient to induce a DDR (Fig. 2.2A). Similarly, BJ fibroblasts depleted of FEN1 and treated with vehicle also displayed a small level of Chk1 phosphorylation and no detectable γ H2AX (Fig. 2.2A). Strikingly, upon treatment with α -amanitin, FEN1-depleted cells showed a robust phosphorylation of Chk1 and strong induction of γ H2AX (Fig. 2.2A).

We also used IF to examine the formation of γ H2AX foci in asynchronous BJ fibroblasts, and IF-FISH to assess the formation of telomere dysfunction-induced foci. Quantification of γ H2AX foci demonstrated that while FEN1 depletion induced foci formation (2.14-fold in shFEN1+vehicle vs. shCtrl+vehicle, $p < 0.0001$), there was no change in γ H2AX foci upon treatment with α -amanitin (Fig. 2.2B,C). Furthermore, we did not observe an increase in telomere dysfunction-induced foci in response to FEN1 depletion or α -amanitin treatment (data not shown). These results indicate first that the amount of DNA damage induced in conditions that increase collision events causes a response only robust enough to be detected by the more sensitive western analysis. Second, they indicate that FEN1 depletion- and Pol II inhibition-induced replication stress and DNA damage is not restricted to telomeres; rather, DNA damage likely occurs throughout the genome wherever collisions occur. Thus, Pol II inhibition alone induces mild replication stress, and the depletion of FEN1 combined with Pol II inhibition results in a DDR that is not observed when FEN1 is depleted alone.

Inhibition of transcription exacerbates the telomere fragility observed upon FEN1 depletion

We next examined whether the replication stress and DDR phenotypes observed in response to Pol II inhibition and FEN1 depletion manifest as telomere fragility. If failure by FEN1 to resolve the structures induced by collision events between the replisome and RNAP results in fragility, then we anticipated the rate of telomere fragility in α -amanitin-treated and FEN1-depleted cells to mirror the replication stress phenotype. As before, we transduced BJ fibroblasts with either shCtrl (control) or shFEN1 and treated the cells with vehicle or α -amanitin for 12 hours prior to collecting metaphase

chromosomes. Consistent with our model, cells expressing shCtrl exhibited an increased rate of telomere fragility upon α -amanitin treatment (1.55-fold in shCtrl+ α -amanitin vs. shCtrl+vehicle, $p = 0.0079$) (Fig. 2.2D,E). When examining only the vehicle-treated cells, we found that as previously demonstrated, FEN1 depletion causes a significant increase in telomere fragility (2.15-fold in shFEN1+vehicle vs. shCtrl+vehicle, $p < 0.0001$) (Fig. 2.2D,E). Strikingly, FEN1-depleted cells treated with α -amanitin displayed a significant 2.76-fold increase in telomere fragility when compared to control, vehicle-treated cells (shFEN1+ α -amanitin vs. shCtrl+vehicle, $p < 0.0001$), and a significant 1.28-fold increase compared to FEN1-depleted, vehicle-treated cells (shFEN1+ α -amanitin vs. shFEN1+vehicle, $p = 0.0017$) (Fig. 2.2D,E). These fragility data mirror the Chk1 phosphorylation phenotype and support a model in which α -amanitin treatment increases co-directional replisome–RNAP collision events that result in structures requiring FEN1 for resolution; without FEN1, the collision events generate replication stress, a DDR, and fragile telomere formation. These experiments suggest that FEN1's role in limiting telomere fragility is dependent upon its ability to resolve structures produced by telomere transcription during DNA replication.

Leading strand-specific telomere fragility is caused by RNA:DNA hybrids

Our data above suggest a role for telomere transcription in telomere fragility induced by FEN1 depletion. Based on findings in prokaryotes, if co-directional collisions occur between the replisome and an RNAP, a structure resembling an Okazaki fragment with a segment of RNA:DNA hybrid would result (20); we postulate that if not resolved, this structure could give rise to fragile telomeres. Indeed, post-collision structures resemble R-loops, which are semi-stable displacement loops in which a nascent mRNA remains

hybridized to its DNA template, while the coding strand DNA remains single-stranded, resulting in replication stress and common fragile site expression (5). At common fragile sites, the enzyme RNase H1 suppresses replication stress phenotypes induced by R-loop formation by hydrolyzing the RNA in RNA:DNA hybrids and thus resolving displacement loops (5). We reasoned that since the post-co-directional collision structure resembles an R-loop, RNA:DNA hybrids might be responsible for telomere fragility, and thus ectopic expression of RNase H1 should resolve the structure and telomere phenotype. Additionally, because our model predicts that the causative structure for fragile telomere formation occurs after the replication fork has passed the locus in question, we wondered if the telomere fragility observed upon FEN1 depletion manifests only on the leading strand, where collisions could occur. This question was especially prescient given that FEN1 is canonically a lagging strand replication protein, and has a previously established role in limiting sister telomere loss at the lagging strand (12).

Following lentiviral transduction with a control hairpin (shCtrl) or FEN1-depleting hairpin (shFEN1), we transduced RPE1 cells with RNase H1 (Ad-RH1) (Fig. 2.3C) and collected cells for protein analysis and metaphase chromosome preparation. To identify if telomere fragility exhibited strand specificity, we used CO-FISH, a technique which exploits the fact that the C-rich and G-rich strands of the mammalian telomere are replicated exclusively by the leading and lagging strand machinery, respectively, allowing the use of strand-specific probes to identify which machinery replicated a given telomere on a metaphase chromosome (28). Strikingly, FEN1 depletion significantly increased leading strand-specific telomere fragility (2.30-fold in shFEN1 vs. shCtrl, $p =$

0.0021) (Fig. 2.3A,B) with no change observed on lagging strand-replicated telomeres (1.26-fold in shFEN1 vs. shCtrl) (Fig. 2.3A,B). Additionally, ectopic expression of RNase H1 rescued fragility on the leading strand-replicated telomere, returning fragility levels to those observed in control cells (1.19-fold in shFEN1+Ad-RH1 vs. shCtrl) (Fig. 2.3A,B). Given the specificity of RNase H1 for RNA:DNA hybrids, these data indicate that RNA:DNA hybrids lead to telomere fragility and suggest that the hybrid/flap structures that arise from co-directional collisions on the leading strand are responsible for the telomere fragility observed upon FEN1 depletion. Furthermore, given that RPE1 cells are telomerase-positive and telomerase expression rescues the sister telomere loss observed upon FEN1 depletion, these data indicate that FEN1's role in limiting telomere fragility at the leading strand is distinct from its known role in limiting sister telomere loss at the lagging strand (12,24).

α -amanitin is known to slow but not disengage the RNAP from the template strand (37,38), and its use would be expected to increase replisome–RNAP collisions and RNA:DNA hybrids. Thus, we next wanted to determine if the fragility we observed upon α -amanitin treatment was also RNA:DNA hybrid-dependent. To address this question, we transduced RPE1 cells with Ad-RH1 (Fig. 2.4C) and treated the transduced cells with α -amanitin for 12 hours prior to metaphase collection. As before, α -amanitin treatment induced an increase in telomere fragility (1.79-fold in α -amanitin vs. vehicle, $p = 0.0008$) (Fig. 2.4A,B). As in the case of telomere fragility following FEN1 depletion, ectopic RNase H1 expression protected α -amanitin-treated cells from telomere fragility, resulting in levels similar to those observed in cells treated with vehicle (1.05-fold in Ad-RH1+ α -amanitin vs. vehicle) (Fig. 2.4A,B). Because α -amanitin treatment exacerbates

telomere fragility in the absence of FEN1 (Fig. 2.2D,E), the ability of RNase H1 to rescue fragility in both α -amanitin-treated (Fig. 2.4A,B) and FEN1-depleted cells (Fig. 2.3A,B) suggests that FEN1's role in limiting telomere fragility is to resolve RNA:DNA hybrid/flap structures that are produced following replisome–RNAP collisions.

FEN1 flap endonuclease activity is required for limiting telomere fragility

Given the unprecedented finding that FEN1 limits leading strand-specific telomere fragility, we sought to identify which of FEN1's known functions were necessary for this activity. FEN1 possesses three unique enzymatic activities: an endonuclease activity on unannealed 5' flaps consisting of either DNA or RNA, a weak exonuclease activity that cleaves nicks, gaps, or recessed 5' ends of double-stranded DNA, and a gap endonuclease activity that cleaves double-stranded DNA at the 3' end of a short single-stranded gap (41–43). FEN1 is also known to interact with PCNA via a PCNA interacting peptide (PIP) box, directly pertaining to its role in DNA replication, and a number of DNA repair proteins via its C-terminus, pertaining to its role in base excision repair (44,45). We utilized a series of previously described FEN1 mutants that impact FEN1's different roles in replication (D181A, Δ P, Δ P Δ C) versus repair (Δ C, D181A, Δ P Δ C, E160D) in genetic knockdown–rescue experiments (12,24) (Fig. 2.5A). To test whether the reduction in telomere fragility mediated by FEN1 requires its DNA repair functions, we used a lentiviral vector to express shCtrl (control) alone, shFEN1 alone, or shFEN1 simultaneously with the wild type (WT), Δ C, or D181A allele of FEN1 (Fig. 2.5A); following transduction we prepared metaphase chromosomes. As before, FEN1 depletion induced leading strand-specific telomere fragility (2.05-fold in shFEN3 vs. shLuc, $p < 0.0001$) (Fig. 2.5B,D). Expression of the WT allele of FEN1 rescued the

leading strand-specific induction of telomere fragility upon endogenous FEN1 knockdown, indicating that the phenotype is specific to FEN1 knockdown (1.18-fold in shFEN1+WT vs. shCtrl) (Fig. 2.5B,D). Unexpectedly, expression of the Δ C allele also rescued FEN1 depletion-induced telomere fragility on the leading strand (1.02-fold in shFEN1+ Δ C vs. shCtrl) (Fig. 2.5B,D). In contrast to the WT and Δ C alleles, the D181A nuclease-dead allele, which is deficient in all known nuclease activities (46,47), failed to rescue the phenotype, instead resulting in an increase in leading strand-specific telomere fragility comparable to the expression of shFEN1 alone (1.83-fold in shFEN1+D181A vs. shCtrl, $p < 0.0001$) (Fig. 2.5B,D). Neither knockdown of FEN1 nor expression of any of the mutant alleles of FEN1 altered the level of telomere fragility on the lagging strand, confirming that FEN1 does not play a role in the phenotype on lagging strand-replicated telomeres (Fig. 2.5B,D). These data indicate that FEN1's flap endonuclease activity is required to limit leading strand-specific telomere fragility, but its interactions with several DNA repair proteins including WRN and BLM (deficient in the Δ C allele), and thus its DNA repair activities, are dispensable for this role. Consequently, FEN1's ability to limit leading strand-specific telomere fragility is distinct from its previously described role in telomere stability, which depends upon FEN1's C-terminally mediated DNA repair activity to suppress sister telomere loss on the lagging strand-replicated telomere (12,24).

Given that FEN1's repair activity is dispensable for its ability to limit telomere fragility, and telomere fragility is associated with replication stress, we next investigated whether FEN1's interaction with PCNA, and thus its replication activity, might be important in this role. To test this possibility, BJ fibroblasts depleted of FEN1 were transduced with the

WT, ΔP , $\Delta P\Delta C$, or E160D cDNA of FEN1 (Fig. 2.5A). Analysis of telomere fragility on metaphase chromosomes revealed that as before, expression of the WT allele rescued the leading strand-specific induction of telomere fragility following FEN1 depletion (1.58-fold in shFEN1 vs. shCtrl, $p < 0.0001$; 0.88-fold in shFEN1+WT vs. shCtrl) (Fig. 2.5C,D). Surprisingly, expression of both the ΔP and E160D constructs also rescued the fragility defect (0.77-fold in shFEN1+ ΔP vs. shCtrl; 1.20-fold in BJ shFEN1+E160D vs. shCtrl) (Fig. 2.5C,D). Only the $\Delta P\Delta C$ allele, a functionally null allele due to its lack of nuclear localization, failed to rescue the leading strand telomere fragility observed upon FEN1 depletion, resulting in an increase similar to that observed upon FEN1 depletion alone (1.61-fold in shFEN1+ $\Delta P\Delta C$ vs. shCtrl, $p < 0.0001$) (Fig. 2.5C,D). As in the previous experiment, none of the FEN1 alleles induced lagging strand-specific telomere fragility (Fig. 2.5C,D). These data indicate that FEN1 requires neither its interaction with PCNA (deficient in the ΔP allele), nor its gap endonuclease and exonuclease activity (deficient in the E160D allele) to limit leading strand-specific fragility. In combination with the data from expression of the ΔC and D181A mutants, our experiments identify FEN1 flap endonuclease activity as necessary for its role in limiting telomere fragility. These data are consistent with FEN1's known activities, as it has previously been shown to cleave flap structures with numerous modifications, including flaps composed of RNA (36,43,48). As such, our data and the literature support a model in which FEN1's flap endonuclease activity could cleave the RNA:DNA hybrid/flap structures produced following a replisome–RNAP collision event (Fig. 2.6).

Discussion

The role of FEN1 described here provides new insights into the breadth of its functions in maintaining genome stability. In addition to known roles in lagging strand DNA replication, base excision repair, and lagging strand telomere stability, we illustrate for the first time a role for FEN1 in leading strand replication. Furthermore, we have identified transcription as an important contributor to telomere fragility, and we have shown that FEN1 may resolve the RNA:DNA hybrid/flap structures resulting from collisions between the transcription and replication machinery. The strand specificity of telomere fragility observed in the absence of FEN1 shows that it has two independent molecular roles for promoting telomere stability: (1) FEN1 limits sister telomere loss at the lagging strand-replicated telomere by facilitating replication fork reinitiation (12), and (2) FEN1 limits telomere fragility at the leading strand-replicated telomere by resolving RNA:DNA hybrid/flap structures produced by co-directional replisome–RNAP collisions (Fig. 2.6).

Though co-directional collisions between the replisome and RNAP are postulated to be less deleterious to DNA replication than head-on collisions, they still necessitate mechanisms to ensure replication fidelity. In bacteria, the primary replicative helicase, DnaB, translocates along the lagging strand template as it unwinds DNA ahead of the replication fork; as such, the helicase can move past an RNAP transcribing from the leading strand, which would result in an inevitable collision between the two polymerases (20). While accessory helicases such as Rep move along the leading strand template, this activity alone cannot prevent co-directional collisions (20,49). Bacteria thus can use a mechanism in which replication restarts on the leading strand template following a co-directional collision using the 3' end of the nascent mRNA as a

primer (20). Collisions between the replisome and RNAP also present a problem to the eukaryotic cell, where highly-transcribed Pol II and Pol III genes are known to impede replication fork progression (33,34). Extremely long genes that require more than a single cell cycle to transcribe are also known to induce collision events; these collisions induce common fragile site expression (5). Observations suggest that even though the eukaryotic replicative helicase, a complex of Cdc45, Mcm2-7, and GINS (CMG), translocates along the leading strand (50), its activity is insufficient to prevent collisions from occurring. Indeed, CMG is unable to bypass both biotin-streptavidin and Qdot (20 nm) roadblocks on the leading strand (50). Even though the eukaryotic replicative helicase translocates along the leading strand, our data suggest that it is unable to bypass an RNAP and/or RNA:DNA hybrid on this strand. Together, these observations suggest that eukaryotes require a similar mechanism to that used by bacteria for the resolution of co-directional replisome–RNAP collisions on the leading strand.

Although FEN1 has no known existing roles in leading strand DNA replication, our results provide an explanation consistent with the enzyme's known substrates and activity. The putative RNA:DNA hybrid/flap structure produced following a co-directional replisome–RNAP collision is similar to the Okazaki fragment flaps FEN1 cleaves during lagging strand replication—differing only in that the flap is composed entirely of ribonucleotides. Thus, our model suggests that human FEN1 acts at the leading strand because co-directional collisions at the telomere only happen on the leading strand template. Because FEN1's ability to limit telomere fragility does not require its C-terminal domain, which interacts with the shelterin protein TRF2 to recruit FEN1 to telomeres during S and G2 phases of the cell cycle (24,51), it is unlikely that FEN1's

ability to process post-collision structures is limited to the telomere. However, in other portions of the genome where replication begins from origins to either side of a particular locus, transcription could be more coordinated with replication to prevent head-on collisions from occurring. Wherever co-directional collisions occur, FEN1 is likely able to process the structures produced.

Because the replication fork replicates the telomere in the centromere-to-telomere direction only, and because mammalian telomeres are only transcribed from the C-rich leading-strand template in the same direction (31,32), replisome–RNAP collisions at the telomere can only occur co-directionally. Our work here, as well as the fact that TERRA depletion induces telomere fragility (39), underscores the role of telomere transcription in fragile telomere formation. Indeed, work in yeast has shown that RNA:DNA hybrids produced by TERRA transcription promote recombination-mediated telomere elongation (40). In ALT-positive cells, RNase H1 has recently been shown to regulate the levels of RNA:DNA hybrids between TERRA and telomeric DNA (16). Like in yeast, TERRA RNA:DNA hybrids are hypothesized to promote recombination between ALT telomeres. In the absence of RNase H1, hybrids accumulate and promote excessive replication stress that causes fragile telomere formation and telomere loss; conversely, overexpression of RNase H1 reduces TERRA hybrids such that they cannot promote recombination, leading to progressive telomere shortening (16). Strikingly, the telomere loss that occurs following RNase H1 depletion in ALT cells is leading strand-specific (16). This work, when combined with ours, strongly implicates transcription-associated RNA:DNA hybrid formation at the telomere as a contributor to telomere fragility.

Despite the recency of telomere fragility as a defined phenotype, it has been identified in reports manipulating the expression of many proteins involved in DNA replication and telomere stability. ATR deficiency or depletion, BRCA2 deletion, RAD51 depletion, and RECQL1 depletion all induce elevated rates of telomere fragility (7–9,52,53). In addition, CTC1 and STN1, both members of the mammalian CST complex, limit telomere fragility (11). Like FEN1, these proteins participate in replication fork progression, replication fork reinitiation, and telomere stability. To our knowledge, no report has identified any perturbation that induces telomere fragility exclusive to the leading or lagging strand, though RNase H1 overexpression has been shown to reduce telomere fragility at the leading strand (16). Indeed, the lack of strand specificity in the telomere fragility produced by TRF1 deletion (7), as well as the involvement of G-quadruplexes (which form exclusively on the lagging strand) in RTEL1 deletion-induced telomere fragility (10), suggests that there are multiple mechanisms leading to fragile telomere formation. Our work underscores the complexity of DNA replication, and in placing the canonical Okazaki fragment-processing protein FEN1 at the leading strand, reveals the first molecular mechanism for fragile telomere formation on the leading strand.

Acknowledgments

We thank Dr. Peter Burgers, Dr. Susana Gonzalo, Dr. Barry Sleckman, Kevin Flanagan, and Megan Ruhland for critical reading. We thank Jingqin Luo for statistical advice. This project was supported, in whole or in part, by National Institutes of Health Grant GM95924 (S.A.S.) and by Training Grant GM007067 from NRSA (D.C.T. and E.A.). This work was also supported by the Siteman Cancer Center at Barnes-Jewish Hospital and Washington University School of Medicine (S.P. and H.R.M.).

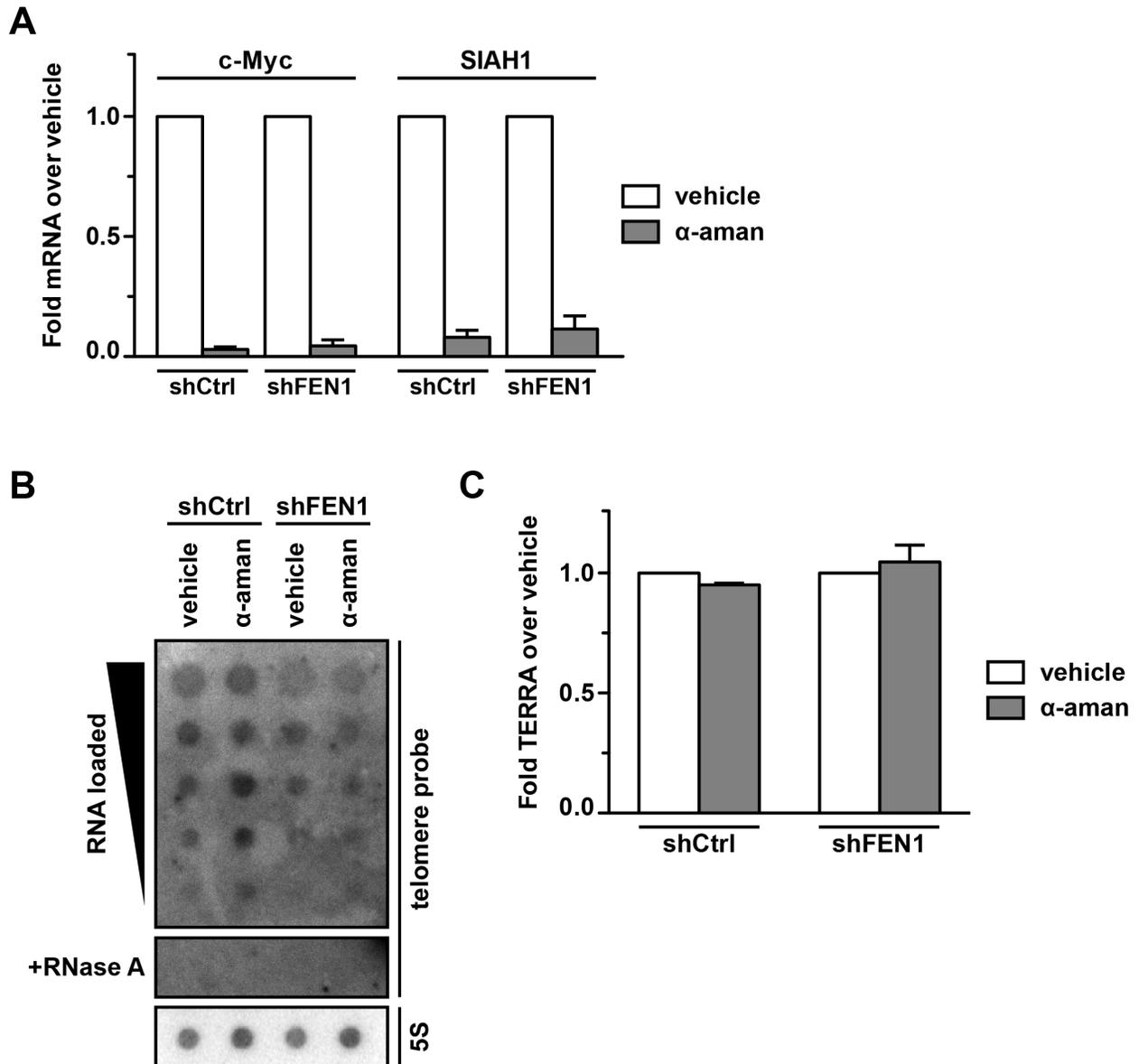


Figure 2.1. α -amanitin treatment abrogates expression of mRNAs with short half lives but does not alter steady-state TERRA levels. (A) qPCR analysis of c-Myc and SIAH1 mRNA expression in cells expressing a control hairpin (shCtrl) or FEN1-depleted cells (shFEN1), treated with either vehicle or α -amanitin (α -aman). mRNA levels in α -amanitin-treated cells are shown as a fold change relative to the vehicle-treated cells. Fold changes were calculated using the $\Delta\Delta C_t$ method; fold changes from two biological replicates were averaged to produce the graph. Error bars represent standard error of

the mean. (B) Northern dot blot to detect TERRA. RNA was isolated from cells expressing a control hairpin (shCtrl) or FEN1-depleted cells (shFEN1) that were treated with either vehicle or α -amanitin (α -aman). Serial dilutions of RNA were loaded onto a membrane. Samples treated with RNase A to control for genomic DNA contamination were also loaded (+RNase A). A telomere repeat DNA probe was hybridized to the membrane (telomere probe) to detect TERRA; the membrane was stripped and re-probed with a 5S rRNA DNA probe (5S) as a loading control. The membrane was visualized with autoradiography. (C) Quantification of TERRA in cells treated with α -amanitin. The northern dot blot in (B) was imaged with a phosphor imager and analyzed by densitometry using Fiji; TERRA levels in α -amanitin-treated cells are shown as a fold change relative to vehicle-treated cells. Two independent experiments were averaged to produce the graph; error bars represent standard error of the mean.

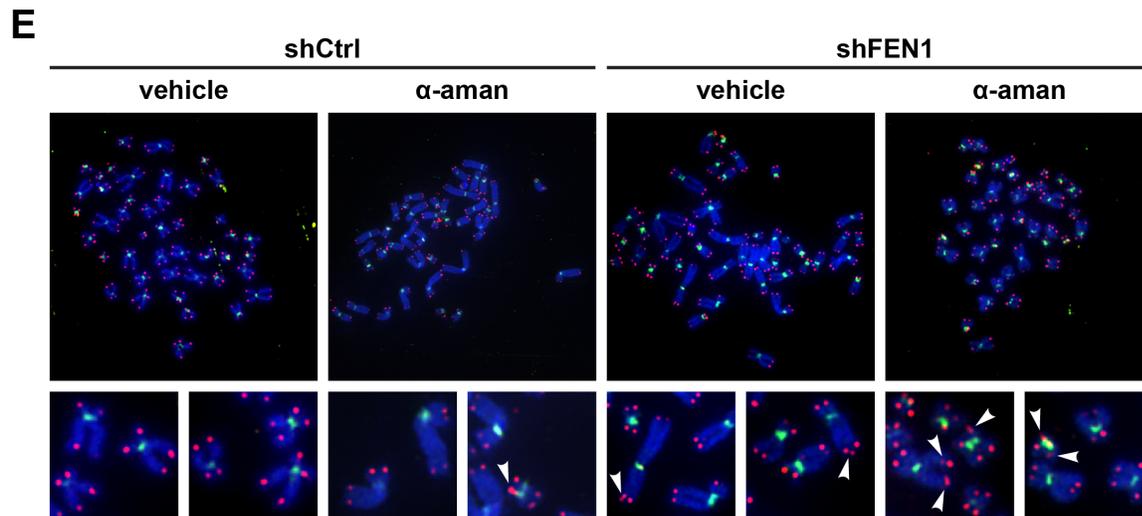
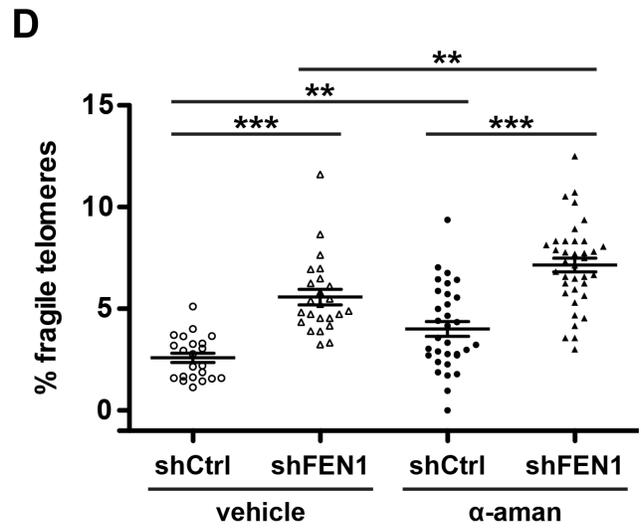
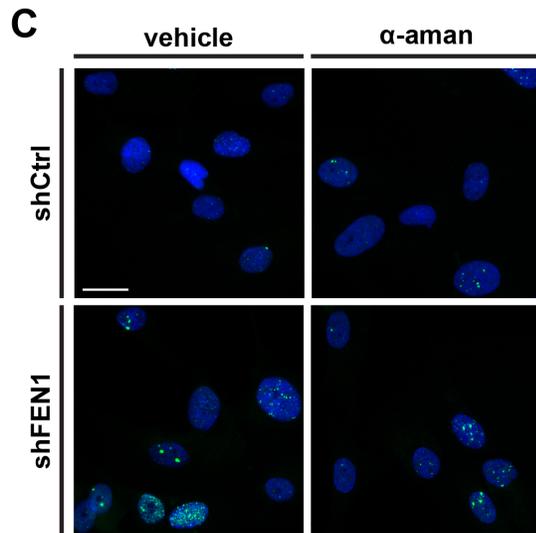
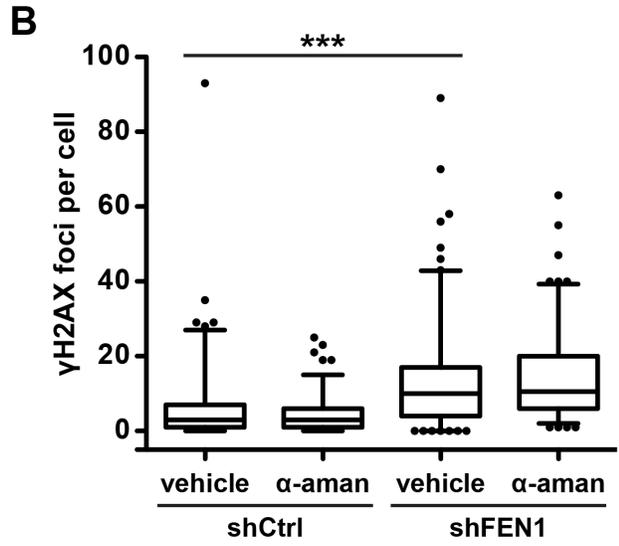
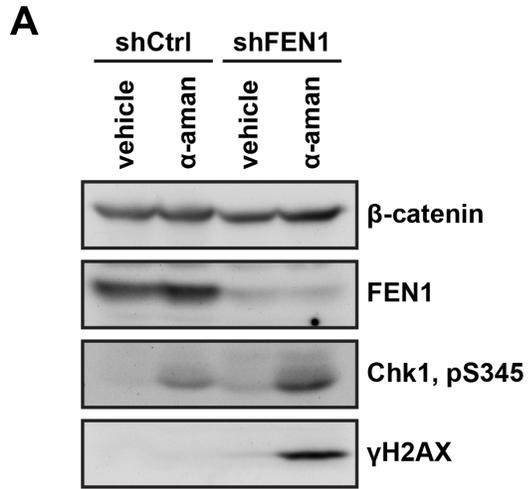


Figure 2.2. FEN1 depletion and transcription inhibition induce replication stress, a DNA damage response, and telomere fragility. (A) Western analysis of FEN1 expression, Chk1 phosphorylation (pS345), and H2AX phosphorylation (γ H2AX) in control (shCtrl) or FEN1-depleted (shFEN1) cells treated with vehicle or α -amanitin (α -aman). β -catenin is shown as a loading control. (B) Quantification of γ H2AX foci per cell. Two independent biological replicates were combined. The box marks the 25th to 75th percentile with the median marked by a horizontal line, whiskers mark the 5th and 95th percentiles, and dots represent values outside the 5–95 percentile range. *p*-values were computed using a two-tailed Mann-Whitney *U* test (***, *p* < 0.001 relative to shCtrl). (C) Representative immunofluorescence images stained with a γ H2AX antibody (green) and DAPI (blue) from BJ fibroblasts expressing a control hairpin (shCtrl) or depleted of FEN1 (shFEN1). Cells were treated with vehicle or α -amanitin (α -aman) as indicated. The scale bar (white) represents 25 μ m. (D) Representative quantification of the rate of telomere fragility. *p*-values were computed using a two-tailed Mann-Whitney *U* test (**, *p* < 0.01; ***, *p* < 0.001). Error bars represent standard error of the mean. (E) Representative metaphase chromosomes processed with FISH from BJ fibroblasts expressing a control hairpin (shCtrl) or depleted of FEN1 (shFEN1). Cells were treated with vehicle or α -amanitin (α -aman) as indicated. Centromeres are green and telomeres are red. Arrowheads mark fragile telomeres in the magnified images.

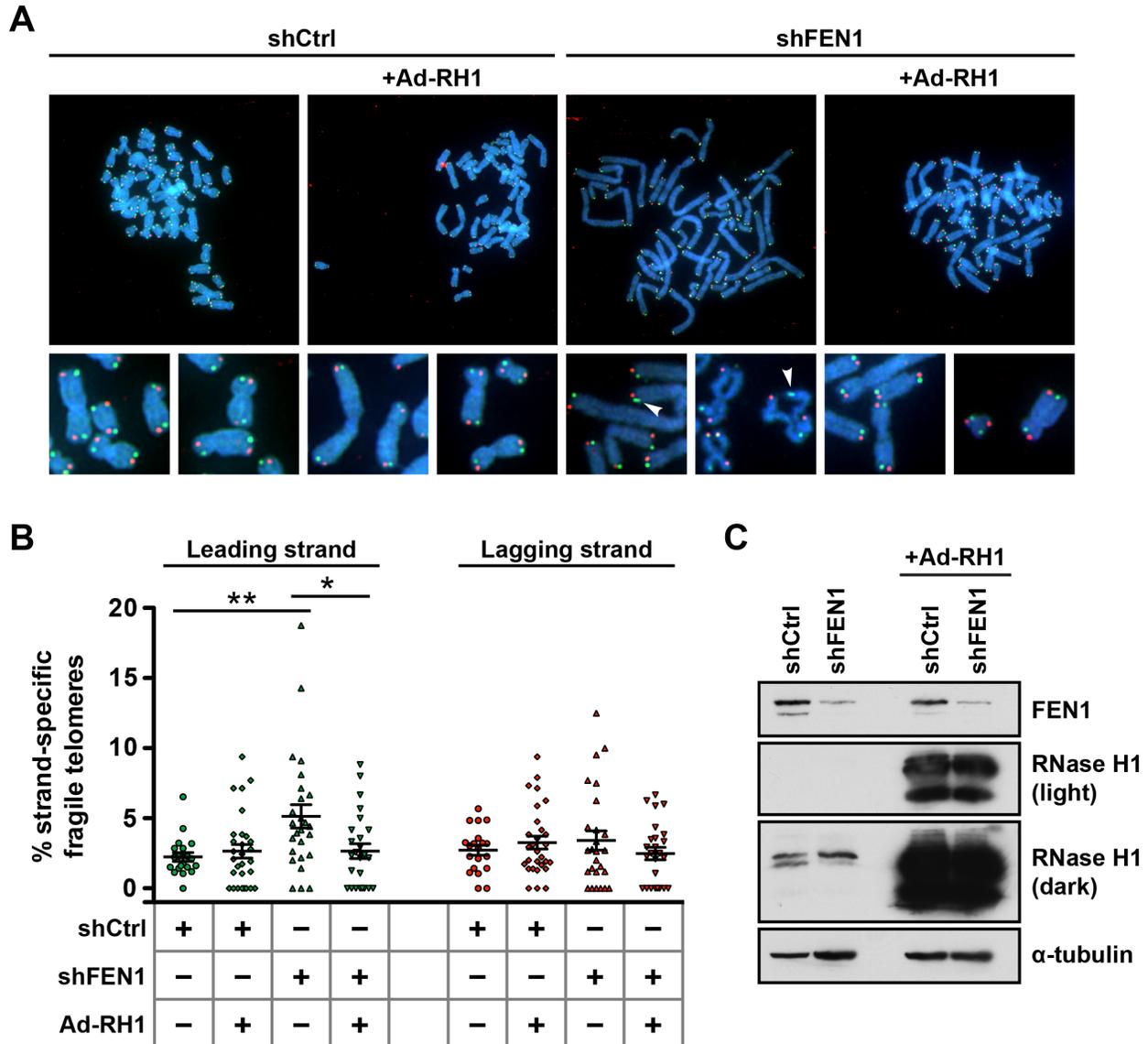


Figure 2.3. RNA:DNA hybrids are responsible for FEN1 depletion-induced leading strand-specific telomere fragility. (A) Representative metaphase chromosomes processed with CO-FISH from RPE1 cells expressing a control hairpin (shCtrl) or depleted of FEN1 (shFEN1), with or without ectopically expressed RNase H1 (Ad-RH1). Leading strand-replicated telomeres are green and lagging strand-replicated telomeres are red. Arrowheads mark fragile telomeres in the magnified images. (B) Representative quantification of the rate of strand-specific telomere fragility, with leading strand-specific telomere fragility shown in green and lagging strand-specific telomere fragility shown in

red. p -values were computed using a two-tailed Mann-Whitney U test (*, $p < 0.05$; **, $p < 0.01$). Error bars represent standard error of the mean. (C) Western analysis of FEN1 and RNase H1 expression in control (shCtrl) or FEN1-depleted (shFEN1) cells, with or without ectopically expressed RNase H1 (Ad-RH1). Two exposures of the same RNase H1 blot are shown. α -tubulin is shown as a loading control.

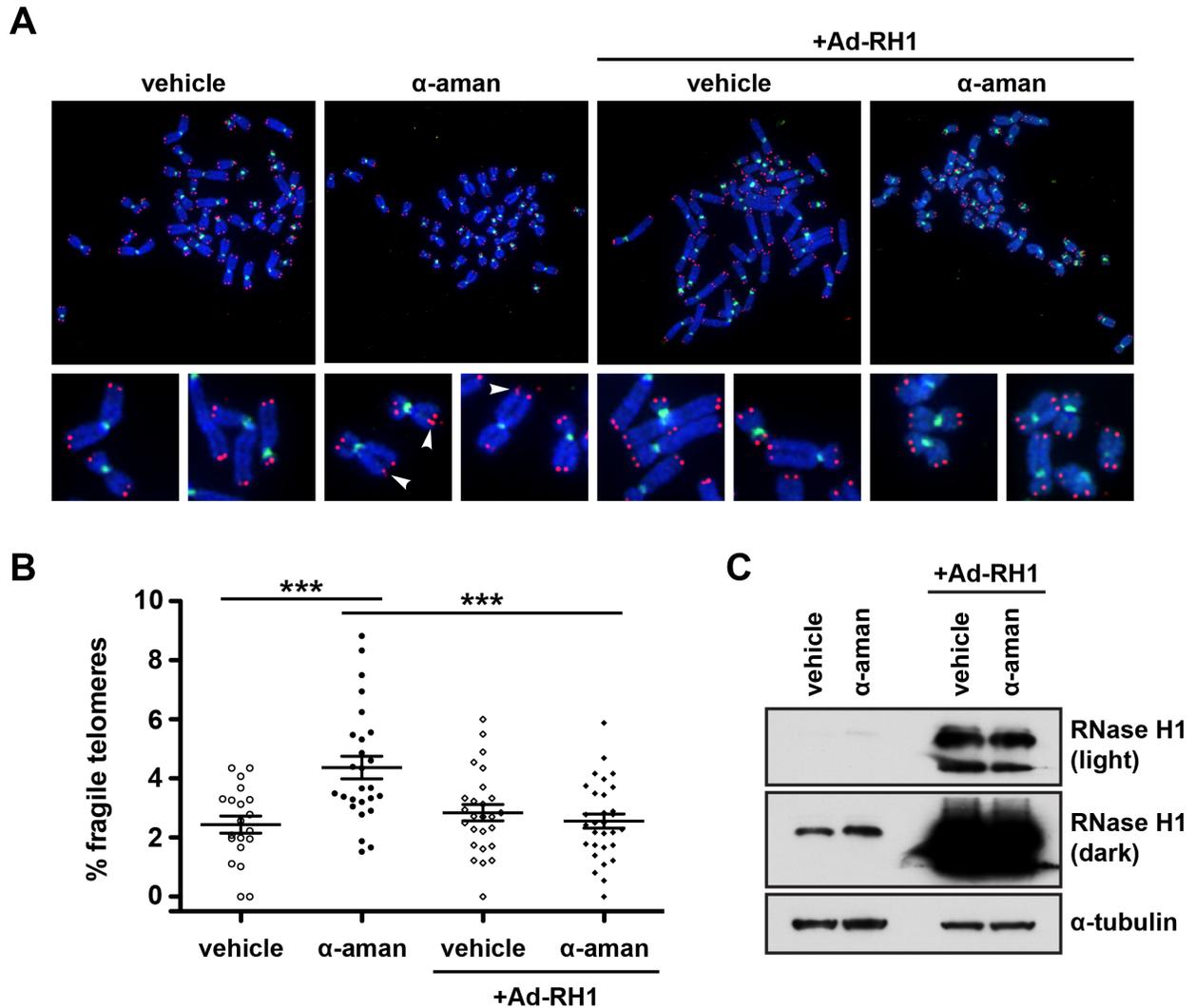


Figure 2.4. RNA:DNA hybrids are responsible for α -amanitin-induced telomere fragility.

(A) Representative metaphase chromosomes processed with FISH from RPE1 cells with or without ectopically expressed RNase H1 (Ad-RH1) and treated with either vehicle or α -amanitin (α -aman). Centromeres are green and telomeres are red. Arrowheads mark fragile telomeres in the magnified images. (B) Representative quantification of the rate of telomere fragility. p -values were computed using a two-tailed Mann-Whitney U test (***, $p < 0.001$). Error bars represent standard error of the mean. (C) Western analysis of RNase H1 expression in cells with or without ectopically expressed RNase H1 (Ad-RH1) treated with vehicle or α -amanitin (α -aman). Two

exposures of the same RNase H1 blot are shown. α -tubulin is shown as a loading control.

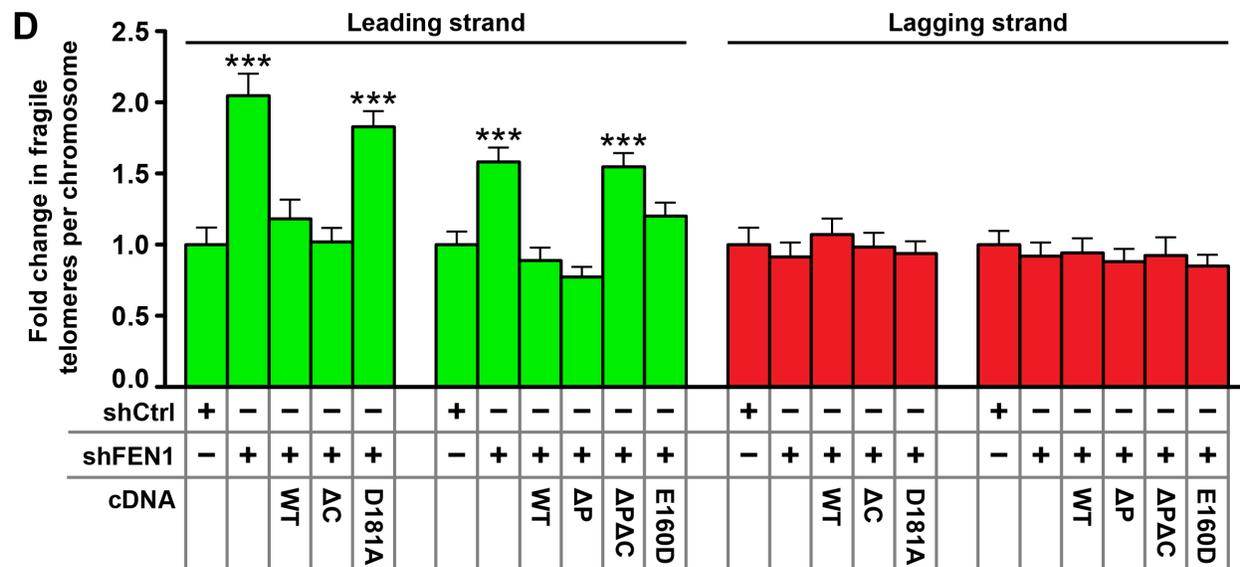
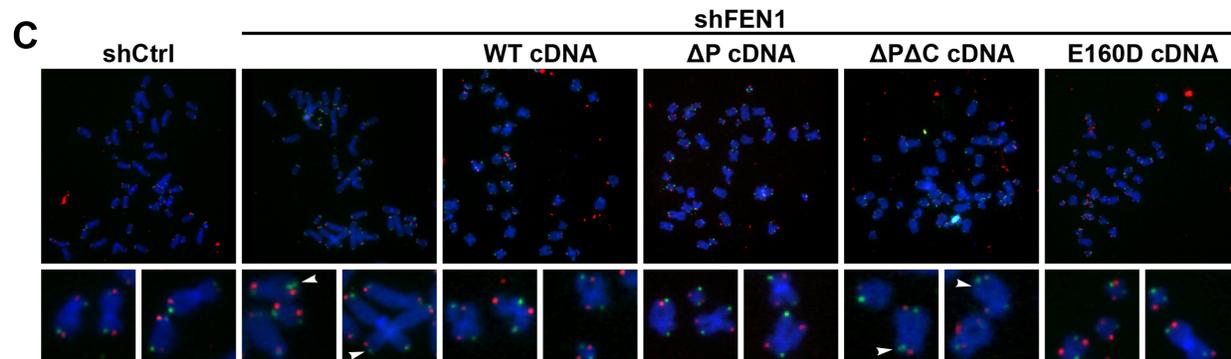
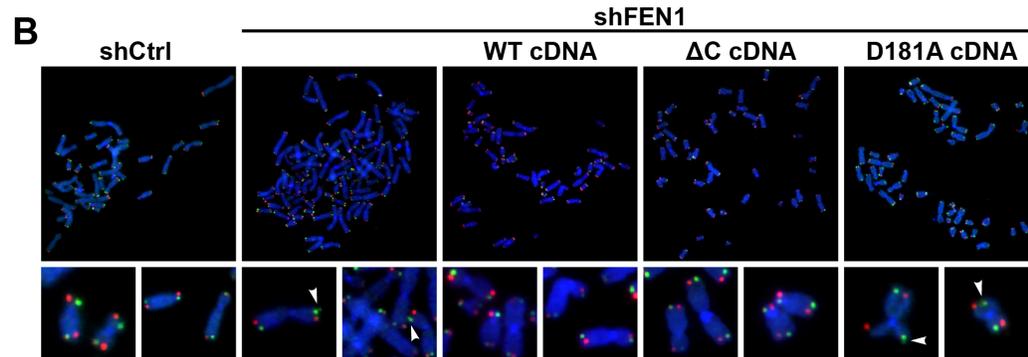
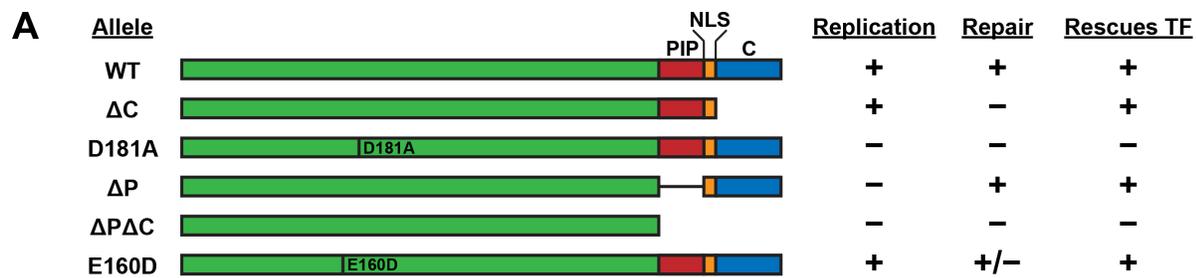


Figure 2.5. FEN1 flap endonuclease activity is required to limit leading strand-specific telomere fragility. (A) Schematic showing FEN1 alleles used in this study. Features indicated include a PIP box (PIP), nuclear localization signal (NLS), C-terminal region (C), and point mutations. The replication competency, repair competency, and ability to rescue telomere fragility (this study) of each allele are shown to the right. (B) Representative metaphase chromosomes processed with CO-FISH from BJL fibroblasts expressing a control hairpin (shCtrl) or depleted of FEN1 (shFEN1). Leading strand-replicated telomeres are green and lagging strand-replicated telomeres are red. FEN1 alleles were ectopically expressed where indicated. Arrowheads mark fragile telomeres in the magnified images. (C) Representative metaphase chromosomes processed with CO-FISH from BJ fibroblasts expressing a control hairpin (shCtrl) or depleted of FEN1 (shFEN1). Leading strand-replicated telomeres are green and lagging strand-replicated telomeres are red. FEN1 alleles were ectopically expressed where indicated. Arrowheads mark fragile telomeres in the magnified images. (D) Quantification of strand-specific telomere fragility per chromosome, with leading strand-specific telomere fragility shown in green and lagging strand-specific telomere fragility shown in red. Two independent biological replicates were analyzed, normalized with shCtrl set to 1 for each mutant group, and combined. p -values were computed using a two-tailed Mann-Whitney U test (***, $p < 0.001$ relative to shCtrl). Error bars represent standard error of the mean.

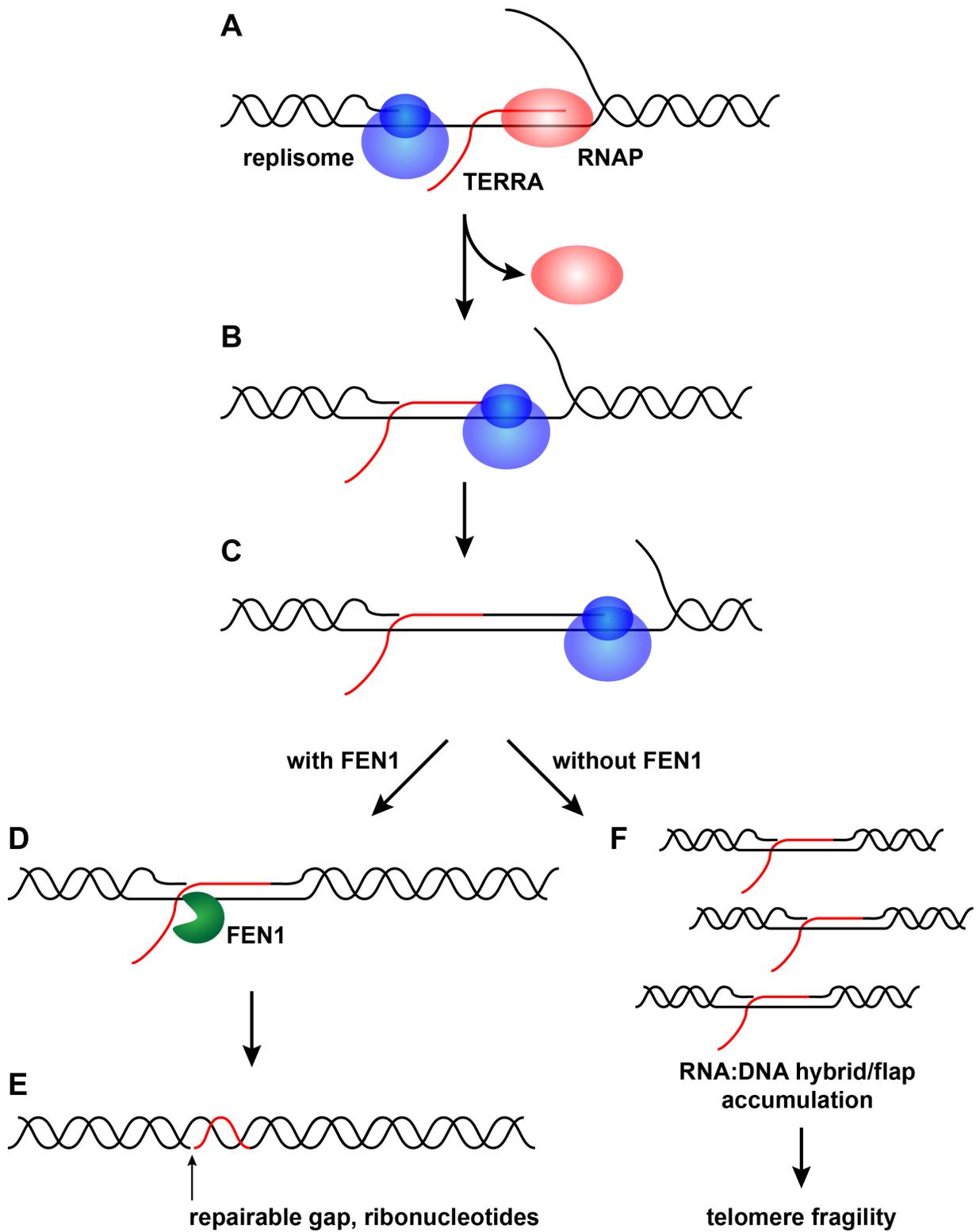


Figure 2.6. A model of FEN1's role following co-directional replisome–RNAP collisions.

(A) RNA Pol II (RNAP) transcribes TERRA from the C-rich leading strand. The

replisome approaches the transcription complex and a co-directional collision occurs. Pol II dissociates from the nascent TERRA. (B) The replisome moves to the 3' end of the TERRA, leaving a 5' RNA flap and RNA:DNA hybrid. (C) The replisome resumes replication of the leading strand using the 3' end of the nascent TERRA as a primer. (D) FEN1 cleaves the 5' RNA flap left behind by the collision. (E) FEN1's cleavage leaves behind a gap and a stretch of RNA:DNA hybrid that can be repaired. (F) In the absence of FEN1, RNA:DNA hybrid/flap structures accumulate and lead to telomere fragility.

References

1. Sidorova JM. Roles of the Werner syndrome RecQ helicase in DNA replication. *DNA Repair (Amst)*. 2008;7(11):1776–86.
2. Singh DK, Ahn B, Bohr VA. Roles of RECQ helicases in recombination based DNA repair, genomic stability and aging. *Biogerontology*. 2009 Jun;10(3):235–52.
3. Gilson E, Géli V. How telomeres are replicated. *Nat Rev*. 2007 Oct;8(10):825–38.
4. Burrow AA, Marullo A, Holder LR, Wang Y-H. Secondary structure formation and DNA instability at fragile site FRA16B. *Nucleic Acids Res*. 2010 May;38(9):2865–77.
5. Helmrich A, Ballarino M, Tora L. Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Mol Cell*. 2011;44(6):966–77.
6. Letessier A, Millot GA, Koundrioukoff S, Lachagès A-M, Vogt N, Hansen RS, et al. Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature*. 2011 Feb;470(7332):120–3.
7. Sfeir A, Kosiyatrakul ST, Hockemeyer D, MacRae SL, Karlseder J, Schildkraut CL, et al. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell [Internet]*. 2009 Jul 10 [cited 2013 Nov 7];138(1):90–103.
8. Martínez P, Thanasoula M, Muñoz P, Liao C, Tejera AM, McNeese CJ, et al. Increased telomere fragility and fusions resulting from TRF1 deficiency lead to degenerative pathologies and increased cancer in mice. *Genes Dev*. 2009 Sep;23(17):2060–75.

9. McNees CJ, Tejera AM, Martínez P, Murga M, Mulero F, Fernández-Capetillo O, et al. ATR suppresses telomere fragility and recombination but is dispensable for elongation of short telomeres by telomerase. *J Cell Biol.* 2010 Mar;188(5):639–52.
10. Vannier J-B, Pavicic-Kaltenbrunner V, Petalcorin MIR, Ding H, Boulton SJ. RTEL1 Dismantles T Loops and Counteracts Telomeric G4-DNA to Maintain Telomere Integrity. *Cell.* Elsevier Inc.; 2012 May;149(4):795–806.
11. Stewart JA, Wang F, Chaiken MF, Kasbek C, Chastain PD, Wright WE, et al. Human CST promotes telomere duplex replication and general replication restart after fork stalling. *EMBO J.* 2012 Aug;31(17):3537–49.
12. Saharia A, Teasley DC, Duxin JP, Dao B, Chiappinelli KB, Stewart SA. FEN1 ensures telomere stability by facilitating replication fork re-initiation. *J Biol Chem.* 2010 Jun;285(35):27057–66.
13. Vallabhaneni H, O’Callaghan N, Sidorova J, Liu Y. Defective repair of oxidative base lesions by the DNA glycosylase Nth1 associates with multiple telomere defects. *PLoS Genet.* 2013 Jan;9(7):e1003639.
14. D’Alcontres MS, Palacios JA, Mejias D, Blasco MA. Topoll α prevents telomere fragility and formation of ultra thin DNA bridges during mitosis through TRF1-dependent binding to telomeres. *Cell Cycle.* 2014 May;13(9):1463–81.
15. Chawla R, Redon S, Raftopoulou C, Wischnewski H, Gagos S, Azzalin CM. Human UPF1 interacts with TPP1 and telomerase and sustains telomere leading-strand replication. *EMBO J.* Nature Publishing Group; 2011 Oct;30(19):4047–58.
16. Arora R, Lee Y, Wischnewski H, Brun CM, Schwarz T, Azzalin CM. RNaseH1

- regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* [Internet]. Nature Publishing Group; 2014 Jan [cited 2014 Oct 22];5:5220. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25330849>
17. Rudolph CJ, Dhillon P, Moore T, Lloyd RG. Avoiding and resolving conflicts between DNA replication and transcription. *DNA Repair (Amst)*. 2007 Jul;6(7):981–93.
 18. Liu B, Wong ML, Tinker RL, Geiduschek EP, Alberts BM. The DNA replication fork can pass RNA polymerase without displacing the nascent transcript. *Nature*. 1993 Nov;366(6450):33–9.
 19. Prado F, Aguilera A. Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO J*. 2005 Mar;24(6):1267–76.
 20. Pomerantz RT, O'Donnell ME. The replisome uses mRNA as a primer after colliding with RNA polymerase. *Nature*. 2008 Dec;456(7223):762–6.
 21. Sordet O, Nakamura AJ, Redon CE, Pommier Y. DNA double-strand breaks and ATM activation by transcription-blocking DNA lesions. *Cell Cycle*. 2010 Jan;9(2):274–8.
 22. Skourti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev* [Internet]. 2014 Jul 1 [cited 2014 Jul 9];28(13):1384–96. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24990962>
 23. Stewart SA, Ben-Porath I, Carey VJ, O'Connor BF, Hahn WC, Weinberg RA.

- Erosion of the telomeric single-strand overhang at replicative senescence. *Nat Genet.* 2003 Apr;33(4):492–6.
24. Saharia A, Guittat L, Crocker S, Lim A, Steffen M, Kulkarni S, et al. Flap endonuclease 1 contributes to telomere stability. *Curr Biol [Internet]*. 2008 Apr 8 [cited 2013 Nov 25];18(7):496–500.
 25. Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, et al. Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA.* 2003 Apr;9(4):493–501.
 26. Honaker Y, Piwnicka-Worms H. Casein kinase 1 functions as both penultimate and ultimate kinase in regulating Cdc25A destruction. *Oncogene [Internet]*. Nature Publishing Group; 2010;29(23):3324–34.
 27. Lansdorp PM, Verwoerd NP, Van De Rijke FM, Dragowska V, Little MT, Dirks RW, et al. Heterogeneity in telomere length of human chromosomes. *Hum Mol Genet.* 1996;5(5):685–91.
 28. Bailey SM, Cornforth MN, Kurimasa A, Chen DJ, Goodwin EH. Strand-specific postreplicative processing of mammalian telomeres. *Science (80-)*. 2001 Sep;293(5539):2462–5.
 29. Duxin JP, Dao B, Martinsson P, Rajala N, Guittat L, Campbell JL, et al. Human Dna2 is a nuclear and mitochondrial DNA maintenance protein. *Mol Cell Biol [Internet]*. 2009 Aug [cited 2014 Aug 6];29(15):4274–82.
 30. Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning: A Laboratory Manual*. 2nd ed. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold

Spring Harbor Laboratory Press; 1989. 7.53-7.55.

31. Schoeftner S, Blasco M a. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* [Internet]. 2008 Feb [cited 2014 Jul 11];10(2):228–36.
32. Azzalin CM, Reichenbach P, Khoriantuli L, Giulotto E, Lingner J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* [Internet]. 2007 Nov 2 [cited 2013 Nov 18];318(5851):798–801.
33. Sabouri N, McDonald KR, Webb CJ, Cristea IM, Zakian VA. DNA replication through hard-to-replicate sites, including both highly transcribed RNA Pol II and Pol III genes, requires the *S. pombe* Pfh1 helicase. *Genes Dev.* 2012 Mar;26(6):581–93.
34. Azvolinsky A, Giresi PG, Lieb JD, Zakian VA. Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. *Mol Cell.* 2009;34(6):722–34.
35. Bah A, Wischnewski H, Shchepachev V, Azzalin CM. The telomeric transcriptome of *Schizosaccharomyces pombe*. *Nucleic Acids Res* [Internet]. 2012 Apr [cited 2014 Aug 6];40(7):2995–3005.
36. Stewart J a, Campbell JL, Bambara R a. Flap endonuclease disengages Dna2 helicase/nuclease from Okazaki fragment flaps. *J Biol Chem* [Internet]. 2006 Dec 15 [cited 2014 Aug 6];281(50):38565–72.
37. Rudd MD, Luse DS. Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage

- modes. *J Biol Chem*. 1996 Aug;271(35):21549–58.
38. Bushnell DA, Cramer P, Kornberg RD. Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 Å resolution. *Proc Natl Acad Sci U S A*. 2002 Feb;99(3):1218–22.
 39. Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM. TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol Cell* [Internet]. 2009 Aug 28 [cited 2013 Nov 25];35(4):403–13.
 40. Balk B, Maicher A, Dees M, Klermund J, Luke-Glaser S, Bender K, et al. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* [Internet]. 2013 Oct [cited 2013 Nov 13];20(10):1199–205.
 41. Parrish JZ, Yang C, Shen B, Xue D. CRN-1, a *Caenorhabditis elegans* FEN-1 homologue, cooperates with CPS-6/EndoG to promote apoptotic DNA degradation. *EMBO J*. 2003;22(13):3451–60.
 42. Zheng L, Zhou M, Chai Q, Parrish J, Xue D, Patrick SM, et al. Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks. *EMBO Rep*. 2005;6(1):83–9.
 43. Liu Y, Kao H-I, Bambara RA. Flap endonuclease 1: a central component of DNA metabolism. *Annu Rev Biochem*. 2004 Jan;73:589–615.
 44. Li X, Li J, Harrington J, Lieber MR, Burgers PM. Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. *J Biol Chem*. 1995 Sep;270(38):22109–12.
 45. Guo Z, Chavez V, Singh P, Finger LD, Hang H, Hegde ML, et al. Comprehensive

- mapping of the C-terminus of flap endonuclease-1 reveals distinct interaction sites for five proteins that represent different DNA replication and repair pathways. *J Mol Biol.* 2008 Mar;377(3):679–90.
46. Shen B, Nolan JP, Sklar LA, Park MS. Essential amino acids for substrate binding and catalysis of human flap endonuclease 1. *J Biol Chem.* 1996 Apr;271(16):9173–6.
 47. Tsutakawa SE, Classen S, Chapados BR, Arvai AS, Finger LD, Guenther G, et al. Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell.* Elsevier Inc.; 2011;145(2):198–211.
 48. Bornarth CJ, Ranalli TA, Henricksen LA, Wahl AF, Bambara RA. Effect of flap modifications on human FEN1 cleavage. *Biochemistry.* 1999 Oct;38(40):13347–54.
 49. Atkinson J, Gupta MK, McGlynn P. Interaction of Rep and DnaB on DNA. *Nucleic Acids Res.* 2011 Mar;39(4):1351–9.
 50. Fu Y V, Yardimci H, Long DT, Ho TV, Guainazzi A, Bermudez VP, et al. Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell.* 2011 Sep;146(6):931–41.
 51. Muftuoglu M, Wong HK, Imam SZ, Wilson DM, Bohr VA, Opresko PL. Telomere repeat binding factor 2 interacts with base excision repair proteins and stimulates DNA synthesis by DNA polymerase beta. *Cancer Res.* 2006 Jan;66(1):113–24.
 52. Badie S, Escandell JM, Bouwman P, Carlos AR, Thanasoula M, Gallardo MM, et al. BRCA2 acts as a RAD51 loader to facilitate telomere replication and capping.

Nat Struct Mol Biol [Internet]. 2010 Dec [cited 2013 Nov 7];17(12):1461–9.

53. Popuri V, Hsu J, Khadka P, Horvath K, Liu Y, Croteau DL, et al. Human RECQL1 participates in telomere maintenance. Nucleic Acids Res [Internet]. 2014 Jan [cited 2014 Aug 29];42(9):5671–88.

Chapter 3:
**Human Ribonuclease H1 resolves R loops and thereby enables progression of
the DNA replication fork**

Shankar Parajuli, Daniel C. Teasley, Bhavna Murali, Jessica Jackson, Alessandro
Vindigni, Sheila A. Stewart

This chapter was originally published in *The Journal of Biological Chemistry*, 2017
(jbc.M117.787473). © the American Society for Biochemistry and Molecular Biology

Introduction

High-fidelity DNA replication is paramount to the maintenance of genome stability. Therefore, cells have evolved various redundant mechanisms to resolve genotoxic challenges including the presence topological structures. If not resolved, these structures can impede replication fork progression, leading to the stalling and eventual collapse of replication forks (1). One such structure is the RNA:DNA hybrid, commonly referred to as an R-loop. Mapping studies have revealed that RNA:DNA hybrid structures are present throughout the genome more frequently than previously appreciated (2, 3). One source of R-loops is transcription during which nascent RNA emanating from RNA polymerase II hybridizes with its template DNA (4). While these topological structures play a vital role in a number of key processes including class switch recombination (CSR) of immunoglobulin genes and transcription termination, their persistence or unscheduled formation and stabilization pose a significant challenge to genome integrity (5, 6). Because DNA replication and transcription occur simultaneously at many regions of genome, hybrids can form in front of the replication machinery and affect its progression. Indeed, these R-loops can lead to increased DNA mutations, unwanted recombination and gross chromosomal aberrations (7). Thus it is not surprising that a number of proteins inhibit the formation of these structures or resolve them once they have formed (4). Topoisomerase I and mRNA export and splicing factors, for example, play an active role in preventing R-loop formation (4, 8). On the other hand, the helicases Senataxin (SETX) and Aquarius (AQR), are tasked with resolving these structures to promote transcriptional termination and maintain genome stability respectively (6, 9). In addition, DNA damage response factors such as

breast cancer susceptibility factors (BRCA1 and BRCA2) are also implicated in preventing R-loop accumulation and the ensuing DNA damage (10, 11).

Ribonuclease H1 (RNH1) is a specialized enzyme that can specifically resolve long RNA:DNA hybrids. A closely related protein complex, ribonuclease H2 (RNH2) is adept at removing single misincorporated ribonucleotides from DNA and is critical for ongoing genomic stability (12). In yeast, RNH1 and RNH2 function redundantly to facilitate efficient double strand break repair during homologous recombination by assisting in the unwinding of DNA strands and RPA binding (13). Ectopic expression of RNH1 in yeast is sufficient to minimize transcription-dependent hyper-recombination, pausing of the replication fork, and HU sensitivity (14, 15, 16). In mammalian cells, RNH1 has an established role in mitochondrial DNA replication, and its deletion is embryonically lethal, demonstrating that RNH2 cannot compensate in this setting (17, 18, 19, 20). RNH1 localizes to the mammalian nucleus (21) and ectopic RNH1 expression is routinely exploited in mammalian cells to resolve RNA:DNA hybrids. Recently RNH1 was shown to prevent unwanted recombination events at telomere by resolving telomeric RNA:DNA hybrids in cells that utilize the ALT mechanism of telomere maintenance (22). Another recent study identified a link between DNA damage and the accumulation of RNA:DNA hybrids at the telomere (23). However, whether RNH1 plays a role in nuclear DNA replication outside of telomeres remains to be explored.

Given the role of RNA:DNA hybrids in replication impairment and the ability of RNH1 to resolve such hybrids, we sought to determine if RNH1 impacts genomic integrity in the

mammalian nucleus and if so how. We depleted RNH1 from human cell lines and found that RNH1 depletion resulted in increased RNA:DNA hybrids, DNA damage response, and slowing of DNA replication forks. Importantly, these phenotypes were dependent upon RNH1 nuclease activity, suggesting that the hybrids were responsible for these phenotypes. Our studies uncover a novel role of RNH1 in the mammalian nucleus and extend its important function in nuclear DNA replication.

Experimental Procedures

Cell culture

Cells were cultured at 37° C in 5% carbon dioxide and atmospheric oxygen, as previously reported (24). 293T cells were obtained from Dr. Robert Weinberg (Massachusetts Institute of Technology) and cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM) containing 10% heat inactivated fetal bovine serum (Δ FBS) and 1% penicillin/ streptomycin (P/S) (Sigma-Aldrich, St. Louis, MO). RPE1 cells were obtained from ATCC and cultured in DMEM: Nutrient Mixture F-12 (DMEM/F12) containing 7.5% Δ FBS and 1% P/S.

siRNA transfection

siRNA transfection was performed using Invitrogen's Lipofectamine RNAiMAX Reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. siRNAs used were control (siCtrl- Cat #4390843) or RNH1-directed (siRNH1- Cat #4390824, ID s48356) from Life Technologies or that directed to the 3' UTR of RNH1 (siRNH1- hs.Ri. RNASEH1.13.1) from Integrated DNA Technologies.

Virus production, infections, and stable cell lines

Lentiviral production and transductions were carried out as previously reported (25). Briefly, 293T cells were transduced with an 8:1 ratio of pHR -CMV-8.2 R packaging plasmid and pCMV-VSV-G and a pLKO.1-puro plasmid carrying an shRNA using TransIT-LT1 (Mirus Bio, Madison, WI). Supernatant-containing virus was collected 48 hours post-transfection and filtered through a 0.45- μm PVDF membrane. RPE1 cells were infected for 4 hours each on 2 consecutive days in the presence of 8 $\mu\text{g/ml}$ protamine sulfate (Sigma). Following infection, transduced cells were selected with 15 $\mu\text{g/ml}$ puromycin sulfate. Stable RPE1 cell lines were prepared by using either a GFP tagged D145N RNH1 construct (a generous gift of Dr. Marteijn) (26) or its wildtype version (modified from D145N construct using site directed mutagenesis, Agilent Technologies).

Western Blot Analysis

Western blot analysis was carried out as previously described with modifications (27). Briefly, cells were washed with PBS and lysed in radio-immunoprecipitation assay buffer (150 mM NaCl, 50 mM 1 M Tris-HCl, pH 8.0, 1.0% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS) supplemented with 1 mM sodium orthovanadate, 1 mM sodium fluoride, 1 mM microcystin-LR, 2 mM phenylmethylsulfonyl fluoride, protease inhibitor mixture (Sigma), and phosphatase inhibitor mixture set I (EMD Millipore, Billerica, MA). Following sonication and centrifugation, supernatant lysate was quantified using the protein assay (Bio-Rad). Lysates were subjected to a SDS-PAGE and transferred to

PVDF membranes for blotting. The following antibodies were used: mouse monoclonal anti RNase H1 (H00246243-M01, Novus Biologicals, Littleton, CO); rat monoclonal anti-tubulin (NB600-506, Novus Biologicals, Littleton, CO); mouse monoclonal anti-phospho (Ser 139) H2AX (05-636, Millipore, CA).

Metaphase Chromosome Preparation

Metaphase chromosomes were prepared as described previously (28). Briefly, cultured RPE1 cells were treated with 0.5 µg/ml colcemid (Sigma) for 6 hours. Arrested metaphase cells were collected by mitotic shake-off, treated with 75 mM potassium chloride, and fixed in 3:1 solution of methanol and acetic acid. Chromosomes were spread by dropping onto glass slides. For analysis via COFISH, 0.3 µg/ml 5-bromo-2'-deoxyuridine (Sigma) and 0.1 µg/ml 5-bromo-2'-deoxycytidine (MP Biomedicals, Santa Ana, CA) were added to the cultured media 18 hours prior to collection of the cells.

Chromosome Orientation - FISH (CO-FISH)

CO-FISH was performed as described previously with modifications (29). Briefly, spread metaphase chromosomes were aged at 65 °C for 18 hours. Aged chromosomes were rehydrated in PBS, treated with 100 µg/ml RNase at 37 °C for 10 min and re-fixed in 4% paraformaldehyde at room temperature (RT) for 10 min. Fixed chromosomes were UV-sensitized in 0.5 µg/ml Hoechst 33258 (Sigma) in 2X SSC at RT for 15 min and exposed to 365 nm UV light for 1 hour using a UV cross-linker (Vilber-Lourmat, Marne-la-Vallée, France). Exposed chromosomes were digested with 3 units/µl Exonuclease III (Promega, Madison, WI) at RT for 15 min, denatured in 70% formamide in 2X SSC,

and dehydrated in cold ethanol before hybridization. Chromosomes were hybridized first using 0.03 µg/ml (PNA Bio, Thousand Oaks, CA) of a leading strand telomere PNA probe (FAM-(TTAGGG)₃) followed by 0.03 µg/ml of a lagging strand PNA probe (Cy3-(CCCTAA)₃). Hybridized chromosomes were mounted using ProLong Gold (Life Technologies, Inc.) with 125 ng/ml DAPI.

Microfluidic- Assisted Replication Track Analysis (maRTA)

maRTA was conducted as described previously (30, 31). Briefly, asynchronous RPE1 cells were labelled for 30 min each with 50 µM CldU and 50 µM IdU with two PBS washes in between. Labelled cells were collected and embedded in agarose plugs for lysis and DNA extraction. DNA was subsequently stretched, denatured and subjected to immunostaining. Antibodies used were rat anti-CldU/BrdU (Abcam, ab6326), mouse anti IdU/BrdU (BD Biosciences, 347580), goat anti-rat Alexa 594 (Invitrogen, A11007), and goat anti-mouse Alexa 488 (Invitrogen, A11001).

Fluorescence Imaging

Metaphase chromosomes from COFISH and labelled DNA tracks from maRTA were imaged on a Nikon 90i epifluorescence microscope using a 100 X 1.40 NA Plan Apo VC objective (Nikon Instruments, Melville, NY) with Cargille Type LDF immersion oil *Cargille Sacher Laboratories, Cedar Grove, NJ). Images were captured using a CoolSnap HQ2 CCD camera (Photometrics, Tucson, AZ), deconvulated with a blind algorithm using NISElements AR (Nikon Instruments) prior to quantification. RPE1 cells

stably expressing GFP tagged WT and D145N RNH1 were visualized and captured without any staining.

DNA immunoprecipitation (DIP)

DIP was performed as described previously with modifications (6). Briefly, 293T cells were pelleted and resuspended in DIP lysis buffer (0.5% NP40, 85 mM potassium chloride, and 5 mM PIPES). Following centrifugation, pelleted nuclei were lysed in DIP nuclear lysis buffer (1% sodium dodecyl sulfate, 25 mM tris-HCl pH 8, and 5 mM EDTA), sheared and digested with two sequential rounds of 100 µg proteinase K for 1.5 hours each at 55 °C. DNA was phenol:chloroform extracted and ethanol-precipitated (32) at which point one half was subjected to an overnight digestion with recombinant ribonuclease H (Roche Applied Science, Penzberg, Germany). Samples were then diluted in DIP dilution buffer (1.1% Triton X-100, 0.01% sodium dodecyl sulfate, 1.2 mM EDTA, 16.7 mM tris-HCl pH 8, and 166.5 mM sodium chloride) and sonicated to generate about 200 bp long DNA fragments. Resulting DNA was quantified using the PicoGreen assay following the manufacturer's protocol (Life Technologies, Grand Island, NY). 10 µg of DNA was immunoprecipitated overnight with 10 µg of S9.6 antibody or mouse IgG. Antibody- DNA complexes were captured by using Protein A magnetic beads (Life Technologies, Grand Island, NY) after equilibration in DIP dilution buffer. After extensive washing, antibody- DNA complexes were eluted from the beads and treated with proteinase K followed by recovery using PCR cleanup columns (Qiagen, Venlo, Netherlands).

S9.6 Immunofluorescence (S9.6 IF)

S9.6 IF was performed as previously described (9). Briefly, RPE1 cells transfected with siCtrl or siRNH1 were fixed with ice-cold methanol for 5 min at -20°C. Fixed cells were blocked in 2% BSA/PBS for an hour at room temperature followed by incubation with the S9.6 primary antibody (1:200 dilution, 1 µg/mL) and goat anti-mouse AlexaFluro-594 conjugated secondary antibody (1:1000, Invitrogen) for 1 hour each at room temperature. Finally, cells were washed in 0.5 µg/ml Hoechst 33258 PBS to label the nuclei and mounted using ProLong Gold. Images were taken at 40X using a Nikon 90i epifluorescence microscope as previously described. Only the nuclear staining of S9.6 signal was considered and analyzed using ImageJ 1.50i.

Genomic Quantitative PCR (qPCR)

Genomic qPCR was performed using Power SYBR Green Master Mix (Life Technologies, Grand Island, NY) following the manufacturer's protocol. 5' region of the Beta-actin pause element (5' pause site), known to form RNA:DNA hybrids, was amplified to assess hybrid formation. Reaction conditions were as described in manufacturer's instructions with 58.7 °C as the annealing temperature. Primers used were: 5'-TTACCC AGA GTG CAG GTG TG-3' (Forward) and 5'-CCC CAA TAA GCA GGA ACA GA-3' (Reverse).

Quantitative Reverse Transcription - PCR (qRT-PCR)

qRT-PCR was performed as described previously (24). Target genes used were RNH1 (Hs00268000_m1, Life Technologies, Inc.) and GAPDH (Hs. PT.39a.22214836, Integrated DNA Technologies).

RESULTS

RNH1 contributes to genome stability and preserves telomere integrity

While RNH2 has well ascribed functions in the mammalian nucleus, RNH1's role has remained more obscure. Because R-loops form throughout the genome and RNH1 can resolve R-loops that would pose barriers to the replication machinery, we hypothesized that RNH1 might play an important role in the nucleus and that its loss might perturb replication fork progression and thus elicit a DNA damage response (4). To test this hypothesis, we depleted RNH1 from normal, checkpoint-competent RPE1 cells and measured the levels of histone H2AX phosphorylation at S139 (YH2AX), a canonical marker of the DNA damage response (33) in control versus RNH1-depleted cells. Following RNH1 depletion, we observed increased levels of YH2AX, demonstrating that RNH1 depletion induces a DNA damage response (Fig 3.1A,B). These data suggest that RNH1 plays an important role in preserving genome stability.

To further interrogate RNH1's function, we first focused our attention to telomeres, chromosomal ends that contain RNA:DNA hybrids (34). Recent work demonstrates that in cells utilizing the alternative lengthening of telomeres (ALT) mechanism, which maintains telomere length independent of telomerase—RNH1 associates with telomeres and regulates the levels of telomeric RNA:DNA hybrids to prevent telomere

loss (22). In these cells, depletion of RNH1 led to hybrid accumulation and abrupt telomere excision. A second study suggested that RNH1 plays an important role in resolving RNA:DNA hybrids at the telomere (23). Because the leading strand replicated telomere is transcribed, RNA:DNA hybrids would be expected to form on the leading strand. Thus, we examined the integrity of the leading strand telomere by performing chromosomal orientation fluorescent in situ hybridization (CO-FISH), which allows one to interrogate the leading versus lagging strand replicated telomere. Surprisingly, CO-FISH analysis revealed no differences in the leading versus lagging strand telomere in control versus shRNH1 cells (data not shown). However, in the RNH1-depleted cells, we observed a significant increase in telomere free ends (TFE), in which both leading and lagging strand telomeres were lost, a phenotype suggestive of DNA replication defects (35) (Fig. 3.1C,D). These data suggest that RNH1 assists the replication machinery by resolving RNA:DNA hybrids that could present a topological barrier to replication fork progression.

Nuclear RNA:DNA hybrid levels increase upon RNH1 depletion

To demonstrate that RNA:DNA hybrids were responsible for DNA damage and telomere loss phenotype upon RNH1 loss, we next measured those hybrid levels in the nucleus. We treated 293T cells with control (siCtrl) or RNH1 directed (siRNH1) siRNAs and collected cells 48 hours later. Transfection with siRNH1 led to a significant (2.5 fold) reduction in RNH1 mRNA levels (Fig. 3.2A) and (3.5 fold) protein levels (Fig. 3.2B) compared to levels present in siCtrl cells. To measure the amount of RNA:DNA hybrids in control versus RNH1-depleted cells, we next extracted nuclear DNA lysate and

subjected it to DNA:RNA immunoprecipitation (DIP) using the well characterized RNA:DNA hybrid antibody, S9.6 (36). We conducted a genomic quantitative PCR on a well-characterized hybrid forming 5' pause site of beta-actin gene as a readout of hybrid signals. As a control for specificity, we also pretreated lysates with recombinant RNaseH enzyme *in vitro* to degrade existing RNA:DNA hybrids in both control and depleted cells. As expected, pretreatment with an *in vitro* RNaseH enzyme led to a 1.8 fold reduction of RNA:DNA hybrids in control and a 3.5 fold in RNH1-depleted cells, confirming the specificity of the S9.6 antibody. Additionally, immunoprecipitation with an IgG control antibody failed to precipitate RNA:DNA hybrids, indicating that the signals we measured were bona fide RNA:DNA hybrids. Analysis of immunoprecipitations from RNH1-depleted cells revealed a significant two-fold increase in the nuclear RNA:DNA hybrids compared to those in control cells (Fig. 3.2C). To further corroborate these findings, we also utilized the S9.6 antibody and carried out immunofluorescence on RPE1 cells. As expected, RNH1-depleted cells showed increased levels of RNA:DNA hybrids as represented by elevated S9.6 signal in the nucleus compared to that in the control cells (Fig. 3.2D,E). Together, these data demonstrate that RNH1 depletion can lead to a significant increase in RNA:DNA hybrids and that this increase correlates with increased DNA damage and telomere loss.

RNH1 depletion results in replication fork slowing and increased termination and stalling

Given the increased RNA:DNA hybrids and DNA damage and loss of both telomeric ends, indicative of a replication defect following RNH1 depletion, we hypothesized that

RNA:DNA hybrids pose barriers to DNA replication forks. This hypothesis was supported by previous studies showing that the removal of RNA:DNA hybrids by ectopically expressed RNH1 can directly affect replication fork movement in yeast (37). To test this hypothesis, we used microfluidic-assisted replication track analysis (maRTA) to directly measure replication fork progression in RPE1 cells depleted of RNH1 (30, 38). RPE1 cells were transduced with siRNAs and cells were collected for western blot analysis 48 hours later. As expected, RNH1 depletion resulted in significant DNA damage as evidenced by increased γ H2AX (Fig. 3.3A). In parallel, we carried out maRTA by plating RNH1-depleted or control cells and labeling them with the nucleotide analogs, CldU (red) and IdU (green), sequentially for 30 minutes each, to allow us to follow replication fork movement (Fig. 3.3B). To restrict our analysis to progressing replication forks, we measured IdU tracks that were directly preceded by a CldU track (Fig. 3.3B). Measuring the lengths of these IdU tracks, we found that RNH1 depleted cells had an average IdU track length of 9.3 (\pm 0.3) μ m while that in the control cells was 14.5 (\pm 0.4) μ m, indicating that the replication forks moved significantly slower in RNH1 depleted cells compared to control cells (Fig. 3.3C). Given the increase in RNA:DNA hybrids associated with RNH1 loss, these data suggest that RNH1 facilitates efficient DNA replication by clearing RNA:DNA hybrids that would otherwise impede replication fork progression during S phase.

To understand how the loss of RNH1 perturbed replication dynamics, we next asked whether other replication parameters including termination, stalling, and origin firing were affected upon RNH1 depletion. Premature termination and stalling events

correspond to CldU (red)-only tracks (Fig. 3.3B). RNH1-depleted cells showed a significant 1.4 fold increase in termination and/or stalling events compared to control cells (Fig. 3.3D). Next, we analyzed the impact of RNH1 depletion on origin firing to address a possibility that slower fork progression triggers the S phase checkpoint and increases the frequency of origin firing as previously reported (39). To measure origin firing, the incidence of tracks with either green-only (IdU) color or red flanked by green on both sides was analyzed (Fig. 3.3B). However, no significant differences in origin firing were observed between RNH1-depleted and control cells (Fig. 3.3E). Collectively, these results indicate that RNH1 plays an important role in assisting fork movement during DNA replication. We suggest that RNH1 does this by resolving RNA:DNA hybrids that pose barriers to a progressing replication fork. Furthermore, our data provide the first evidence of RNH1's role in global DNA replication in the mammalian nucleus.

Nuclease activity of RNH1 is required for efficient replication fork movement

To further characterize RNH1's role in DNA replication, we tested whether the nuclease function of RNH1 was required for this activity. To do this, we created a series of RPE1 cell lines ectopically expressing either a GFP-tagged wild type (WT) or a well characterized GFP-tagged nuclease-dead (D145N) form of RNH1 (26, 40). To establish a direct role in the nucleus, these RNH1 constructs lacked the mitochondrial targeting sequence present in the endogenous gene, thereby restricting their expression to the nucleus. We confirmed the nuclear localization of ectopically expressed proteins by visualizing GFP expression only in the nucleus (Fig. 3.4A). These stable RPE1 cells were transfected with an siRNH1 directed towards the 3' untranslated region (3'UTR)

that did not target the ectopically expressed protein. We observed that RNH1 depletion using a 3' UTR siRNA was comparable to that of a previously used coding sequence (CDS) targeting siRNA (Fig. 3.4B). We also observed robust expression of our ectopically tagged RNH1 proteins and significant depletion of endogenous RNH1 levels (Fig. 4B). Next, using maRTA, we again measured replication fork movement and found that ectopic expression of WT RNH1 restored fork movement to the levels observed in siCtrl cells (Fig. 3.4C). Indeed, while RNH1 depleted forks moved an IdU length of 8 μm , ectopic expression of WT RNH1 increased this to 12 μm , levels we observed in siCtrl cells. In contrast, ectopic expression of the catalytically-dead D145N allele of RNH1 failed to rescue the replication fork movement defect (fork movement was 7.9 μm , nearly identical to that observed in RNH1 depleted cells). Together, these findings demonstrated that the nuclease activity of RNH1 is required for the unperturbed movement of replication forks in mammalian cells. Similarly, we also measured fork termination and stalling events upon ectopic expression of WT and D145N alleles in RNH1 depleted cells. As expected, both events were reversed by ectopic expression of WT RNH1 but not the nuclease dead allele, thereby reiterating the importance of the nuclease function of RNH1 in the fidelity of replication fork progression (Fig. 3.4D). Neither WT nor D145N alleles of RNH1 affected the levels of origin firing in RNH1 depleted cells (Fig. 3.4E). These data suggest that the nuclease activity of RNH1 is required for resolution of RNA:DNA hybrids and therefore efficient movement of replication forks during nuclear DNA replication.

Discussion

Our study establishes a role for RNH1 in genomic DNA replication. Indeed, we illustrate for the first time that RNH1 nuclease activity is required for efficient fork movement during nuclear DNA replication. Furthermore, we have established a correlation between the accumulation of RNA:DNA hybrids and replication defects observed upon RNH1 depletion. Taken together, we propose a model wherein RNH1 resolves RNA:DNA hybrids to assist the replication machinery in its uninterrupted movement during DNA replication.

The unscheduled formation and stabilization of RNA:DNA hybrids have been postulated to be detrimental to the replication machinery. The importance of resolving these structures is probably best underscored by the multitude of proteins that act on RNA:DNA structures. Indeed, helicases such as SETX, AQR, and DHX9 in mammalian cells and Sen1 and PIF1 in yeast have all been shown to resolve RNA:DNA hybrids (9, 41, 42, 43). Here we add RNH1 to a growing list of proteins and show that endogenous RNH1 is required to similarly remove RNA:DNA hybrids and if these hybrids are not removed, replication is significantly impacted. Further our work demonstrates that these other proteins are unable to compensate for loss of RNH1 in replication fork progression. However, how RNH1 is regulated in the nucleus and how its activity is coordinated with the replication machinery remains unclear. A recent study from Nguyen et al., elegantly demonstrated that RPA can interact with RNH1 and stimulate its activity raising the possibility that RNH1 is tightly regulated by the DNA replication and repair machinery (44).

Given that RNH1 loss elicits replication defects such as fork slowing and termination and fork stalling, it will be critical to determine how this impacts checkpoint activation and cell cycle progression. Furthermore, understanding the fate of accumulated RNA:DNA hybrids upon RNH1 depletion is another interesting avenue worth pursuing. As previously noted, RNA:DNA hybrids arising from different sources can be processed via separate mechanisms (9). For example, those involved in CSR are not processed via nucleotide excision repair (NER) whereas those arising from loss of some RNA processing factors or CPT treatment are processed by NER. It is also worth evaluating if redundant nucleases and helicases including RNaseH2, SETX, and AQR could rescue effects of RNH1 loss.

The study of R-loops and their resolution have sparked more attention in recent years due to the fact that R-loops are associated with a number of diseases including cancers and several neurodegenerative disorders (45). This underscores a need for understanding these structures, their origins, stabilization and resolution along with their impact on cellular processes. By revealing RNH1's function in R-loop resolution in the nucleus, our study adds to the diversity of mechanisms targeting such structures. Furthermore, our study identifies a previously unknown function of RNH1 in nuclear DNA replication. Together, our work broadens the understanding of RNA:DNA structures and places RNH1 as a novel mechanism to resolve those structures and assist in nuclear DNA replication.

Acknowledgements: We thank Dr. Kevin C. Flanagan for helpful comments and Sandra Crocker and Shashikant Kulkarni for examination of chromosomes. We also thank Zhongsheng You for antibodies. This work was supported by the Predoctoral Cancer Biology Pathway Grant, Siteman Cancer Center/ Barnes Jewish Hospital (SP), the Molecular Oncology Training Grant NIH T32CA113275 (BM), NIH grant R01GM108648 (AV), DOD BRCP Breakthrough Award BC151728 (AV), NIH grant CA130919 (SAS), and American Cancer Society Research Scholar Award (SAS).

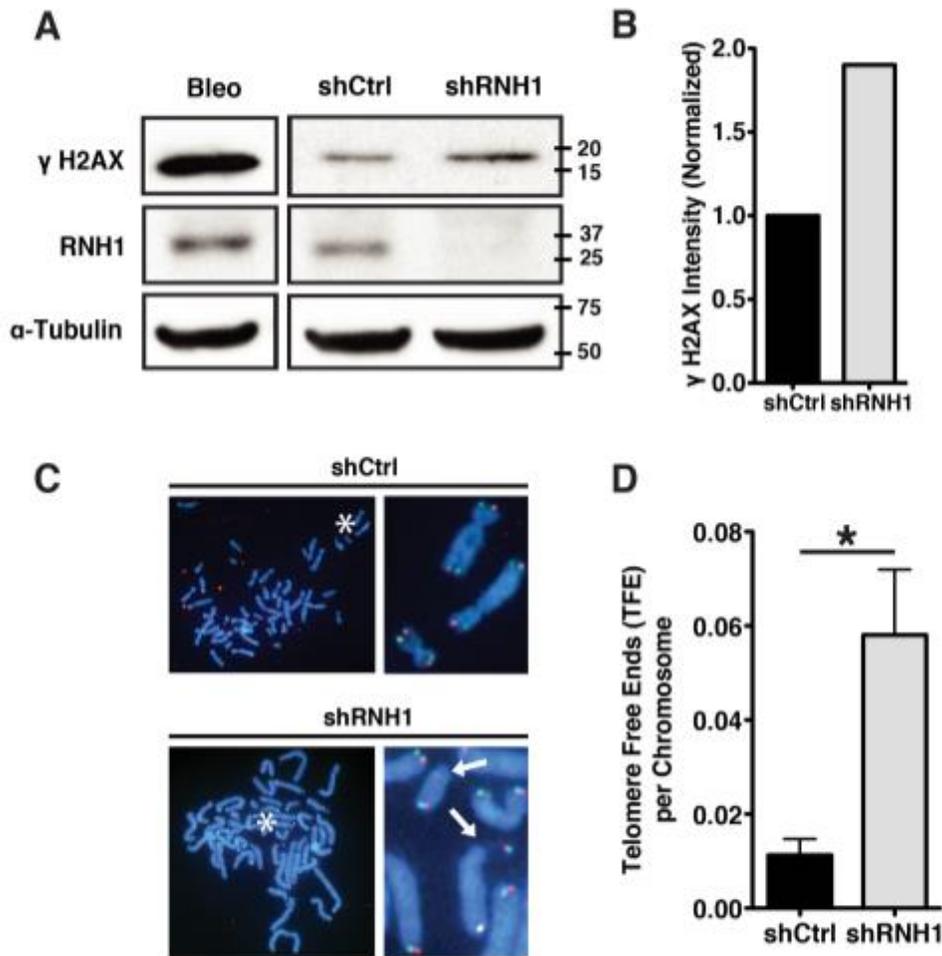


Figure 3.1. RNH1 contributes to genome stability and preserves telomere integrity

A. Western analysis of RNH1 expression (RNH1) and H2AX phosphorylation (γ H2AX) in control (shCtrl) and RNH1 depleted RPE1 cells (shRNH1). Bleo treated cells (Bleo) is a positive control for γ H2AX. α Tubulin is shown as a loading control. Molecular weight in kilodaltons is marked to the right for reference. **B.** Quantification of γ H2AX intensity in shCtrl and shRNH1 cells from western blot in A. **C.** Representative metaphase chromosomes processed with CO-FISH from shCtrl or shRNH1 RPE1 cells. Leading strand-replicated telomeres are green, and lagging strand-replicated telomeres are red. Regions marked by white asterisks are magnified; white arrowheads indicate telomere free ends (TFE) in magnified images. **D.** Representative quantification of telomere loss

in shCtrl and shRNH1 RPE1 cells. A minimum of 700 metaphase chromosomes were analyzed. p values were computed using a two-tailed Student's t test (*, $p < 0.05$). Error bars represent standard error of the mean.

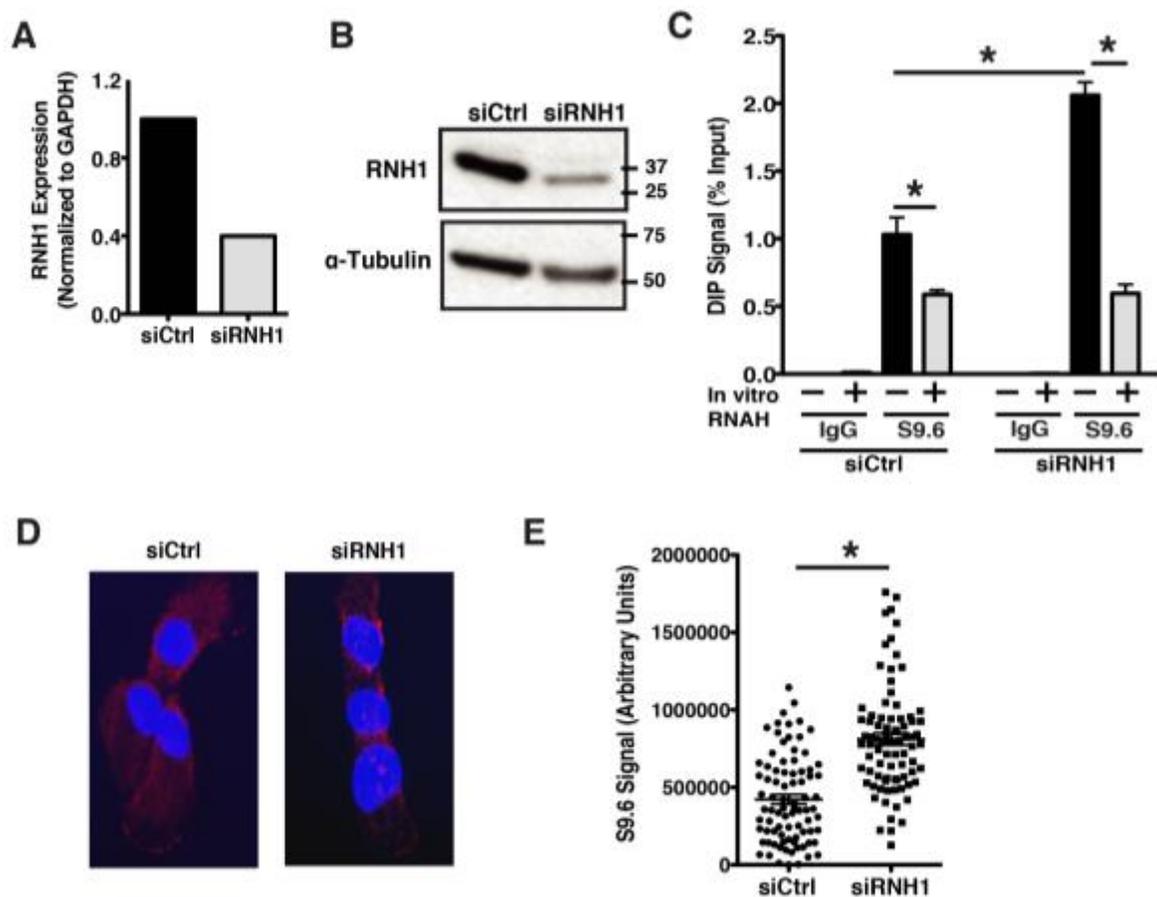


Figure 3.2. Nuclear RNA:DNA hybrid levels increase upon RNH1 depletion

A. qRT-PCR analysis of RNH1 mRNA in 293 T cells transfected with a control siRNA (siCtrl) or RNH1-targeted siRNA (siRNH1). Expression levels were calculated using the $\Delta\Delta C_t$ method and normalized relative to GAPDH expression. **B.** Western analysis of RNH1 expression (RNH1) in control (siCtrl) and RNH1 depleted 293T cells (siRNH1). α Tubulin is shown as a loading control. Molecular weight in kilodaltons is marked to the right for reference. **C.** Quantification of the DNA:RNA immunoprecipitation (DIP) signal shown as a percent of input in siCtrl and siRNH1 293T cells. Pre-treatment of lysate with *in vitro* RNaseH (In vitro RNAH) enzyme serves as a control for RNA:DNA hybrids.

IgG is a non-specific antibody whereas S9.6 is a RNA: DNA hybrid-specific antibody. Analysis of 3 technical repeats from a representative experiment is shown. p values were computed using a 3-way ANOVA's Sidak's multiple comparisons test (*, $p < 0.05$). Error bars represent standard error of the mean. D. Representative images of S9.6 immunofluorescence on RPE1 control (siCtrl) and RNH1-depleted (siRNH1) cells. Blue staining marks the nuclei and red is S9.6 signal (RNA:DNA hybrids). E. Quantification of S9.6 signal (raw integrated density) (arbitrary units) for siCtrl and siRNH1 cells and shown is one of three independent experiments where a minimum of 80 nuclei were analyzed per sample. p values were computed using a non-parametric Mann-Whitney test (*, $p < 0.05$). Error bars represent standard error of the mean.

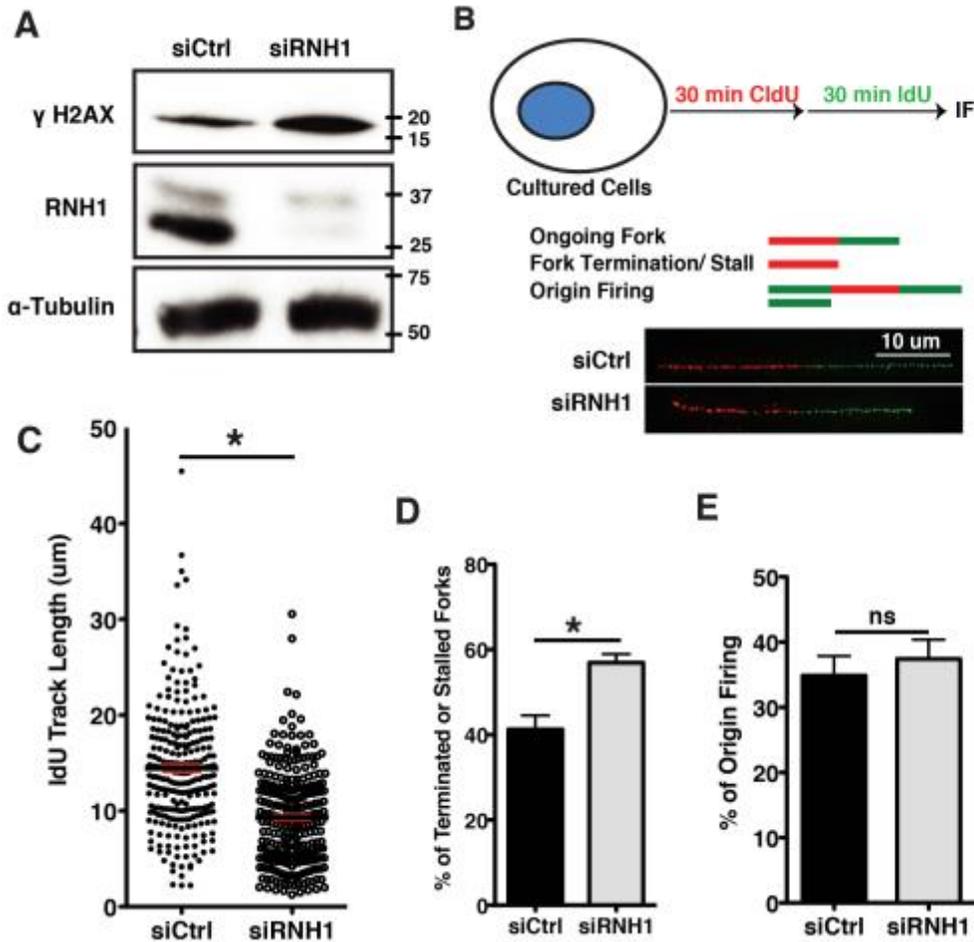


Figure 3.3: RNH1 depletion results in replication fork slowing and increased termination and stalling

A. Western analysis of RNH1 expression (RNH1) and H2AX phosphorylation (γ H2AX) in RPE1 cells transfected with siCtrl and siRNH1. α Tubulin is shown as a loading control. Molecular weight in kilodaltons is marked to the right for reference. **B.** Schematics showing labeling of cells for microfluidic- assisted replication track analysis (maRTA). Transfected RPE1 cells were labeled with base analogs CldU and IdU for 30 min each and subjected to the maRTA protocol and DNA visualization in red (CldU) and green (IdU) by immunofluorescence (IF). Ongoing forks were marked by a red track

(IdU) followed by green (CldU); terminated and/or stalled were red-only tracks; origin firings were both green-only and red flanking green on either side. Representative DNA tracks for siCtrl and siRNH1 samples are shown. **C.** A representative quantification of three independent biological experiments showing the IdU track length (μm) preceded by a CldU track. Analysis included a minimum of 260 two-color DNA tracks (moving fork) isolated from siCtrl and siRNH1 cells each. p values were computed using a two-tailed Student's t test (*, $p < 0.05$). Error bars represent standard error of the mean. **D.** Quantification of percentage of termination and stalling events in DNA isolated from siCtrl and siRNH1 samples. Mean from three independent experiments were analyzed and each analysis included between 210 and 260 DNA tracks per sample. p values were computed using a two-tailed Student's t test (*, $p < 0.05$). Error bars represent standard error of the mean. **E.** Quantification of percentage of origin firings in DNA isolated from siCtrl and siRNH1 samples. Graph represents combined means from three independent experiments that included between 275 and 350 DNA tracks per sample. p values were computed using a two-tailed Student's t test (ns, $p > 0.05$). Error bars represent standard error of the mean.

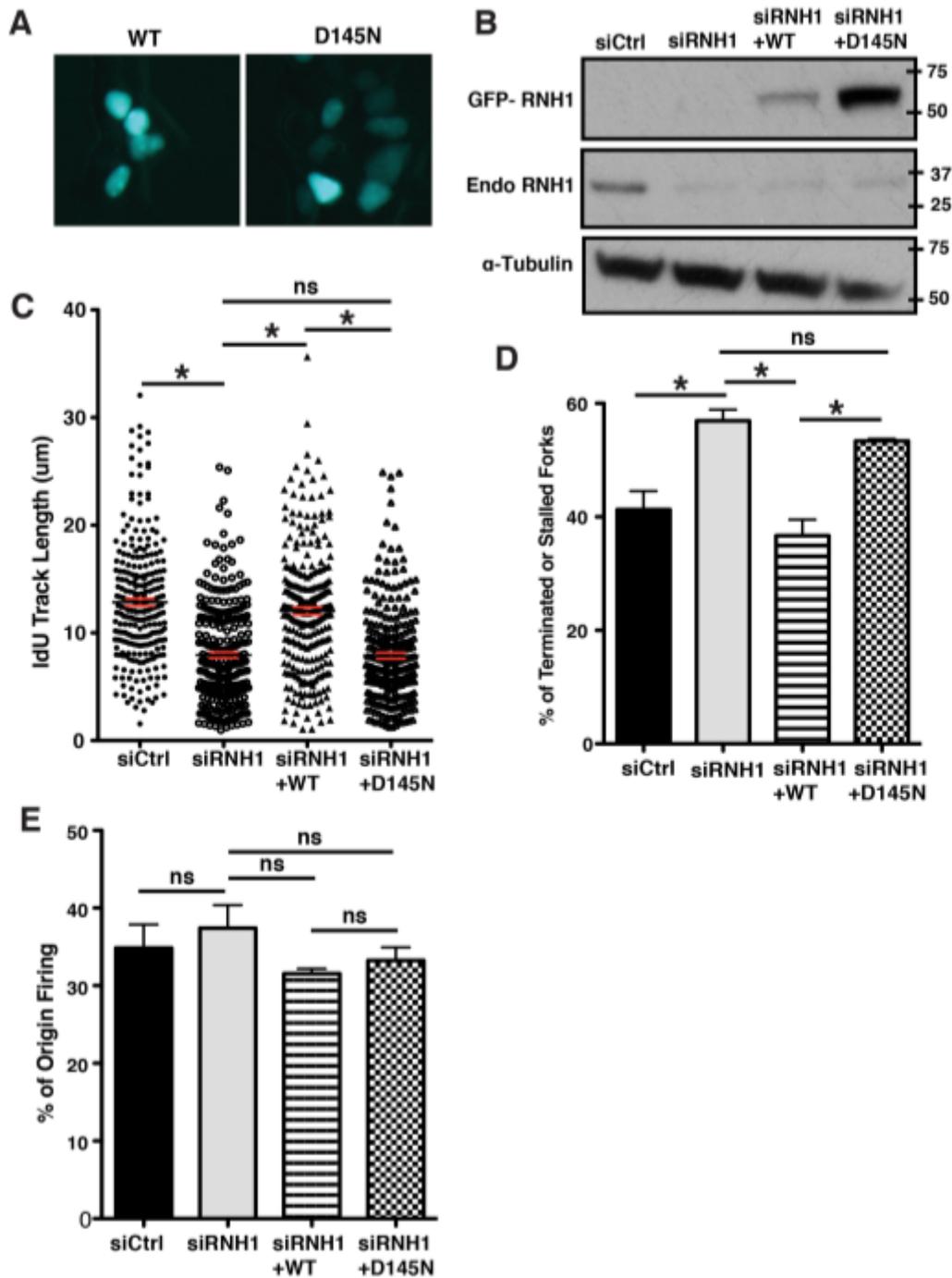


Figure 3.4: Nuclease activity of RNH1 is required for efficient replication fork movement

A. Representative images verifying the nuclear localization (green) of ectopically expressed RNH1 in 293T cells transfected with a GFP-tagged wildtype RNH1 (WT) or

nuclease dead (D145N) allele. **B.** Western analysis of RNH1 expression (endogenous and ectopic) in RPE1 cells transfected with a control siRNA (siCtrl) or RNH1-directed (siRNH1) with or without ectopic expression of either GFP tagged wildtype (siRNH1+ WT) or nuclease dead (siRNH1+ D145N) RNH1. α Tubulin is shown as a loading control. Molecular weight in kilodaltons is marked to the right for reference. **C.** A representative quantification of three independent biological experiments showing the IdU track length (μm) preceded by a CldU track. Analysis included 265 to 280 two-color DNA tracks (ongoing fork) isolated from each of the four samples. p values were computed using a one-way ANOVA with Bonferroni multiple comparisons test (*, $p < 0.05$; ns, $p > 0.05$). Error bars represent standard error of the mean. **D.** Quantification of percentage of termination and stalling events in isolated DNA from all four samples. Mean from three independent experiments were analyzed and each analysis included between 240 and 350 DNA tracks per sample. Error bars represent standard error of the mean. **E.** Quantification of percentage of origin firings in DNA isolated from all four samples. Graph shown represents combined means from three independent experiments that included between 225 and 250 DNA tracks per sample. Error bars represent standard error of the mean.

References

1. O'Driscoll, M. The pathological consequences of impaired genome integrity in humans; disorders of the DNA replication machinery. *J. Pathol.* 1–16 (2016).
2. Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I. & Chédin, F. GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *TL - 23. Genome Res.*, 1590–1600 (2013).
3. Wongsurawat, T., Jenjaroenpun, P., Kwoh, C. K. & Kuznetsov, V. Quantitative model of R-loop forming structures reveals a novel level of RNA-DNA interactome complexity. *Nucleic Acids Res.* **40**, (2012).
4. Santos-Pereira, J. M. & Aguilera, A. R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* **16**, 583–597 (2015).
5. Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T. E. & Lieber, M. R. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* **4**, 442–451 (2003).
6. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).
7. Aguilera, A. & García-Muse, T. R loops: from transcription byproducts to threats to genome stability. *Mol. Cell* **46**, 115–24 (2012).
8. Li, M., Pokharel, S., Wang, J.-T., Xu, X. & Liu, Y. RECQ5-dependent SUMOylation of DNA topoisomerase I prevents transcription-associated genome instability. *Nat. Commun.* **6**, 6720 (2015).

9. Sollier, J. *et al.* Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Mol. Cell* **777–785** (2014).
10. Bhatia, V. *et al.* BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* (2014).
11. Hatchi, E. *et al.* BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol. Cell* **57**, 636–647 (2015).
12. Reijns, M. a M. *et al.* Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* **149**, 1008–1022 (2012).
13. Ohle, C. *et al.* Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. *Cell* **167**, 1001–1013.e7 (2016).
14. Huertas, P. & Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* **12**, 711–721 (2003).
15. Wellinger, R. E., Prado, F. & Aguilera, a. Replication Fork Progression Is Impaired by Transcription in Hyperrecombinant Yeast Cells Lacking a Functional THO Complex. *Mol. Cell. Biol.* **26**, 3327–3334 (2006).
16. Gómez-González, B., Felipe-Abrio, I. & Aguilera, A. The S-phase checkpoint is required to respond to R-loops accumulated in THO mutants. *Mol. Cell. Biol.* **29**, 5203–13 (2009).
17. Ruhanen, H., Ushakov, K. & Yasukawa, T. Involvement of DNA ligase III and ribonuclease H1 in mitochondrial DNA replication in cultured human cells. *Biochim. Biophys. Acta* **1813**, 2000–7 (2011).
18. Lima, W. F. *et al.* Viable RNaseH1 knockout mice show RNaseH1 is essential for

- R loop processing, mitochondrial and liver function. *Nucleic Acids Res.* **44**, 5299–5312 (2016).
19. Cerritelli, S. M. *et al.* Failure to Produce Mitochondrial DNA Results in Embryonic Lethality in *Rnaseh1* Null Mice. *Mol. Cell* **11**, 807–815 (2016).
 20. Cerritelli, S. M. & Crouch, R. J. Ribonuclease H: The enzymes in eukaryotes. *FEBS J.* **276**, 1494–1505 (2009).
 21. Suzuki, Y. *et al.* An upstream open reading frame and the context of the two AUG codons affect the abundance of mitochondrial and nuclear RNase H1. *Mol. Cell. Biol.* **30**, 5123–5134 (2010).
 22. Arora, R. *et al.* RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat. Commun.* **5**, 5220 (2014).
 23. Sagie, S. *et al.* Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. *Nat. Commun.* **8**, 14015 (2017).
 24. Teasley, D. C. *et al.* Flap endonuclease 1 limits telomere fragility on the leading strand. *J. Biol. Chem.* jbc.M115.647388 (2015).
 25. Stewart, S. A. *et al.* Lentivirus-delivered stable gene silencing by RNAi in primary cells Lentivirus-delivered stable gene silencing by RNAi in primary cells. *Rna* **9**, 493–501 (2003).
 26. Tresini, M. *et al.* The core spliceosome as target and effector of non-canonical ATM signalling. *Nature* **523**, 53–58 (2015).
 27. Honaker, Y. & Piwnica-Worms, H. Casein kinase 1 functions as both penultimate and ultimate kinase in regulating Cdc25A destruction. *Oncogene* **29**, 3324–34 (2010).

28. Lansdorp, P. M. *et al.* Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* **5**, 685–691 (1996).
29. Bailey, A. S. M., Cornforth, M. N., Kurimasa, A., Chen, D. J. & Goodwill, E. H. Strand-Specific Postreplicative Processing of Mammalian Telomeres Published by : American Association for the Advancement of Science Linked references are available on JSTOR for this article : Strand-Specific Strand-Specific Postreplicative Postreplicativ. **293**, 2462–2465 (2017).
30. Berti, M. *et al.* Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nat. Struct. Mol. Biol.* **20**, 347–54 (2013).
31. Schlacher, K. *et al.* Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* **145**, 529–542 (2011).
32. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual. Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989).
33. Siddiqui, M. S., François, M., Fenech, M. F. & Leifert, W. R. Persistent γ H2AX: A promising molecular marker of DNA damage and aging. *Mutat. Res. - Rev. Mutat. Res.* **766**, 1–19 (2015).
34. Azzalin, C. M., Reichenbach, P., Khoraiuli, L., Giulotto, E. & Lingner, J. and RNA Surveillance Factors at. *Science* **318**, 798–801 (2007).
35. Wasmuth, E. V & Lima, C. D. Fission yeast Stn1 is crucial for semi-conservative replication at telomeres and subtelomeres. *Nucleic Acids Res.* **45**, 1–15 (2016).
36. Boguslawski, S. J. *et al.* Characterization of monoclonal antibody to DNA:RNA and its application to immunodetection of hybrids. *J. Immunol. Methods* **89**, 123–

- 130 (1986).
37. Jones, R. M. *et al.* Increased replication initiation and conflicts with transcription underlie Cyclin E-induced replication stress. *Oncogene* **32**, 3744–3753 (2013).
 38. Sidorova, J. M., Li, N., Schwartz, D. C., Folch, A. & Monnat, R. J. Microfluidic-assisted analysis of replicating DNA molecules. *Nat. Protoc.* **4**, 849–61 (2009).
 39. Zhong, Y. *et al.* The level of origin firing inversely affects the rate of replication fork progression. *J. Cell Biol.* **201**, 373–383 (2013).
 40. Wu, H., Lima, W. F. & Crooke, S. T. Investigating the Structure of Human RNase H1 by Site-directed Mutagenesis. *J. Biol. Chem.* **276**, 23547–23553 (2001).
 41. Chakraborty, P. & Grosse, F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst)*. **10**, 654–665 (2011).
 42. Mischo, H. E. *et al.* Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell* **41**, 21–32 (2011).
 43. Boué, J. B. & Zakian, V. A. The yeast Pif1p DNA helicase preferentially unwinds RNA-DNA substrates. *Nucleic Acids Res.* **35**, 5809–5818 (2007).
 44. Nguyen, H. D. *et al.* Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Mol. Cell* **in-press**, 832–847.e4 (2017).
 45. Richard, P. & Manley, J. L. R loops and links to human disease. *J. Mol. Biol.* (2016).

Chapter 4:
Conclusions and future directions

Shankar Parajuli

Conclusions

DNA replication and DNA repair are vital cellular processes that maintain genome stability. While numerous redundant mechanisms have evolved to ensure that these processes are faithfully executed, they are not immune to dysfunction and failure. When severe, such problems can result in genome instability, cell cycle arrest, senescence, and apoptosis (1), (2). Importantly, such manifestations are associated with a wide range of human pathologies from autoimmune diseases and cardiovascular diseases to muscular dystrophy and cancer. Therefore, understanding the molecular mechanism of DNA replication and DNA repair is critical for better understanding human diseases and developing therapeutics.

To that end, my thesis work uncovers novel functions of two DNA replication and repair proteins. First, we show that human flap endonuclease 1 (FEN1), a structure specific endonuclease, is required for maintaining telomere stability. In particular, FEN1 loss elicits telomere fragility on the leading strand replicated telomere, a phenotype uncharacteristic of a lagging strand protein. Second, we discover that human ribonuclease H1 (RNH1), an RNA:DNA hybrid specific nuclease, facilitates nuclear DNA replication. We propose that RNH1 accomplishes this by resolving RNA:DNA hybrids that can present roadblocks to a moving replisome complex.

FEN1 limits telomere fragility on the leading strand

As described in chapter 2, we provide the first evidence that a lagging strand protein, FEN1, functions in leading strand DNA replication by preventing leading strand-specific

telomere fragility. This finding was in a stark contrast to a lagging strand specific telomere phenotype that our laboratory had previously reported where FEN1 prevented a sister telomere loss (STL) phenotype at the lagging strand (3), (4). Because RNA:DNA hybrids form at the leading strand telomere, we proposed that the fragility phenotype could be a result of a collision between the replication and transcription machineries (5). To test this hypothesis, we chemically inhibited transcription and found that telomere fragility was exacerbated suggesting that transcription was an important element of the fragility phenotype at the leading strand. Furthermore, we also identified that this phenotype was a result of RNA:DNA hybrids as evidenced by the rescue of the phenotype upon ectopic expression of RNH1 in FEN1-depleted cells. Finally, we showed that FEN1's flap endonuclease activity, not its interaction with PCNA nor its gap endonuclease and exonuclease activity, was required to limit leading strand telomere fragility phenotype. Taken together, we propose a model wherein FEN1 limits leading strand fragility by using its flap endonuclease activity to cleave the RNA:DNA hybrid structures produced as a result of a replisome-transcription collision event.

Human RNH1 limits R loops and facilitates DNA replication

We report in chapter 3 a novel role of human RNH1 in the nuclear DNA replication. While human RNH1 is required for mitochondrial DNA replication, its nuclear function has remained unclear except at the telomere of ALT cells (6), (7). Given that RNA:DNA hybrids were abundant in the nucleus and that these topological structures could pose a challenge to a moving replication fork, we hypothesized that RNH1 loss may elicit replication defects. Upon RNH1 depletion, we first found a DNA damage response and

accumulation of RNA:DNA hybrids, suggesting that these structures are responsible for inducing DNA damage. We also analyzed metaphase chromosomes and found that RNH1-depleted cells had elevated levels of telomere loss in comparison to the control cells. This result was not surprising because of the important roles that RNA:DNA hybrids play to maintain telomere stability (8), (6), (9). This also suggested that this could be a result of replication fork collapse at the difficult-to-replicate telomere template. Thus, we proposed a hypothesis that RNH1 might be important in DNA replication outside of telomeres. To test that hypothesis, we utilized a single molecule DNA replication technique to examine DNA replication efficiency (10). We discovered that the replication forks slowed significantly upon RNH1 depletion. Similarly, replication termination and stalling events also increased upon RNH1 loss. Although we observed significant DNA replication defects, we were surprised that cells failed to activate the S phase checkpoint and cell cycle arrest in RNH1 depleted cells. This suggests that either RNH1 depletion is not significant enough to warrant a checkpoint arrest or that the RNH1 function is compensated by other ribonucleases. It is also possible that the checkpoint activation is at a level undetectable by our current tools.

Future Directions

Structure and physiology of telomere fragility

While telomere fragility results from several manipulations including TRF1 loss in mice and FEN1 loss in human cells as described in chapter 2, its structure composition and physiological relevance are largely unknown (11). First, to date, fragile telomeres are only described as abnormal telomeres that appear as multiple or smeared telomeres as

visualized on metaphase chromosomes. However, how such structures are formed and what proteins decorate and stabilize them remain unclear. Because visualization of some of the fragile telomeres using telomere specific PNA probes reveal a non-telomeric sequences interspersed between two or more telomeric sequences, one might speculate such structures result from recombination between telomeric and non telomeric DNA sequences. It is also possible that those non-telomeric sequences are the sub-telomeric regions. Furthermore, whether such non-telomeric sequences are sequences of the same chromosome or different chromosomes remains to be seen. One way to address these questions would be to combine a telomere PNA FISH with a sub telomere DNA FISH on metaphase chromosomes. Because of the currently availability of chromosome-specific sub telomeric DNA probes, these experiments can reveal whether fragile telomeres are a result of intra-chromosomal or inter-chromosomal recombination or both (12). These experiments would also uncover if those non-telomeric regions are sub telomeric sequences. Similarly, a lot of these structural identities of fragile telomeres could be revealed by a new generation of super-resolution microscopy. Another important question to investigate telomere fragility is to further explore the roles of RNA:DNA hybrids. Although we suggested that telomere fragility induced upon FEN1 depletion is RNA:DNA hybrid-dependent as evidenced by rescue upon ectopic RNH1 expression, we did not directly measure those levels. Future experiment could employ S9.6, an RNA:DNA hybrid specific antibody, to carry out an immunoprecipitation to directly measure hybrid levels.

A second set of intriguing questions regarding telomere fragility revolve around its physiological relevance. In other words, what is the significance of fragile telomere, if any? To begin to address this broad question, we could begin to ask specific questions. Are fragile telomeres stalled intermediates of a repair process? Or are these end products of a failed repair? Similarly, why would a cell form such structures? Is it a preferred damage state for cell survival? Some of these questions can be answered by beginning to closely look at the repair systems. One way to do this would be to utilize a genetic approach and knockdown or knockout proteins of both DNA replication and repair. However, such genetic knockdown experiments should be conducted with proper controls to avoid confounding results. For example, knockdown of Rad51, an essential homologous recombination protein, led to an increase in telomere fragility (12). This is counter-intuitive given telomere fragility is thought to be a recombination-dependent phenotype. However, it turns out that Rad51 is also required for telomere replication and capping. Furthermore, because telomere fragility occurs through several means as discussed earlier, it is possible that the mechanism of formation of such structures may dictate their physiological roles thereby further complicating attempts to discern such structures. To understand if fragile telomeres were stable or transient structures, our laboratory conducted a series of time-course experiments by utilizing low doses of aphidicolin to induce telomere fragility. Our unpublished and preliminary data show that this phenotype was rescued 48 hours later suggesting that cells have mechanisms to repair or resolve such structures. However, whether these findings apply to fragility induced by other mechanisms such as TRF1 loss remains to be seen.

Regulation of RNH1 during DNA replication

Although we identified a novel function of RNH1 in DNA replication as discussed in chapter 3, a lot remains unclear about how RNH1 is regulated and how it functions in the nucleus. First, we show that RNH1 depletion slows down replication fork movement and increases termination and stalling events but how RNH1 is assisting the replication machinery is unclear. Is RNH1 always associated with the fork during replication? Or is RNH1 recruited to assist the fork only when the RNA:DNA hybrids are encountered? One way to address these questions would be to conduct an isolation of proteins on nascent DNA (iPOND) technique (13). This technique allows for the labeling of nascent DNA followed by the purification of all proteins associated with an active replication complex. PCNA, as an essential DNA replication protein, can serve as a positive control to identify if RNH1 behaves in a similar fashion. In fact, using iPOND, our preliminary data show that RNH1 associates with a moving replication fork. Regardless of whether RNH1 associates with an active replisome complex or not, identifying RNH1's interaction partners might shed light on how it is regulated.

Another important question revolves around how RNH1 functions in the nucleus to facilitate DNA replication. Although we show that RNH1 depletion elicits several DNA replication defects, how RNH1 does this is unclear. We only provide a correlation between increased hybrid levels and replication defects upon RNH1 depletion. It is possible that there could be indirect effects of RNH1 that we have yet to consider. One way to address this would be to inhibit replication by using aphidicolin, for example, and assess replication defects in control and RNH1-depleted cells. If replication defects we

observed upon RNH1 depletion are indirect effects, we can test a possibility that RNH1 may work together with other nucleases to co-ordinate its function. Furthermore, it is worth exploring if depletion of other ribonucleases such as RNH2 or helicases such as SETX and AQR would elicit such replication defects. This would require conducting DNA replication studies as described in chapter 3 after depletion of these factors. Lastly, understanding how a cell responds to RNH1 depletion and associated replication defects can provide insights into its function. Because global DNA damage response is observed upon RNH1 depletion as evidenced by increase in γ H2AX, it will be interesting to investigate the cell cycle checkpoint response. Examining checkpoint markers such as Chk1, RPA, Chk2, ATM, and ATR may open insight into the cellular stress response upon RNH1 depletion. It is possible that other nucleases or helicases discussed above may compensate for RNH1 function thereby preventing checkpoint response. Alternatively, accumulated hybrids upon RNH1 loss may be processed by DNA repair system without invoking checkpoint response in a manner slightly different than reported with hybrids that accumulate upon AQR depletion.

References

1. Harper JW, Elledge SJ. The DNA Damage Response: Ten Years After. *Mol Cell*. 2007;28(5):739–45.
2. Toueille M, Hübscher U. Regulation of the DNA replication fork: A way to fight genomic instability. *Chromosoma*. 2004;113(3):113–25.
3. Saharia A, Guittat L, Crocker S, Lim A, Steffen M, Kulkarni S, et al. Flap endonuclease 1 contributes to telomere stability. *Curr Biol [Internet]*. 2008 Apr 8 [cited 2013 Nov 25];18(7):496–500.
4. Saharia A, Teasley DC, Duxin JP, Dao B, Chiappinelli KB, Stewart S a. FEN1 ensures telomere stability by facilitating replication fork re-initiation. *J Biol Chem [Internet]*. 2010 Aug 27 [cited 2013 Nov 25];285(35):27057–66.
5. Cusanelli E, Chartrand P. Telomeric repeat-containing RNA TERRA: a noncoding RNA connecting telomere biology to genome integrity. *Front Genet [Internet]*. 2015;6(April):1–9.
6. Arora R, Lee Y, Wischnewski H, Brun CM, Schwarz T, Azzalin CM. RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun [Internet]*. Nature Publishing Group; 2014 Jan [cited 2014 Oct 22];5:5220.
7. Cerritelli SM, Crouch RJ. Ribonuclease H: The enzymes in eukaryotes. *FEBS J*. 2009;276(6):1494–505.
8. Balk B, Maicher A, Dees M, Klermund J, Luke-Glaser S, Bender K, et al. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence.

- Nat Struct Mol Biol [Internet]. Nature Publishing Group; 2013;20(10):1199–205.
9. Xu Y, Komiyama M. Structure, function and targeting of human telomere RNA. Methods [Internet]. Elsevier Inc.; 2012;57(1):100–5.
 10. Sidorova JM, Li N, Schwartz DC, Folch A, Monnat RJ. Microfluidic-assisted analysis of replicating DNA molecules. Nat Protoc [Internet]. 2009;4(6):849–61.
 11. Sfeir A, Kosiyatrakul ST, Hockemeyer D, Macrae SL, Schildkraut CL, Lange T De. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. 2010;138(1):90–103.
 12. Ravnan JB, Tepperberg JH, Papenhausen P, Lamb a N, Hedrick J, Eash D, et al. Subtelomere FISH analysis of 11 688 cases: an evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. J Med Genet [Internet]. 2006;43(6):478–89.
 13. Sirbu BM, Couch FB, Cortez D. Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA. Nat Protoc [Internet]. Nature Publishing Group; 2012;7(3):594–605.