IDENTIFYING AND DEFINING THE GENOME MAINTENANCE FUNCTIONS OF SMARCAL1

By

Carol Bansbach Robbins

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirments

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

August, 2012

Nashville, Tennessee

Approved:

Professor David Cortez

Professor Scott Heibert

Professor Brandt Eichman

Professor Christine Eischen

To my parents and my family

and

To my favorite man, Matthew Robbins

ACKNOWLEDGEMENTS

I would first like to acknowledge the source of financial support that contributed to the completion of my thesis work. I received funding from the Center in Molecular Toxicology through a National Institute of Environmental Health Sciences (NIEHS) training grant (NIH T32 ES007028, Training Program in Environmental Toxicology). Also, I am honored that the faculty of the biochemistry department distinguished me with the 2009 Leon W. Cunningham Award.

I am especially indebted to Dr. David Cortez, my mentor, who has been instrumental in my development as a scientist. I have tried to absorb as much of his intelligence as possible in the five years I have been a part of his laboratory. I really admired the skill with which he ran a laboratory, mentored, and found time to work at the bench. When we found out that multiple laboratories were studying SMARCAL1, he left his office on many occasions to help with experiments that were crucial to advancing publication of the SMARCAL1 story. I am very thankful that Dr. Cortez provided me with the independence to direct my project. I feel equipped to conduct independent research in the future.

I would like to acknowledge the members of my thesis committee for their input and support throughout my graduate career. Discussions with Drs. Brandt Eichman, Christine Eischen, Scott Heibert, and Larry Marnett were essential to the progress of my dissertation research. They helped to identify critical experiments that were lacking and confirm those that supported my conclusions.

I would also like to thank Dr. Jennifer Pietenpol and members of her laboratory. The group meetings with the Pietenpol laboratory provided another opportunity to discuss my project and hear feedback that generated new ideas and perspectives.

I am grateful to every member of the Cortez Laboratory, past and present. I am confident it is the best laboratory to work in because of the people. Thank you Gloria Glick for placing our orders, ensuring the lab runs orderly, and keeping all of us in order. I cannot type or say a big enough THANK YOU to Courtney Lovejoy. Working along side of her my first year and a half in the lab was the best! I learned so much from her and we had a blast screening genes together. She will always be a dear friend and colleague. To Edward Nam, Bianca Sirbu, and Jami Couch, thanks for helping me get through the ups and downs of graduate school. Eddie, it was a sad day when you graduated and left me with no one to fill the drama-filled void in my life. You also gave me great science advice and taught me the precious skill of phosphopetide mapping. Bianca, you always found a way to keep things positive and always had an encouraging word to inspire me to get up and try again. In addition, your dancing and iPOND skills will forever amaze me. And Jami, you were like a mini-Dave to me. Somehow you know everything and that made you a great resource for everything. All three of you made lab more than a job; you made it an experience that I will always cherish. Thanks.

Finally, I would like to acknowledge all of my friends and family. They keep me sane and provide a never-ending source of encouragement, love, and

fun times. To my husband, who decided to marry me even though I was in lab constantly and stressed out most of the time, thank you for telling me I could do it. Even though I did not know you when I started this journey, it is hard to picture finishing it without you by my side. I love you.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AH2	annealing helicase 2
APH	aphidicolin
ATM	Ataxia Telangiectasia Mutated
ATP	adenosine tri-phosphate
ATR	ATM- and RAD3-Related
ATRIP	ATR-interacting protein
BrdU	bromodeoxyuridine
BSA	bovine serum albumin
CHK1	checkpoint kinase 1
CID	C-terminal inhibition domain
CDK	cyclin dependent kinase
CDT1	CDC10-dependent transcript 1
CPT	camptothecin
DAPI	4’6’-diamidino-2-phenylindole
DDR	DNA damage response
DDRi	DDR kinase inhibitors
DNA	deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
dNTP	deoxyribonucleotide triphosphate
dsDNA	double stranded DNA
G1/G2	gap phase 1/gap phase 2

GFP	green fluorescent protein
γH2AX	H2AX phospho-serine 139
H1/H2	HARP1/HARP2
HARP	Hep-A related protein
HR	homologous recombination
HU	hydroxyurea
ICL	interstrand crosslink
IP	immunoprecipitation
IR	ionizing radiation
KAP1	KRAB domain-associated protein 1
Ku70/80	Lupus Ku autoantigen protein p70/p80
MCM	minichromosome maintenance
MRN	MRE11-RAD50-NBS1
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NLS	nuclear localization signal
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PIKK	PI-3 kinase-related kinase
RBD	RPA binding domain
RDR	recombination dependent repair
RPA	replication protein A

RFC	replication factor C
SDS	sodium dodecyl sulfate
SIOD	Schimke immuno-osseous dysplasia
SMARCAL1	SWI/SNF, matrix associated, actin dependent regulator of chromatin A-like 1
SNF2	sucrose non-fermenting 2
siRNA	small interfering RNA
ssDNA	singe stranded DNA
S/TQ	serine/threonine glutamine
TIPIN	timeless interacting protein
TLS	translesion synthesis
UV	ultraviolet radiation

CHAPTER I

INTRODUCTION

Approximately 10,000 trillion cell divisions happen in the typical human lifetime. Cell division occurs via an ordered set of events collectively termed the cell division cycle. During each cell division cycle, over 6 billion base pairs of DNA must be replicated during DNA synthesis (S-phase) and then segregated to each daughter cell during mitosis (M-phase). Cellular energy is focused on growth and preparation for S and M phases during the Gap 1 (G1) and Gap 2 (G2) phases. Furthermore, the cell division cycle happens in the context of thousands of DNA lesions that occur in each cell each day. In all but a few specialized cases, the goal is to create and maintain an identical genome for each cell in the body. The faithful replication and maintenance of the genome is essential for cellular viability and disease prevention.

Maintaining the genome requires the concerted actions of cellular metabolism, cell cycle, and DNA repair activities that together constitute genome maintenance pathways (Figure 1.1). Enzymes like superoxide dismutase protect cellular molecules from the action of reactive oxygen species generated as a byproduct of cellular metabolism, thereby reducing the DNA damage burden in the cell (1). DNA lesions that do occur are identified and repaired by multiple DNA repair systems that scan the genome looking for imperfections (2). Many proteins work to ensure a single round of DNA replication is completed with

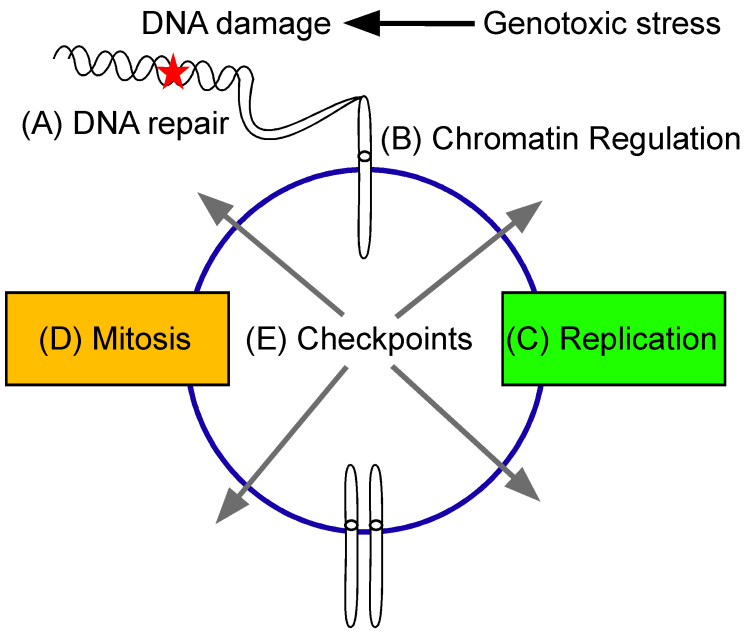

Figure 1.1. Genome maintenance requires the coordination of multiple cellular activities. (A) Multiple DNA repair pathways operate to remove DNA lesions caused by endogenous and exogenous genotoxic agents. (B) DNA repair and metabolism occurs in the context of chromatin. Chromatin modifications regulate protein access to the DNA as well as signaling responses to DNA damage. (C) DNA replication must faithfully duplicate the DNA and chromatin structure once and only once per cell division cycle. (D) Proper spindle assembly and chromosome segregation during mitosis ensures each daughter cell receives a complete copy of the genome. (E) Cell cycle checkpoints monitor DNA damage, replication, and mitosis.

minimal errors each cell cycle (3). Proper chromosome segregation during mitosis requires spindle assembly, kinetochore attachment, and physical separation (4). All of these genome maintenance activities happen in the context of chromatin necessitating a large cohort of enzymes that control chromatin dynamics (5).

Finally, many of these activities are coordinated through signaling pathways that constitute the DNA damage response (DDR). DDR signaling regulates cell cycle checkpoints, DNA repair, DNA replication, and can promote the elimination of damaged cells from the cycling population through either apoptosis or senescence (6).

In this chapter, I will briefly introduce the regulators of DDR signaling and how they are activated to coordinate genome maintenance pathways. Next, I will focus on the replication stress response, a DDR pathway that is required during DNA replication. Many types of DNA damage present challenges to the completion of DNA replication. The essential function of the replication stress response is to promote activities that overcome these challenges to complete faithful duplication of the genome. Finally, I will introduce the SNF2 family of proteins. The ATP-dependent activities of many SNF2 proteins are required in genome maintenance pathways. Importantly, defects in SNF2 genome maintenance functions are associated with human disease. My thesis work led to the identification of the SNF2 protein SMARCAL1 as a genome maintenance protein. I further characterized SMARCAL1 and determined that it functions as a component of the replication stress response. *SMARCAL1* mutations cause the

human disease Schimke immunosseous dysplasia (SIOD). The data presented in this thesis suggests that loss of SMARCAL1 function in patients may cause DNA replication-associated genome instability that contributes to the phenotypes of SIOD.

The DNA Damage Response

The DDR is an evolutionarily conserved network of genome surveillance systems that senses aberrant DNA structures, transmits the DNA damage signal, and responds to DNA damage by coordinating cellular responses that promote genome maintenance (6-8). The cellular response to DNA damage varies depending on the type of DNA lesion. The DNA damage-sensing component of DDR pathways dictates which DNA repair pathway will be used and recruits DDR kinases to initiate a DNA damage signal cascade. In fact, artificially concentrating sensor proteins in cells is sufficient to activate a DNA damage signal in the absence of DNA damage (9). After sensing the damage, transmitting the DNA damage signal throughout the nucleus is critical to activate the proteins that carry out the response to DNA damage.

DNA damage response kinases

The major regulators of DDR signaling in mammalian cells are the protein kinases ataxia-telangiectasia mutated (ATM), ATM and Rad3-related (ATR), and DNA-dependent protein kinase (DNA-PK). The DDR kinases ATM, ATR, and DNA-PK belong to a conserved family of phosphoinositide 3-like (PI3K)-related

protein kinases (PIKKs). All three share significant sequence homology and target an overlapping set of substrates involved in the DDR. However, distinct types of DNA damage activate these kinases. This distinction is dependent on the recruitment of DNA-PK, ATM, and ATR to DNA lesions by specific DNA damage-sensing proteins (Figure 1.2). Localization to the site of DNA damage is required for activation of DDR kinases.

DNA-PK and ATM are activated in response to rare occurrences of double-stranded DNA breaks (DSBs). The Mre11-Rad50-Nbs1 (MRN) complex and Ku70/80 heterodimer both bind DNA ends at DNA double stranded breaks (DSBs) and recruit ATM and DNA-PK, respectively. DNA-PK activation at DSBs promotes non-homologous end-joining repair (NHEJ) of DSBs (10). NHEJ repair seals breaks by ligating DNA ends at the expense of creating microdeletions. NHEJ activity is required to rejoin breaks generated during immunoglobulin class switch recombination. This process is essential for lymphocyte development and requires both DNA-PK and ATM (11).

ATM was first identified as the product of the gene that is mutated or lost in the human genetic disorder ataxia-telangiectasia (A-T). Cells from A-T patients exhibit extreme sensitivity to DSB inducing agents. In response to DSBs, ATM enforces cell cycle checkpoints, promotes DSB repair, and when DSBs persist unrepaired, activates cell death pathways (12).

ATR responds to DNA damage types that cause single-stranded DNA (ssDNA). The repair of base adducts, DNA crosslinks, and DSBs includes ssDNA intermediates that can activate ATR. However, ssDNA is most frequently

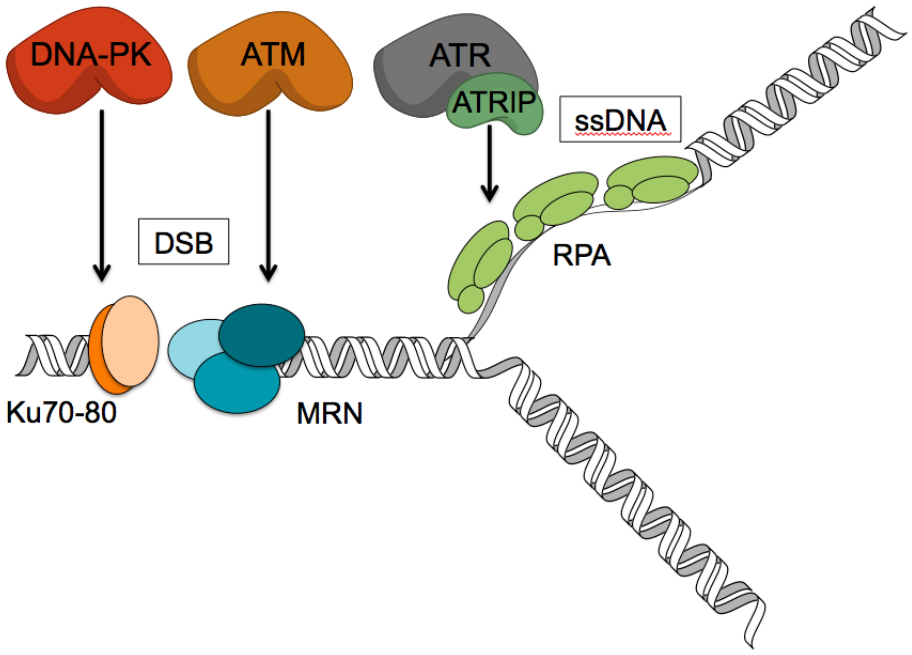

Figure 1.2. Localization to sites of DNA damage activates DDR kinases. DNA-PK, ATM, and ATR are protein kinases that regulated the activities of the DDR. By phosphorylating hundreds of overlapping substrates, they coordinate many cellular responses that preserve genome integrity. DNA-PK and ATM are activated in response to DSB. Ku70/80 and MRN bind DNA double-stranded ends and recruit DNA-PK and ATM to the damage site, respectively. ATR responds to damage that causes ssDNA, most frequently generated during DNA replication. RPA is a ssDNA binding protein. RPA-ssDNA recruits ATR through an interaction with ATRIP.

generated during DNA replication stress (mechanism discussed below). RPA is a ssDNA binding protein that consist of three subunits, RPA1, RPA2, and RPA3. RPA binds ssDNA in the cell and interacts with many DNA processing proteins (13). RPA-ssDNA binds ATR through its obligate binding partner, ATR-interacting protein (ATRIP) and this is important for the localization of ATR to sites of DNA damage (14). RPA-ssDNA is not sufficient for ATR activation at replication forks but it does coordinate the loading of the 9-1-1 complex, which is required for localizing the protein activator of ATR, topoisomerase-binding protein-1 (TOPBP1). Once activated, ATR signaling coordinates many of the same activities as ATM, controlling cell cycle, repair, and when necessary activating pathways of apoptosis.

Unlike DNA-PK and ATM, the ATR/ATRIP complex is essential for the viability of replicating cells and is activated every S-phase to respond to DNA replication stress (14-17). The requirements for ATR activity during S phase emphasize that the genome is particularly sensitive to ssDNA accumulation during the process of DNA replication. Furthermore, it indicates that ATR signaling is especially important to coordinate DDR activities that maintain the genome during replication.

The Replication Stress Response

Proliferating cells must complete S phase under challenging conditions where DNA lesions, DNA-protein barriers, and DNA secondary structures slow or inhibit replication fork progression causing an event referred to as replication

stress. The replication stress response is a component of the DDR that acts during every cell division cycle to overcome complex replication fork challenges and ultimately promote faithful and complete duplication of the genome. Failures in the response to replication stress result in the accumulation of mutations and genomic rearrangements. In this section, I will introduce the basic components of DNA replication and discuss possible mechanisms that preserve the integrity of replication forks that are challenged.

DNA replication

DNA replication initiates from replication origins dispersed throughout the genome. Not all origins are used or 'licensed' for DNA replication in every cell cycle. Thus, marking replication origins competent for replication initiation requires an extensively regulated and ordered series of events (18-20). For those origins that do initiate (or fire) each round of DNA replication, DNA unwinding and synthesis proceeds bi-directionally, forming a replication bubble with two diverging replication forks (Figure 1.3A). DNA synthesis begins at a free 3'OH group of a short RNA primer synthesized by DNA polymerase alpha (α) on both strands. The two parental DNA strands are synthesized by different mechanisms due to the directionality of DNA synthesis (5'→3'). This means that nucleotides are only added to the 3' end of elongating primers. The leading strand refers to the template strand that is synthesized in the direction of DNA unwinding or fork progression. Synthesis on the lagging strand (parental strand complementary to the leading strand) is discontinuous and primers are

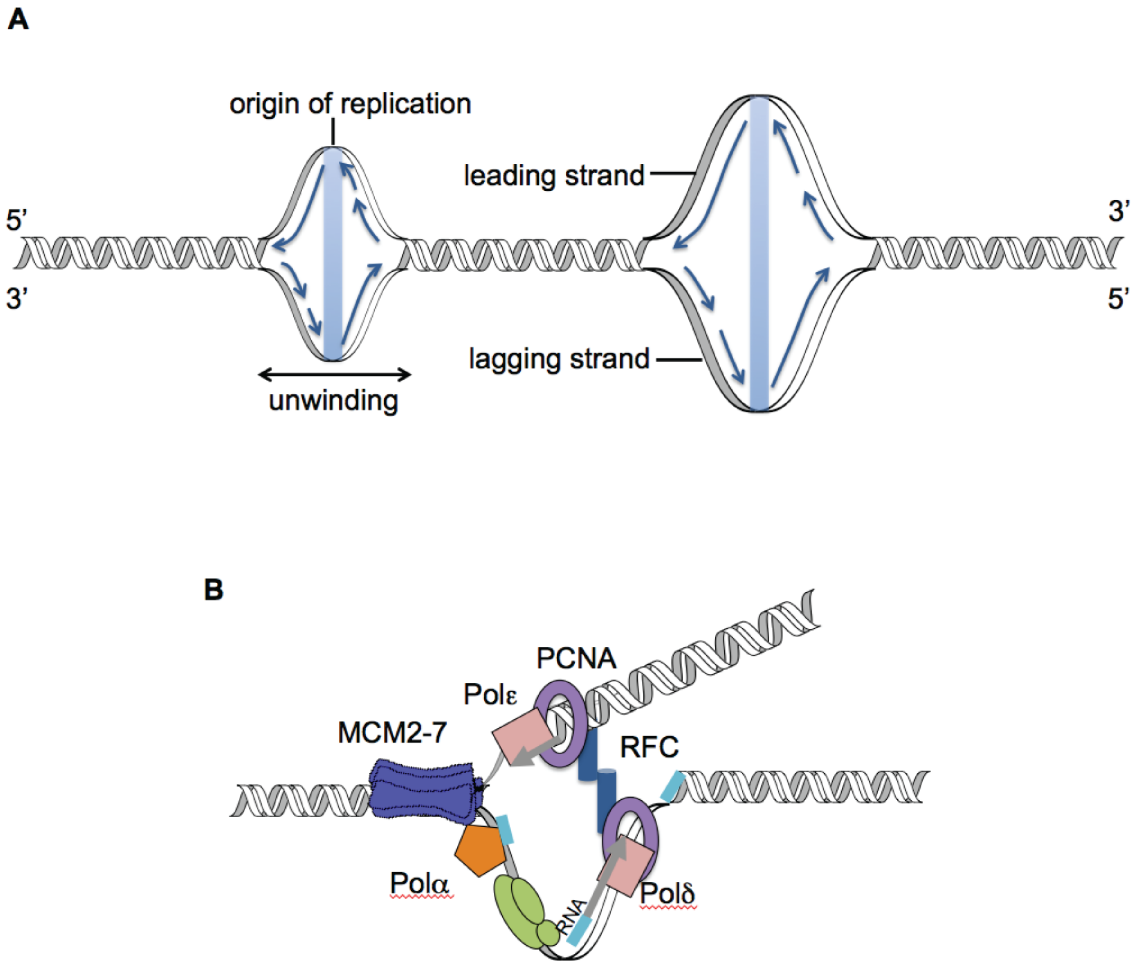

Figure 1.3. Progression of bi-directional replication forks requires a functional replisome. (A) DNA replication initiates from origins of replication and proceeds bi-directionally. Replication forks move in the direction of helicase unwinding. The two parental DNA strands at a replication fork are synthesized by different mechanisms. Leading strand synthesis is continuous in the direction of DNA unwinding. Lagging strand synthesis is discontinuous. Primers are elongated to form Okazaki fragments that are processed and ligated to one another. (B) Replication fork progression depends on maintaining a functional replisome.

repeatedly synthesized by polymerase-α and elongated by replicative polymerases. This process forms Okazaki fragments that must be eventually ligated to complete DNA synthesis on the lagging strand.

Replication fork progression depends on maintaining a functional complex of proteins that promote coupled DNA unwinding, DNA priming, and DNA synthesis. The basic complex, referred to as the replisome (Figure 1.3B), contains the replicative helicase mini-chromosome maintenance proteins 2-7 (MCM2-7), which unwinds the parental DNA in an ATP-dependent manner (21); the clamp loader, replication factor C (RFC) that loads the sliding clamp proliferating cell nuclear antigen (PCNA), which coordinates DNA polymerases at the fork enabling processive DNA replication (22); and three DNA polymerases, the polymerase-α primase and the replicative polymerases, polymerase-ε and polymerase-δ. In addition to these core replisome components other proteins travel with the fork and function in fork progression (23-25).

Sources of replication stress

Once DNA replication initiates, obstructions on the DNA template that oppose replisome activities, either DNA synthesis or unwinding, halt the progression of the replication fork and cause fork stalling. Replicative polymerases can halt synthesis when faced with insufficient dNTP pools or lesions in the DNA template. In the case of limited dNTPs, polymerases cannot incorporate the next nucleotide leading to what is termed ‘uncoupling’ of helicase and polymerase functions. DNA unwinding of parental strands continues and

ssDNA accumulates since parental strands are unwound but not copied (Figure 1.4A). Which strand (leading or lagging) is inhibited likely determines the extent to which ssDNA accumulates (Figure 1.4B). Constant re-priming of synthesis on the lagging strand facilitates bypass synthesis downstream of the lesion. Thus significant uncoupling does not occur. In Xenopus egg extracts, the bulk of new primers synthesized in response to chemical inhibitors of fork progression depend on the lagging strand polymerase (26). The leading strand polymerase also supported primer synthesis but to a lesser extent, suggesting there may be larger stretches of ssDNA on the leading strand. DNA polymerases can also stall temporarily at DNA lesions like ultraviolet (UV) induced pyrimidine dimers and base adducts, and subsequently replicate past the lesion either by polymerase exchanges or template switching (Figure 1.4C)(reviewed in (27)).

In some cases, replication stress inhibits DNA unwinding by the MCM2-7 helicase. These blocks completely stall the replisome and do not generate significant amounts of ssDNA (Figure 1.4D). Chemical crosslinks between DNA strands that prevent DNA unwinding are detrimental to genome stability if not resolved. Removing and restarting DNA synthesis from such complex lesions requires a specialized DNA repair mechanism termed interstrand crosslink (ICL) repair that uses proteins involved in nucleotide excision repair, HR, and translesion synthesis (28).

Although DNA repair pathways remove the majority of lesions before replication forks encounter them, the variety and frequency of DNA lesions insist that some DNA lesions will escape repair and encounter a replication fork. Thus,

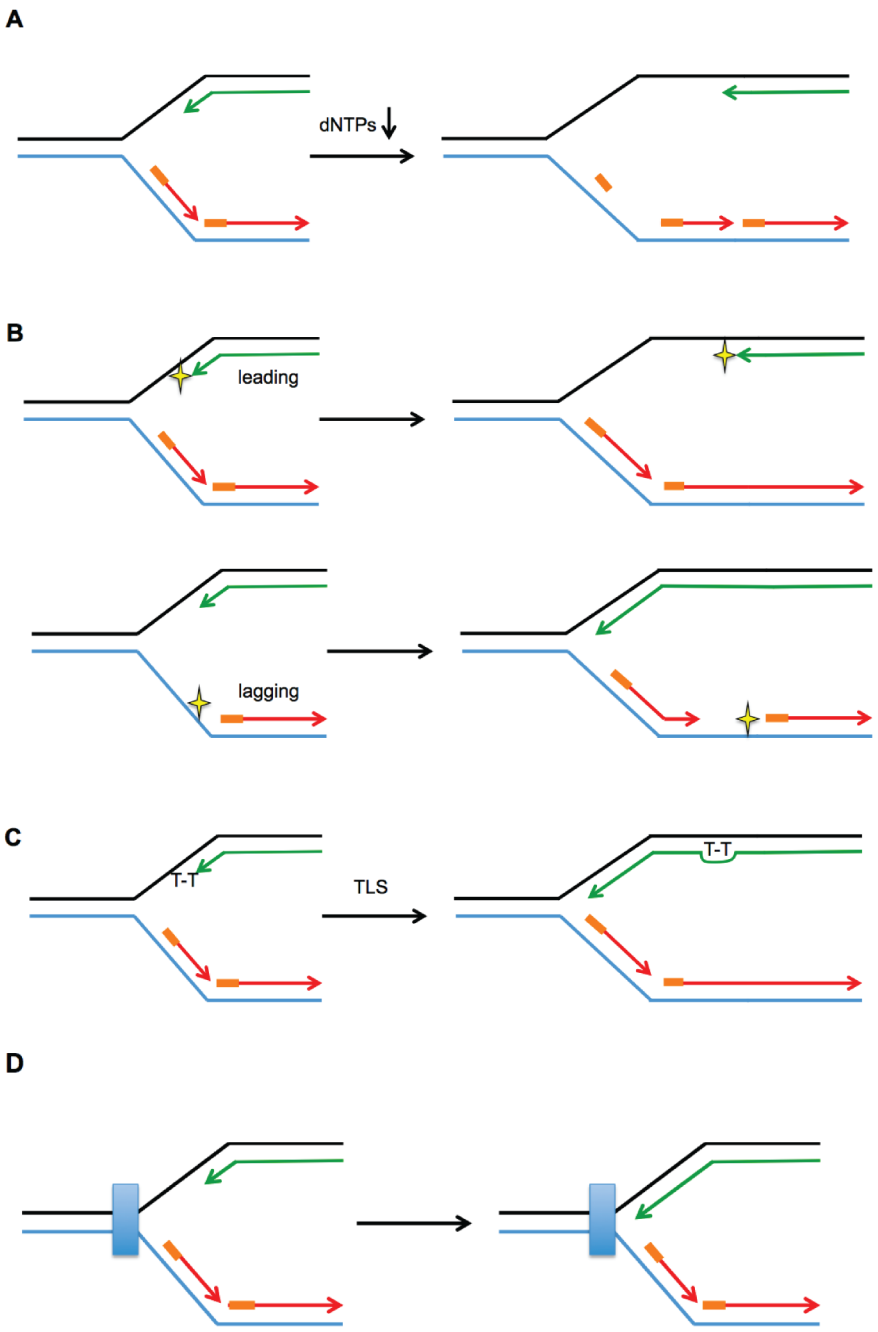

Figure 1.4. Mechanisms of replication fork stalling. Fork stalling occurs when either DNA unwinding or DNA synthesis is inhibited. (A) When nucleotides (dNTPs) are low this can stall both polymerases and lead to uncoupling of helicase and polymerase activities. This generates ssDNA. (B) DNA lesions have different affects depending which strand they block. (C) Translesion synthesis (TLS) can promote synthesis past thymidine dimers caused by ultraviolet irradiation (T-T). (D) Depicts a chemical crosslink that inhibits MCM2-7 unwinding.

cells developed strategies to respond to fork stalling that are largely dependent on the nature of the lesion. I will focus my introduction on the response to replication fork challenges that lead to functional uncoupling of helicase and polymerase functions.

Replication fork stability

In the absence of replication stress response pathways, stalled replication forks typically do not restart and accumulate aberrant DNA replication structures that eventually form DSBs. These aberrant DNA structures were directly observed by electron microscopy in replication checkpoint deficient yeast mutants (29). Numerous studies analyzing the genetic requirements for avoiding such structures and DSB accumulation during S phase have established the concept of replication fork stability. Stalled forks are actively stabilized by replication stress response factors. Stable replication forks are replication forks that retain the ability to restart replication once replication blocks are removed.

Stalled replication forks that do not restart are defined as collapsed forks. The frequency of fork collapse is much higher when fork-stabilizing functions are impaired. These forks are considered unstable and generally result in MUS81-EME1 dependent DSBs. MUS81-EME1 is a structure specific endonuclease complex that cleaves replication- and recombination-associated DNA structures (30). Although DSBs are associated with fork collapse, they are not the cause. In fact, DSB formation is an integral step in promoting the repair and restart of a subset of stalled replication forks. The current model is that aberrant replication

fork remodeling is the cause of fork collapse and that DSBs occur as a result of MUS81-EME1 cleavage of these structures.

The underlying mechanisms of fork stability have mostly been derived from experiments using the fork stalling drug hydroxyurea (HU). HU inhibits global replication by depleting cellular nucleotide pools, which stalls the DNA polymerase without affecting the DNA template. The advantages of using HU over other DNA damaging agents are twofold. One, by inhibiting global replication HU potentially mimics the effects of all lesions that inhibit polymerases. And two, every 'lesion' can be effectively removed by simply removing the HU. These characteristics of HU provide the opportunity to examine general requirements for fork stability when forks are stalled over time. Importantly, forks that are stalled for 2 hours are processed differently than forks stalled for 24 hours (31). It is not clear mechanistically when or how cells transition from one fork restart pathway to the other.

Replication fork restart

Replication fork restart pathways are best characterized in bacterial replication systems. This is primarily due to the fact that in combination with the ease of genetics in this system, DNA replication in bacteria is simple. Duplication of the circular genome in bacteria initiates at one sequence-specific site. If either replication fork is blocked irreversibly, DNA replication fails. For this reason, bacteria must restart stalled or collapsed forks by any means, mutagenic or not. In *E.coli* the formation of Holliday junctions, or 'chicken feet', occurs at stalled

replication forks (Figure 1.5)(32). A Holliday junction is a mobile junction between four strands of DNA. This structure is also observed when pathways of replication fork stability are inhibited in eukaryotic cells, suggesting this may be a critical intermediate for many restart pathways. These structures are then processed by various mechanisms that reform the necessary DNA structures for replisome reloading and subsequent fork restart. This evidence from bacteria indicates that extensive remodeling of stalled forks is necessary to facilitate fork restart.

Fork remodeling

In general, fork remodeling is an ATP-dependent process that changes either the structure of template DNA or the newly synthesized (nascent) DNA, or both at stalled replication forks. Replication fork reversal is the general definition for the remodeling activities that generate chicken feet structures at stalled replication forks. Reversed forks were found (ex. Figure 1.6A) in yeast, frog, mouse, and human cells that were exposed to low doses of the topoisomerase I (TopI) inhibitor camptothecin (CPT) (33). CPT inhibition of Top1 induces torsional strain during DNA unwinding that slows replication fork progression. Inhibiting the fork reversal in these conditions results in fork collapse. Thus, active fork reversal may be a mechanism for stabilizing forks during times of DNA torsional stress.

Fork reversal requires the coordinated activities of DNA helicases and DNA translocases to unwind nascent and parental strands, re-anneal the nascent

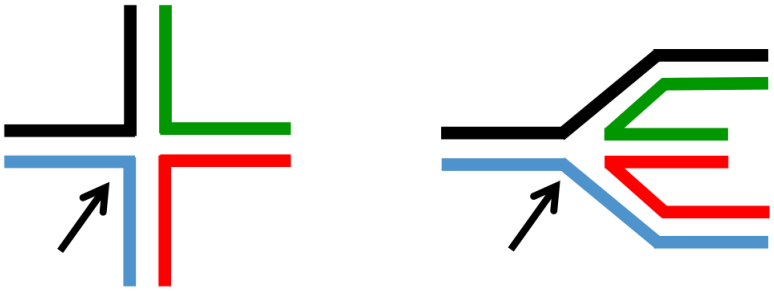

Figure 1.5. Holliday junctions form at stalled replication forks. Holliday junctions are DNA structures that can form between to DNA duplexes that are complementary. The result is a 4-way DNA junction. They are commonly depicted as cruciform-like structures (left). The same structure can form at stalled replication forks, referred to as a chicken foot (right). Arrow points to the junction that is common between both models.

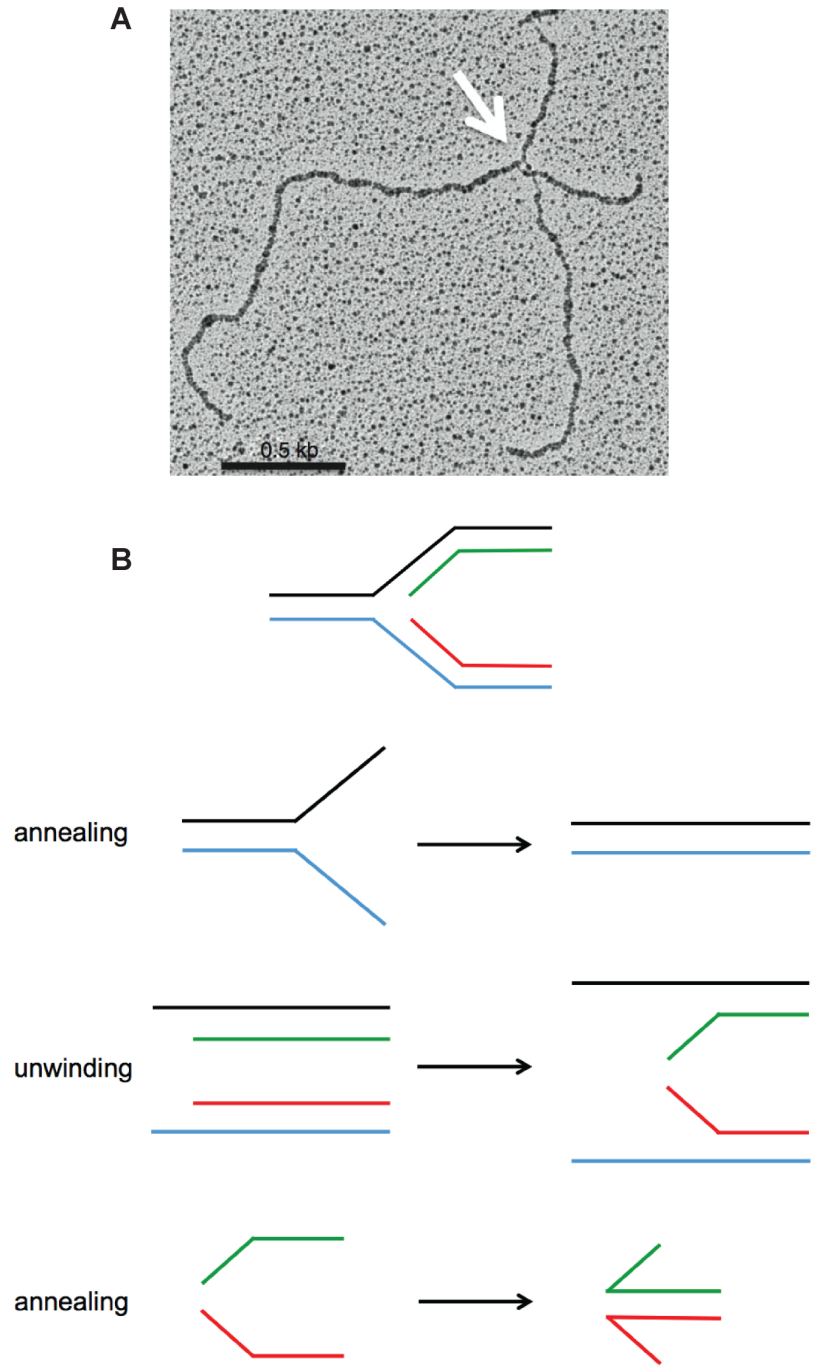

Figure 1.6. Remodeling activities required for fork reversal. (A) Image borrowed from (33) showing reversed fork by EM. (B) Fork remodeling by fork reversal is an ATP-dependent process requiring the activities of DNA translocases and DNA helicases. Reversal of the fork shown (top) requires coordinated steps of annealing and unwinding. Strand annealing is required to reanneal parental strands, unwinding of the nascent strands facilitates continued annealing of parental strands, and finally nascent strands are reannealed.

strands, and re-anneal parental strands (Figure 1.6B). Fork reversal into a Holliday junction occurs in two steps and there is the potential to restart the fork at either step (Figure 1.7). The first step in fork reversal involves remodeling theexcess parental ssDNA generated by uncoupling (Figure 1.7A). Many DNA binding proteins can support the annealing of two complementary ssDNA molecules *in vitro*. But at a replication fork, ssDNA is bound by RPA, which must be removed to allow base pairing. DNA translocases like RAD54 can remove RPA to promote loading of RAD51 (a necessary step in recombination pathways). However, displacing RPA must also be coupled to re-annealing of the parental strands.

Until recently, it was not clear whether an enzyme could catalyze this reaction at forks. Re-annealing of RPA-ssDNA is catalyzed by at least two closely related enzymes of the SNF2 family (discussed below), SMARCAL1 (SWI/SNF, matrix associated, actin-dependent regulator of chromatin, A-like 1) and AH2 (annealing helicase 2), which have been deemed ‘annealing helicases’. These proteins may function in cells to re-anneal parental strands and ‘undo’ the effects of helicase and polymerase uncoupling. Evidence provided in this thesis supports that the annealing helicase activities of SMARCAL1 are indeed required to limit ssDNA accumulation during replication stress, and deficiencies in annealing helicase activity cause fork collapse. Although SMARCAL1 is likely involved in later steps of reversal, it seems that its re-annealing activity is an essential step in promoting fork stability.

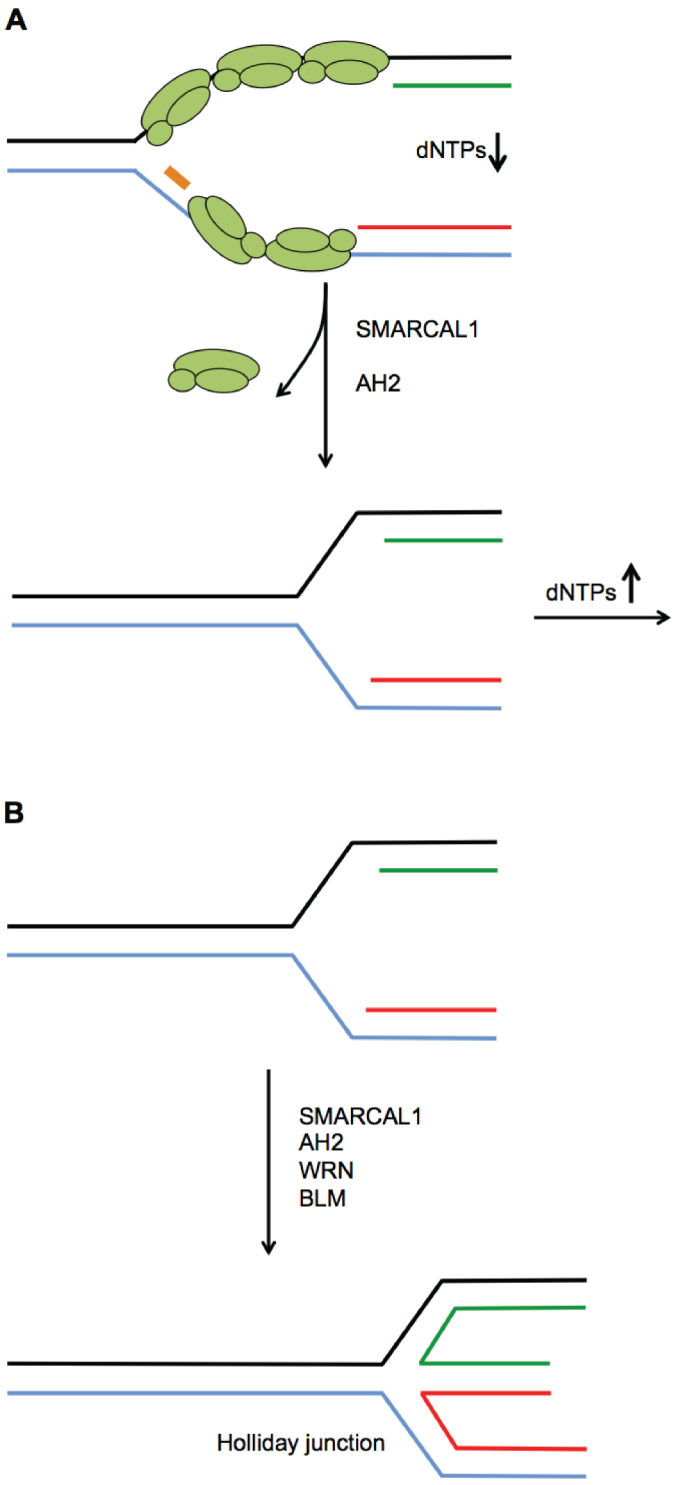

Figure 1.7. Fork reversal requires coordinated steps of re-annealing and unwinding. (A) The first step at a fork stalled by low dNTPs or HU is to remodel excess RPA-ssDNA generated by uncoupling. This requires strand annealing coupled to RPA removal. (B) The second step requires helicase activities to unwind nascent strands and this is coordinated with continued re-annealing of parental strands. The result is formation of a Holliday junction.

Assuming the DNA lesion or other replication inhibitor has been removed, fork re-annealing to the site of the block could result in recoupling of replisome activities and fork progression would resume. However, if the block to replication persists fork reversal likely continues. Step two (Figure 1.7B) requires the unwinding of nascent-parental duplexes. SMARCAL1 and AH2 may cooperate with DNA helicases at this step to promote reannealing of the nascent strands as they are unwound. The helicases involved at this step are likely the Werner syndrome helicase (WRN) and Bloom syndrome helicase (BLM). Both can promote fork reversal on model replication fork substrates in vitro (34,35). And, in cells WRN and BLM are required for fork progression and fork restart following replication stress (36,37).

The fork reversal presented here is specific to forks where the functional uncoupling of helicase and polymerase activity has generated long stretches of RPA-ssDNA. However, fork reversal is not thought to be limited to these structures and can likely initiate at step two in the absence of excessive ssDNA.

Restarting from reversed forks

How does fork reversal promote replication restart? In some instances, fork reversal facilitates lesion repair or bypass. For example, certain DNA lesions may only be recognized by DNA repair machinery in the dsDNA context. Fork reversal may be a mechanism to promote repair of DNA lesions (Figure 1.8A). Fork reversal may also promote bypass of the lesion. In this model, the nascent strand from the lesion-free strand is used as an alternate template for the

replication past the lesion (Figure 1.8B). A necessary step in both pathways is the resolution of remodeled substrates to restore an active replication fork. The DNA ends formed by nascent strand re-annealing may be prone to nuclease activities. Degradation of nascent strands could restore the fork structure (Figure 1.8C). Alternatively, the reversed fork could be 'reversed'. This would require branch migration of the Holliday junction (Figure 1.8D). Among others, SMARCAL1, WRN, and BLM promote branch migration on Holliday junctions in vitro. However, in this scenario their annealing and unwinding activities would necessarily be in the reverse direction. Recent in vitro evidence suggests that some helicases have a directional preference for translocation on reversed fork substrates (38). Fork reversal enzymes may exhibit similar directionality in cells allowing the opportunity to reverse fork remodeling.

Recombination-dependent restart

If fork reversal does not lead to lesion bypass, another mechanism for restart of stalled forks involves the components of homologous recombination and is referred to as recombination-dependent restart (RDR). The Holliday junction formed by fork reversal can be processed via two proposed mechanisms (Figure 1.9)(32,39,40). First, the double-stranded DNA end resulting from fork reversal can recombine with the homologous sequences in the parental strands to form a displacement loop (D-loop) and subsequently another Holliday junction (Figure 1.9, left). As long as the original block is removed, cleavage of both Holliday junctions by endonucleases would generate a new fork structure

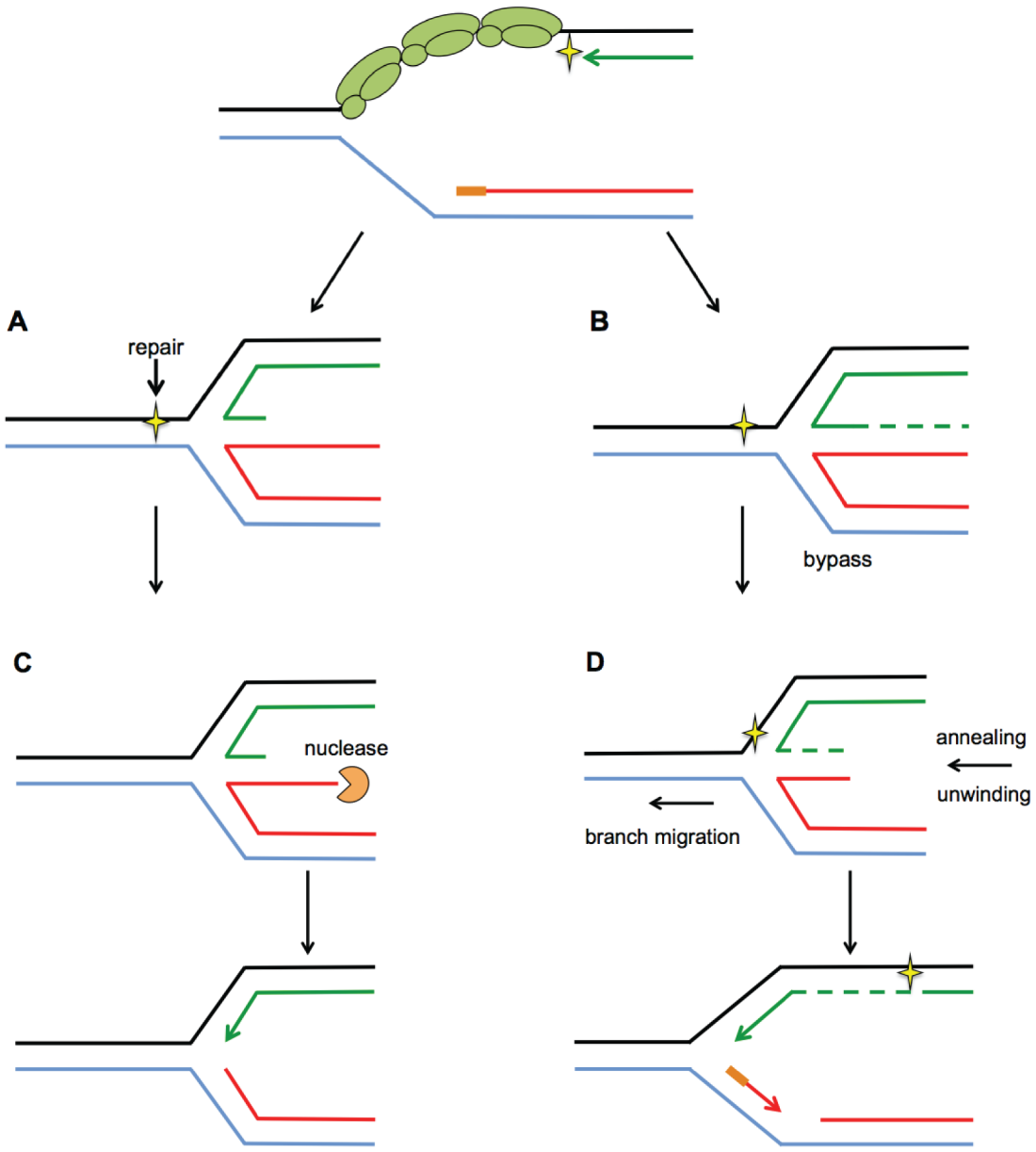

Figure 1.8. Fork reversal promotes repair or bypass of DNA lesions. (A) DNA repair proteins recognize specific lesions only in the context of dsDNA. Fork reversal returns the lesion to the parental duplex and may facilitate repair. (B) Another option following fork reversal is to use the lesion-free strand as a template for DNA synthesis. When the fork is restored as in (D), replication past the lesion can resume. (C,D) A necessary step in restart after fork reversal is resolution of the reversed fork. (C) Nascent-strand ends are subject to nuclease activity and may be degraded to restore the fork structure. (D) Branch migration of the Holliday junction can restore the fork without degrading DNA.

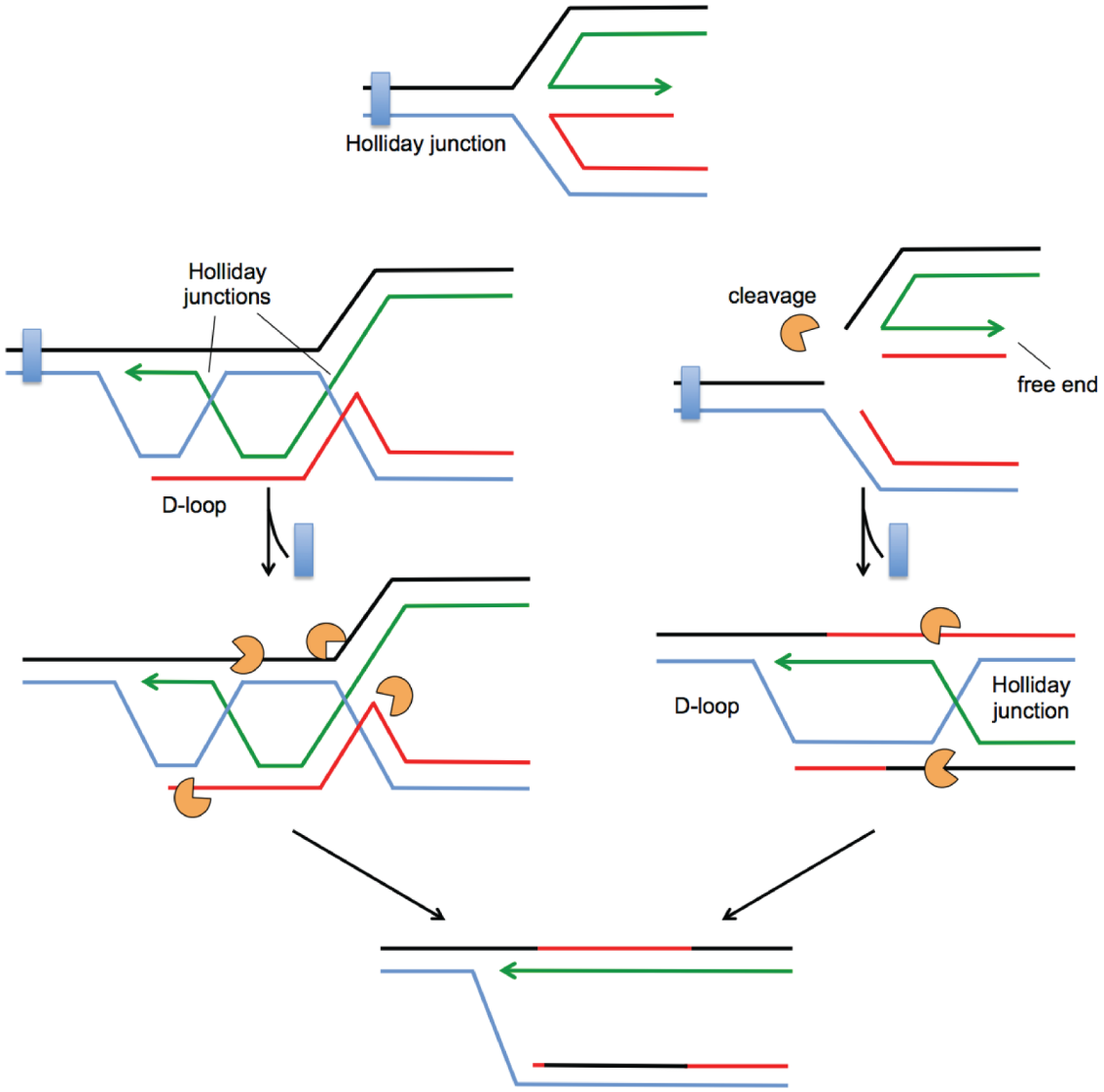

Figure 1.9. Recombination-mediated restart. The Holliday junction formed by fork reversal is a substrate for recombination enzymes. There are two options for processing the fork via RDR. (Left) The dsDNA end formed by nascent strand annealing can recombine with the parental duplex. The free 3' end of the leading strand invades the duplex and displaces the homologous parental strand, forming a D-loop. This forms a second Holliday junction. The 4-way junction is a substrate for endonucleases and cleavage occurs at both junctions. This removes the connection between the sister duplexes and results in crossover between nascent and parental DNA. (right) Alternatively a cleavage event happens at the initial Holliday junction. This generates a free arm in the top duplex that can invade the bottom duplex and form a D-loop and Holliday junction. This is cleaved and resolved as described above. Both result in a fork structure from which replication can resume. Figure adapted from (40).

on which replisome components could be reassembled and replication restarted. Second, endonucleases may cleave at the stalled fork prior to recombination from the dsDNA end. The outcome is the same, the dsDNA end resulting from cleavage forms a D-loop with the homologous duplex, and additional cleavage of the Holliday junction generates the fork structure (Figure 1.9, right).

Reassembling the replisome

Fork remodeling requires the disassembly of replisome components. So, to resume replication replisome components must be reassembled. PCNA can reload polymerase on 3' ends of DNA, which would be available after reversing fork remodeling. However, it is unclear how the MCM helicase would be recoupled. This remains a speculative step in the process. Determining whether or not MCM helicase activity can be recoupled to stalled forks would be a major discovery and clarify the overarching mystery in the field of whether fork reversal is an artifact or an active pathway of fork restart. Despite the uncertainties in mechanisms of fork reversal and restart, the number of enzymes required for replication fork stability in cells indicates some requirement for extensive fork processing.

Fork rescue

When considering replication fork restart pathways in eukaryotes, one needs to appreciate that eukaryotic DNA replication initiates from many origins. If a replication fork stalls, DNA replication can be rescued from the progression of a

nearby fork (Figure 1.10A). Also, only a subset of origins are fired each round of replication. After prolonged exposure to HU, most stalled forks collapse and replication reinitiates from new origin firing (31). This evidence implies that excess origins function as an alternative for restarting forks and completing replication (Figure 1.10B). This pathway may be preferred over restart pathways like RDR, which involves cleavage events that risk inappropriate recombination. However, in certain genomic locations origins are scarce and there may be no potential for fork rescue (41). Stalled forks under these conditions may require fork restart mechanisms like RDR. Indeed, studies of mammalian fork restart indicate that replication inhibitors activate recombination pathways (42,43).

Evidence demonstrating the genetic requirements of fork remodeling enzymes in replication fork restart pathways primarily comes from single molecule analysis of replication dynamics (44). In these studies, fork restart was analyzed in cells following exposure to HU for various amounts of time. Because every replication fork is stalled in the presence of HU, HU-stalled forks mimic stalled forks that cannot be rescued by origin firing or nearby forks. By measuring the resumption of DNA synthesis at individual forks after the HU is removed, we have made interpretations about the enzymatic activities required to maintain viable forks during prolonged fork blocks. Prolonged fork stalling is likely rare in cells, and most forks directly restart without significant fork remodeling. However, defects in replication fork remodeling enzymes are associated with human syndromes and cancer suggesting that prolonged fork stalling is a challenge to genome stability however infrequent it may be (45-47).

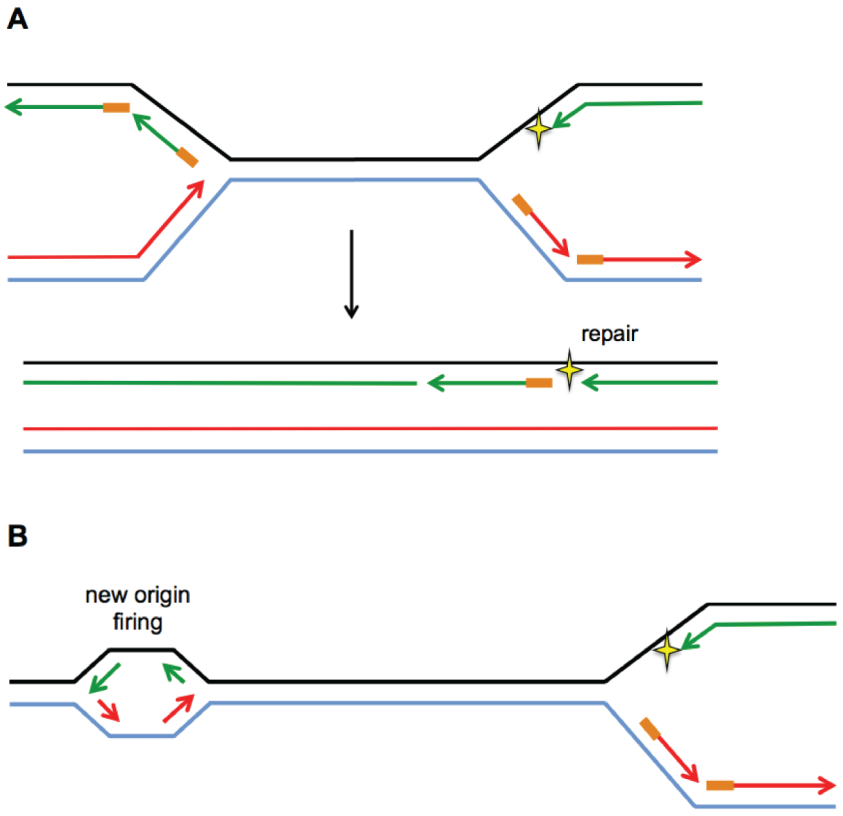

Figure 1.10. Replication fork rescue. (A) Stalled forks have the potential to be rescued by the progression of neighboring forks. This rescues the requirement for the stalled fork to restart and complete replication. The DNA lesion can be repaired post-replication. (B) Excess origins are licensed each round of replication. After prolonged stalling, new origins can fire new replication forks to rescue stalled forks as in (A).

Fork restart and repair pathways rely on the activities of DNA helicases, DNA translocases, recombinases, and nucleases. Further characterizing the function of these enzymes and how they cooperate at stalled forks will help elucidate the molecular mechanisms that promote fork restart and genome stability during DNA replication.

The SNF2 Family

Members of the SNF2 family of ATPases typically function within multi-protein complexes that remodel chromatin. However, SNF2 proteins have chromatin independent functions and are more generally appreciated as DNA translocases. Using the energy from ATP hydrolysis, they translocate along DNA rearranging nucleosomes or other DNA-interacting proteins to facilitate gene regulation, replication, recombination, and DNA repair.

ATPase motor

All SNF2 proteins contain seven motifs that are similar to those found in helicases. These motifs are divided between two domains, referred to as ATPase_N and ATPase_C, and must be in close proximity to each other to hydrolyze ATP (48,49). The motor domain enables SNF2 proteins to convert the chemical energy from ATP hydrolysis into mechanical force for translocating along DNA substrates.

Based on similarities in their helicase-like domains, SNF2 family members are grouped into 24 subfamilies (Figure 1.11A). The organization of their helicase-like domains likely confers the functional specificity of these enzymes on DNA substrates. SNF2 proteins contain additional sequences outside of this domain that further define the biological context in which the specific ATPase motor is needed.

DNA remodeling activities

The DNA remodeling activities of SNF2 proteins in the context of chromatin have been appreciated for decades. SNF2 proteins translocate DNA with enough force to disrupt DNA-protein interactions, like those between DNA and nucleosomes. By doing so, they remodel the accessibility of DNA to transcription factors and DNA repair factors. What has become obvious in recent years is that SNF2 proteins have very important chromatin-independent functions, especially at DNA replication forks. The yeast protein Rad5 can reverse model replication forks in vitro (50). In combination with its ubiquitin ligase activity, Rad5 promotes template switching to bypass lesions at stalled replication forks. Helicase-like Transcription Factor (HLTF) is one of two mammalian proteins closely related to yeast Rad5, and promotes similar activities (51). The annealing helicases SMARCAL1 and AH2 use their ATP-driven motor to rewind RPA-bound complementary ssDNA strands (52,53). This activity can promote fork reversal as well (54).

Defects in SNF2 proteins cause genome maintenance syndromes

Importantly, mutations in SNF2 genes cause human disease. ATR-X (a-thalassemia/mental redardation, X-linked) syndrome (MIM 301040) is caused by mutations in *ATRX*. Patients with ATR-X syndrome exhibit severe developmental delay, mental retardation, and abnormal hemoglobin. ATRX interacts with the histone variant macroH2A and may regulate its distribution in the genome. Deposition of macroH2a demarks transcriptionally silenced chromatin. Therefore, deregulated macroH2A distribution especially at specific loci may contribute to the phenotypes of ATRX syndrome (55).

Defects in *Excision repair cross-complementing rodent repair deficiency, complementation group 6* (*ERCC6*), gene function cause Cockayne syndrome B (MIM 133540). ERRC6 is essential for transcription-coupled excision repair, a process that removes bulky DNA lesions that stall transcription (56). Cockayne syndrome patients exhibit progressive multisystem degeneration, premature aging, and extreme UV sensitivity.

HLTF is frequently inactivated in human colon cancers (57). Evidence from *Hltf*-deficient mice suggests that loss of function promotes transformation by inducing genomic instability in the colon. HLTF can catalyze the removal of proteins from DNA and may do so at replication forks to maintain genome stability. The diseased states resulting from alterations in ATR-X, ERRC6, and HLTF functions emphasize the importance of SNF2 family activities in genome maintenance.

SMARCAL1

SMARCAL1 defines a distant subfamily of SNF2 proteins (Figure 1.11A). This subfamily includes two subtypes: SMARCAL1 and AH2. SMARCAL1 and AH2 have highly similar ATPase motors but are flanked by completely different accessory domains. The SMARCAL1 subtypes include at least one HARP (HepA-related protein) domain N-terminal to the ATPase domain. Human SMARCAL1 is 954 amino acids and its defined domains consist of two HARP domains, HARP1 and HARP2, the highly conserved ATPase domain, and an RPA binding domain, which we defined in Chapter IV (Figure 1.11B). Evidence presented in this thesis and elsewhere (58-61) identified the RPA binding domain as an essential domain for SMARCAL1 function in cells. The HARP domains are required for the annealing helicase activity of SMARCAL1 and chimeric proteins with HARP1/2 fused to other ATPase motors can support annealing helicase activity in vitro and in cells (62). Furthermore, HARP2 is as part of a minimal catalytic core that can accomplish all in vitro activities defined for full-length SMARCAL1 (54).

Like other SNF2 proteins, SMARCAL1 ATPase activity is stimulated by DNA binding. When SMARCAL1 was initially cloned and characterized (63), it was classified as a ssDNA stimulated ATPase. Further characterization of SMARCAL1 DNA binding has defined a much broader list of SMARCAL1 DNA substrates (54). SMARCAL1 binds a number of DNA substrates with a general requirement for ssDNA-dsDNA transitions (Figure 1.12). However, SMARCAL1 also binds complex substrates like 3-way and 4-way junctions that do not have

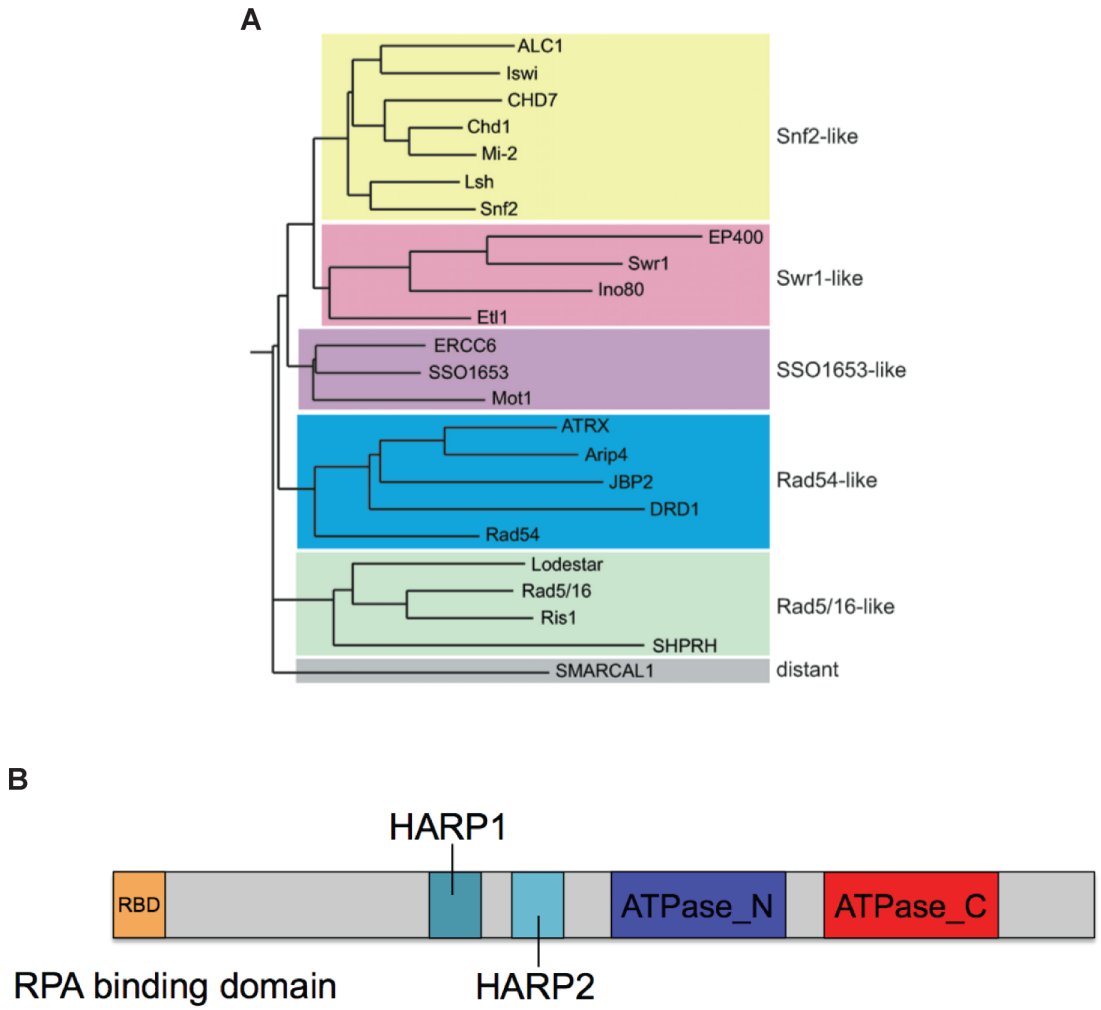

Figure 1.11. SMARCAL1 is defines a distant subfamily of SNF2 proteins. (A) Borrowed from (64)This tree is a representation of the subfamilies defined based on full-length alignments of the ATPase domains. SMARCAL1 is distantly related to other subfamilies in the SNF2 superfamily. SMARCAL1 family members have at least one HARP domain N-terminal to the ATPase domain. Depicted in (B) human SMARCAL1 has a defined RPA binding domain (RBD), two HARP domains, and a highly conserved ATPase domain, which consists of the ATPase_N and ATPase_C domains.

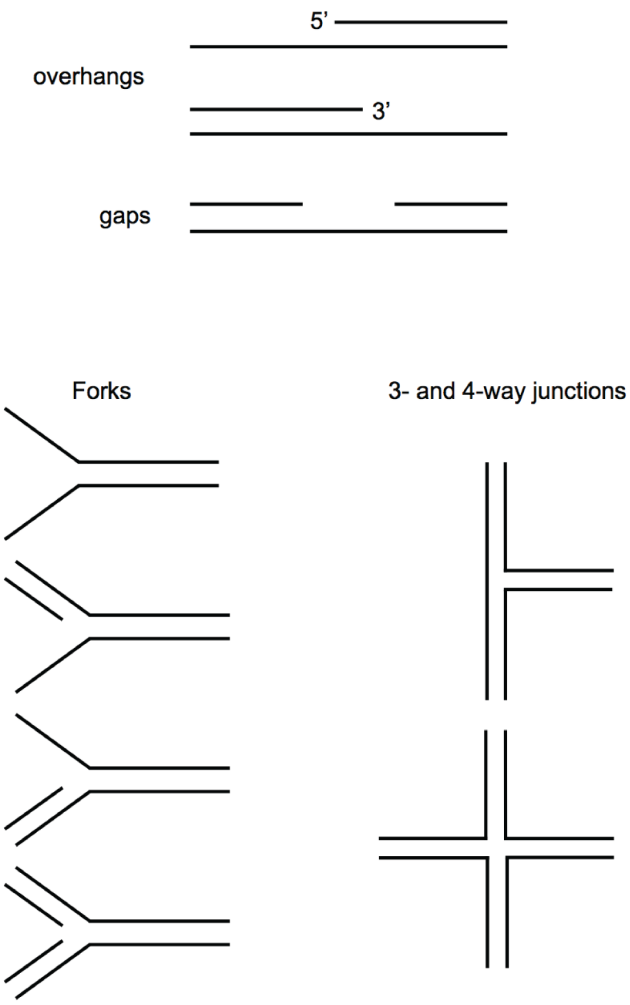

Figure 1.12. DNA substrates that stimulate SMARCAL1 ATPase activity.

any intended ssDNA regions. Importantly, SMARCAL1 also reverses model replication fork substrates, indicating that the enzymatic capabilities of SMARCAL1 reach farther than re-annealing RPA-ssDNA (54).

Loss of function mutations in *SMARCAL1* cause Schimke immuno-osseous dysplasia (SIOD) (65). SIOD clinical phenotypes include growth retardation, renal failure, recurrent infections, cerebral ischemia, and skin pigmentation beginning in early childhood. Disease severity varies with some patients experiencing *in utero* onset and death within 5 years, while others have a later onset of symptoms and live until their early twenties. We provide evidence that the phenotypes of SIOD are due in part to defects in SMARCAL1 annealing helicase activities at stalled replication forks. Thus, SIOD represents another genome maintenance syndrome.

The cellular processes that require annealing helicase activity were largely unknown until the identification of SMARCAL1 as a replication stress response protein. We identified SMARCAL1 in a functional genomic screen for genome maintenance genes (66). The details of this screen will be presented in Chapter III. After identifying SMARCAL1 as a genome maintenance protein, I further discovered that the annealing helicase activities of SMARCAL1 promote genome integrity specifically at DNA replication forks (Chapter IV) (58). SMARCAL1 is a substrate for the DDR kinases ATR, ATM, and DNA-PK. In Chapters V and VI, I present evidence demonstrating that DDR-dependent and independent phosphorylation regulates SMARCAL1 replication stress response activities.

CHAPTER II

MATERIALS AND METHODS

Antibodies

Antibodies were obtained as follows: KAP1 pS824 (Bethyl Laboratories, A300-767A) and SMARCAL1 (Bethyl Laboratories, A301-616A), γH2AX (Upstate Biotechnology, 05-636, clone JBW301), HA.11 (Covance, MMS-101P, clone 16B12), PCNA-PC10 (Santa Cruz Biotechnology, sc-56) and FLAG/OctA-Probe D-8 (Santa Cruz Biotechnology, sc-807), cyclin A Ab-7 (Thermo-Scientific, RB-1548-PO) and RPA/p34 Ab-1 (Thermo-Scientific, MS-691-PO, clone 9H8), BrdU (Becton Dickenson, 555627, clone 3D4), FLAG M2 (Sigma, F3165, clone M2). Additionally, custom SMARCAL1 antibodies targeting the N or C-terminus of SMARCAL1 and two different phosphopeptides were produced by Covance and Open Biosystems. See Table 2.1 for a comprehensive list of SMARCAL1 antibodies and the conditions for various applications.

ATPase assay

The ATPase assay was performed using purified protein from three different conditions: (1) Flag-SMARCAL1 and Flag-SMARCAL1-Δ32 were purified from baculovirus-infected insect cells essentially as previously described (52) with the following modifications. Cells were lysed in 20mM Tris pH 7.5, 150mM NaCl, 0.1mM EDTA, 1mM DTT 0.2mM PMSF, 1μg/ml leupeptin, 1μg/ml

Table 2.1. SMARCAL1 antibodies and conditions for various applications

Antibody	Antigen	IF	WB	IP	Notes
192 (Covance Custom)	purified fragment 1-101	nd	nd	nd	Serum from additional rabbit; not as specific as 193
193 (Covance Custom)	purified fragment 1-101	no	1:2000; 1hr	nd	note this is serum and not an affinity purified antibody; works well for WB detection; use this for ΔC applications
909 (Open Biosystems Custom)	KKRRFEFFDNWDSFTSPL	1:1000	1:5000; 1hr	2μL:25μL beads per 20mg lysate	IF is best for overexpressed protein; there is a lot of background in the cytoplasm which makes it difficult to detect endogenous SMARCAL1; use this antibody for WB and IP; 909 peptide does not elute SMARCAL1 from the 909-Ab efficiently
Bethyl (A301-616A)	904-954	1:50	1:1000; 1hr	nd	Used for detecting endogenous SMARCAL1 foci; requires triton extraction.
p-S652 (Bethyl custom)	KSDVLpSQLPAK	no	1:500; 1hr	nd	This antibody is not very specific and requires a lot of protein. It does detect DDRi sensitive phosphorylation at pS652. Background can be reduced by adding the NP-peptide to reduce background. Add peptide (5mg/mL) 1:5000 to Ab dilution and pre-bind at least 1hr. Peptide is stored with Ab aliquots in -80
p-S889-1 (Thermo custom_ animal PA4315)	KIYDLFQKpSFEKE	nd	1:500; 1hr	nd	This antibody is specific for pS889. Works well on FLAG-purified and eluted protein. Detection of endogenous is dirty because of background from IP. Elution would clean this up. Also pre-incubate the primary with NP-peptide (5mg/mL) before using; 1:5000. Peptide stored with antibody in -80.
p-S889-2 (Thermo custom_ animal PA4277)	KIYDLFQKpSFEKE	nd	1:500	nd	Has not been characterized extensively; may work just as well as pS889-1

aprotinin, and 0.1% triton. A final CM-sepharose step was added after elution of protein from the Flag-beads.

(2) HEK-293T cells were transfected at the 10cm dish scale with pLPCX-Flag-SMARCAL1 plasmids using Lipofectamine 2000 (Invitrogen). The day after transfection cells were passaged to one, 15cm dish. Seventy-two hours after transfection, the cells were lysed in NETN buffer (150mM NaCl, 1mM EDTA, 20mM Tris, and 0.5% Igepal) for 30 min on ice. After high-speed centrifugation, the cleared lysates were incubated with 30μL of EZview Red ANTI-Flag-M2 Affinity Gel beads (Sigma, F2426) for 3 h at 4°C. The beads were washed three times in NETN and twice in SMARCAL1 buffer (20 mM HEPES at pH 7.6, 20% glycerol, 0.1 M KCl, 1.5 mM MgCl2, 0.2 mM EDTA, 1 mM DTT, 0.2 mM PMSF, 0.01% IGEPAL CA-630). The bound proteins were eluted in 75μL of SMARCAL1 buffer containing 0.3 mg/ mL Flag peptide on ice for 1 hour. Eluted protein was aliquoted and stored at -80°C.

(3) Endogenous SMARCAL1 was purified from 18- and 9-15 cm dishes for 16hr HU and NT conditions, respectively. 80mg and 40mg of lysate was input into 909-IPs on protein A magnetic dynabeads to purify the HU and NT protein. Following the IP, beads were washed three times with NETN buffer and two times with SMARCAL1 buffer then eluted with 909 peptide 1:5 in 50μL of SMARCAL1 buffer for 1hr at RT. Following elution 1μL of purified protein was used in each reaction in triplicate.

The ATPase assay was performed as follows. The day before the ATPase reaction, chromatography buffer (1M Formic acid, 0.5M LiCl) was prepared and

added to the chromatography tank to pre-equilibrate. The buffer does not need to be prepared fresh each time. The fork DNA substrate (15μM final concentration) was prepared by mixing 3μL of top (100μM) and 3μL of bottom-fork (100μM) DNA primers to 12μL of H_2O and 2μL of NaCl (0.5M) for a final volume of 20μL. The primers were annealed using the RB anneal program. Conditions for each 10μL ATPase reaction were as follows: 0.24μL of ATP, [γ-^{32}P] (Perkin Elmer, BLU502Z250UC, 10mCi/mL), 1μL of cold ATP (1μM), 1μL of diluted fork DNA (50nM or 12.5nM for 5 and 1.25nM final concentration) or H_2O for 0nM DNA, 1μL of 50nM SMARCAL1, and 6.76μL of ATPase buffer (20mM HEPES pH 7.6, 0.1M KCl, 5mM $MgCl_2$, 3% glycerol, 0.25mg/mL BSA, 0.05mM EDTA, 0.5mM DTT, 0.01% Igepal). Master mixes for 0, 1.25, and 5nM DNA conditions were prepared for all samples with the exception of adding SMARCAL1. After adding hot ATP to the reaction mix, 9μL was aliquoted to individual sample tubes on ice. After adding 1μL of purified SMARCAL1 to all tubes, tubes were transferred from ice to the heat block set at 30°C for a reaction time of 30 min. To stop the reaction, samples were transferred back to ice. An aliquot (1μL) of each reaction was spotted on PEI cellulose plates (Selecto Scientific, 05-719-854) and allowed to dry for at least 5 minutes. The plates were subjected to thin-layer chromatography for 45 min. The plates were dried and quantitated using a phosphorimager. The results are presented as the ratio of ATP hydrolyzed to total ATP during the reaction. ATPase assays were performed a minimum of three times each, and graphs depict means and standard deviation error bars. Fork DNA sequence used:

Top-5'-CCAGTGAATTGTTGCTCGGTACCTGCTAAC-3'
Bot-fork-5'-GACATTTGATACCGAGCAACAATTCACTGG-3'

Cell culture and transfections

HEK293T and U2OS cells were cultured in DMEM + 7.5%FBS. SIOD patient fibroblasts (SD31) were cultured in DMEM + 15%FBS. Transfections were performed by Lipofectamine 2000 (Invitrogen), Fugene6, or FugeneHD (Roche). siRNA transfections were performed with HiPerfect (Qiagen) using siRNAs purchased from Qiagen or Dharmacon. Sequences for Qiagen siRNA used in this thesis were as follows:

SMARCAL1_1: CAGCTTTGACCTTCTTAGCAA
SMARCAL1_3: TTGAGTTATGAGTTAGGTCAA
SMARCAL1_5: TTGATTGGGTACAATGCGCAA

A Dharmacon ON-TARGET plus siRNA was used for silencing SMARCAL1 when not otherwise specified. The siRNA-resistant SMARCAL1 mutants used for complementation experiments contain 10 wobble base pair mutations in the S6-targeting sequence. Sequence of the S6 or SMARCAL1_6 siRNA: GCUUUGACCUUCUUAGCAA.

Standard siRNA transfections were performed either at a 10cm or 6-well scale in a reverse-transfection format. For transfections in 10cm dishes, 2.4μL of siRNA (20μM) were added to 960μL of Opti-Mem media and 40μL of HiPerfect reagent. This transfection mix was incubated for 10 min at room temperature and then added to $2.3x10^6$ U2OS cells in 8mLs of media. For the 6-well format, 1.2μL of siRNA (20μM) were added to 400μL of Opti-Mem media and 12μL of HiPerfect.

After 10 min, transfection mixes were added to $3x10^5$ U2OS cells in 2mL of media. Assays were typically performed 48-72 hours following transfection.

Chromatin fractionation

Chromatin fractionation was performed as described previously (67).

DDR activation assay

Cells were transfected with Fugene HD at 12-well scale. 0.5μL HD + 45μL Opti-Mem + 250ng of pLEGFP-SMARCAL1 DNA were incubated for 15 minutes at RT and then resuspended in in 1mL of media containing 150,000 cells. 24 hours later cells were seeded into 96 well CellCarrier plates (Perkin Elmer, 6005550) that are compatible with Opera imaging. At least 2 samples in the group were counted to estimate cell number and plate 8000 cells per well. 48hrs post tranfection cells were washed once with 100μL of PBS and fixed for 10 min with 100μL of 3%-paraformaldehyde and 2% sucrose solution. All washes were 100μL of PBS dispensed by a 96 well plate dispenser and discarded by dumping the plate and tapping the excess on a blue pad. The IF protocol consisted of the following: after fixation, 3 PBS washes; add 100μL of a triton-X-100 solution (20mM HEPES, 50mM NaCl, 3mM $MgCl_2$, 10.269g of sucrose, 0.5% Triton X-100) for 10min at 4°C; 4 PBS washes; block with 5% BSA-PBS in 100μL; 1 PBS wash; aspirate PBS with multi-channel aspirator and add primary antibody (mouse anti-γH2AX, 1:10,000 diluted in 1%BSA-PBS) 50μL per well with a multi-channel pipette; after a 30 min incubation at 30°C + humidity, 3 PBS washes;

aspirate PBS and add secondary (Cy5 goat anti mouse; Invitrogen A10524; diluted 1:500 in 1% BSA-PBS) for 20 min at RT in the dark; 3 PBS washes; add 100μL of DAPI diluted in PBS for 3 min in dark; 3 PBS washes. Leave cells in 100μL PBS for imaging. Images were collected by the Opera on the same day as staining.

Opera acquisition was set up as follows using the Opera Filter Selection Tool: Camera 1: 540/75; Camera 2: N/A; Camera 3: 690/50; Camera 4: select all; Detection Dichro: empty; Primary Dichro: 405/488/640. Select Camera 3 and 4 for exposure 1 (Cy5 and DAPI) and select Camera 1 for exposure 2 (GFP). Samples were visualized with the 20X water objective and approximately 12 subfields were imaged in each well. Samples were seeded into 3 wells each time for experimental replicates. Analysis: Columbus analysis software was used to define nuclei using detection method B, and then calculate mean intensity per nucleus for GFP and H2AX. See Chapter V text for a more detailed description of further data analysis.

Immunofluorescence

With the exception of the quantitative DDR activation assay, immunofluorescent images were obtained with a Zeiss Axioplan microscope equipped with a Zeiss camera. In Chapter IV, the percent of cells staining with both cyclin A and γH2AX predicted by chance was calculated as follows: (% of cyclin A^+ & $\gamma H2AX^+$ cells) ÷ (% cyclin A^+ cells × % $\gamma H2AX^+$ cells). A staining

protocol similar to that described for the DDR activation assay was used for all immunofluorescence on coverslips.

Lysis conditions

NETN buffer: 150mM NaCl, 1mM EDTA, 20mM Tris, and 0.5% Igepal. This buffer supplemented with 1mM PMSF, 5μg/mL aprotinin, 5μg/mL leupeptin, 1mM NaF, 1mM Na_2VO_4, 1mM DTT, and 50mM B-glycerophosphate was used for all methods with a cell lysis step. For HEK293T cells, cells were collected by scraping and washed twice with PBS. In many cases, cell pellets were not lysed immediately but stored after collection at -80°C. Cells were lysed for 30 minutes on ice.

Mass spectrometry

Endogenous SMARCAL1 was purified from HEK293T following 16hr HU, separated by SDS-PAGE and coomassie blue stained. The SMARCAL1 band was excised from the gel and mailed to Dr. Jun Qin at Baylor College of Medicine for analysis. Samples were run on two machines, LTQ and Orbi. (Dr. Qin emailed a powerpoint file with spectra on October 2, 2009)

Phosphopeptide analysis

For endogenous maps, HEK293T cells were plated in 10cm dishes the day before labeling. 1.5 x 10^7 cells were plated in each of 2-10cm dishes for untreated; 2.5 x 10^7 cells were plated in each of 2-10cm dishes for HU samples.

2mM HU was added for 16hrs before cells were washed twice with equilibrated phosphate-free DMEM (Cellgro, DMEM, 1X with 4.5g/L glucose and sodium pyruvate without L-glutamine and phosphate, 17-206-Cl, 6x100mL). Cells were incubated in 5 mLs of equilibrated phosphate-free DMEM containing 10% dialyzed FBS (Gibco, 26400-036) with or without HU for 30 minutes. During this time, a betashield box was also placed in the incubator to equilibrate. To label cells, ^{32}P-orthophosphoric acid (Perkin Elmer, Phosphorous-32 radionuclide, 25mCi or 10mCi, NEX053025MC or NEX053010MC) was added to the media for a final concentration of 0.5mCi/mL for 30 minutes. Cells were washed once with PBS and 700µL of trypsin added to each plate for 5 minutes. 700µL of media was added to trypsinized cells, cells were transferred to a 1.5mL screw cap microcentrifuge tube and pelleted at 1000 rpm for 5 minutes. Pellets were washed once with PBS and combined into a single tube and pelleted. Cells were lysed in 1mL of NETN lysis buffer containing 1mM PMSF, 5µg/mL aprotinin, 5µg/mL leupeptin, 1mM NaF, 1mM Na_2VO_4, 1mM DTT, and 50mM B-glycerophosphate for 30 minutes on ice and then cleared at 4^{o}C for 20 minutes at 16.1xg.

During the clearing step, Protein A dynabeads were prepared for immunoprecipitation. Per sample, 50µL of the resuspended bead mixture was washed one time with lysis buffer and then incubated with 2µL of the 909 antibody in 500µL of lysis buffer for 10 minutes, rotating at room temperature. Beads were always prepared as a mix for all samples and divided equally into tubes following antibody binding. After the pre-binding, beads were washed once

in lysis buffer then resuspended in buffer again to divide the beads into two tubes. 1mL of lysate was added to the 909-beads and rotated at room temperature for 30 minutes behind the shield. After the IP, beads were washed 3 times with NETN buffer and once with NETN containing 500mM LiCl. 30µL of 2X sample buffer was added to beads and boiled for 10 minutes. 909-IPs were separated by SDS-PAGE on an 8% acrylamide gel for 1 hour at 120V and transferred to methanol-activated PVDF at 0.2 Amps for 6 hours.

Usually the following day, PVDF membranes were exposed to film for 1 hour to visualize proteins with ^{32}P incorporated. Film was aligned to the membrane, and a syringe was used to perforate the membrane around radiolabeled SMARCAL1. Membrane was cut and transferred to a 1.5mL centrifuge tube. 1mL of 100% methanol was added to the membrane for 1 minute and then removed. Membranes were washed with 1mL of 0.05M of ammonium bicarbonate containing 0.1% Tween-20 once and then incubated in 1mL of the solution for 30 minutes at 37^{o}C. Membranes were washed onced with 1mL of 0.05M ammonium bicarbonate without Tween-20. 30µL of 0.05M ammonium bicarbonate was added to cover the membrane, and 10µL of sequencing grade trypsin (resuspended at 1mg/mL in trypsin buffer, Promega, V5111, 100µg) was added. Proteins were digested at 37^{o}C for 3 hours, and another 10µL of trypsin was added for an additional 3 hours. 400µL of milliQ water was added to digested proteins and dried in a speed vacuum over night (no heat).

The next day, 400µL of milliQ water was added to dried samples, vortexed well, and then dried in a speed vacuum for 3 hours or until completely dry. Digested proteins were then resuspended in 400µL of pH 1.9 electrophoresis buffer (2.5% formic acid, 7.8% glacial acetic acid), vortexed well, transferred to a new tube, and lyophilized again. After lyophilizing, centrifuge tubes were placed with caps open in a dry scintillation vial and ^{32}P was counted in a scintillation counter. Peptides were resuspended in pH 1.9 buffer to equalize counts/volume. Usually 500 counts were then spotted onto cellulose coated glass thin layer chromatography plates according the standard protocol (68). Plates were electrophoresed at 1000V for 30 minutes on a Hunter Box in pH 1.9 buffer and then completely dried in the fume hood for several hours. Plates were then transferred to a chromatography tank containing phosphochromatography buffer (7.5% glacial acetic acid, 25% pyridine, 37.5% *n*-Butanol) for 16hrs. Plates were dried and then exposed to film for 4 days (42). This time can be adjusted if fewer counts are spotted.

The only difference for GFP-SMARCAL1 maps is that cells are transfected two days before labeling. The transfection protocol was as follows: 60µL of LF2000+1.5mL of Opti-mem was incubated for 5 min at RT then mixed with 900ng of pLL5.0-GFP-GW-SMARCAL1 DNA +23µg of salmon sperm DNA in 1.5mL of Opti-mem. 20 min later LF2000-DNA complexes were resuspended 8 mL of DMEM containing $2x10^7$ cells and plated in 10cm dishes. 24 hours later cells were passaged to 2 10-cm dishes and in HU conditions, HU was added for 16hrs before labeling, approximately 48hrs after transfection.

Phospho-shift assay

HEK293T cells were transfected with Lipofectamine 2000 at a 6-well scale. LF2000 complexes were prepared as follows: 5μL of LF2000 was added to 125μL Opti-Mem and incubated for 5 minutes before adding it to 125μL of Opti-Mem containing 100ng of SMARCAL1 in the pLL5.0-GFP-GW vector and 2μg of sheered salmon sperm DNA. Complexes were incubated for 20 minutes. During this time cells were prepared at a density of $9x10^5$ per mL. LF2000-DNA complexes were resuspended in 2mL of cells and plated in 6-well dishes. The day after transfection cells were split from 1:2 into 2 wells. 2mM HU was added to one of the wells for 16-18hrs. Cells were harvested and lysed in 60μL of NETN buffer plus additives for 30 minutes. After clearing protein concentrations were measured and 2X sample buffer added. Samples were boiled for 10 minutes and then separate on 6% acrylamide SDS-PAGE for 2 hours at 100V. Gels were transferred onto nitrocellulose for 6hrs at 0.2A. Standard 909-immunoblotting conditions were used with the exception of supplementing the milk-TBST with NaF (1:1000).

Protein interactions

The two-hybrid screen was performed by Hybrigenics using their ULTImate Y2H methodology. Immunopurifications of endogenous SMARCAL1 complexes from nuclear extracts was performed with the 909 antibody. Associated proteins were identified by tandem-mass spectrometry by the Vanderbilt Mass Spectrometry Center. Flag-HA dual epitope purifications were

performed according to published protocols (69) using 293T cells stably expressing Flag-HA-SMARCAL1.

Replication stress agent sensitivity assays

Cells transfected with siRNAs using HiPerfect (Qiagen) were treated with drug (HU, 3mM; APH,15μM; CPT, 150nM) for 24 hours, washed, and then allowed to recover for 24 hours. Cell viability was measured using the WST-1 reagent (Roche). All statistical tests were performed on the log of the ratio of treated/untreated measurements using a two-tailed unpaired t-test.

Vectors

The details of all vector constructs are available upon request. The pLEGFP-C1 vector was used to drive levels of GFP-SMARCAL1 expression sufficient to cause DNA damage. pLL5.0-GFP-GW was used for attenuated expression levels that did not induce DNA damage. siRNA-resistant mutants and phospho-mutants were generated using site-directed mutagenesis and primers and diagnosis are listed in Table 2.2.

Table 2.2. Primers for SMARCAL1 mutagenisis and sequencing.

Mutation	Primer sequence	Primer No.	diagnosis
11wobble (for S6 siRNA resistance)	ctggcctgatcaacattgtctcgttcgatttattgtcgaagttggaaaaacagc	P136/P137	none
S173A	ctggccaaaccaaagagtgcacaagagacaccagctc	P41/P42	adds ApaLI
S173D	cctctggccaaaccaaagagtgatcaagagacaccagctca	P43/P44	adds BclI
S652A	caagtccgacgtccttgctcagctgcctgccaagc	P61/P62	adds BclI
S652D	gcctcaagtccgacgttctagatcagctgcctgccaagc	P45/P46	lose AatII
S889A	ctacgacctattccagaaggcctttgagaaagaaggaag	P102/P103	adds StuI
S889D	ctacgacctattccagaaggactttgagaaagaaggaag	P114/P115	none
S889E	ctacgacctattccagaaggaatttgagaaagaaggaag	P100/P101	none
S919A	gaagtgcttcaggaacatctggatccagtgcccagaacatgggagaca	P21/P22	adds BamHI
S919D	tcaggaacatctggaagtagtgatcagaacatgggagacaccctg	P118/P119	none
S55A	ttccaggccaagcaagggccggcccaaaatttcccaagg	P143/P144	adds NaeI
S115A	caggccacagtccacgtgcacaaatggctctcactgg	P37/P38	adds ApaLI
S236A	gggtcctcagtccaaaaaggagtgaatgcgcagaagggaaagtg	P92/P93	adds FspI
S350A	gcctacttcgaggcagacatctcatatgcacaggaccttattgcgcttt	P82/P83	adds NdeI
S407A	tcaccctggcgtttgctgcgcagctcaagaagacatc	P27/P28	adds FspI
T592A	cccgcagagctctacgcccagatcatcgcagtc	P141/P142	none
T742A	ccataaggtggtcctggatgcaattgcgcaagagcttgag	P39/P40	lose HgaI
ΔC (delete 861-954)	cacctgggagcaggccttccttcgg	P106/P107	size
Seq. primer 2 (300-)	gccagaagaaatgcccacagcc	P73	
Seq. primer 3 (750-)	gttgattgggtacaatgcgg	P74	
Seq. primer 4 (1311-)	gccctttcagagagctgg	P75	
Seq. primer 5 (1803-)	cttcccccagtttcatgcc	P76	
Seq. primer 6 (2310-)	gttccaactgtcggagaggc	P77	

CHAPTER III

A cDNA OVEREXPRESSION SCREEN FOR IDENTIFYING DEREGULATED GENOME MAINTENANCE ACTIVITIES[1]

Introduction

ATR and other DDR pathways are not only activated by endogenous and exogenous sources of DNA damage (such as radiation and reactive oxygen species), but also respond to tumorigenic gene function defects that promote genome instability (70). The overexpression or activation of oncogenes and inactivation of some tumor suppressors induce DDR activation in pre-malignant lesions, presumably due to aberrant cell division cycles and replication stress (71-75). These findings suggest a model whereby the ATM/ATR-mediated DDR serves as a barrier to constrain tumor development by limiting cellular proliferation, inducing apoptosis, and promoting DNA repair (70).

Several oncogenes and tumor suppressors that induce DDR activation when deregulated have been identified; however, the heterogeneity of cancer suggests there are many others that challenge genome integrity and promote tumorigenesis. To better understand the variety of biological processes required for genome maintenance and determine the extent to which deregulation of these

[1] This Chapter includes excerpts from reference 66. Lovejoy, C.A., Xu, X., Bansbach, C.E., Glick, G.G., Zhao, R., Ye, F., Sirbu, B.M., Titus, L.C., Shyr, Y. and Cortez, D. (2009) Functional genomic screens identify CINP as a genome maintenance protein. *Proc Natl Acad Sci USA*, **106**, 19304-19309.

pathways could also activate the DDR, Courtney Lovejoy and I completed a functional genomic screen for genes that when silenced or overexpressed cause activation of the DDR. Using the screen methodology developed by Courtney, I identified genes that activate the DDR when overexpressed.

Results

A three step methodology was used to identify genes that when overexpressed cause activation of the DDR in the absence of any added genotoxic agents (Figure 3.1A). First, pools of three cDNAs were co-transfected into HeLa cells with a GFP expression vector to identify successfully transfected cells. A total of 5,796 cDNAs expressed from the pCMV-SPORT6 vector were analyzed. Two days after transfection, fixed cells were monitored for DDR activation resulting from gene overexpression by immunofluorescence staining using a phospho-peptide-specific antibody to the ATM substrate KAP1 (KRAB domain associated protein 1) (76). Secondly, the vectors from cDNA pools that caused KAP1 phosphorylation were deconvoluted and the KAP1 phosphorylation assay was repeated. Finally, individually positive cDNA vectors were transfected into a distinct cell type (U2OS) and DDR activation was monitored by immunofluorescence analysis of γH2AX.

A total of 97 genes were identified that caused DDR activation when overexpressed in both HeLa and U2OS cells (Figure 3.1B and Appendix A). Gene products with biological functions linked to gene expression, cell cycle

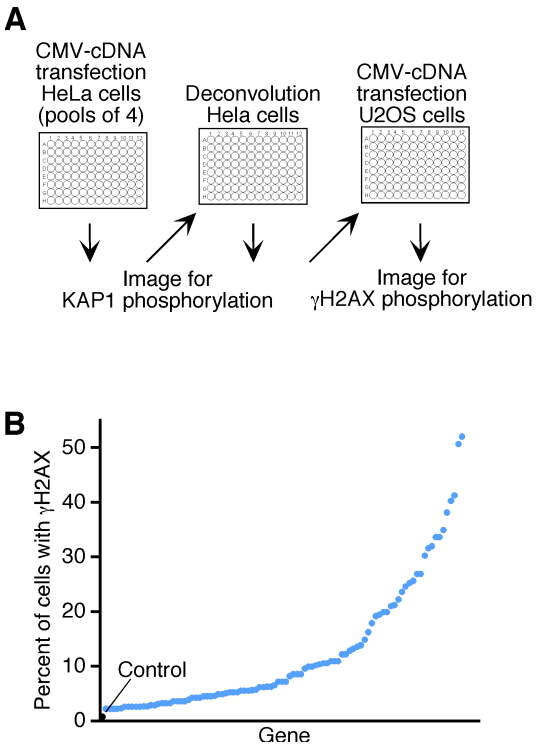

Figure 3.1. A cDNA overexpression screen identifies genes that cause DDR activation. (A) Schematic of the cDNA overexpression screen. (B) The average of the γH2AX count from each of the 97 positive cDNAs is graphed and compared to the control empty vector.

regulation, nucleic acid metabolism, and cancer were strongly over-represented when compared to the biological functions present within the cDNA screening library (Figure 3.2). Some of these genes have been linked to cancer (Table 3.1), and represent candidates for the oncogenes that promote replication stress and DDR activation in pre-malignant lesions.

Among the genes linked to cancer are several Ets family transcription factors that act as oncogenes in Ewing sarcoma, prostate cancer and myeloid leukemia (77). We confirmed that their overexpression activates the DDR even in untransformed epithelial cells (Figure 3.3)

In some cases both overexpression and silencing of a gene caused DDR activation. For example, we identified *SMARCAL1* and members of the *H2AFY* histone family in both the siRNA and cDNA overexpression screens. We also found that *PLK1* overexpression activated the DDR, while silencing *PLK1* was previously reported to cause DNA damage (78,79).

To further understand the functional relationships between the genes identified in the cDNA overexpression screen, we performed an extensive bioinformatics analysis using published literature and functional annotation programs. This analysis placed many of the genes into four major functional groups: the ATM/ATR related DNA damage response, mitosis, chromatin regulation, and RNA metabolism (Figure 3.4). In order to visualize the combined efforts of our genome maintenance screens, both RNAi and cDNA overexpression genes are depicted within the four groups.

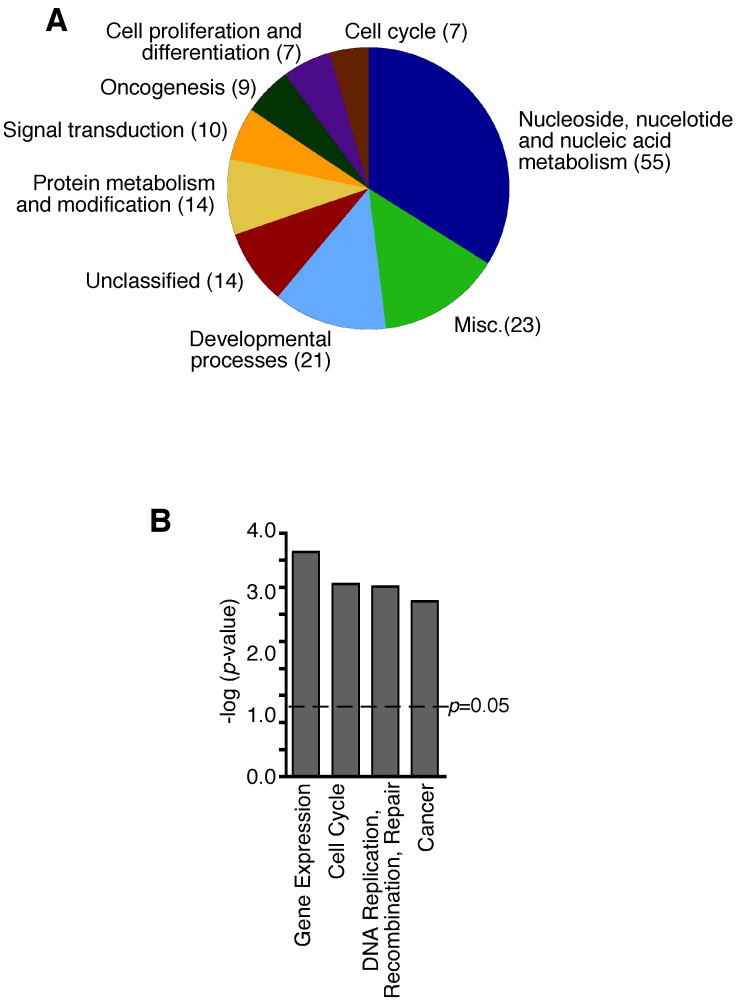

Figure 3.2. Biological classifications of genes reproducibly activating the DDR after overexpression. Classifications were assigned using (A) PANTHER or (B) Ingenuity pathways.

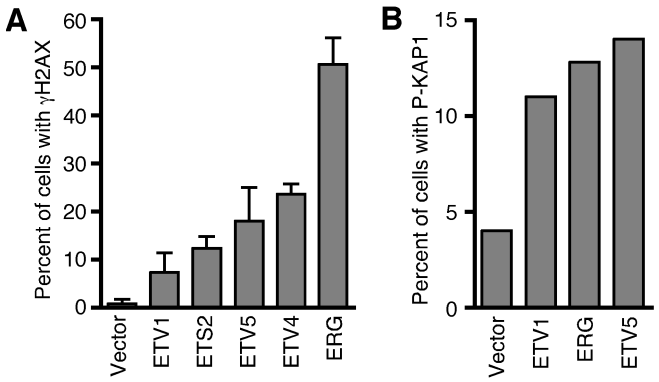

Figure 3.3. Activation of the DDR by overexpression of *ets* family. DDR activation by overexpression of ets family members was monitored in U2OS cells (A) or hTERT-immortilized retinal epithelial cells (B). Error bars are standard deviation (n=3).

In some cases the functional relationships can be understood mechanistically based on previously published research. For example, *DDB1* was identified as a genome maintenance gene in the RNAi screen and *CDT1* was identified in the overexpression screen (Figure 3.4C). DDB1 regulates replication origin firing by promoting the ubiquitin-dependent degradation of CDT1. Previous publications demonstrated that an accumulation of CDT1 in DDB1-silenced cells causes re-replication, and consequently, DDR activation (80,81). Another functional relationship that is apparent in the connectivity diagrams involves three mitotic protein kinases: CHK1, PLK1, and WEE1 (Figure 3.4B). Aberrant mitotic regulation caused by either too much PLK1 activity or too little CHK1 or WEE1 activity may cause premature entry into mitosis and an accumulation of DNA damage (82).

Discussion

As expected, many of the genes identified are suspected or known oncogenes (Table 3.1). The cDNA overexpression screen found several oncogenes that induce genetic instability when overexpressed, including *PLK1* which causes chromosomal instability due to its function in centrosome and mitotic control (83). Another interesting example is the *DEK* oncogene. DEK regulates chromatin and DNA topology, and was originally identified in a translocation in acute mylogenous leukemia (84). *DEK* overexpression also suppresses the phenotypic defects of an ataxia-telangiectasia cell line defective

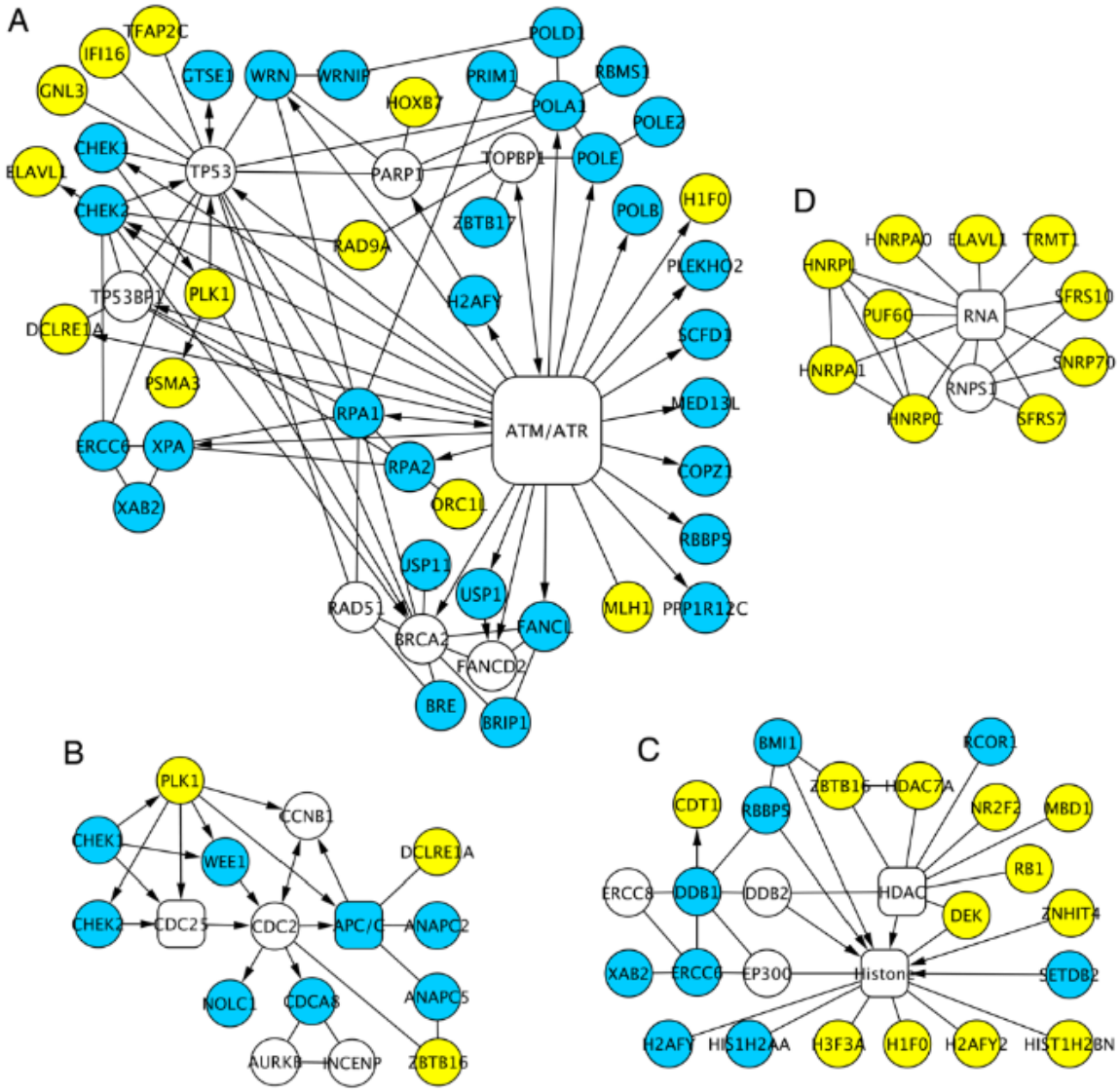

Figure 3.4. Network modeling of gene products identified in the RNAi and cDNA genome maintenance screens. In each network, the genome maintenance genes identified in the RNAi screen are colored in blue and genes identified in the cDNA overexpression screen are colored in yellow. A solid line between two gene products indicates a protein-protein interaction; an arrow indicates that one gene product acts on the gene product to which the arrow is drawn (for example, a kinase-substrate relationship). Cellular activities containing more than one gene product, such as the anaphase promoting complex (APC/C), are diagramed as rounded rectangles. Uncolored circles were not identified in the screens but were added to the network to establish connections. Transcriptional relationships are not diagrammed. The predominant pathways identified include (a) DNA damage response (b) mitotic control (c) chromatin regulation, and (d) RNA metabolism.

Table 3.1. Genes identified in the cDNA overexpression screen that are published oncogenes.

Entrez ID	Gene Name	Gene Symbol	chromosome location	Cancer Types	Example references
7913	DEK oncogene (DNA binding)	DEK	6p22.3	leukemia	1549122
4286	microphthalmia-associated transcription factor	MITF	3p14.2-p14.1	melanoma	16001072
2150	coagulation factor II (thrombin) receptor-like 1	F2RL1	5q13	pancreas, colon, breast	15280447
7022	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)	TFAP2C	20q13.2	breast	784604
7704	zinc finger and BTB domain containing 16	ZBTB16	11q23.1	leukemia	8384553, 8387545
3217	homeobox B7	HOXB7	17q21-q22	melanoma, ovarian	17308091
2115	ets variant gene 1	ETV1	7p21.3	multiple types	18563191
3280	hairy and enhancer of split 1, (Drosophila)	HES1	3q28-q29	ovarian, breast	18719287
5347	polo-like kinase 1 (Drosophila)	PLK1	16p12.1	multiple types	16557283
2114	v-ets erythroblastosis virus E26 oncogene homolog 2 (avian)	ETS2	21q22.3; 21q22.2	multiple types	18563191
2119	ets variant 5 (ets-related molecule)	ETV5	3q28	multiple types	18563191
2118	ets variant gene 4 (E1A enhancer binding protein, E1AF)	ETV4	17q21	multiple types	18563191
1994	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 (Hu antigen R)	ELAVL1	19p13.2	multiple types	17132932
81620	chromatin licensing and DNA replication factor 1	CDT1	16q24.3	lung	18508524
54828	breast carcinoma amplified sequence 3	BCAS3	17q23	breast	16617102
2078	v-ets erythroblastosis virus E26 oncogene homolog (avian)	ERG	21q22.3	multiple types	18563191

in ATM activity (85). This suppression is cell-line specific and unique to a relatively mild ATM mutation (deletion of amino acids 2427 and 2428). Thus, *DEK* overexpression may cause chromatin changes that increase the activity of the mutant ATM protein, leading to partial suppression of the DNA damage-sensitivity phenotype.

Several Ets family transcription factors also caused DDR activation when overexpressed. Since these proteins are frequently overexpressed and/or amplified in human tumors (77), they may contribute to tumorigenesis by promoting genome instability (70). This is also true for other known or putative oncogenes identified in the overexpression screen. In some cases, gene overexpression may directly perturb genome maintenance activities such as DNA repair. For example, HOXB7 is linked to DNA repair through interactions with PARP and DNA-PK proteins (86) and its overexpression is linked to melanoma. Overexpression of the DNA repair proteins DCLRE1A and MLH1 may cause DNA damage via dominant negative effects since these proteins function as part of protein complexes.

Other major categories of genes identified in the screen include regulators of mitosis (PLK1) and several chromatin regulatory proteins (DEK, H3F3A, H1F0, H2AFY2). In some cases the effect on genome maintenance may be indirect through transcriptional changes. In other cases, the chromatin regulation may directly affect DNA repair. Chromatin regulation and DNA repair were also prominent categories of genes found in an *S. cerevisiae* screen that used increased spontaneous Rad52 foci as an assay (87).

We also identified several proteins that are phosphorylated by ATM or ATR in response to DNA damage but with unknown functions including PLEKHO2, SCFD1, MED13L, COPZ1, RBBP5, and PPP1R12C (88). Our data confirm their placement in a DDR pathway, and further analyses will be required to understand their specific genome maintenance activity.

Finally, for some genes identified it is unclear why deregulation would lead to DDR activation. For example, a large number of RNA binding proteins not previously linked to DNA metabolism was found (Figure 3.10D). In some of these cases, the effect may be indirect or could be because of induction of apoptosis. We excluded obvious apoptotic cells in our analysis. We did identify *CIDEC*, *CASP10,* and *CRADD* that participate in apoptotic pathways. Thus, in some cases the γH2AX staining could be a result of the initiation of early stages of an apoptotic program.

In conclusion, our RNAi and cDNA overexpression screening methodologies have proven successful in identifying new genome maintenance activities. In Appendix B, I provide an extensive review of other functional genomic approaches to identify and characterize genome maintenance pathways.

CHAPTER IV

SMARCAL1 FUNCTIONS AT STALLED REPLICATION FORKS TO MAINTAIN GENOME INTEGRITY[2]

Introduction

Genome maintenance requires the coordinated activities of cell cycle, DNA repair, and DNA replication proteins. Defects in any of these activities can cause genome instability and disease including developmental disorders, premature aging, and cancer predisposition (89-92). The ataxia-telangiectasia mutated (ATM), ATM and Rad3-related (ATR), and DNA-dependent protein kinase (DNA-PK) are apical protein kinases activated by DNA damage or replication stress that function to coordinate cell cycle transitions with DNA metabolism including DNA repair and replication (93,94). These kinases are activated in premalignant lesions presumably due to replication stress caused by activated oncogenes or inactivation of tumor suppressors (71,73).

Accurate replication of the genome and continuous surveillance of its integrity are essential for cell survival and the avoidance of diseases such as cancer. The genome is constantly exposed to environmental and endogenous genotoxic insults that challenge DNA replication. The replication stress response

[2] The majority of this chapter has previously been published in reference 58. Bansbach, C.E., Bétous, R., Lovejoy, C.A., Glick, G.G. and Cortez, D. (2009) The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes Dev*, **23**, 2405-2414..

is a subset of the DNA damage response that acts during every cell division cycle to deal with these challenges and promotes the faithful duplication of the genome. The accumulation of single-stranded DNA at stalled replication forks as a consequence of polymerase and helicase uncoupling is an important signal to activate replication stress response pathways including the ATR checkpoint (95). The heterotrimeric single-stranded DNA binding protein replication protein A (RPA) is a mediator of multiple protein-protein interactions at stalled replication forks that promote signaling and repair (13,94,96).

All nucleic acid metabolism including replication happens in the context of chromatin. Chromatin regulation is an integral part of the DNA damage and replication stress responses. Perhaps the best understood example is phosphorylation of histone H2AX by the ATM/ATR kinases (97). This histone modification regulates the recruitment of both checkpoint and repair proteins to DSBs. In addition, everything from the location of replication origins to the elongation rate of replication forks is influenced by chromatin structure (98).

The SNF2 family of ATPases act in the context of chromatin to regulate transcription, replication, repair, and recombination. Sequence analysis has defined 24 SNF2 subfamilies (64). Many of these proteins act as chromatin remodelers to alter DNA-protein interactions. In yeast, the activities of SNF2 family members are important for replication initiation and replication through heterochromatin (99-101). They are also directly involved in responding to replication stress. For example, mutations in the yeast Ino80 complex cause hypersensitivity to replication stress agents such as hydroxyurea (HU).

Furthermore, this complex accumulates at stalled replication forks and is essential to restart replication (102).

Defects in SNF2 genes cause many human diseases including Schimke immunoosseous dysplasia (SIOD) due to loss of function mutations in the SNF2 protein SMARCAL1 (also called HARP and DNA-dependent ATPase A) (65). SIOD is a multi-system disorder characterized by renal failure, growth defects, immune deficiencies and other complex phenotypes.

In vitro, SMARCAL1 is a DNA-dependent ATPase that binds and is stimulated by DNA containing fork structures (103). SMARCAL1 has not been shown to remodel chromatin; however, it was recently demonstrated to be an ATP-dependent annealing helicase *in vitro* (52). SMARCAL1 reanneals single-stranded DNA bubbles in plasmid DNA that were stabilized by RPA. Prior to the evidence presented in this Chapter, the cellular context for where and when this activity might be important had not been previously described.

We identified SMARCAL1 in two functional genetic screens for genome maintenance activities in human cells (66). These screens, introduced in Chapter III, utilized markers of activated DNA damage response pathways to find genes whose deregulation either by RNAi or overexpression cause spontaneous DNA damage in the absence of any genotoxic agents. SMARCAL1 was identified in both screens. Subsequent experiments have defined SMARCAL1 as a replication stress response protein that acts to maintain genome integrity at stalled replication forks.

Results

Changes in SMARCAL1 expression cause activation of the DNA damage response in cells undergoing DNA replication

In the RNAi screen cells were transfected with one RNAi molecule per well in 96-well plates. Initially, DNA damage response activation resulting from RNAi-mediated gene silencing in HeLa cells was assayed by immunofluorescence staining using a phospho-peptide-specific antibody to the ATM substrate KAP1. The damage-induced phosphorylation of this transcriptional co-repressor is necessary for ATM-mediated chromatin relaxation following DNA double-strand break (DSB) formation (76). To confirm the DNA damage response activation phenotype and ensure it was neither cell type nor marker specific, four individual siRNA oligonucleotides for each candidate gene were transfected into a second cell type (U2OS), and DNA damage response activation was monitored by examining H2AX S139 phosphorylation (γH2AX). The ATM/ATR family of kinases phosphorylate γH2AX at sites of DSBs resulting in the formation of discrete, nuclear foci (104,105).

A cDNA overexpression screen was completed using a similar methodology using cDNAs expressed from a CMV promoter. In this case, small pools of cDNA expression vectors were co-transfected with a GFP vector to mark successfully transfected cells and DDR activation was monitored with antibodies to phosphorylated KAP1. Positive pools were then deconvoluted and tested in both HeLa and U2OS cells using both the KAP1 and γH2AX markers.

SMARCAL1 was identified in both the RNAi and cDNA overexpression screens. In the RNAi screen, three of four siRNAs targeting SMARCAL1 caused substantial activation of the DNA damage response in the absence of any added genotoxic agents (Figure 4.1A,B). We confirmed the siRNAs successfully silenced SMARCAL1 expression and the amount of H2AX phosphorylation correlated with the level of knockdown (Figure 4.1C).

SMARCAL1 silencing and overexpression cause distinct γH2AX phenotypes. γH2AX localizes to foci in SMARCAL1-silenced cells but exhibits pan-nuclear staining in cells overexpressing SMARCAL1 (Figure 4.1B). γH2AX foci are thought to represent sites of DSBs and the foci induced in SMARCAL1-silenced cells did co-localize with other markers of DSBs including Rad51 (data not shown). The cause of the pan-nuclear γH2AX staining is unclear, but we also observed this staining pattern in a large percentage of cells transfected with RNAi to silence the single-stranded DNA binding protein RPA. However, we were unable to detect a significant increase in Rad51 foci formation or evidence of DSBs using Comet assays in the overexpression context raising the possibility that the DNA damage response activation is not associated with DSBs in this case. Neither of the staining patterns was associated with apoptotic nuclei and both patterns were due to chromatin-associated γH2AX (data not shown). Neither overexpression nor siRNA silencing of SMARCAL1 caused a discernable change in cell proliferation rates but we did observe a slight increase in the percentage of cells in the G2/M phase of the cell cycle (Figure 4.2).

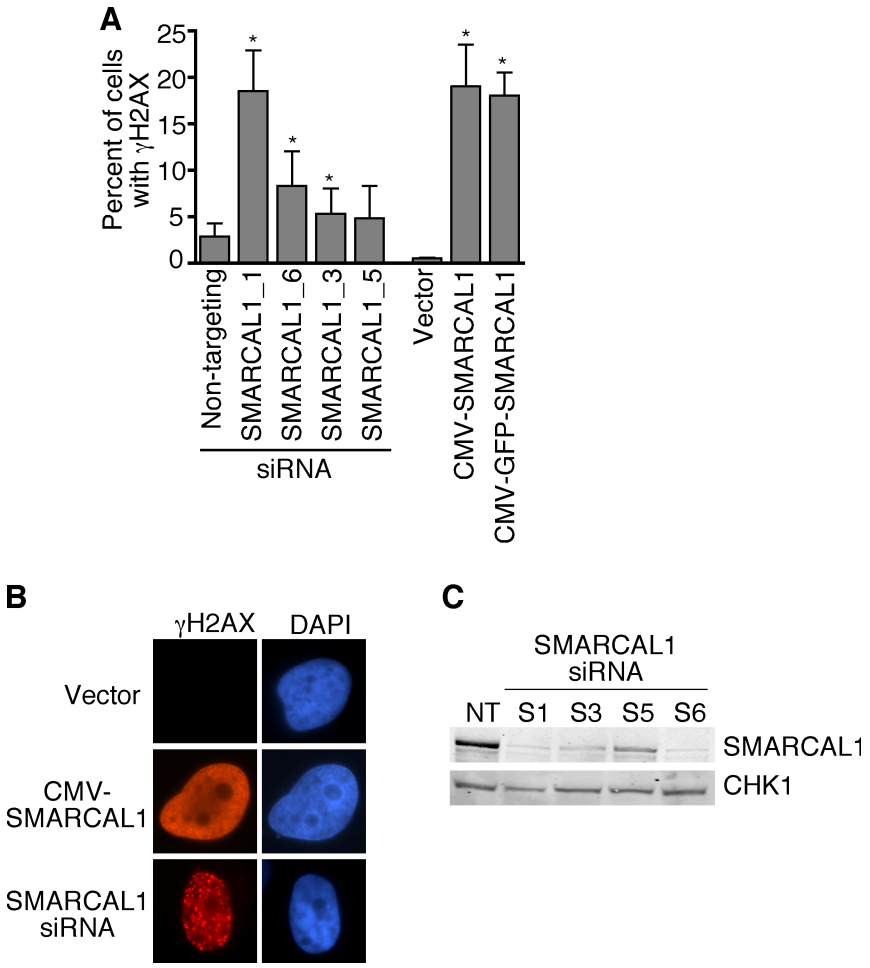

Figure 4.1. Deregulation of SMARCAL1 expression causes activation of the DNA damage response. (A) U2OS cells were transfected with siRNA targeting SMARCAL1 or expression vectors encoding either SMARCAL1 or GFP-SMARCAL1. Three days after transfection, cells were stained with antibodies to γH2AX and appropriate secondary antibodies. The percentage of cells with γH2AX staining was scored. Error bars are standard deviation (n≥3). (*, $P<0.05$) (B) Representative images of γH2AX staining. (C) Immunoblots of U2OS cell lysates with antibodies to SMARCAL1 or CHK1 after transfection with the indicated SMARCAL1 siRNAs. (NT, non-targeting).

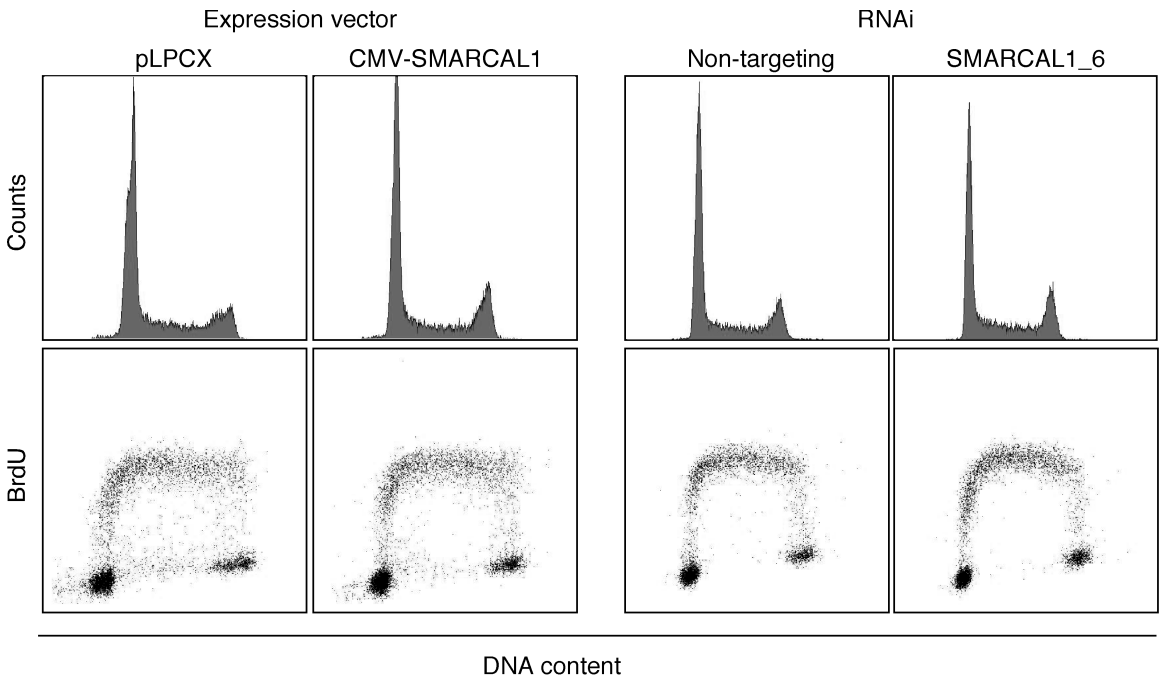

Figure 4.2. Overexpression and silencing of SMARCAL1 do not affect cell proliferation. Cell cycle profiles of cells transfected with SMARCAL1 cDNA or siRNA. HeLa cells were transfected with control or SMARCAL1 expression vectors, or with non-targeting or SMARCAL1 siRNA. Three days after transfection, the cells were labeled with BrdU, stained with fluorescein-conjugated anti-BrdU antibodies and propidium iodide, and processed by flow cytometry.

Approximately 20% of the SMARCAL1 overexpressing cells exhibited the pan-nuclear γH2AX-staining pattern. These cells were exclusively in S-phase since they incorporated BrdU and were cyclin A-positive (Figure 4.3A,B). RNAi silencing of SMARCAL1 also caused DNA damage predominantly in replicating cells. Co-staining SMARCAL1-silenced cells with antibodies to both γH2AX and cyclin A demonstrated that 60% of the $\gamma H2AX^+$ cells stained strongly for cyclin A (Figure 4.3C). This is 1.75-fold higher than predicted by chance and similar to what is observed when the replication checkpoint protein CHK1 is silenced (Figure 4.3D). In contrast, there is no preference for γH2AX staining in any phase of the cell cycle after ionizing irradiation, which causes DNA breaks irrespective of cell cycle position. We also did not observe DNA damage response activation in G_0-arrested, SMARCAL1-depleted cells (data not shown). Thus, the DNA damage response activation caused by both overexpression or silencing of SMARCAL1 is associated with cells undergoing DNA replication.

SMARCAL1 localizes to stalled replication forks via an interaction with RPA34

Adding an N-terminal green fluorescent protein (GFP) tag to SMARCAL1 did not alter the γH2AX phenotype caused by overexpression (Figure 4.1A) but did allow us to correlate the γH2AX with both SMARCAL1 expression levels and localization. GFP-SMARCAL1 accumulated in nuclear foci in 20% of the GFP-SMARCAL1 over-expressing cells (Figure 4.4A). The GFP-SMARCAL1 overexpression foci perfectly correlated with the pan-nuclear γH2AX phenotype.

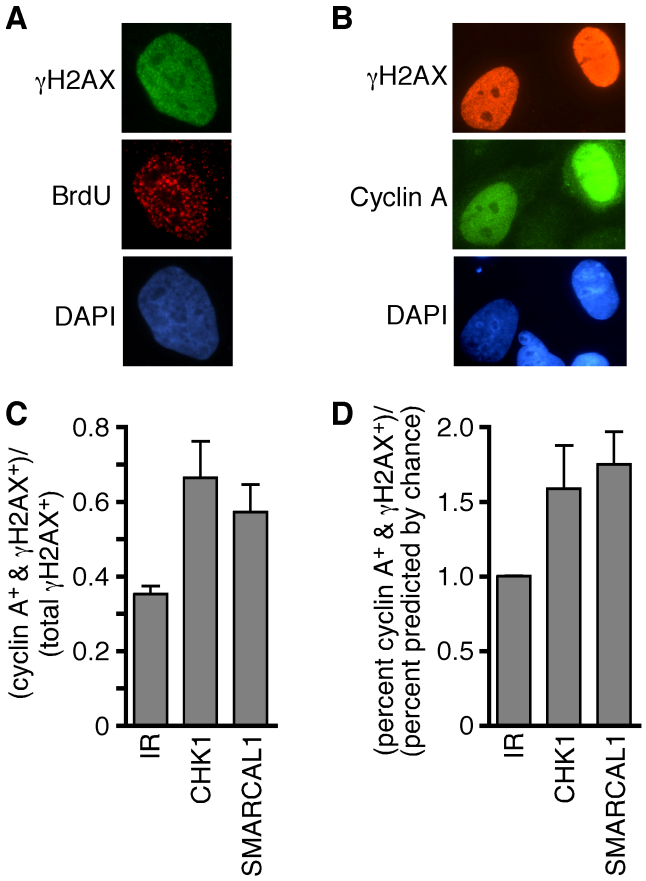

Figure 4.3. DNA damage response activation caused by SMARCAL1 deregulation occurs in S phase cells. (A and B) U2OS cells were transfected with a SMARCAL1 overexpression vector then fixed and stained with the indicated antibodies and appropriate fluorophore-conjugated secondary antibodies. In (A) the cells were pulsed with BrdU for 20 minutes prior to staining. (C and D) U2OS cells transfected with siRNA targeting CHK1 or SMARCAL1_06 siRNA were stained with antibodies to γH2AX and cyclin A. Cells were scored as positive for one or both proteins by immunofluorescence imaging. Untransfected cells were also irradiated with 5Gy of ionizing radiation for comparison.

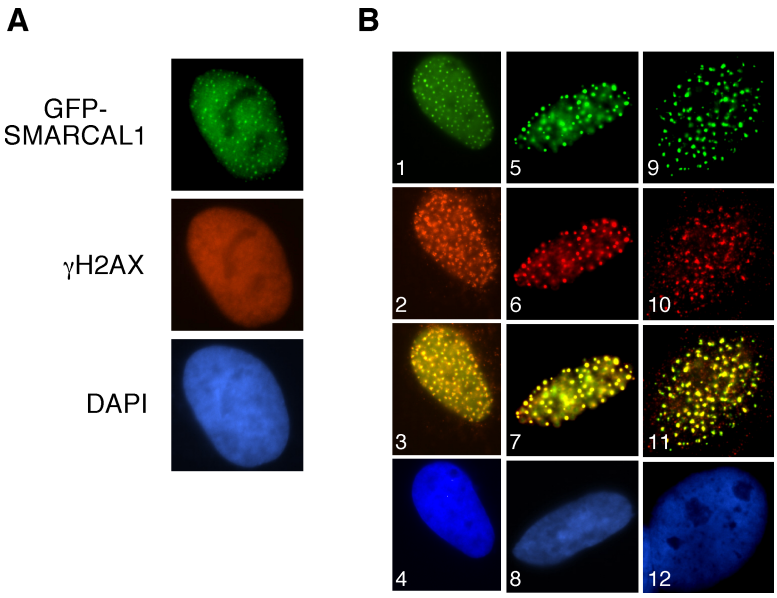

Figure 4.4. Overexpressed GFP-SMARCAL1 accumulates at replication centers. U2OS cells were transfected with the pLEGFP-SMARCAL1 expression vector then fixed and stained. Staining in (B) is as follows: (1-4 = SMARCAL1 (green) and BrdU (red); 5-8 = SMARCAL1 (green) and RPA (red); 9-12 = SMARCAL1 (green) and PCNA (red)). 3, 7, and 11 are merged images; 4, 8, 12 were stained with DAPI.

Cells in which GFP-SMARCAL1 did not accumulate into foci did not have detectable γH2AX staining.

The focal localization pattern and its restriction to S-phase cells suggested SMARCAL1 might be localized to replication factories. Indeed, labeling replication factories with BrdU, RPA, or PCNA demonstrated co-localization of SMARCAL1 foci with sites of DNA replication (Figure 4.4B). Treating cells with HU induced a marked increase in the percentage of cells containing SMARCAL1 foci (Figure 4.5A) Importantly, we also found that endogenous SMARCAL1 localizes to stalled replication forks (marked by the ATR-interacting protein ATRIP) following treatment with HU or ultraviolet (UV) radiation (Figure 4.5B,C). The percentage of cells containing SMARCAL1 foci increased with the time of HU-exposure (Figure 4.5D). Endogenous SMARCAL1 foci were rarely observed in undamaged cells. These localization data place SMARCAL1 at sites of replication stress. We also observed SMARCAL1 foci in response to ionizing radiation treatment. These foci appeared several hours after irradiation in only the S/G2 phase cells that contain cyclin A (data not shown). This pattern is consistent with a requirement for end-resection at a DSB.

To identify SMARCAL1-interacting proteins we performed both immunopurifications combined with mass-spectrometry and a two-hybrid screen. Immunopurifying endogenous SMARCAL1 complexes with antibodies to the C-terminus of SMARCAL1 or using a tandem-epitope approach combined with mass spectrometry identified peptides from all three subunits of the single-stranded DNA binding protein replication protein A (RPA) (Figure 4.6A). The

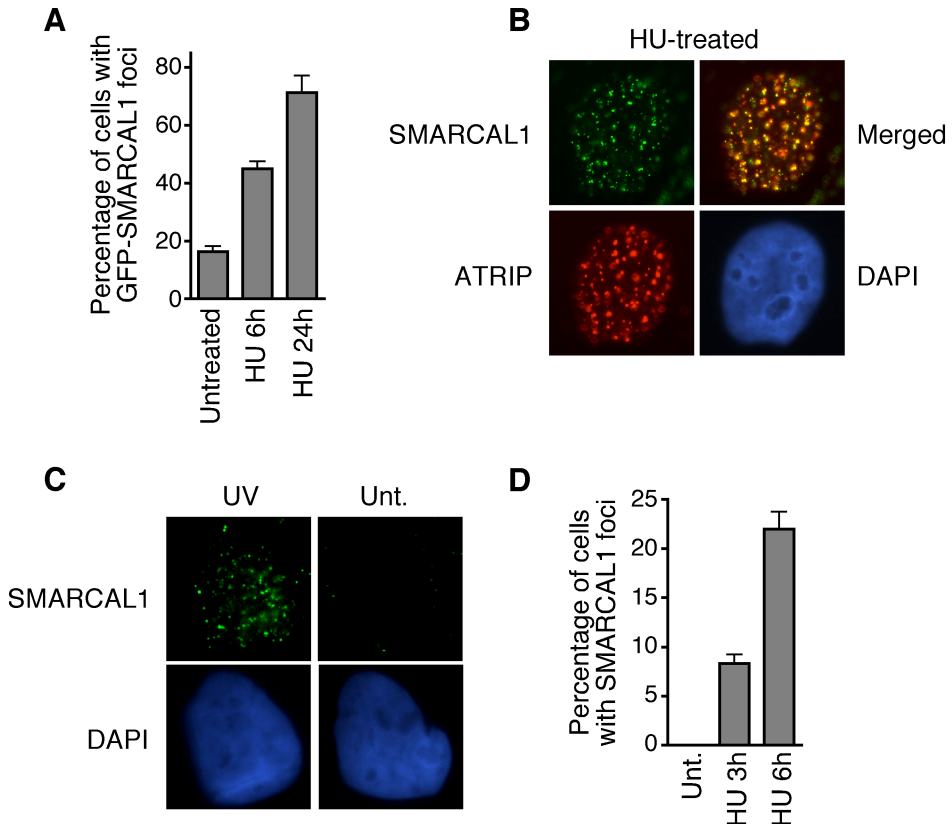

Figure 4.5. Endogenous SMARCAL1 localizes to stalled replication forks. In (A) the cells containing SMARCAL1 localized to foci were scored after addition of 2mM HU. (B,C) Localization of endogenous SMARCAL1 was examined by indirect immunofluorescence with affinity purified anti-SMARCAL1 antibody in HU, UV radiation-treated (1hr after 50J/m^2), or (Unt.) untreated U2OS cells. These cells also stably express HA-ATRIP, which permitted analysis of co-localization using anti-HA monoclonal antibody. Specificity of the SMARCAL1 antibody was confirmed in SMARCAL1-silenced cells. Cells were pre-extracted with 0.5% triton X-100 solution for 10min prior to fixation so only chromatin bound-SMARCAL1 is observed. (C) The percentage of cells containing endogenous SMARCAL1 in foci was scored in cells treated with 2mM HU for the indicated times. Error bars in all graphs are standard deviation (n=3).

two-hybrid screen using a full-length SMARCAL1 protein bait identified ten interacting fragments of the 32kDa subunit of RPA. The minimal interacting region contained RPA32 amino acids 156-267 corresponding to the winged-helix protein-interaction domain (32C) (Figure 4.6B). This region of RPA32 binds to several DNA repair proteins including UNG2, XPA, and RAD52 (106). Co-immunoprecipitation of endogenous proteins confirmed the SMARCAL1-RPA interaction and indicated it is stimulated by replication stress (Figure 4.6C,D).

While a C-terminal antibody to SMARCAL1 efficiently co-immunoprecipitates all three RPA subunits (Figure 4.5A) an antibody raised to the N-terminus of SMARCAL1 fails to co-immunoprecipitate RPA (data not shown) suggesting it may compete for the same binding surface as RPA. Sequence analysis of this SMARCAL1 region identified a highly evolutionarily conserved sequence (amino acids 1-32) that has similarity to the Timeless-Interacting Protein (TIPIN) (Figure 4.7A). Notably, this region of TIPIN was previously demonstrated to bind RPA32 (107) suggesting it might be the RPA32-interacting surface on SMARCAL1. Indeed, deletion of the first 32 amino acids of SMARCAL1 severely impairs its ability to interact with RPA (Figure 4.7B). Furthermore, this SMARCAL1-ΔN32 protein fails to localize to intranuclear foci either in the absence or presence of HU (Figure 4.8C).

To determine whether this N-terminal region of SMARCAL1 is also sufficient to bind RPA32 and localize SMARCAL1 to stalled forks, we fused it to GFP. A GFP protein containing only the first 32 amino acids of SMARCAL1 (GFP-SMARCAL1 1-32) bound poorly to RPA and localized to foci in only a small

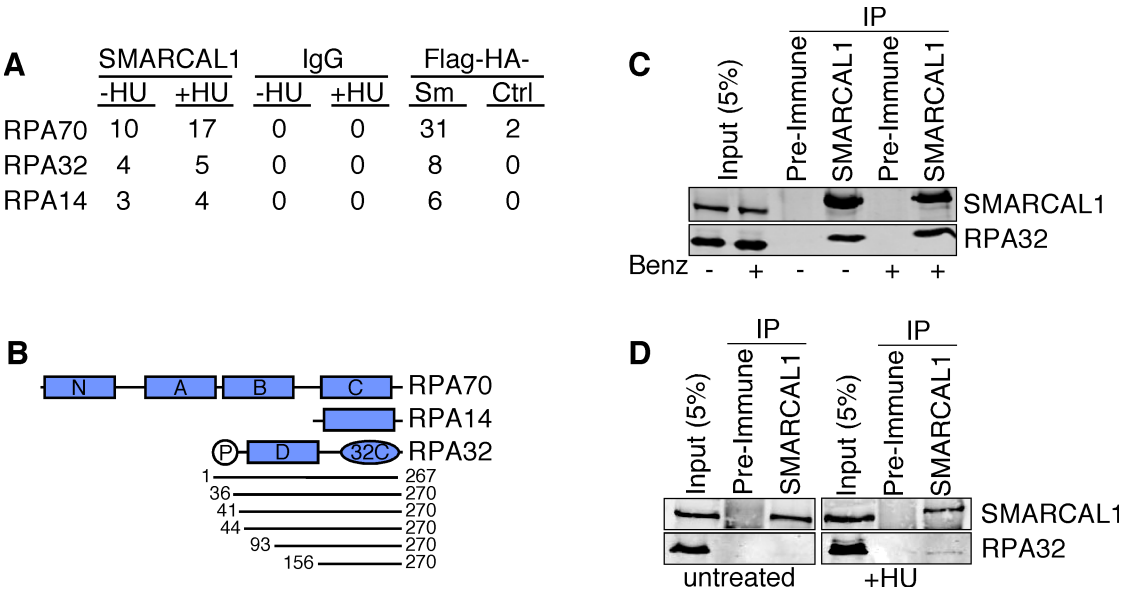

Figure 4.6. Immunopurifications and two-hybrid screen identify RPA as an interacting partner of SMARCAL1. (A) Endogenous SMARCAL1 or Flag-HA-SMARCAL1 was immunopurified from nuclear cell extracts and the resulting protein complexes were analyzed by mass spectrometry. The table indicates the number of peptides identified for each RPA protein subunit. IgG is a control immunoprecipitation; Sm = Flag-HA-SMARCAL1; Ctrl=Flag-HA empty vector cells. Where indicated, the cells were treated with 1mM HU for 16 hours prior to the purification. (B) A schematic diagram of RPA subunits. The lines under RPA32 indicate the sizes (with amino acid numbers) of the different RPA32 protein fragments identified in the unbiased two hybrid screen using full-length SMARCAL1 as a bait. (C,D) HeLa cell nuclear extracts were used for immunoprecipitation with anti-SMARCAL1 or pre-immune antibodies. In (C), the lysates were treated with benzonase nuclease (Benz.) as indicated to ensure the interaction is not dependent on DNA. Nuclear extracts from untreated or cells treated with HU for 8 hours were used for immunoprecipitation in (D). Immunoprecipitated proteins were separated by SDS-PAGE and immunoblotted with RPA or SMARCAL1 antibodies. Data for panels A, C, and D were produced by Remy Betous.

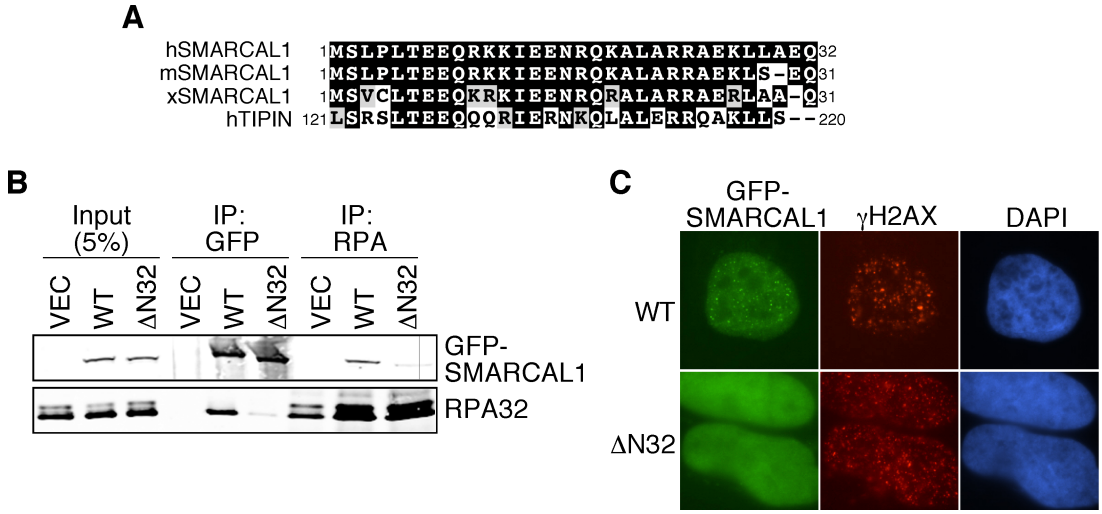

Figure 4.7. The N-terminal region of SMARCAL1 is necessary for RPA binding and localization to stalled replication forks. (A) Sequence alignment of human-, mouse-, and *Xenopus* SMARCAL1 with human TIPIN. (B) HEK293 cells were transfected with wild-type GFP-SMARCAL1 (WT) or GFP-SMARCAL1 lacking the first thirty-two amino acids (Δ32). Cells transfected with an empty vector (VEC) were also prepared as a control. Anti-GFP or anti-RPA immunoprecipitated proteins were separated by SDS-PAGE and blotted with anti-GFP or anti-RPA2 antibodies. (C) U2OS cells were transfected with GFP-SMARCAL1-WT or -Δ32 vectors and treated with 2mM HU for 6 hours. Approximately 30% of the wild-type SMARCAL1 expressing cells had SMARCAL1 foci while we never observed the SMARCAL1-Δ32 protein in foci. Representative images are shown. Data for panel B was produced by David Cortez.

percentage of HU-treated cells (Figures 4.8A,B). However, fusion of a slightly longer portion of the SMARCAL1 N-terminus (amino acids 1-115) to GFP created a protein that binds to RPA and localizes to stalled replication forks as efficiently as full-length SMARCAL1 (Figure 4.8A-C). In these experiments we used a GFP-expression vector (pLL5.0-GFP-GW) with attenuated expression levels to avoid spontaneously activating the DNA damage response in the absence of HU that is observed with wild-type GFP-SMARCAL1 expressed at higher levels. Thus, the N-terminus of SMARCAL1 encodes a binding surface for RPA32 that is both necessary and sufficient to localize SMARCAL1 to stalled replication forks.

An SIOD-associated patient mutation and an RPA-binding mutant are both defective in the cellular functions of SMARCAL1

To determine whether the spontaneous DNA damage response activation observed after high levels of SMARCAL1 overexpression is due to too much enzymatic activity we examined whether an SIOD-associated patient mutation (R764Q) in SMARCAL1 that perturbs its ATPase and annealing helicase activities alters the ability of overexpressed SMARCAL1 to cause DNA damage. We also tested the ability of SMARCAL1 lacking its RPA binding domain to cause DNA damage response activation to determine whether localization to replication factories and/or RPA binding is essential. GFP-SMARCAL1-R764Q is deficient in activating the DNA damage response when overexpressed in cells compared to wild-type SMARCAL1 (Figure 4.9A) Similarly, the GFP-SMARCAL1 ΔN32 protein completely fails to activate the DNA damage response although it is expressed at equal levels as wild-type SMARCAL1 (Figure 4.9A,B). The SIOD

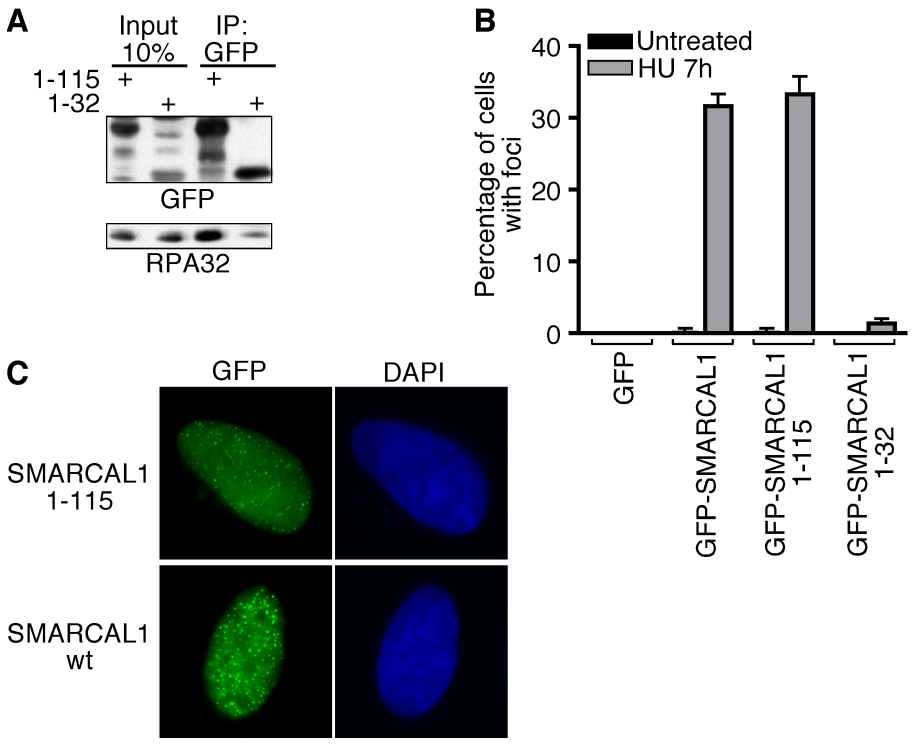

Figure 4.8. Amino acids 1-115 of SMARCAL1 are sufficient to bind RPA and localize to stalled replication forks. (A) GFP-SMARCAL1 containing only the first 32 or 115 amino acids of SMARCAL1 was transfected into cells, immunoprecipitated and immunoblotted. (B) The indicated expression vectors (pLL5.0-GFP-GW backbone for attenuated expression levels) were transfected into HeLa cells. The cells were treated for 7h with 1mM HU or left untreated and the percentage of cells containing foci of the indicated proteins was scored. Error bars are standard deviation (n=3). (C) Representative images of the localization of SMARCAL1-1-115 and wild-type SMARCAL1. David Cortez performed all experiments for this figure.

mutant SMARCAL1 protein can be recruited to stalled replication forks as efficiently as wild-type SMARCAL1 in HU-treated cells (Figure 4.9C). In contrast, as noted previously, the SMARCAL1-ΔN32 protein does not localize to foci. These data indicate that both the RPA binding and enzymatic activities of SMARCAL1 are necessary to cause replication-associated DNA damage response activation. Overexpression of SMARCAL1 likely deregulates its activity perhaps allowing aberrant access to replication forks and causing problems during DNA replication.

We next asked whether the SIOD patient mutation or the RPA-binding mutation would impair the function of endogenous SMARCAL1 in cells. Specifically, we tested the ability of wild-type and mutant SMARCAL1 proteins to functionally complement the spontaneous appearance of γH2AX foci in SMARCAL1-silenced cells. U2OS cells co-transfected with siRNA-resistant pLL5.0-GFP-GW-SMARCAL1 cDNAs and either non-targeting siRNA as a control or SMARCAL1 siRNA were stained with antibodies to γH2AX. GFP-positive cells were scored for γH2AX foci. As expected, transfection of the GFP empty vector by itself did not prevent the appearance of γH2AX foci in SMARCAL1-silenced cells (Figure 4.9D). Expression of wild-type GFP-SMARCAL1 completely suppressed this defect. However, neither expression of the R764Q SIOD mutant nor the ΔN RPA-binding mutant SMARCAL1 protein was capable of preventing the appearance of γH2AX foci when endogenous SMARCAL1 was silenced (Figure 4.9D). Thus, RPA-binding is essential for the

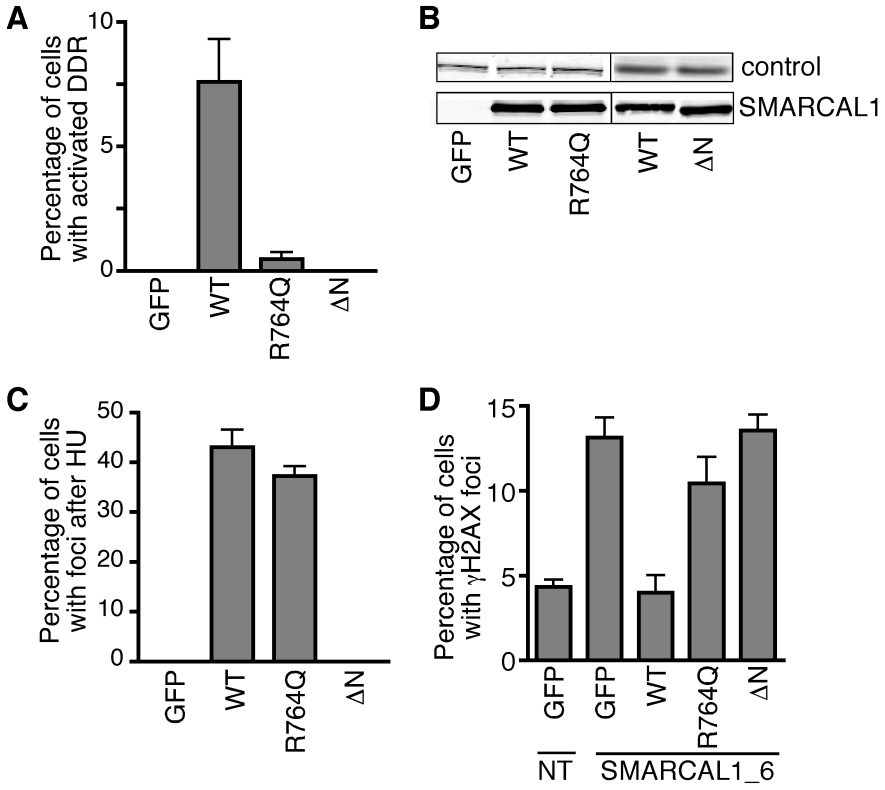

Figure 4.9. RPA binding and annealing helicase activities of SMARCAL1 are required for its cellular functions. (A) HeLa cells were transfected with expression vectors (pLEGFP backbone for high levels of expression) encoding GFP-tagged wild-type SMARCAL1 (WT), the annealing helicase deficient SIOD patient mutant protein R764Q, or SMARCAL1-Δ32. The amount of DNA damage response activation was scored by monitoring KAP1 phosphorylation. (B) Expression levels of the GFP-SMARCAL1 proteins were determined by immunoblotting. The loading control was either ATM or RPA32. (C) The percentage of HU-treated cells in which GFP-SMARCAL1 wild-type or mutant proteins localized to intra-nuclear foci was scored. Cells were treated with HU for 8 hours. (D) U2OS cells were co-transfected with the indicated pLL5.0-GFP-GW plasmids that encode siRNA-resistant SMARCAL1 cDNAs and siRNA targeting SMARCAL1. GFP-positive cells were scored for γH2AX staining. (GFP, GFP empty vector; NT, non-targeting siRNA control). Error bars in all panels are standard deviation (n=6).

function of SMARCAL1 *in vivo*. These data also link the SIOD phenotype caused by the R764Q mutation directly to a genome maintenance defect.

The ability of SMARCAL1 to bind to RPA raises the question of whether this binding only serves to regulate the localization of SMARCAL1 or if it also is important for the annealing helicase activity previously described for SMARCAL1 (52). The annealing helicase assay utilizes the ability of RPA to bind to supercoiled plasmid DNA and form bubbles by melting the DNA strands. When SMARCAL1 is added, the DNA strands are re-annealed in an ATP-dependent reaction, thereby displacing RPA. We tested SMARCAL1- ΔN32 to determine if this RPA-binding mutant maintains its ATPase and annealing helicase activities. We found no difference between wild-type and mutant SMARCAL1 in either assay (Figure 4.10). Fork DNA stimulates the ATPase activity of SMARCAL1-ΔN32 equivalently to wild-type SMARCAL1 (Figure 4.10B). Furthermore, it acts as an annealing helicase as efficiently as wild-type SMARCAL1 (Figure 4.10C,D). The small differences observed at some concentrations were not reproducible. Thus, RPA binding primarily serves to regulate the localization of SMARCAL1, which is critical for SMARCAL1 function in cells but not essential for its enzymatic activity.

SMARCAL1 is regulated by the DNA damage response

In the course of studying SMARCAL1, we noticed that it often migrates as multiple bands on SDS-PAGE gels. Closer examination indicated that SMARCAL1 exhibits a phosphorylation-dependent gel-mobility shift when cells

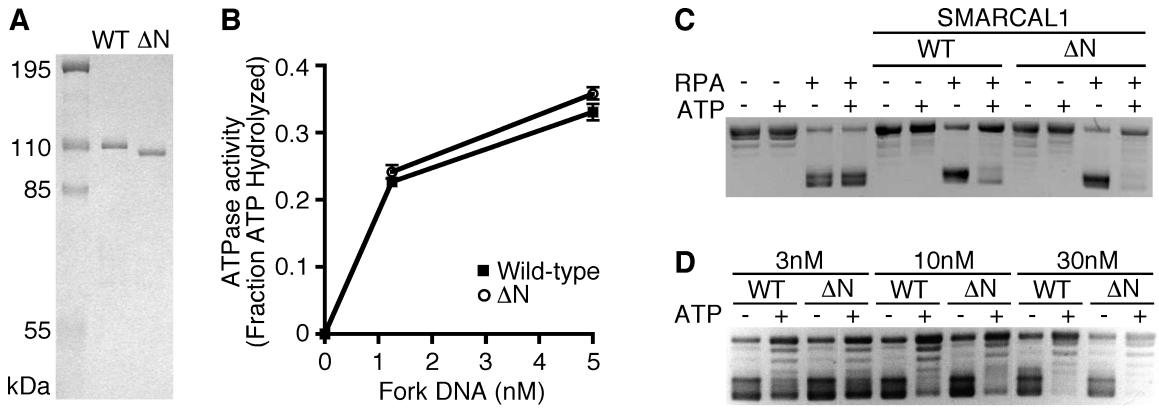

Figure 4.10. The SMARCAL1 RPA binding domain is not necessary for ATPase or annealing helicase activities *in vitro*. (A) Coomassie stained gel of the Flag-SMARCAL1 and Flag-SMARCAL1-D32 purified from baculovirus-infected insect cells. (B) ATPase activity of the wild-type and D32 SMARCAL1 proteins was measured in the presence of the indicated concentrations of fork DNA. Error bars are standard deviation (n=3). (C and D) Annealing helicase activities of wild-type and D32 SMARCAL1 proteins. The concentration of SMARCAL1 protein in (C) is 300nM. Remy Betous performed all experiments for this figure.

are treated with replication stress and DNA damaging agents including HU, IR, and UV-radiation (Figure 4.11A,B).

SMARCAL1 contains several consensus motifs (SQ/TQ) for phosphorylation by the DNA damage and replication stress-activated ATM/ATR family of kinases. To determine whether these kinases can phosphorylate SMARCAL1 we used *in vitro* kinase assays. Purified wild-type ATM and ATR but not kinase-dead proteins efficiently phosphorylated SMARCAL1 *in vitro* (Figure 4.12A). Purified DNA-PK also phosphorylated SMARCAL1 in a DNA-stimulated reaction (Figure 4.12B).

To determine which of these kinases may be responsible for the phosphorylation of SMARCAL1 in cells, each was depleted using siRNA. However, we failed to detect any difference in the HU-dependent phosphorylation of SMARCAL1 when individual kinases were silenced (Figure 4.13A). We then examined combinations of specific kinase inhibitors and RNAi. Specific ATM and DNA-PK inhibitors had no appreciable effect on SMARCAL1 phosphorylation after HU even when used together (Figure 4.13B, compare lanes 2 and 5). The inhibitors also had only a minimal effect after IR (Figure 4.13B, compare lanes 6 and 9). Silencing ATR by itself also had minimal effects after HU treatment, but it did reduce phosphorylation after IR (Figure 4.13B, compare lanes 6 and 16). Silencing ATR and treating cells with the specific ATM and DNA-PK inhibitors significantly reduced the phosphorylation of SMARCAL1 both after HU and IR treatment (Figure 4.13B, compare lanes 5 and 15, 9 and 19). The residual phosphorylation of SMARCAL1 under these circumstances in which all three

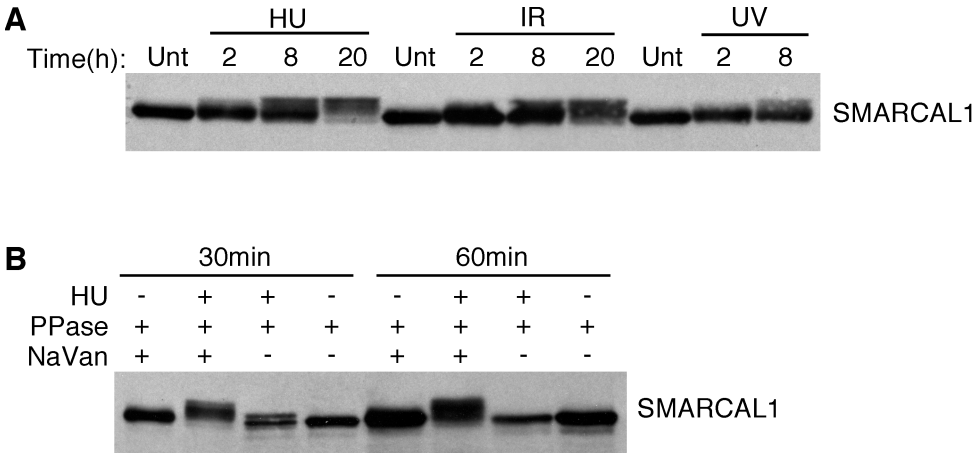

Figure 4.11. SMARCAL1 is phosphorylated in response to replication stress or DNA damage. (A) Cell lysates from cells treated with 1mM HU, 10Gy of IR, or 50J/m^2 of UV-radiation, were separated by SDS-PAGE and immunoblotted. The time points represent the duration of the HU treatment or the length of time the cells were allowed to recover following the radiation exposure. (Unt, untreated). (B) Cell lysates from untreated or HU-treated cells were incubated in the presence of lambda phosphatase and the phosphatase inhibitor sodium vanadate as indicated for 30 or 60 minutes prior to separation by SDS-PAGE and immunoblotting. David Cortez performed all experiments for this figure.

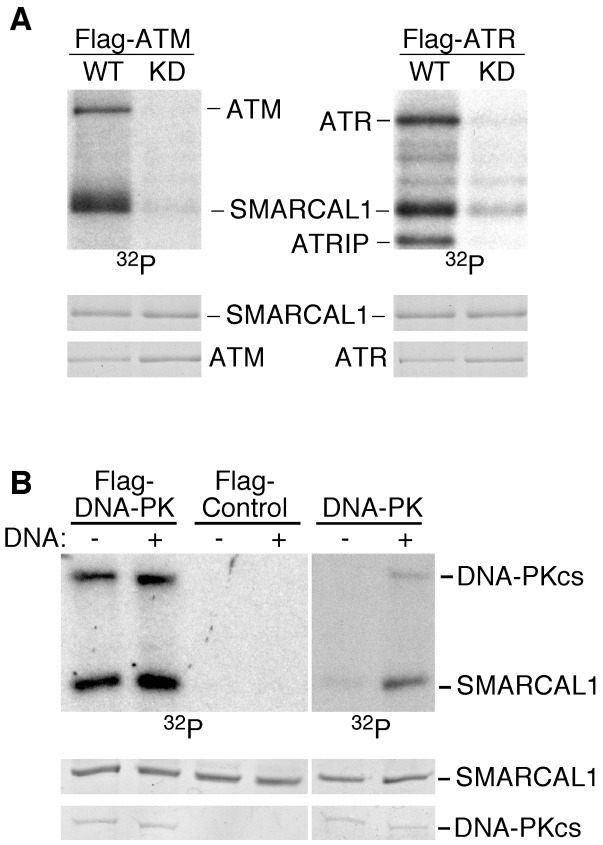

Figure 4.12. ATM, ATR, and DNA-PK phosphorylate SMARCAL1 *in vitro*. (A and B) Flag-ATM or Flag-ATR wild-type (WT) or kinase dead (KD) proteins or wild-type Flag-DNA-PKcs were immunopurified from transfected cells. Wild-type DNA-PK was also purchased from Promega (B, right panel). The kinases were incubated with purified SMARCAL1 in the presence of γ-^{32}P-ATP. Sonicated salmon sperm DNA was added to the DNA-PK reactions as indicated. Reactions were separated by SDS-PAGE and subjected to autoradiography (^{32}P). The levels of SMARCAL1, ATM, ATR, and DNA-PKcs proteins in each reaction were visualized by coomassie staining. David Cortez performed all experiments for this figure.

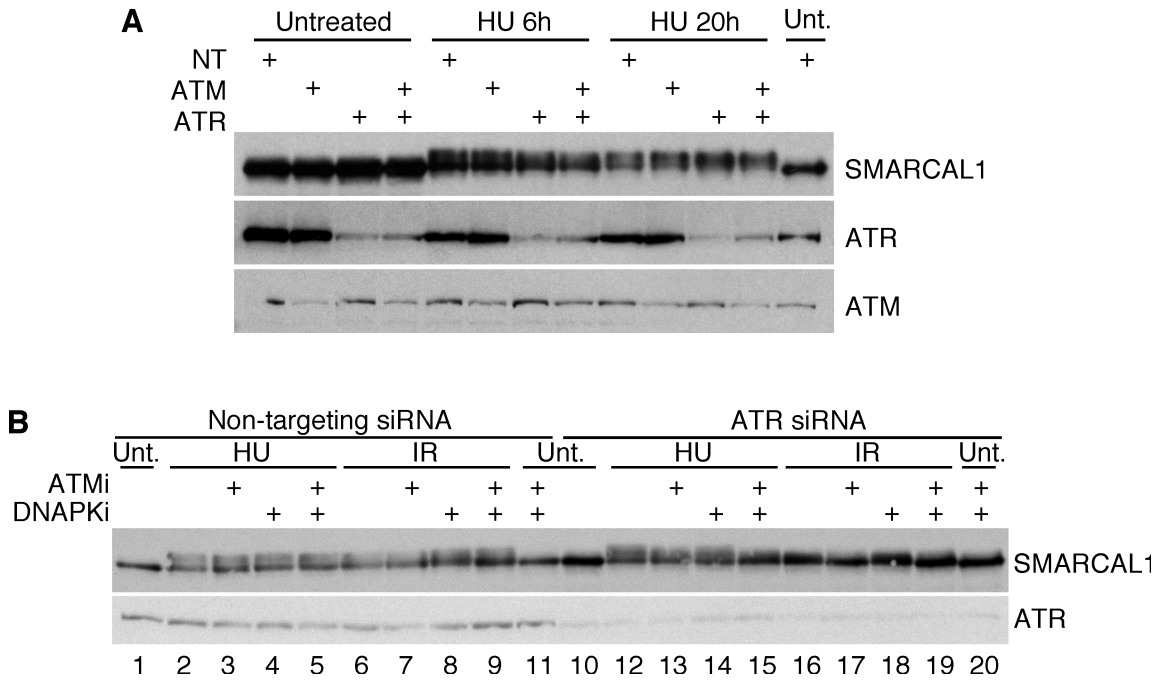

Figure 4.13. ATM, ATR, and DNA-PK phosphorylate SMARCAL1 in cells. (A) 293T cells transfected with the indicated siRNAs (NT, non-targeting) were treated with HU for 6 or 20h. Lysates were separated by SDS-PAGE and immunoblotted with antibodies to SMARCAL1, ATR, or ATM. (B) 293T cells were transfected with either non-targeting (NT) or ATR siRNA. Cells were also treated with specific ATM (KU55933**,** 10μM**)** (108) or DNA-PK (KU57788, 1μM**)** (109) inhibitors as indicated and exposed to 1mM HU or 10Gy of IR. Cell lysates were separated by SDS-PAGE and immunoblotted for SMARCAL1 or ATR. David Cortez performed all experiments for this figure.

kinases are inhibited may be due to incomplete silencing of ATR, inhibition of ATM or DNA-PK, or potentially the involvement of additional kinases.

This type of phosphorylation pattern is common to many substrates of the ATM/ATR family of kinases such as RPA32 that are phosphorylated by multiple family members. SMARCAL1 is likely phosphorylated on several sites since the banding pattern on SDS-PAGE gels is complex. To identify sites of phosphorylation, we immunopurified SMARCAL1 from HU-treated cells and subjected it to mass spectrometric analysis. Several phosphorylation sites were observed including S173 that lies in a good consensus sequence for the ATM/ATR family (SQE) (see appendix for full list and spectra). Our data indicate that SMARCAL1 functions downstream of the ATM/ATR/DNA-PK kinases in response to replication stress. I present further analysis of SMARCAL1 damage-dependent phosphorylation in Chapter V.

Loss of SMARCAL1 function causes persistent RPA phosphorylation, RPA loading onto chromatin and hypersensitivity to replication stress agents.

The localization of SMARCAL1 to stalled replication forks, appearance of γH2AX foci in S-phase of SMARCAL1-silenced cells, phosphorylation by the ATM/ATR/DNA-PK kinases, and its ability to act as an annealing helicase *in vitro* suggests that SMARCAL1 may have a critical function to anneal DNA strands at stalled replication forks. If this were true, we would expect that SMARCAL1-deficient cells should accumulate more RPA bound to single-stranded DNA than control cells. Consistent with this hypothesis, more RPA is loaded onto chromatin in SMARCAL1-deficient cells than control cells and this difference is

even more pronounced after exposure to HU (Figure 4.14A). SMARCAL1 is also loaded onto chromatin under these conditions. RPA loaded at stalled replication forks is rapidly phosphorylated by the ATM/ATR family of kinases (96). SMARCAL1-silenced cells treated with HU exhibited hyper-phosphorylation of RPA that also persisted after the HU was removed (Figure 4.15B). These data are consistent with a model in which the annealing helicase activity of SMARCAL1 functions to reduce the amount of single-stranded DNA available for RPA binding at stalled forks.

If SMARCAL1 activity is important for the proper nucleic acid metabolism at sites of replication stress, then cells lacking SMARCAL1 should be hypersensitive to replication stress agents. Indeed, silencing SMARCAL1 causes a marked hypersensitivity to HU, the DNA polymerase inhibitor aphidicolin, and the topoisomerase I inhibitor camptothecin (Figure 4.15A-C). Thus, SMARCAL1 activity is important to maintain viability when cells are challenged with drugs that induce uncoupling of helicase and polymerases at replication forks.

Discussion

The SIOD disease is caused by loss of function mutations in *SMARCAL1* (65). This disease has a complex phenotype affecting multiple organs and tissues. There is also variability in phenotypic severity among patients. These characteristics have made it difficult to propose a unifying model for the

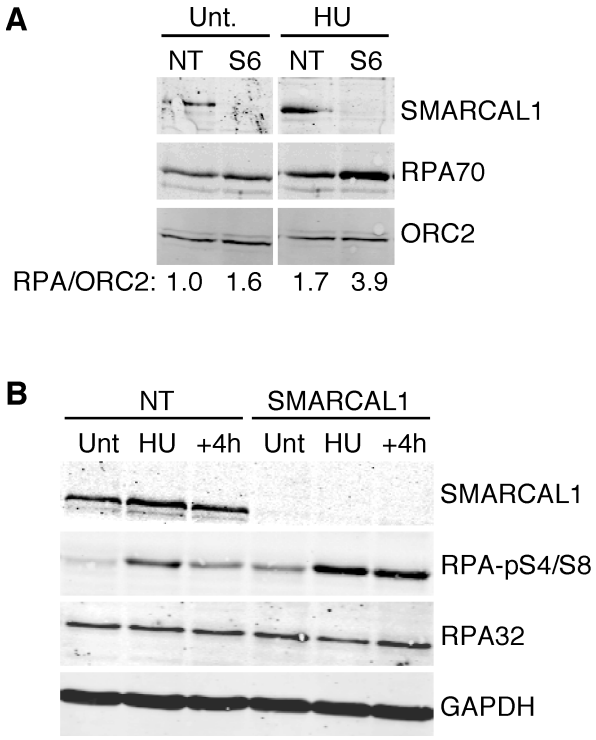

Figure 4.14. SMARCAL1-deficient cells exhibit increased RPA-loading onto chromatin and persistent RPA phosphorylation after replication stress. (A) The chromatin fraction of cells transfected with non-targeting (NT) or SMARCAL1 (S6) siRNA was isolated and immunoblotted with the indicated antibodies. Quantitation of the amount of RPA versus ORC2 in the chromatin fraction was performed with an Odyssey instrument (arbitrary units). (B) Cells transfected with non-targeting or SMARCAL1 siRNA were treated with HU for 5h (HU) then the HU was removed and the cells were allowed to recover for 4h (+4h). Total cell lysates were prepared and immunoblotted with the indicated antibodies. (Unt, untreated).

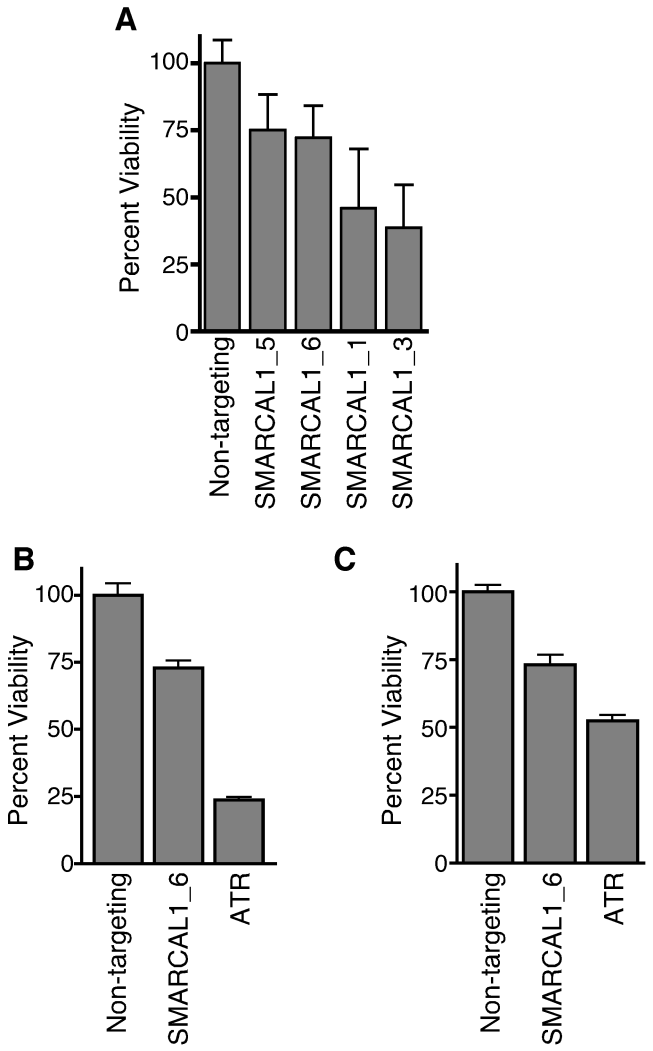

Figure 4.15. SMARCAL1-deficient cells are hypersensitive to replication stress agents. (A) Cells transfected with the indicated siRNA were split into four plates, two of which were treated with 2mM HU for 24hrs. The HU was removed and replaced with standard growth media for an additional 24h before cell viability was quantitated using the WST-1 cell proliferation reagent (Roche). Reported values are the mean and standard deviation of three experiments. All four SMARCAL1 siRNAs yielded significant hypersensitivity to HU with p values ranging from 0.02 to 10^{-4}. (B and C) Cells transfected with the indicated siRNAs were treated with 5mM aphidicolin or 150nM camptothecin and cell viability compared to untreated controls. P value of non-targeting compared to SMARCAL1-silenced for both aphidicolin (n-=12) and camptothecin (n=24) treatments are less than 10^{-5}.

molecular functions of SMARCAL1 that would explain the disease. Our mechanistic analysis of SMARCAL1 suggests at least part of the underlying molecular cause of SIOD is a defect in cellular responses to replication stress.

SMARCAL1 is a member of the SNF2 family of ATPases that function in a variety of chromatin-associated processes including transcription, replication, and DNA repair (64). These proteins often alter DNA-protein interactions as a consequence of hydrolyzing ATP. A single SMARCAL1 protein is found in humans and it is evolutionarily conserved in invertebrates including *C. elegans* and *D. melanogaster*. However, yeast orthologues have not been identified.

A major advance in understanding SMARCAL1 activity was made with the identification of its annealing helicase activity *in vitro* (52). While many helicases can act to reverse a fork under some *in vitro* conditions, SMARCAL1 is the only protein identified thus far that primarily acts to anneal two DNA strands without exhibiting any helicase activity. The context of where and when this activity is important in the cell has not been defined.

We have now defined SMARCAL1 as a replication stress response protein that acts at stalled replication forks to limit replication-associated DNA damage. SMARCAL1 contains an RPA32 interacting domain at its N-terminus similar to a domain in the replication stress response protein TIPIN (107). This domain is both necessary and sufficient to localize SMARCAL1 to stalled forks. RPA binding is required for the genome maintenance activity of SMARCAL1 during DNA replication, but it is not required for its annealing helicase activity *in vitro*.

Like many replication stress response proteins, SMARCAL1 is phosphorylated by the ATM/ATR/DNA-PK family of protein kinases. SMARCAL1 contains multiple consensus sequences for phosphorylation and its complex gel mobility pattern indicates that several may be phosphorylated. Our mass spectrometry analysis indicates one of these is S173. The ATM/ATR/DNA-PK kinases have somewhat overlapping and redundant activities on SMARCAL1 in response to replication stress since all three kinases must be inhibited to observe a significant reduction in phosphorylation after HU. This is not unusual for substrates of these kinases since HU can activate all three kinases and there is significant cross-talk between them.

Both silencing and overexpression of SMARCAL1 causes activation of the DNA damage response kinases in replicating cells. Activation of the DNA damage response upon overexpression of SMARCAL1 is strictly limited to S-phase cells. The pattern of γH2AX staining in these SMARCAL1 overexpressing cells is unusual in that it is pan-nuclear and never localized in discrete foci. This pattern has previously been noted in cells in which ATR is activated in the absence of DNA damage (110,111). However, we have found no evidence that SMARCAL1 functions as a direct regulator of ATR activation. We have also observed this pan-nuclear pattern when RPA is depleted from cells with RNAi. However, in this circumstance a significant portion of cells in the population also exhibit focal γH2AX staining. As yet, the molecular basis of the SMARCAL1 overexpression-associated pan-nuclear γH2AX staining is unclear, but our data does indicate that it is dependent on both the ATPase and RPA binding activities

of SMARCAL1. Multiple proteins bind to the same region of RPA32 as SMARCAL1 (106). A competition between interacting proteins is likely to prevent SMARCAL1 from gaining access to the limited amount of RPA at normal elongating forks. However, when SMARCAL1 is overexpressed, it may out-compete other proteins, gain access to normal replication forks, and then use its enzymatic activity to reanneal DNA strands inappropriately. RPA silencing may phenocopy SMARCAL1 overexpression in some cells when RPA levels are insufficient to stabilize the ssDNA at elongating forks.

Silencing SMARCAL1 increases both the amount of RPA bound to chromatin and its phosphorylation following a challenge with HU. RPA phosphorylation also persists after removing HU from the growth media in SMARCAL1-silenced cells. These results are consistent with the idea that SMARCAL1 acts as an annealing helicase at stalled forks to limit the amount of single-stranded DNA that accumulates. Since the DNA at a replication forks is not bound into nucleosomes, it may have an increased propensity to form bubbles or flaps that would be substrates for the annealing helicase activity of SMARCAL1. DNA polymerase and helicase activities are uncoupled at replication forks in response to DNA lesions or agents like HU that interfere with DNA synthesis (95). Uncoupling causes the formation of regions of single-stranded DNA that is rapidly coated with the single-stranded DNA binding protein RPA. We envision at least four possible substrates for the annealing helicase activity of SMARCAL1 at a stalled fork (Figure 4.16). (a) The uncoupled replication fork itself could be a substrate, although this may require dissociation

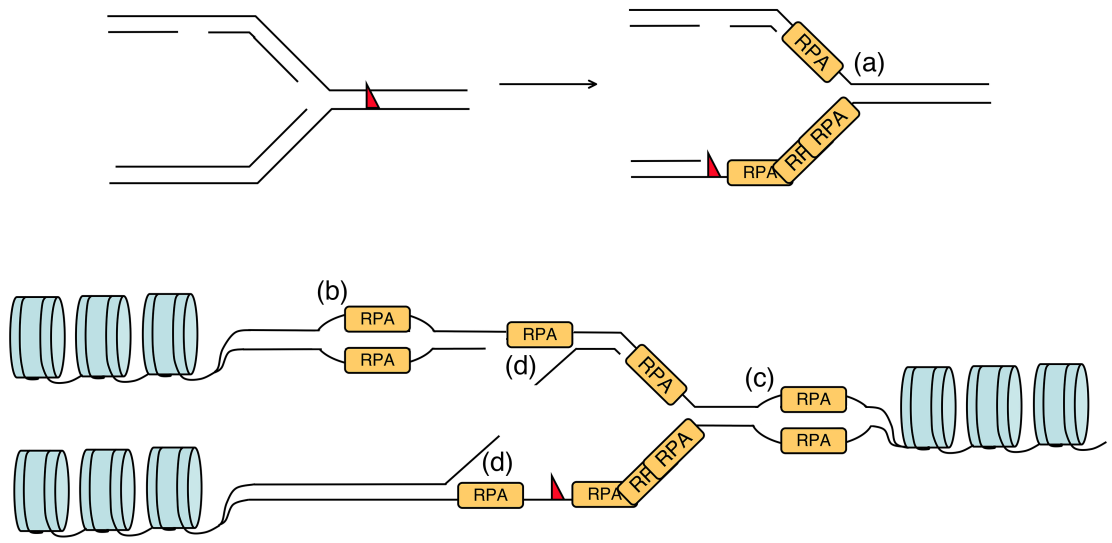

Figure 4.16. Model of possible substrates for SMARCAL1 annealing helicase activity at stalled replication forks. (see text for details)

of the MCM helicase complex for SMARCAL1 to obtain access. (b or c) Single-stranded bubbles may form in the DNA at the stalled fork since it is not stabilized into nucleosome structures. The positive supercoiling ahead of the fork would likely inhibit bubble formation at position (c). However, the pre-catanenes behind the fork may not prevent ssDNA bubbles from forming especially in genomic sequences that are AT rich or prone to secondary structure formation. (d) SMARCAL1 may anneal DNA flaps that could form when replicative polymerases dissociate from the DNA. Flaps could form on either the leading strand or lagging strand at the 5' or 3' end of an Okazaki fragment. If SMARCAL1 prevents flap formation, then it may act to inhibit a process of fork-reversal (112,113), which is energetically favorable but thought to be inhibited by the tethering of the nascent strand ends to their parental templates by the replisome during normal replication (114). Finally, it is possible that SMARCAL1 acts at replication origins to limit replication origin firing. Of course, these potential activities are not mutually exclusive.

SMARCAL1 deficiency causes the appearance of DSB markers at replication factories, increased RPA loading and phosphorylation in response to replication stress, and decreased viability after treatment of cells with agents that stall replication forks. Cells expressing only an SIOD-associated mutant SMARCAL1 protein accumulate γH2AX foci even in the absence of a replication stress agent. While it is difficult to link these effects directly to the symptoms in SIOD patients, these data suggest that loss of SMARCAL1 function in patients

may cause DNA replication-associated genome instability that contributes to the pleiotropic phenotypes of this disease.

Chapter V

CHARACTERIZATION OF DDR-REGULATED SMARCAL1 PHOSPHORYLATION

Introduction

The DDR kinases ATR, ATM, and DNA-PK phosphorylate hundreds of overlapping substrates at serine and threonine residues that are followed by glutamine (S/TQs). The substrates identified thus far, mostly by massive proteomic efforts (88,115,116), function in diverse cellular pathways. How DDR-dependent phosphorylation regulates these pathways remains vastly uncharacterized.

During DNA replication, DDR kinases are important for responding to DNA damage. ATR activation is especially important during S phase to coordinate DNA replication with cell cycle transitions (94,117). ATR signaling promotes fork stability, preventing the accumulation of aberrant fork structures and DNA damage during replication (29). In the absence of fork stabilizing activities, DSBs regularly form at collapsed replication forks and activate ATM and DNA-PK to promote fork repair (44). It is not clear how ATR signaling stabilizes forks but the most likely targets are proteins that function directly at stalled replication forks. Investigating the functional consequences of DDR-dependent phosphorylation for this particular group of proteins is critical for understanding the mechanisms that promote replication fork stability.

SMARCAL1 is a substrate of the DDR kinases ATM, ATR, and DNA-PK and is phosphorylated by all three kinases in cells after DNA damage. SMARCAL1 travels with at least some replication forks in untreated cells and accumulates at stalled replication forks when cells are treated with replication inhibitors (54,58). Importantly, SMARCAL1 deficiencies cause fork restart defects after replication stress, sensitivity to replication stress agents, and aberrant fork structures that result in DNA double-strand breaks (DSB) (54,58,59). Recent biochemical evidence suggests that SMARCAL1 can catalyze fork regression and remodel a variety of fork-like structures; activities that can both potentially promote replication restart and prevent replication-associated DNA double strand breaks *in vivo* (54).

The functional consequences of SMARCAL1 phosphorylation are unknown. We hypothesized that SMARCAL1 is targeted by DDR kinases to promote replication fork stability. We determined that SMARCAL1 is phosphorylated at three major sites following DNA damage. Importantly, SMARCAL1 localization to stalled forks and subsequent DNA binding are required for phosphorylation. DNA binding stimulates SMARCAL1 ATPase activity and SMARCAL1 fork remodeling activities. Our data suggest that DDR-dependent phosphorylation regulates activated SMARCAL1 at replication forks. Furthermore, we provide functional evidence to support that phosphorylation in the ATPase hinge domain of SMARCAL1 limits the activities of SMARCAL1 at stalled replication forks.

Results

DDR-regulated phosphorylation of SMARCAL1 requires DNA binding at stalled replication forks

SMARCAL1 exhibits a phosphorylation-dependent gel mobility shift when cells are treated with the replication stress agent hydroxyurea (HU). Though less complex than the endogenous SMARCAL1 shift, we were able to detect an HU-induced phospho-shift of GFP-SMARCAL1 when we use conditions that express low levels of protein (Figure 5.1A). SMARCAL1 localizes to HU-stalled replication forks through an interaction with RPA. To determine whether SMARCAL1 was phosphorylated directly at stalled forks, we examined the phospho-shift of the ΔN mutant, which cannot bind RPA or localize to stalled replication forks (58). The ΔN mutant does not exhibit a detectable phospho-shift (Figure 5.1B, lane 6) suggesting that DDR-regulated phosphorylation depends on SMARCAL1 localization to forks.

The RPA interaction likely functions to concentrate SMARCAL1 near its *in vivo* substrates, replication fork structures. We expect that once localized SMARCAL1 binds DNA to promote fork restart. SMARCAL1 has two HARP domains that are critical for SMARCAL1 activities in cells (62). We recently defined the HARP domains as DNA binding domains in SMARCAL1, with HARP2 being essential for DNA binding (54). We analyzed the shift pattern of HARP mutants. Mutants lacking the entire HARP1 domain or containing two point mutations in conserved HARP residues (ΔH1 and H1-WF) exhibit a decreased shift (Figure 5.1B, lanes 8,10). Similar mutants of HARP2 (ΔH2 and

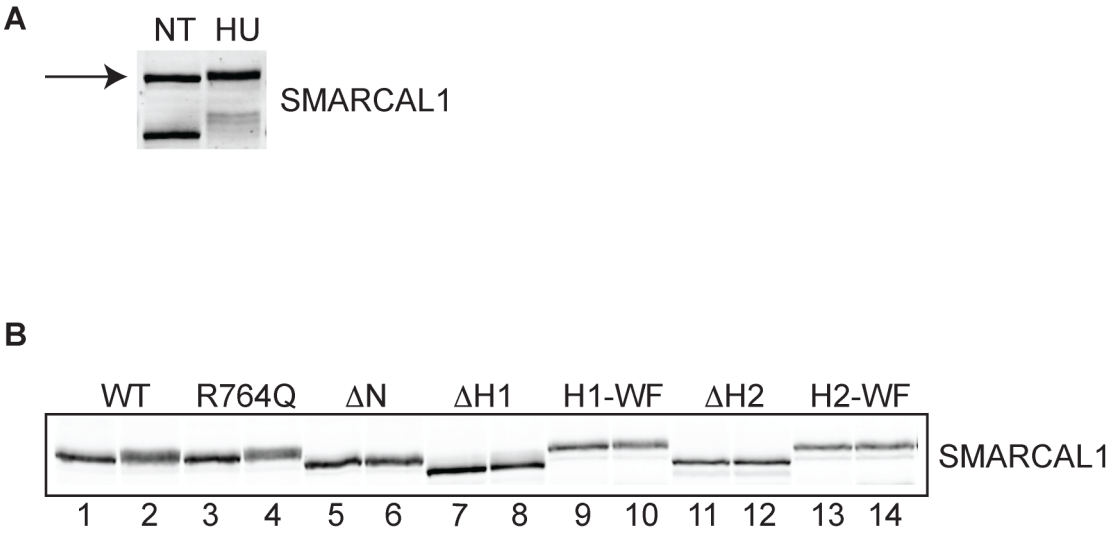

Figure 5.1. The SMARCAL1 DDR-regulated phospho-shift depends on localization and DNA binding. (A) HEK293T cells were transfected with expression vectors (pLL5.0-GFP-GW backbone for low expression) encoding GFP-tagged wild-type SMARCAL1. Cell lysates from transfected cells treated with 2mM HU for 16hr were separated by SDS-PAGE and immunoblotted with SMARCAL1 antibodies. The arrow marks the GFP-SMARCAL1 protein, which is larger and runs above endogenous SMARCAL1, lower band. (NT) No treatment. (B) Using the same conditions as in (A) cells were transfected with various GFP-SMARCAL1 mutants. Image was cropped to show only GFP-SMARCAL1 bands. Odd lanes (NT), even lanes (HU).

H2-WF) resemble the ΔN mutant and exhibit no HU-induced shift (Figure 5.1B, lanes 12,14).

These observations correlate well with our HARP DNA binding studies. HARP1 mutants retain some DNA binding while HARP2 mutants are deficient in DNA binding (54). Since the HARP2 mutants are also inactive ATPases, it is possible that SMARCAL1 ATPase function affects whether or not SMARCAL1 is a substrate for DDR kinases. However, the ATPase-dead mutant R764Q, which retains DNA binding activity, shifts similar to WT protein (Figure 5.1B, compare lanes 2 and 4). Importantly, all mutants except ΔN co-localize with RPA foci at stalled replication forks (Figure 5.2). Together these results suggest that SMARCAL1 DDR-regulated phosphorylation is dependent on localization to stalled replication forks and subsequent DNA binding. Consistent with our results, Postow and colleagues observed an increase in the caffeine-sensitive phosphorylation of frog SMARCAL1 when it was bound to DNA (118). This also implies that DDR-dependent regulation of SMARCAL1 is conserved.

Identifying DDR-regulated phosphorylation sites in SMARCAL1 by phosphopeptide mapping

Identifying phosphorylation sites in SMARCAL1 is important for establishing the functional significance of SMARCAL1 phospho-regulation following DNA binding at stalled replication forks. We initially analyzed SMARCAL1 purified from HU-treated cells by mass spectrometry (see Table 6.1 for results). We detected phosphorylation on two DDR kinase consensus sites,

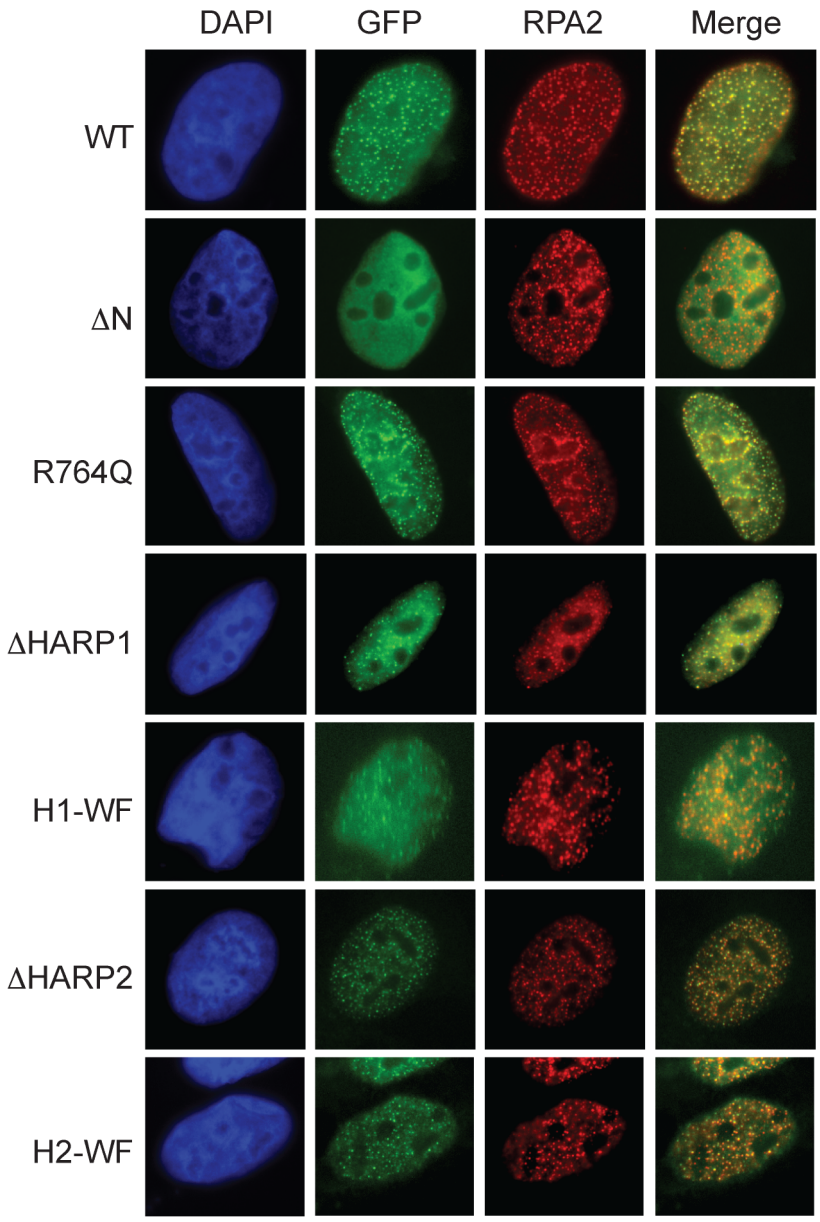

Figure 5.2. HARP mutants co-localize with stalled replication forks. U2OS cells were transfected with pLL5.0-GFP-SMARCAL1 vectors, either WT or mutant. Two days later cells were treated with HU for 5 h, then fixed and stained with RPA2 antibodies and DAPI.

S173 and S652. We had no indication from this analysis to what extent SMARCAL1 was phosphorylated in response to damage. Our phospho-shift studies suggest that SMARCAL1 is likely phosphorylated on multiple sites since the banding pattern on SDS-PAGE gels is complex.

To better characterize the full extent to which SMARCAL1 is phosphorylated after DNA damage, we examined SMARCAL1 phosphopeptides obtained from ^{32}P-orthophosphate metabolically labeled cells. Phosphopeptide maps of SMARCAL1 show differences in phosphorylation before and after 16hrs in HU (Figure 5.3A). When comparing the no treatment and 16hr HU map, we note the appearance of phosphopeptide-*a, -d,* and-*e,* along with increases in -*b* and -*c*. Phosphopeptide-*1* and -*2* represent major sites of phosphorylation in SMARCAL1 during conditions of normal cell growth. As discussed below, DDR-regulated phosphorylation is likely responsible for the change in mobility observed for phosphopeptide-*1* and -*2* (denoted by larger circle) in the HU map.

Because we are interested in identifying DDR-regulated phosphorylation, we determined whether any of the HU-specific phosphopeptides were sensitive to DDR kinase inhibition. Cells were exposed to HU for 16hrs. DMSO or DDRi (a mix of individual specific inhibitors for ATM, ATR, and DNA-PK) was added to cells during the ^{32}P metabolic labeling step. Inhibition of all three DDR kinases led to the loss of phosphopeptide-*a* and -*d* and somewhat restored the mobility of phosphopeptides-*1* and -*2* (Figure 5.3B). Phosphopeptide-*e* was not detectable in the control so we could not determine the DDR dependence of its

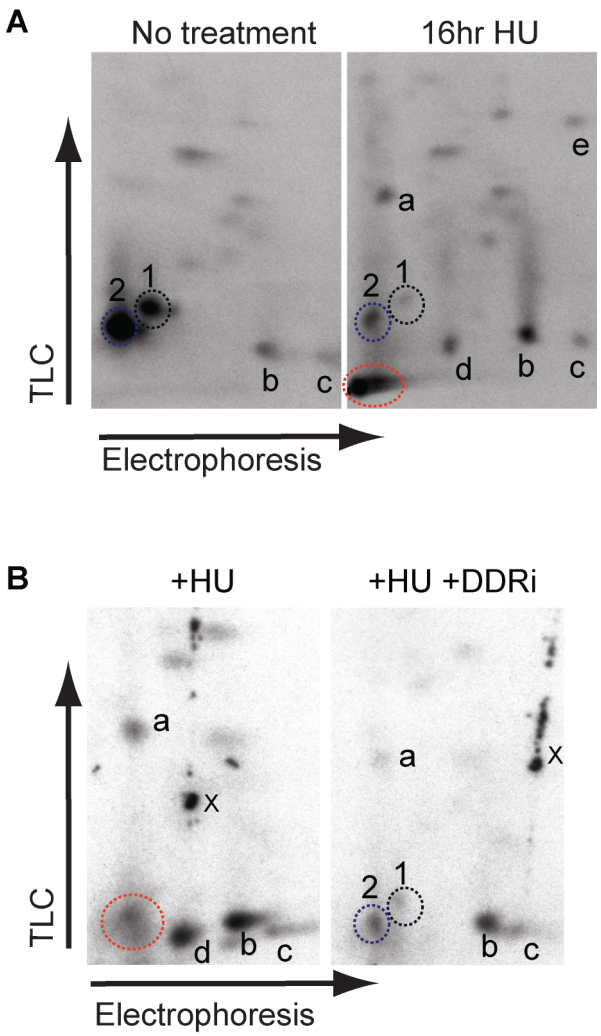

Figure 5.3. HU-induced phosphorylation of SMARCAL1 depends on DDR kinases. (A,B) HEK293T cells were treated with HU for 16hr and then metabolically labeled in the presence of HU with $^{32}P_i$ for 30min. SMARCAL1 immunoprecipitations were prepared, separated by SDS-PAGE, and transferred to PVDF membranes. SMARCAL1 bands incorporating ^{32}P were excised from membranes and trypsin-digested to isolate phosphopeptides. Equal counts of ^{32}P were spotted on chromatography plates and subjected to two-dimensional thin layer chromatography analysis. (A) Phosphopeptides that change between no treatment and 16hr HU treatment are labeled [1,2, a, b, c, d, e]. A change in the mobility of 1 and 2 is marked with a red circle. (B) DMSO or a mixture of kinase inhibitors for ATR, ATM, and DNA-PK (DDRi) was added for 30min prior to labeling and was present during the labeling for (+HU) and (+HU +DDRi) respectively. (x) marks spots on the film caused by condensation and do not represent phosphopeptides.

phosphorylation. These results demonstrate that the majority of HU-induced phosphopeptides that we can detect by mapping depend on DDR kinase activity.

Mapping SMARCAL1 mutants to identify DDR-regulated phosphopeptides

We next determined whether phosphopeptide maps of transiently expressed GFP-SMARCAL1 protein were similar to endogenous SMARCAL1 maps. One concern with transient expression of exogenous protein is that the overexpressed protein will not be regulated similarly. To address this issue, we used transfection conditions that resulted in very low levels of GFP-SMARCAL1 expression. These conditions were optimized for analyzing the phospho-shift of GFP-SMARCAL1 mutants (Figure 5.1). Even though the phospho-shift for GFP-SMARCAL1 is not as robust as the endogenous protein shift, we observed a similar pattern of phosphopeptides in GFP-SMARCAL1 phosphopeptide maps indicating the GFP protein is regulated similarly (Figure 5.4). We assigned the same phosphopeptides (*1,2; a-d*) with exception of phosphopeptide-*e* that was not detectable. We also note the appearance of phosphopeptide-*3,* which is likely related to the mobility change of phosphopeptide-*1* and -*2* observed in the endogenous HU-map.

Overexpression of SMARCAL1 activates a DNA damage response (58). We attribute the increases of phosphopeptide-*a*, -*b*, and -*c* in the untreated GFP-SMARCAL1 maps to low levels of DDR activation due to SMARCAL1 overexpression. We observe an increase in these phosphopeptides after HU-

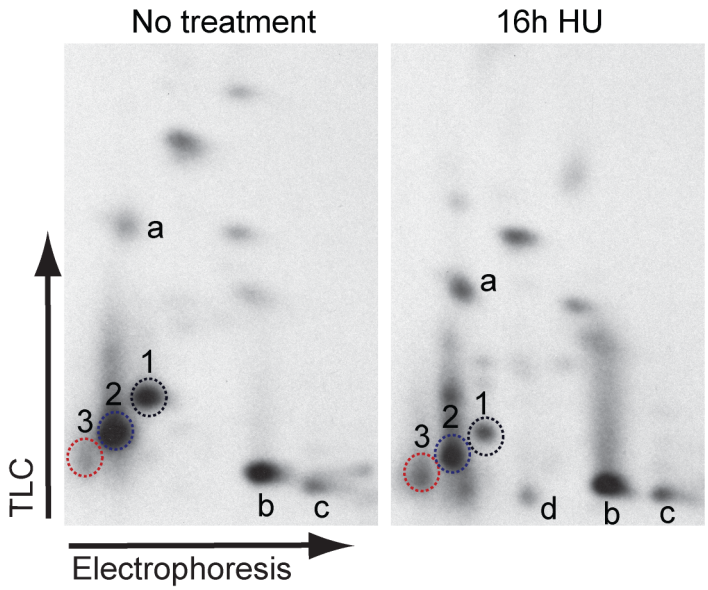

Figure 5.4. GFP-SMARCAL1 is regulated similarly to endogenous SMARCAL1. HEK293T cells were transfected with expression vectors (pLL5.0-GFP-GW backbone for low expression) encoding GFP-tagged wild-type SMARCAL1. The day after transfection cells were passaged and HU was added for 16hrs. Cells were metabolically labeled and peptides processed as in (Figure 5.3). GFP-SMARCAL1 was purified using SMARCAL1 antibodies. Samples were separate on a lower percentage SDS-PAGE to resolve the GFP-SMARCAL1 band from the endogenous SMARCAL1 band. GFP-SMARCAL1 bands were excised instead of endogenous SMARCAL1. DDR-regulated phosphopeptides are marked [1-3, a-d].

treatment indicating that more molecules of SMARCAL1 are phosphorylated at these sites following robust activation of the DDR. Similar to the endogenous HU map, phosphopeptide-*d* appears following HU treatment. The altered mobility of GFP-SMARCAL1 phosphopeptide-*1* and -*2* is less drastic following HU. Instead, we observe the appearance of phosphopeptide-*3* in the GFP-SMARCAL1 HU map. The interpretation of this phosphopeptide pattern is discussed below.

Since we were able to identify DDR-induced phosphopeptides in GFP-SMARCAL1, we used GFP-SMARCAL1 phospho-mutants to elucidate the identity of specific phosphorylation sites. SMARCAL1 contains 10 S/TQ DDR kinase consensus sites. Initially, we mutated 6 of these sites (S115, S173, S350, S652, T742, S919) to alanine. The phospho-shift of this 6A mutant is considerably reduced compared to WT (Figure 5.5A). This mutant retains the ability to localize to stalled forks, so the 6 alanine mutations do not affect RPA binding (Figure 5.5B). Consistent with the reduction in phosphorylation by phospho-shift analysis, the 6A phosphopeptide map reveals loss of the most abundant DDR-regulated phosphopeptides. Phosphopeptide-*3* and -*d* are absent along with substantial decreases in -*b* and -*c* (Figure 5.6A, compare panels 1 and 2). To determine the identity of phosphopeptide-*3*, -*b, -c, and -d* we analyzed single mutants for each site in the 6A mutant. A single S173A mutant resulted in the loss of phosphopeptide-*d* and the decrease in -*b*, and -*c* (Figure 5.6A, panel 3). Phosphopeptide-*3* was absent in the S919A map (Figure 5.6A, panel 4). S115A, S350A, T742A, and S652A maps showed no detectable changes in phosphopeptides compared to WT (Figure 5.6B and C).

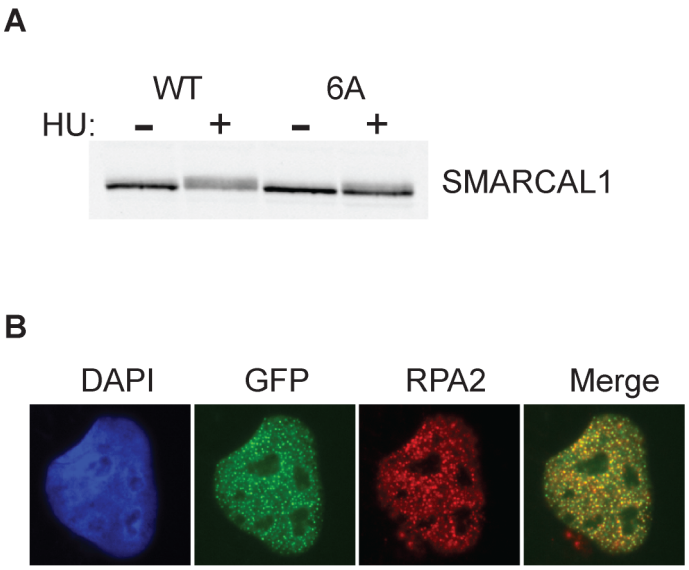

Figure 5.5. A 6A phospho-mutant reduces SMARCAL1 phosphorylation but does not affect localization to stalled forks. (A) HEK293T cells were transfected with pLL5.0-GFP-WT or 6A mutants and analyzed for phosphoshift after 16hr in 2mM HU. (B) U2OS cells were transfected with the same expression vectors. Two days later, cells were treated with HU for 5 h, then fixed and stained with RPA2 antibodies and DAPI.

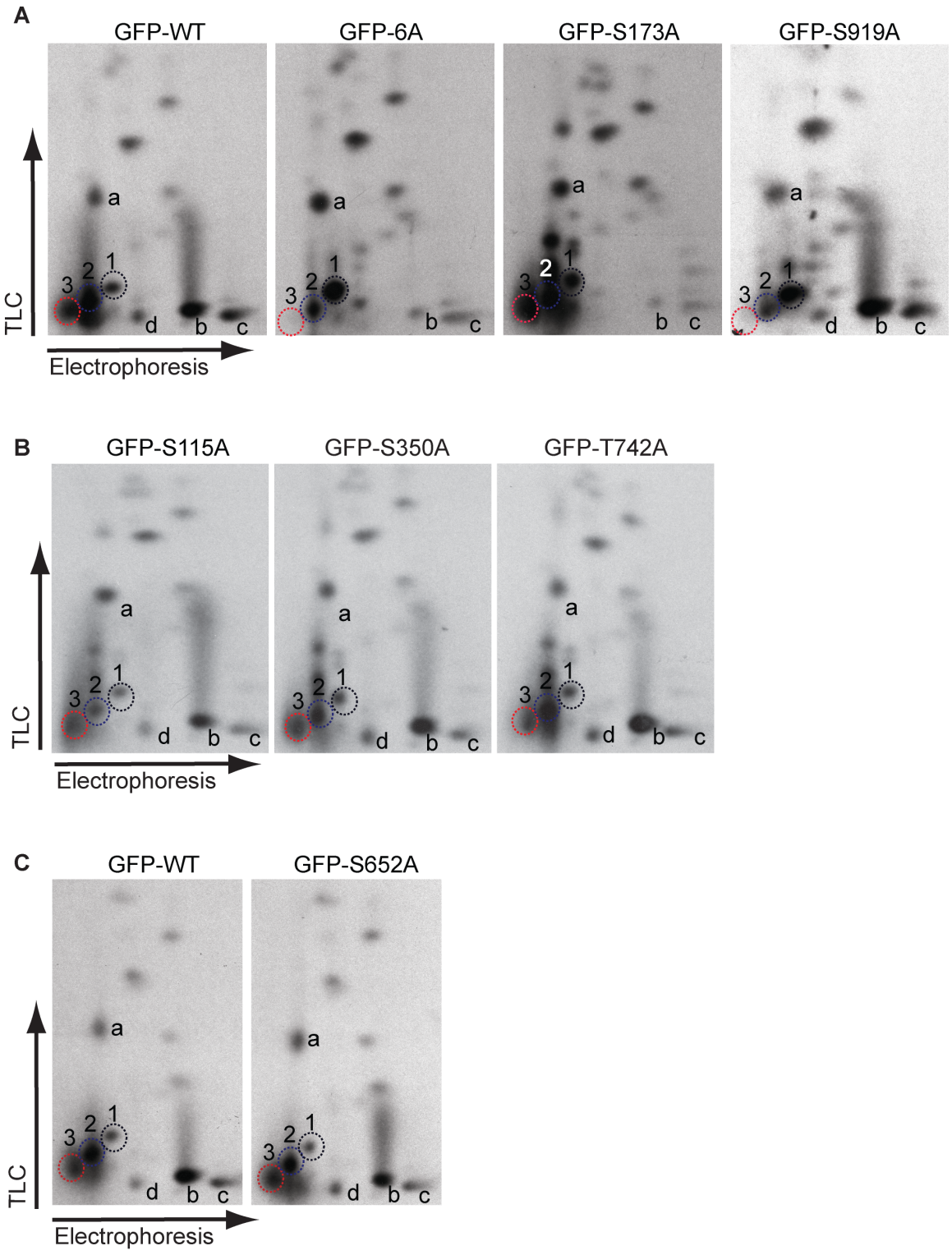

Figure 5.6. S173 and S919 are major DDR-regulated phosphorylation sites in SMARCAL1. All maps were prepared as in Figure 5.4 from cells treated with HU for 16hrs. DDR-regulated peptides are marked.

S173 is a major site of DDR-regulated phosphorylation

We further investigated why the S173A mutant affected three phosphopeptides. Trypsin cleavage occurs at specific sites in a peptide sequence, preferentially at Arg and Lys that are not followed by Pro. Using the ExPASy peptide cutter tool we generated a list of SMARCAL1 tryptic peptides. The S173 peptide generated by this tool (Figure 5.7A, peptide (ii)) has the potential to be inefficiently cleaved. Phosphorylation near trypsin cleavage sites has been demonstrated to reduce the efficiency of cleavage (Benore-Parsons 1989). Additionally, trypsin cleavage is inefficient at Arg followed by Asp. Based on these characteristics of trypsin cleavage and the sequence surrounding tryptic peptide (ii), two additional phosphopeptides containing S173 could result from inefficient trypsin cleavages at the N- and C-terminal ends (Figure 5.7A, peptide (i) and (iii)). These peptides vary in size, charge, and hydrophobicity and would have distinct mobilities in both dimensions of separation. Although we can speculate on which products correlate with phosphopeptide-*b*, -*c*, and -*d*, the important observation is that S173 is present in each phosphopeptide and represents the major site of DDR-regulated phosphorylation.

In our mass spectrometry investigations it was unclear whether S172 or S173 was the phosphorylated residue. We did identify a peptide that was clearly phosphorylated at S173, but did not rule out the possibility that S172 could be phosphorylated as well. Additionally, a phosphopeptide containing both S172 and S181 phosphorylations was identified among 494 other unique phosphopeptides in a massive mass spectrometry study that profiled phosphorylation events in

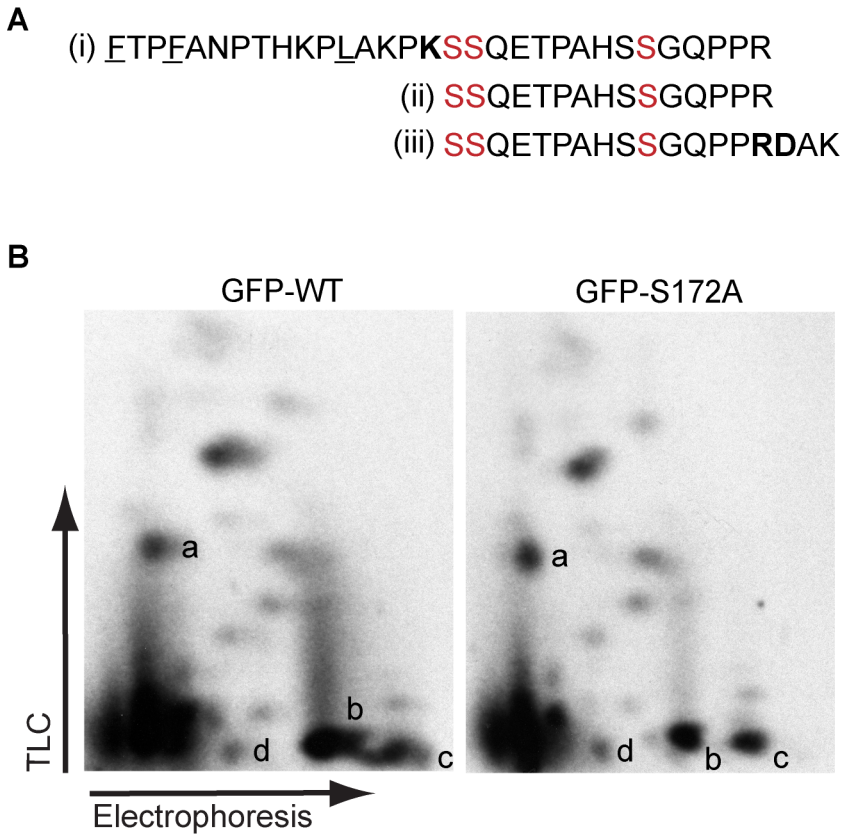

Figure 5.7. Phosphopeptide-*b*,-*c*, and -*d* are tryptic products of a S173-containing phosphopeptide. (A) Possible tryptic cleavage products surrounding the S173 peptide. Serines in red denote sites of phosphorylation (S172, S173, and S181 from left to right). (i) Phosphorylation of S172 or S173 may cause inefficient cleavage at the K in bold. This cleavage product adds a number of hydrophobic residues (underlined). (ii) Represents the fully cleaved tryptic product. (iii) Inefficient cleavage can occur at Rs followed by D. (B) Maps were prepared as in Figure 5.4 from cells treated with HU for 16hrs.

human skin fibroblasts (119). An alternative interpretation of the S173A mutant map is that S173 is required for DDR-regulated phosphorylation of S172. The loss of phosphopeptide signal could bedue to loss of S172 phosphorylation instead of S173 directly. To address this discrepancy, we mapped the S172A mutant. The S172A map exhibits some reduction in phosphopeptide-*b* and -*c* signal compared to WT, but overall does not appreciably inhibit S173 phosphopeptides (Figure 5.7B). We conclude that phosphorylation of S173 accounts for the majority of the damage-inducibility of phosphopeptide-*b, -c,* and -*d*.

The S919 peptide is phosphorylated at a minimum of three sites

Our map of the S919A mutant revealed that phosphopeptide-*3* contains phosphorylated S919 (Figure 5.6A, panel 4). S919 is located on a large tryptic peptide containing 14 potential sites of phosphorylation (Figure 5.8A). Peptides migrate in the electrophoresis step according to size and charge. Phosphopeptide-*1, -2,* and -*3* migrate very little from the origin in the horizontal direction suggesting these peptides are large and negatively charged. Also, the mobility pattern of phosphopeptides *1-3* resembles the diagonal pattern associated with a single peptide containing multiple phosphorylation states (120). In the organic chromatography buffer, peptides migrate further based on hydrophobicity. The addition of phosphate (decreasing hydrophobicity) retards the mobility of phosphopeptides in the chromatography dimension, resulting in a downward diagonal with increasing phosphates. We therefore assigned

A

EGSDMELLEAAESFDPGSASGTSGSSSQNMGDTLDESSLTASPQK

B

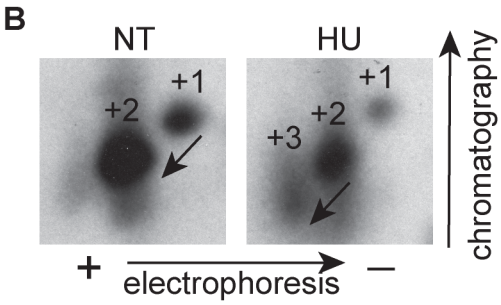

Figure 5.8. Phosphopeptides *1-3* represent different phosphorylation states of a single peptide. (A) Tryptic peptide containing S919 (underlined). Amino acids in red represent possible sites of phosphorylation. (B) Images are cropped from maps of GFP-WT protein, purified from either untreated cells (NT) or HU treated cells (HU). +1, +2, +3 mark the addition of phosphates on a single peptide. The addition of phosphates retards the mobility of the peptide as indicated by the arrow. Peptides move in the electrophoresis phase based on size and charge, more positively charged peptides moving farther right. Peptides move up in the chromatography phase based on hydrophobicity.

phosphopeptide -*1*, -*2*, and -*3* as +1, +2, and +3 phosphates on a single peptide respectively (Figure 5.8B). The appearance of phosphopeptide-*3* reflects the damage-induced phosphorylation of S919 on phosphopeptide-*2*. Further mutational analysis is required to determine what additional sites are phosphorylated in this large peptide. We speculate that these peptides likely contain DNA-damage independent phosphorylation sites since phosphopeptide-*1* and -*2* are dominant in the endogenous map of SMARCAL purified from untreated cells (Figure 5.3A).

A phospho-specific antibody identifies S652 as a DDR-regulated phosphorylation site

We identified S652 phosphorylation by mass spectrometry. However, we were not able to detect loss of any phosphopeptides in S652A phosphopeptide maps (Figure 5.6C). To characterize phosphorylation of S652 further, we generated a phospho-specific antibody to a peptide phosphorylated at S652. The pS652 antibody recognizes FLAG-SMARCAL1 on immunoblots following immunoprecipitation (Figure 5.9). The antibody most strongly recognizes WT protein purified from cells treated with HU (Figure 5.9A, lane 5), and the HU-induced signal is phosphatase sensitive. The detection of WT protein purified from untreated cells and of S652A protein purified from HU treated cells is relatively similar (compare lanes 1 and 3). This suggests that the antibody weakly detects a non-pS652 epitope in addition to pS652. The recognition of this epitope is partly sensitive to phosphatase but does not change in response to DNA damage. Although this antibody is not entirely specific to pS652, we were

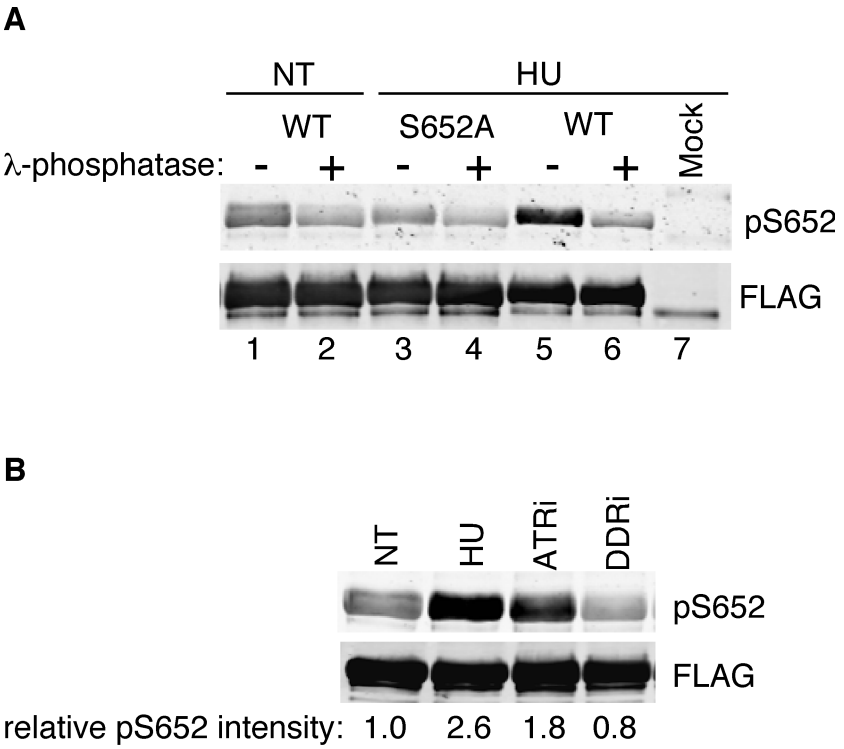

Figure 5.9. A phospho-specific S652 antibody detects DDRi-sensitive increases in S652 phosphorylation in response to HU. HEK293T cells were transfected with FLAG expression vectors containing WT or S652A SMARCAL1. FLAG immuno-preciptations using FLAG-beads were prepared, separated by SDS-PAGE, and transferred to membranes for immunoblotting with pS652 and FLAG antibodies. (A) Following FLAG purifications, beads were incubated in the presence of l-phosphatase to remove phosphates for 30min before separating by SDS-PAGE. (B) Cells were treated with HU for 6hrs with the addition of either DMSO (HU), ATR inhibitor (ATRi, vertex, 5μM), or DDRi. Intensity of pS652 and FLAG signal were determined by Odyssey quantitation and relative pS652 intensity was calculated by normalizing the pS652 intensity according to protein levels for each sample, i.e. FLAG intensity. The corrected pS652 intensity was then used to calculate the pS652 intensity relative to the NT sample.

confident that the bulk of signal detected in SMARCAL1 following HU is due to phosphorylation at S652. We next determined whether S652 phosphorylation depends on DDR kinases. Using the ATR inhibitor (ATRi) was not sufficient to reduce pS652 in SMARCAL1 purified from HU-treated cells (Figure 5.9B). However, inhibition of all three kinases with DDRi completely reduced pS652 to untreated levels. Thus, S652 is another DDR-regulated phosphorylation site in SMARCAL1.

A S652 phospho-mimetic mutant has decreased ATPase activity

From our mass spectrometry and phosphopeptide mapping studies, we conclude that the majority of SMARCAL1 DDR-regulated phosphorylation occurs at 3-SQ consensus sites: S173, S652, and S919. To determine the functional consequences of phosphorylation at these sites we generated mutants that mimicked constitutive phosphorylation (S to D) to analyze alongside of the (S to A) mutants we used in our mapping experiments. We first confirmed that all phosphorylation mutants localize to stalled replication forks (Figure 5.10). All mutants are capable of co-localizing with RPA at stalled replication forks.

High levels of SMARCAL1 cause replication-associated DNA damage response activation, primarily in the form of pan-nuclear H2AX phosphorylation (pan-γH2AX). Both RPA binding and enzymatic activities of SMARCAL are required to efficiently induce pan-γH2AX (58). The negative effects of SMARCAL1 overexpression suggest that regulating SMARCAL1 activity at replication forks is necessary for genome maintenance. Left unchecked,

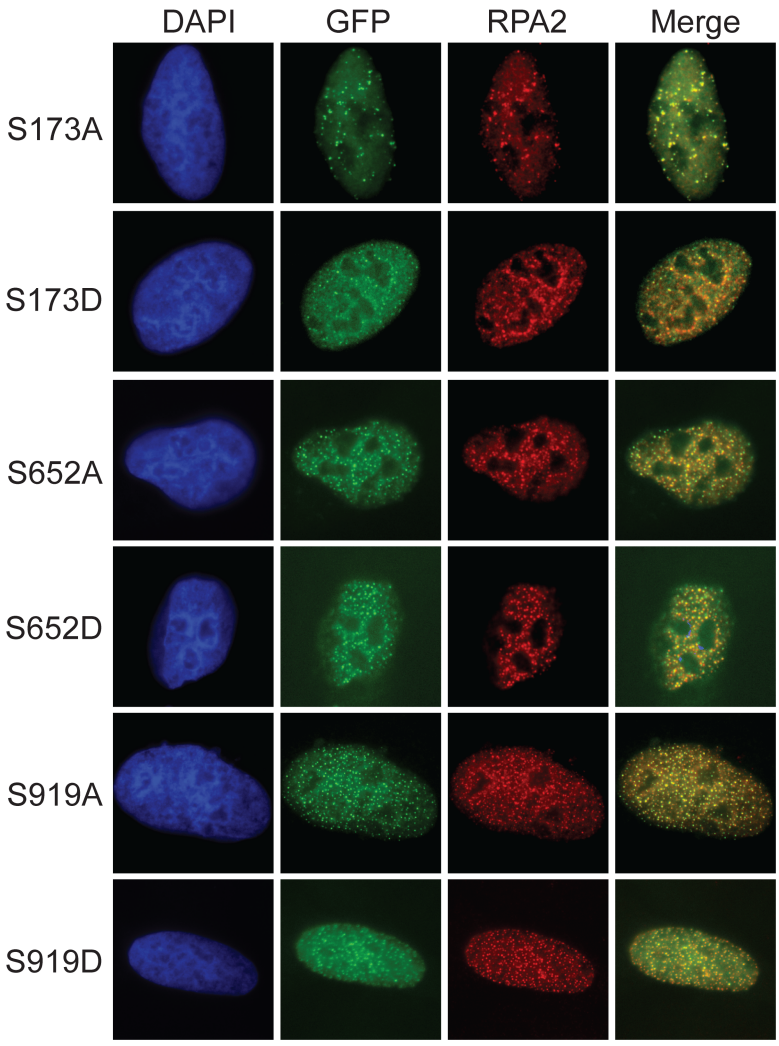

Figure 5.10. Phospho-mutants localize to stalled replication forks. U2OS cells were transfected with pLL5.0-GFP-SMARCAL1 vectors, either WT or mutant. Two days later cells were treated with HU for 5 h, then fixed and stained with RPA2 antibodies and DAPI.

SMARCAL1 annealing activity could potentially reverse every active replication fork. Since we determined that the phospho-mutants do not affect RPA-dependent localization, we hypothesized that another possible function of DDR-regulated phosphorylation could be to limit SMARCAL1 enzymatic activity at stalled replication forks. We purified phospho-mutants from cells and tested the fork-stimulated ATPase activity of each (Figure 5.11). S173A/D and S919A/D mutants hydrolyzed ATP similar to WT (Figure 5.11A and B). However, the S652D mutant exhibited a significant decrease in fork-stimulated ATPase activity (Figure 5.11C). The S652A mutant retained WT activity suggesting that the serine itself is not critical for ATP hydrolysis.

The fork-stimulated ATPase assay measures the ability of SMARCAL1 to bind a simple forked-DNA molecule and hydrolyze ATP. In cells, the substrate at a stalled replication fork is more complex and any remodeling that SMARCAL1 promotes would require removing other DNA-binding proteins at the fork. Thus, remodeling replication forks in cells likely requires more efficient enzymatic activity than would be required in vitro. The pan-γH2AX caused by SMARCAL1 overexpression reflects high levels of SMARCAL1 annealing helicase activity at replication forks. We reasoned that we could measure the activity of phospho-mutants in cells by determining the efficiency of pan-γH2AX induction. If S652 is important for regulating SMARCAL1 activity at replication forks we expect that a constitutively limited SMARCAL1 (i.e. S652D) will also be limited in causing pan-γH2AX when overexpressed.

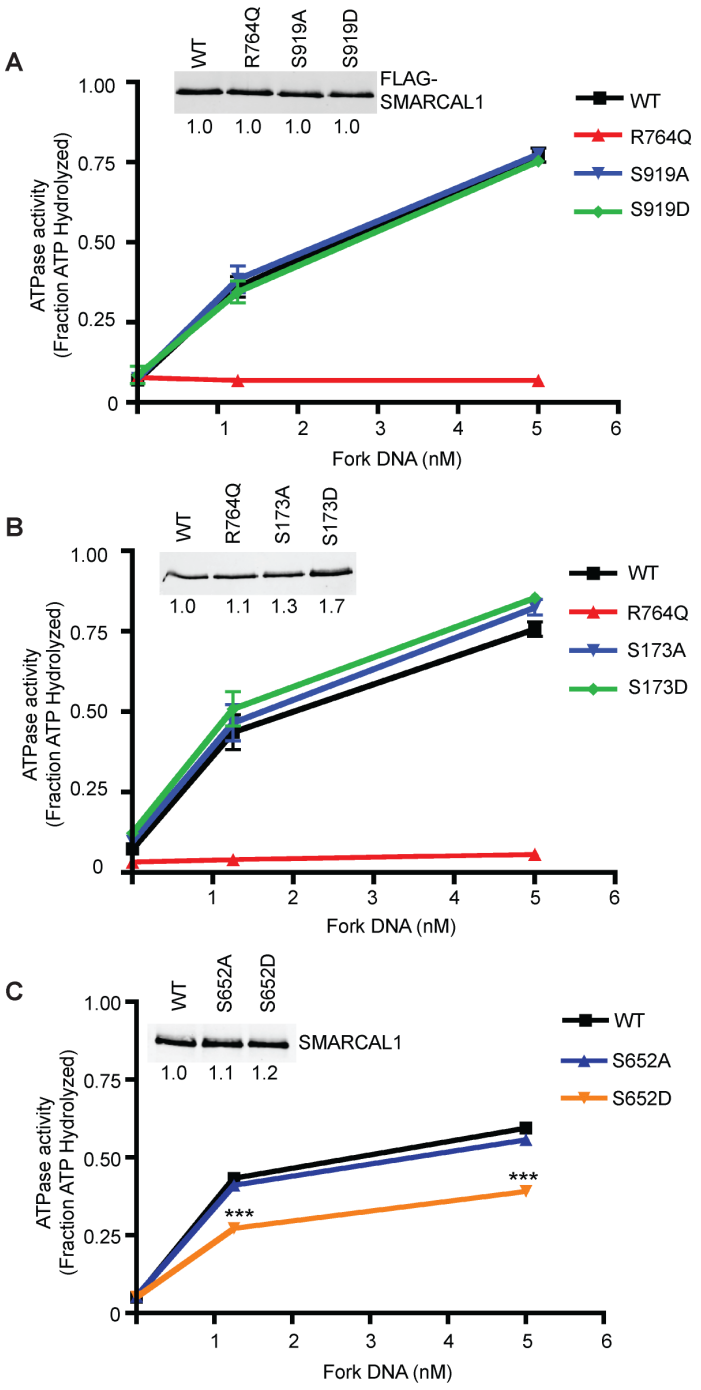

Figure 5.11. The S652D mutant has decreased ATPase activity *in vitro*. Protein was purified from HEK293T cells that were transiently overexpressing FLAG-SMARCAL1. ATPase activity of the WT and phospho-mutants was measured in the presence of the indicated concentrations of fork DNA. Error bars are standard deviation (n=3). Quantification of FLAG-SMARCAL1 by WB represents an average of protein input across replicates. (C) (***) *P* values for 1.25nM DNA and 5nM DNA were P<0.0001 and *P*=0.0002 respectively. Two-tailed unpaired t-test.

We developed our previous qualitative overexpression assay (used in Figure 4.9) to be highly quantitative using the Opera Imaging platform and Columbus software. SMARCAL1 expression levels largely determine the phenotype in this assay. Our new methodology allows us to compare hundreds of individual nuclei across all samples based on their GFP-SMARCAL1 intensities. To determine percent DDR activation, we first defined an upper and lower cutoff for GFP-SMARCAL1 intensity. The R764Q mutant can activate the DDR but requires much higher levels than WT. Before analyzing the entire data set, we sorted nuclei from WT and R764Q control samples based on γH2AX intensity and determined the average GFP intensities for nuclei that had DDR activation (γH2AX intensity >1000) (Figure 5.12A). This initial analysis demonstrates that the average GFP-R764Q expression levels required for DDR activation are higher than for GFP-WT. Additionally, there were not considerable differences in the number of nuclei imaged, so the decreased number of nuclei observed is a consequence of R764Q decreased activity. We set the upper GFP intensity limit at the average intensity for GFP-WT cells with DDR activation.

For this analysis we analyzed nuclei with GFP intensities between 500 and 2500. Percent DDR activation is calculated from the percent of these nuclei with pan-nuclear γH2AX. When we apply these cutoffs across all samples, we observe significant decreases in the abilities of the R764Q mutant and S652D mutant to induce DDR activation (Figure 5.12B). Similar to our in vitro results, the S652D mutant retains some enzymatic activity but exhibits reduced activity

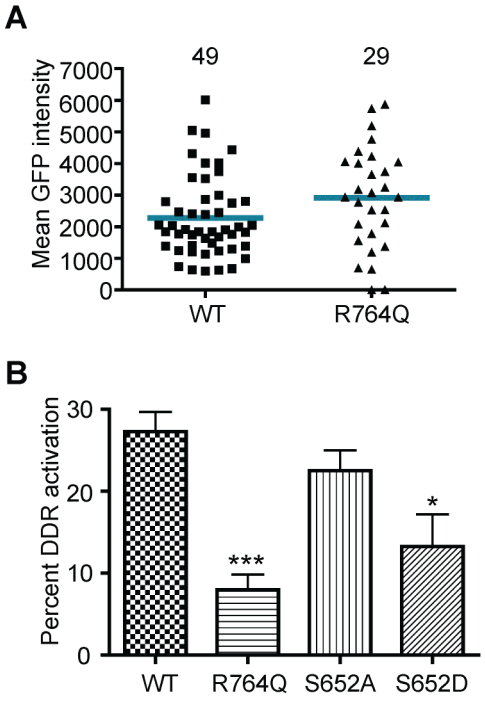

Figure 5.12. A quantitative assay for DDR activation following SMARCAL1 over-expression demonstrates that the S652D mutant is less active in cells. U2OS were transfected with high expressing GFP-SMARCAL1 vectors. Cells were plated into 96-well plates for imaging by the Opera, 3 wells per sample. 48hrs after transfection cells were fixed and stained for γH2AX and appropriate secondary. DAPI was added for detecting nuclei. 12 subfields were analyzed per well. Images collected were analyzed for mean γH2AX intensity and mean GFP intensity per nuclei by the Columbus analysis software. (A) The GFP intensities for individual nuclei (n=49 and n=29) that had mean γH2AX intensities greater than 1000. Data is from one well and was used to set a GFP cutoff. Bar represents average intensity. (B) Percent pan-γH2AX for S652 phospho-mutants using the cutoff defined in (A). (***) *P*=0.0007. (*) *P*=0.0231. Two-tailed unpaired t-test.

when compared to WT. Again, the S652A mutant does not exhibit any significant differences in activity.

Endogenous SMARCAL1 purified from HU-treated cells is less active

Our S652 mutant studies suggest that DDR-regulated phosphorylation may be a mechanism for limiting SMARCAL1 enzymatic activity at replication forks. To test whether phosphorylated SMARCAL1 is indeed less active, we purified endogenous SMARCAL1 from untreated cells or cells treated with HU for 16 hours and measured its ATPase activity. We selected a dose and time for HU treatment in which we observed 100% of SMARCAL1 in the shifted form (Figure 5.1A). We found that SMARCAL1 purified from HU treated cells is less active, exhibiting nearly a two-fold decrease in ATPase activity (Figure 5.13). This indicates that SMARCAL1 is less active following long exposures to replication stress. Although the role of phosphorylation in this observation is circumstantial, our evidence for S652 phosphorylation suggests that this decrease in activity is mostly due to the DDR-dependent phosphorylation of S652.

Discussion

We set out to identify DDR-regulated phosphorylation sites in SMARCAL1 and found that the majority of DDR-dependent phosphorylation events occur at S173, S652, and S919. The identities of two additional DDR-regulated phosphopeptides have yet to be determined. Importantly, we discovered that

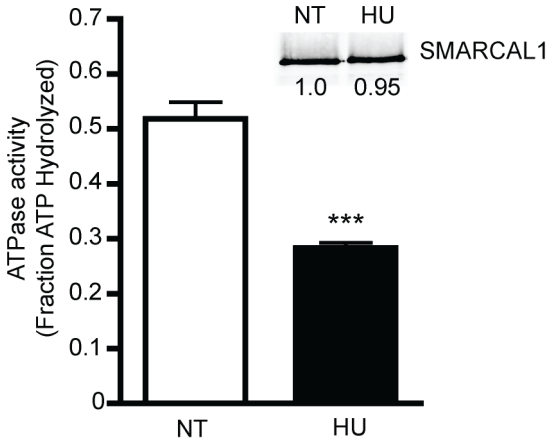

Figure 5.13. Endogenous SMARCAL1 purified from HU-treated cells exhibits decreased ATPase activity. Endogenous SMARCAL1 was purified from HEK293T cells that were untreated (NT) or treated with HU for 16hrs. ATPase activity of SMARCAL1 from NT and HU conditions was measured in the presence of 5nM fork DNA. Error bars are standard deviation (n=3). Quantification of SMARCAL1 by WB represents an average of protein input across replicates. Purified protein for WB was treated with phosphatase before input to collapse the shifted band for more accurate quantification (***) *P*=0.0007. Two-tailed unpaired t-test.

DDR-regulated phosphorylation has functional consequences. Evidence from a SMARCAL1 mutant mimicking constitutive phosphorylation at S652 suggests that S652 may be targeted by DDR kinases to limit SMARCAL1 enzymatic activity at stalled forks.

Unidentified phosphopeptides

The 6A phospho-mutant led to the identification of two major DDR regulated phosphorylation sites in SMARCAL1. We failed to identify phosphopeptide-*a*, which decreased after cells were treated with DDR kinase inhibitors and was detectable in GFP-SMARCAL1 maps (Figure 5.3B and 5.4). To determine whether phosphopeptide-*a* contained one of the remaining 4 S/TQs in SMARCAL1, we generated and mapped single mutants for each (S55A, S236A, S407A, T592A). However, none of these mutations led to the disappearance of any additional phosphopeptides (Figure 5.14). Phosphopeptide-*a* may be phosphorylated by DDR kinases at a non-consensus serine or threonine. Alternatively, it contains a site that is phosphorylated by another kinase whose activity is dependent on DDR kinase activity.

Phosphopeptide-*e* was an HU-induced peptide not detectable in the GFP-SMARCAL1 maps (Figure 5.4). We identified S652 phosphorylation by mass spectrometry and confirmed it as a DDR-regulated site using a pS652 antibody. However, we were not able to detect loss of any phosphopeptides in S652A phosphopeptide maps (Figure 5.6C). It is possible that phosphopetide-*e* is a peptide containing p-S652. To validate the identity of phosphopeptide-*e*, we have

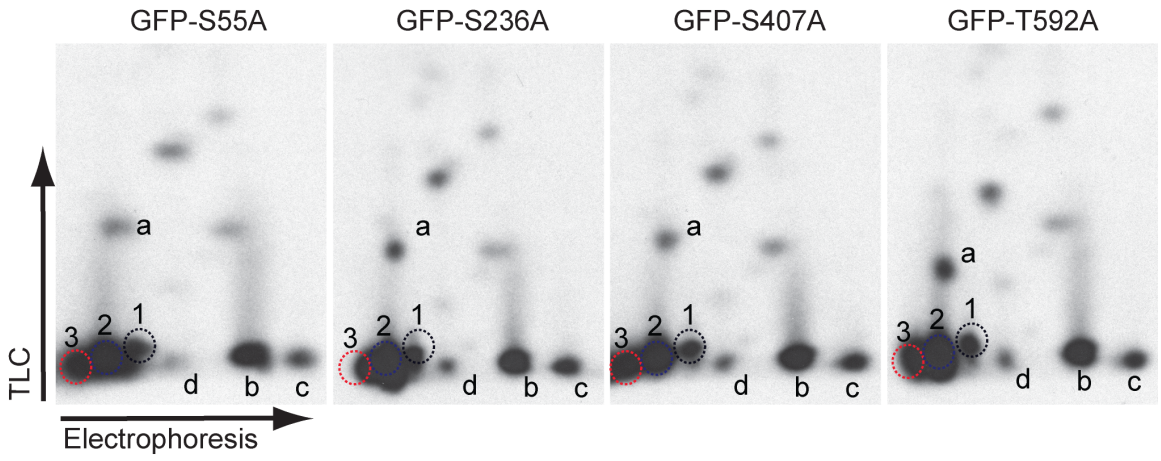

Figure 5.14. Mutating the 4 remaining S/TQ sites in SMARCAL1 does not affect DDR-regulated phosphopeptides. Maps were prepared as in Figure 5.4. from cells treated with HU for 16hrs.

synthesized the tryptic peptide and plan to phosphorylate it *in vitro* with DDR kinases in the presence of gamma-^{32}P-ATP. Phosphorylated S652 peptides will be spotted on plates with SMARCAL1 phosphopeptides purified from HU-treated cells to determine if the two phosphopeptides overlap.

A model for limiting SMARCAL1 activity at stalled replication forks.

S652 resides in a highly conserved hinge domain located between the ATPase_N and ATPase_C domains of SMARCAL1 (Figure 5.15A). Based on a crystal structure of the SNF2 protein SsoRad54 bound to DNA, these domains likely undergo rotational changes that position them in a conformation required for ATP binding and hydrolysis (121-123). This conformational change involves a 180° rotation of the ATPase_C domain to reposition necessary active site motifs from an “open-” inactive orientation, to a “closed-” active orientation (Figure 5.15B). Nucleotide binding is thought to initiate the switch from open to closed (124). The hinge domain likely provides flexibility for the conformational change that occurs between the ATPase_N and ATPase_C domains. In support of this model, a mutation in the hinge of the yeast SNF2 protein Mot1 abolishes ATPase activity (122). The authors propose that this mutation impairs the flexibility of the linker and inhibits the transition from the open to closed conformation. We hypothesize that alterations affecting flexibility of the hinge domain result in loss of SMARCAL1 enzymatic functions. Thus, phosphorylation of S652 has the potential to disrupt critical structural components of the enzymatic core and render SMARCAL1 a less active enzyme.

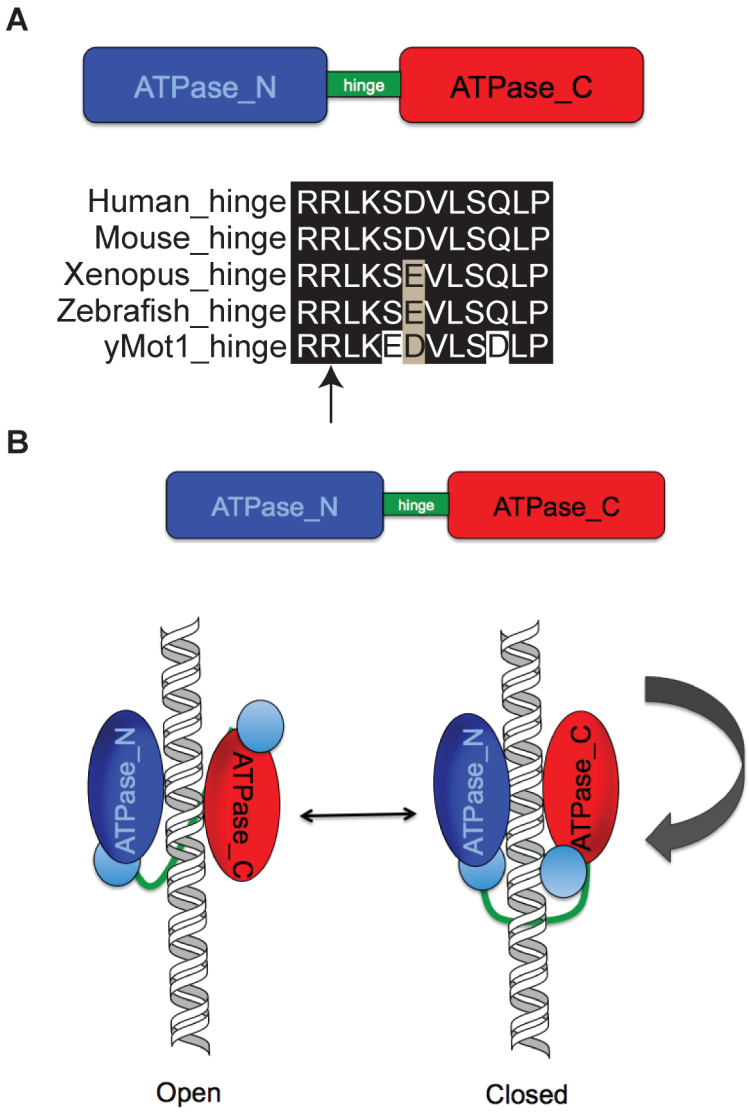

Figure 5.15. A conserved hinge domain in SMARCAL1 and other SNF2 family members provides the flexibility for rotation of ATPase_N and ATPase_C domains. (A) A hinge domain is located between the ATPase_N and ATPase_C domains and is highly conserved in various species of SMARCAL1. The hinge domain of yMot1 is almost identical to SMARCAL1. Arrow marks the site of the R1507K mutation in yMot1 that abolishes ATPase activity. (B) Model for the conformational change that positions active site motifs (light blue spheres) to the closed-active state. Arrow shows direction of 180° rotation (adapted from (122)).

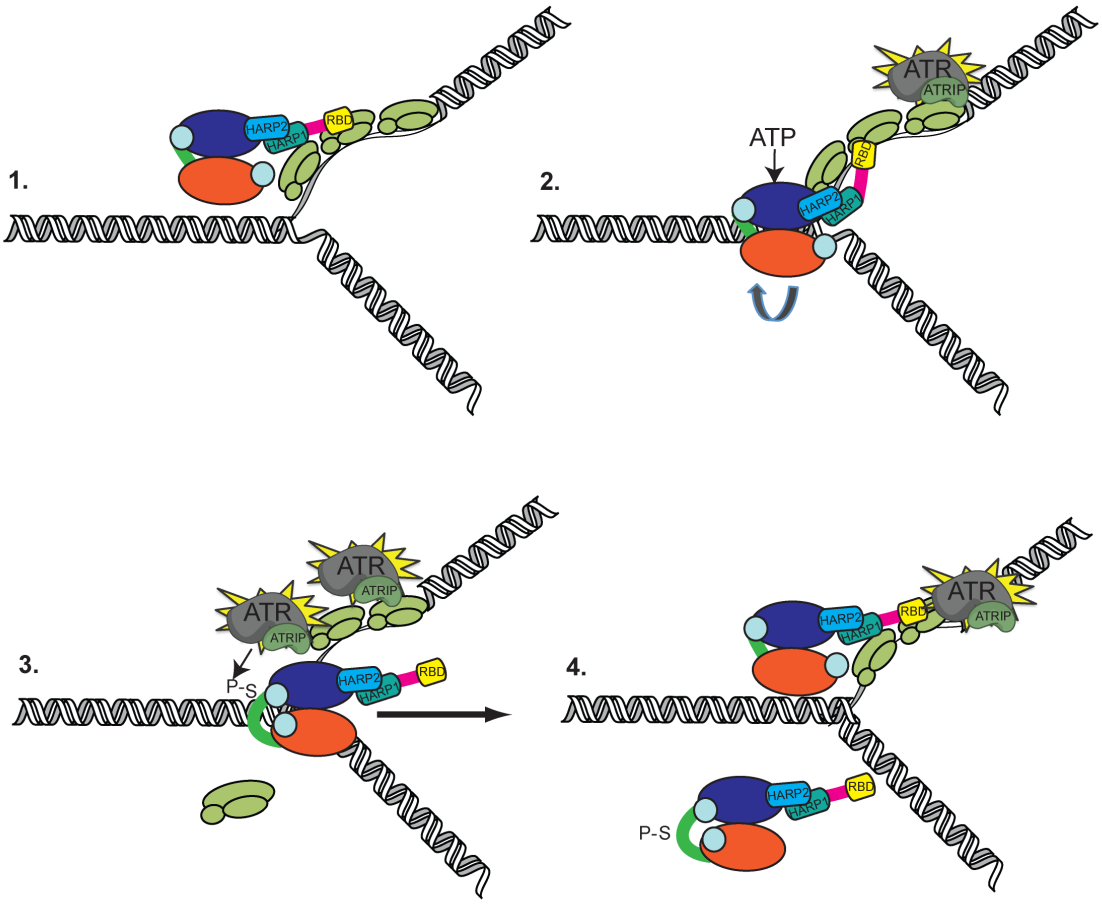

Figure 5.16. A model for limiting SMARCAL1 activity at stalled replication forks. 1. In the presence of replication stress, SMARCAL1 accumulates at stalled replication forks via its N-terminal RPA binding domain (RBD). (2) Localization to forks concentrates SMARCAL1 near its DNA substrates and subsequently SMARCAL1 binds the fork DNA. (3) DNA binding promotes ATP binding and transition from the open-inactive state to the closed-active state. ATP hydrolysis drives SMARCAL1 translocation on the DNA to reanneal the parental strands, displacing RPA. The conformational change exposes S652 to DDR kinases like ATR, which is activated by stalled replication forks. Phosphorylation of S652 in the hinge domain hinders ATPase domain flexibility thereby inhibiting the next enzymatic cycle of ATP binding and hydrolysis. Presumably this inactivated molecule dissociates and a new molecule of SMARCAL1 can bind. This mechanism provides the opportunity for SMARCAL1 to catalyze necessary fork remodeling activities, but with limited processivity.

We propose the following model for how DDR-dependent phosphorylation of S652 regulates SMARCAL1 activities (Figure 5.16). In the presence of replication stress, SMARCAL1 accumulates at stalled replication forks via an interaction with RPA. Localization to forks concentrates SMARCAL1 near its DNA substrates and subsequently SMARCAL1 binds DNA. DNA binding promotes ATP binding and transition from the open-inactive state to the closed-active state. This conformational change facilitated by the SMARCAL1 hinge domain exposes S652 to DDR kinases like ATR, which is activated by stalled replication forks. Phosphorylation of S652 in the hinge domain hinders ATPase domain flexibility thereby inhibiting the next enzymatic cycle of ATP binding and hydrolysis. This mechanism provides the opportunity for SMARCAL1 to catalyze necessary fork remodeling activities, but with limited processivity.

Other enzymes function to remodel replication forks. For example, WRN and BLM are DNA helicases required for fork restart following replication stress (36,37). WRN and BLM are also substrates for DDR kinases (125-127). The DDR-dependent phosphorylation of replication fork remodeling enzymes may serve to coordinate the cooperation of these enzymes at forks. The fork remodeling that promotes fork restart requires unwinding and rewinding of DNA at multiple steps. Little is know about the exchange of these factors on their DNA substrates in cells. Our studies of SMARCAL1 phospho-regulation indicate that phosphorylation could be a mechanism to regulate their processivity and influence their residence times on fork substrates. Characterizing the functional

significance of DDR-dependent phosphorylation events on fork remodeling enzymes is critical for understanding the dynamics of fork stabilization and restart.

Chapter VI

PHOSPHORYLATION REGULATES AN AUTO-INHIBITORY DOMAIN IN SMARCAL1

Introduction

SMARCAL1 is phosphorylated at multiple sites during normal and replication-stressed conditions. We identified a total of 7 sites as phospho-sites by mass spectrometry (Table 6.1). The SMARCAL1 sample used for this study was purified from cells after 16 hours in HU with the intention of identifying sites that contributed to the DDR-dependent regulation of SMARCAL1. Indeed, we identified S173 and S652 and determined that they are both phosphorylated by DDR kinases following replication stress (Chapter V). We also identified a cluster of SP sites in the N-terminus of SMARCAL1 (S112, S123, S129, and S198). SP is a consensus for cyclin dependent kinases, which are activated cell cycle dependent manner (128). We aligned the protein sequences of this region between the RPA binding domain and the first HARP domain of human, mouse, xenopus, and zebrafish SMARCAL1 to determine the conservation of these sites (Figure 6.1A). We find there is little sequence conservation, but there are number of SP sites in this region in all species. The cluster of phosphorylation sites in this region may have potential to regulate cell cycle specific activities of SMARCAL1.

Table 6.1. Sites identified by mass spectrometry

Phoshorylated Site	Peptide identified by MS
S112	KPEEMPTACPGHS*PR
S123, S129	SQMALTGIS*PPLAQS*PPEVPK
S129	SQMALTGISPPLAQS*PPEVPK
S172 or S173	S*S*QETPAHSSGQPPR
S198	AS*PSGQNISYIHSSSESVTPR
S652	LKSDVLS*QLPAK
S889	IYDLFQKS*FEK

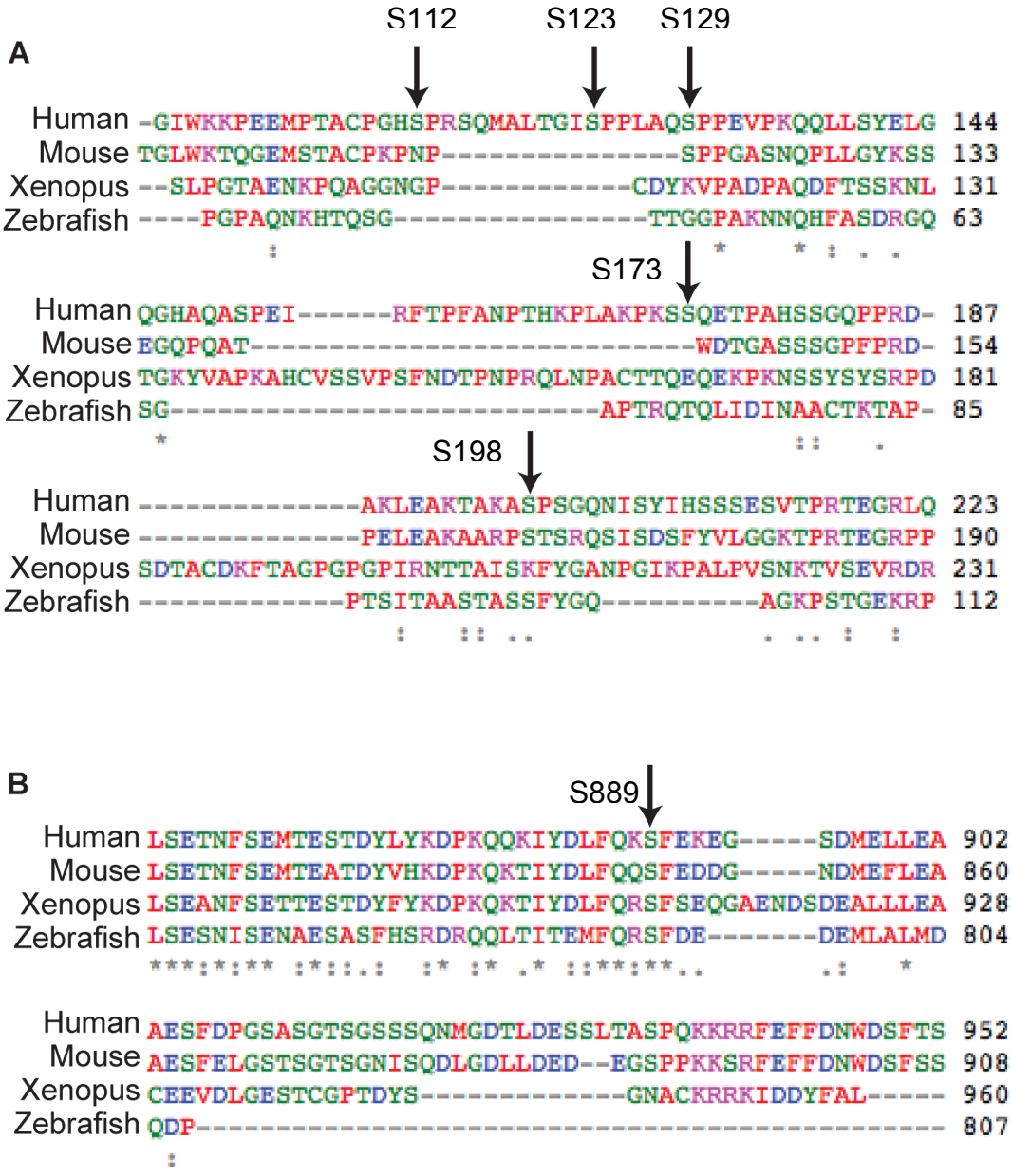

Figure 6.1. Conservation of SMARCAL1 phospho-sites. Alignments of human, mouse, xenopus and zebrafish SMARCAL1 were performed using ClustalW. (A) Alignment for the N terminus of SMARCAL1 spanning aa96-223. Arrows indicate S112, S123, S129, S173, and S198. (B) Alignment for the C-terminus of SMARCAL1 spanning aa858-952. Arrow indicates S889.

We also identified S889. The SMARCAL1 C-terminus is more conserved than the N-terminal region defined above. Likewise, we find that S889 is well conserved across SMARCAL1 species (Figure 6.1B). Because of its strong conservation, we expected that this site might have a conserved function in regulating SMARCAL1 activities. We initially investigated S889 phosphorylation to determine its contribution to DDR-dependent regulation of SMARCAL1. Instead, we discovered that S889 phosphorylation regulates an auto-inhibitory domain in the C-terminus. Our results indicate that phosphorylation of S889 is critical for priming SMARCAL1 activation at stalled replication forks.

Results

S889 phosphorylation regulates SMARCAL1 enzymatic activity

S889 is located C-terminal to the core enzymatic domain of SMARCAL1. A fragment of SMARCAL1 consisting of amino acids 317-870 is a functional annealing helicase *in vitro* (54). Therefore, we did not expect S889 to be required for SMARCAL1 enzymatic activities. We purified FLAG-SMARCAL1 phospho-mutants (S889A and S889D) from cells and tested whether these mutants retained fork-stimulated ATPase activity. Surprisingly, the S889A mutant is significantly less active than WT at both concentrations of fork DNA tested (Figure 6.2A). The opposite effect was observed for the phospho-mimetic mutant. S889D activity was significantly higher than WT when stimulated by 1.25nM fork-DNA. To further confirm our results, we measured ATP hydrolysis at various

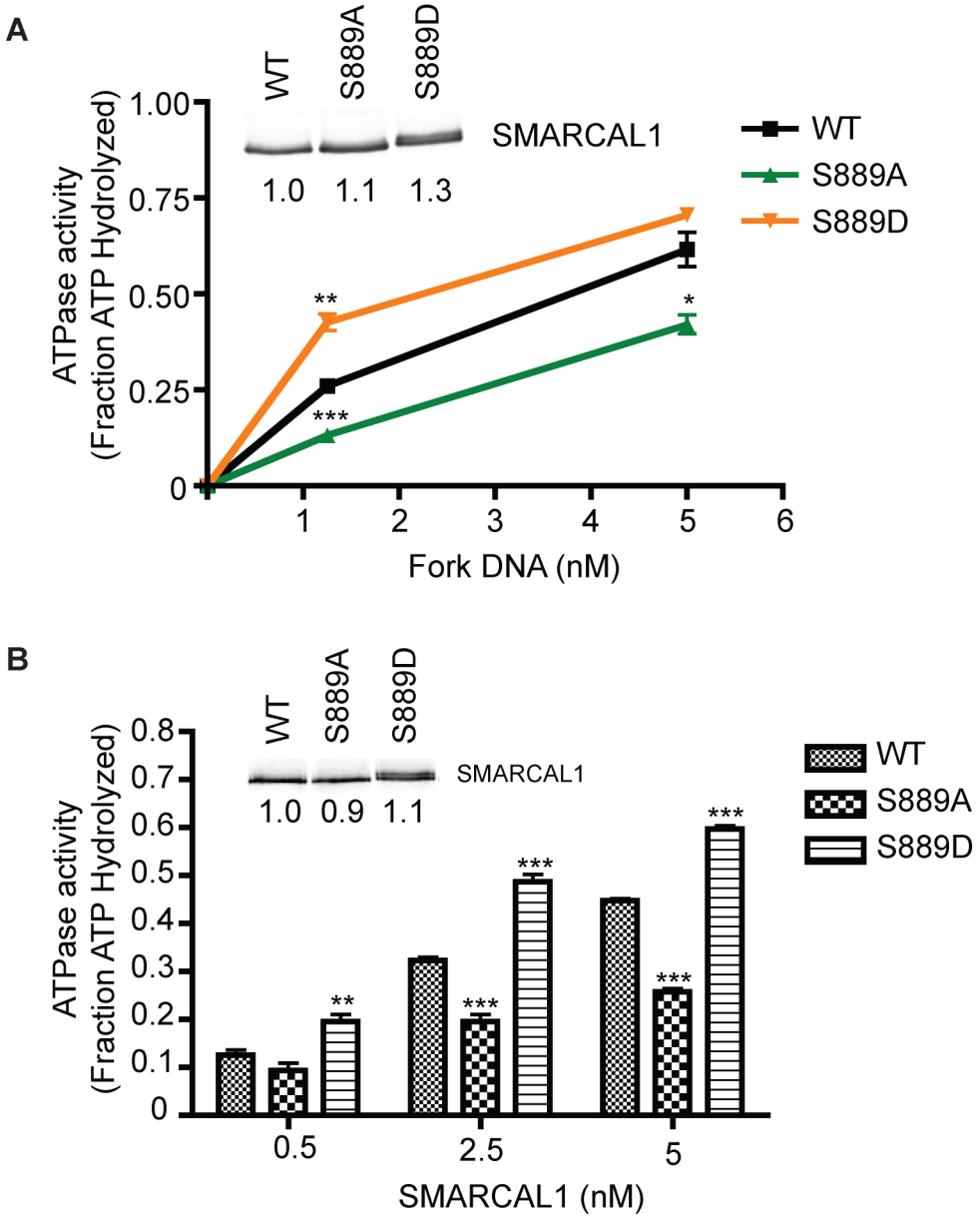

Figure 6.2. S889 mutants have altered ATPase activity. Protein was purified from HEK293T cells that were transiently overexpressing FLAG-SMARCAL1. (A) ATPase activity of the WT and phospho-mutants was measured in the presence of the indicated concentrations of fork DNA. (B) ATPase activity was measured at the indicated concentrations of FLAG-SMARCAL1 in the presence of 1.25nM fork DNA. Error bars are standard deviation (n=3). Quantification of FLAG-SMARCAL1 by WB represents an average of protein input across replicates. For (A), (*)*P*=0.0186, (**)*P*=0.0018, (***)*P*=0.0003. For (B), (**)*P*=0.0031, (***)*P*<0.0001. Two-tailed unpaired t-test calculated between WT and mutant in each condition.

concentrations of SMARCAL1 in the presence of 1.25nM fork DNA. At very low concentrations the S889D mutant was still able to hydrolyze more ATP than WT (Figure 6.2B). This trend was true for all SMARCAL1 concentrations tested. The S889A mutant was consistently less active.

We next asked whether this change in activity in vitro translated into a change in activity in cells. We first confirmed that mutating S889 does not compromise localization to stalled replication forks. Both mutants co-localize with γH2AX foci after HU treatment (Figure 6.3A). We then measured the ability of S889A and S889D to activate a DDR when overexpressed. Using the quantitative assay described in Chapter 5 (Figure 5.12), we determined that in cells the activity of the S889A mutant resembles the ATPase-dead mutant R764Q (Figure 6.3B). And consistent with in vitro observations, the S889D was overactive compared to WT.

Phosphorylation of S889 affects DDR-dependent phosphorylation

In the course of studying S889 mutants, we noticed that the S889D mutant usually migrates as two bands on SDS-PAGE gels (Figure 6.2). Since S889D is an overactive form of SMARCAL1 and activates a DDR more readily than WT when overexpressed, we reasoned that this shifted form could be due to DDR-regulated phosphorylation. To confirm that this shift was due to phosphorylation we purified protein from cells overexpressing FLAG-SMARCAL1 and treated it with phosphatase. The phosphatase treatment collapsed S889D into a single band (Figure 6.4A). Due to high expression levels, we observed a mild shift of

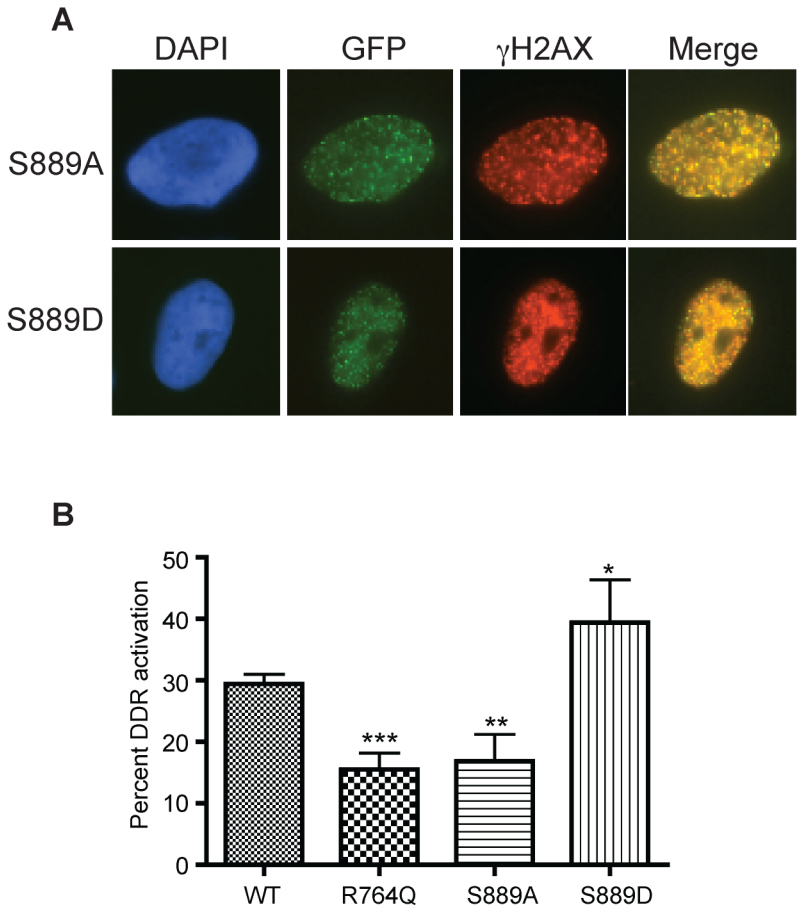

Figure 6.3. S889 phosphorylation is required for SMARCAL1 activity in cells. (A) U2OS cells were transfected with lower expressing pLL5.0-GFP-SMARCAL1 vectors. The day after transfection cells were seeded onto coverslips. The next day cells were treated with HU for 5 h, then fixed and stained with γH2AX antibodies and mounted on slides with ProLongGold +DAPI. (B) Higher expressing pLEGFP-SMARCAL1 vectors were transfected into U2OS. The day after transfection, cells were seeded into 96 well plates. The next day cells were fixed in 96 well plates and stained with γH2AX antibodies and DAPI. Cells were imaged using the Opera and individual nuclei were measured for mean GFP intensity and mean γH2AX intensity using the Columbus analysis software. Data represents the percent of cells (500<GFP<2500) with mean γH2AX >1000. Error bars are standard deviation (n=7 wells, 3 biological replicates). (***) *P*=0.008. (**)*P* =0.0027; (*)*P* =0.0409. Two-tailed unpaired t-test calculated between WT and mutant.

WT protein that was also reduced by phosphatase treatment. This further supports that a mutant mimicking constitutive phosphorylation of S889 is hyperactive.

When we transfect lower expressing vectors of GFP-SMARCAL1 we generally observe a phospho-shift only in HU-treated cells. In these conditions, we see a striking difference in the shift pattern of WT and S889D proteins. The S889D mutant is super-shifted even in the untreated samples (Figure 6.4B). The phospho-shift increases with the addition of HU. This may indicate that the S889D is not only more active, but is also a better substrate for phosphorylation by DDR kinases. Importantly, we observe the opposite effect for S889A. We reproducibly observe a reduced shift of the S889A protein; however, it shifts slightly more than the ΔN mutant indicating DDR phosphorylation is not completely abrogated. This distinguishes the S889A mutant from the R764Q mutant, which is also inactive but phospho-shifts similar to WT.

Phosphorylation of S889 regulates an auto-inhibitory domain in the C-terminus of SMARCAL1

The S889A mutant exhibits reduced activity compared to WT in all assays tested thus far. The phospho-shift result suggests that S889A more closely resembles a DNA binding mutant. However, S889 is not located in a region of SMARCAL1 that would obviously contact DNA and promote enzymatic function. If S889 did participate in some step of catalysis we would expect a truncation of resembles a DNA binding mutant. However, S889 is not located in a region

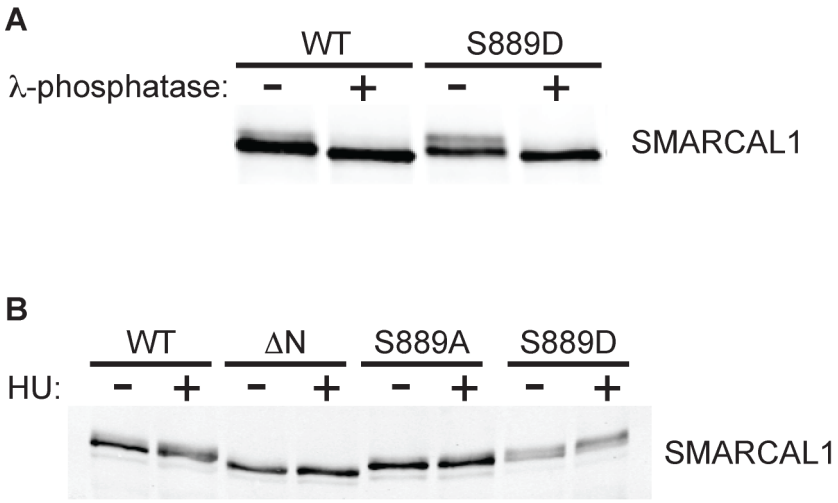

Figure 6.4. S889D mutant is hyper-phosphorylated. Samples were prepared, separated by SDS-PAGE, transferred to nitrocellulose and immunoblotted with SMARCAL1 antibodies. (A) FLAG-SMARCAL1 was overexpressed in HEK293T cells and purified from cell lysates. Following immunopurification, protein was incubated in the presence of λ-phosphatase 1X buffer (-/+) enzyme for 45min at 30°C. (B) HEK293T cells were transfected with the lower expressing pLL5.0-GFP-SMARCAL1 and treated two days later with 2mM HU for 16hrs.

of SMARCAL1 that would obviously contact DNA and promote enzymatic function. If S889 did participate in some step of catalysis we would expect a truncation of the C-terminal portion containing S889 to result in decreased activity as well. We generated a C-terminal truncation mutant that removed amino acids 861-954 (ΔC). This region is not essential for SMARCAL1 annealing helicase activity, but we wanted to confirm the activity of a ΔC mutant alongside S889 mutants in the ATPase assay. We repeated the fork-stimulated ATPase assay with all mutants. The ΔC mutant shows increased activity similar to the S889D mutant (Figure 6.5A). Remarkably, removing the C-terminus containing S889 has the same affect on ATPase activity as mimicking phosphorylation at S889.

When we analyzed the localization of the ΔC mutant in cells, we found that the majority of the protein is cytoplasmic. Since we use protein purified from human cells in our ATPase assays we wanted to confirm that localization of the ΔC mutant did not impact our result. We added an NLS and confirmed that this protein localizes to the nucleus (NLS-ΔC). We then compared the activities of the ΔC and NLS-ΔC mutant. When localized properly a ΔC mutant is still significantly more active than WT protein (Figure 6.5D). We also note that the NLS-ΔC mutant exhibits a slight phospho-shift compared to ΔC. This further supports that the NLS-ΔC localizes properly and causes DDR activation similar to S889D. The ΔC data confirms that S889 is not directly involved in the enzymatic functions of SMARCAL1. Furthermore, it indicates that the C-terminus contains

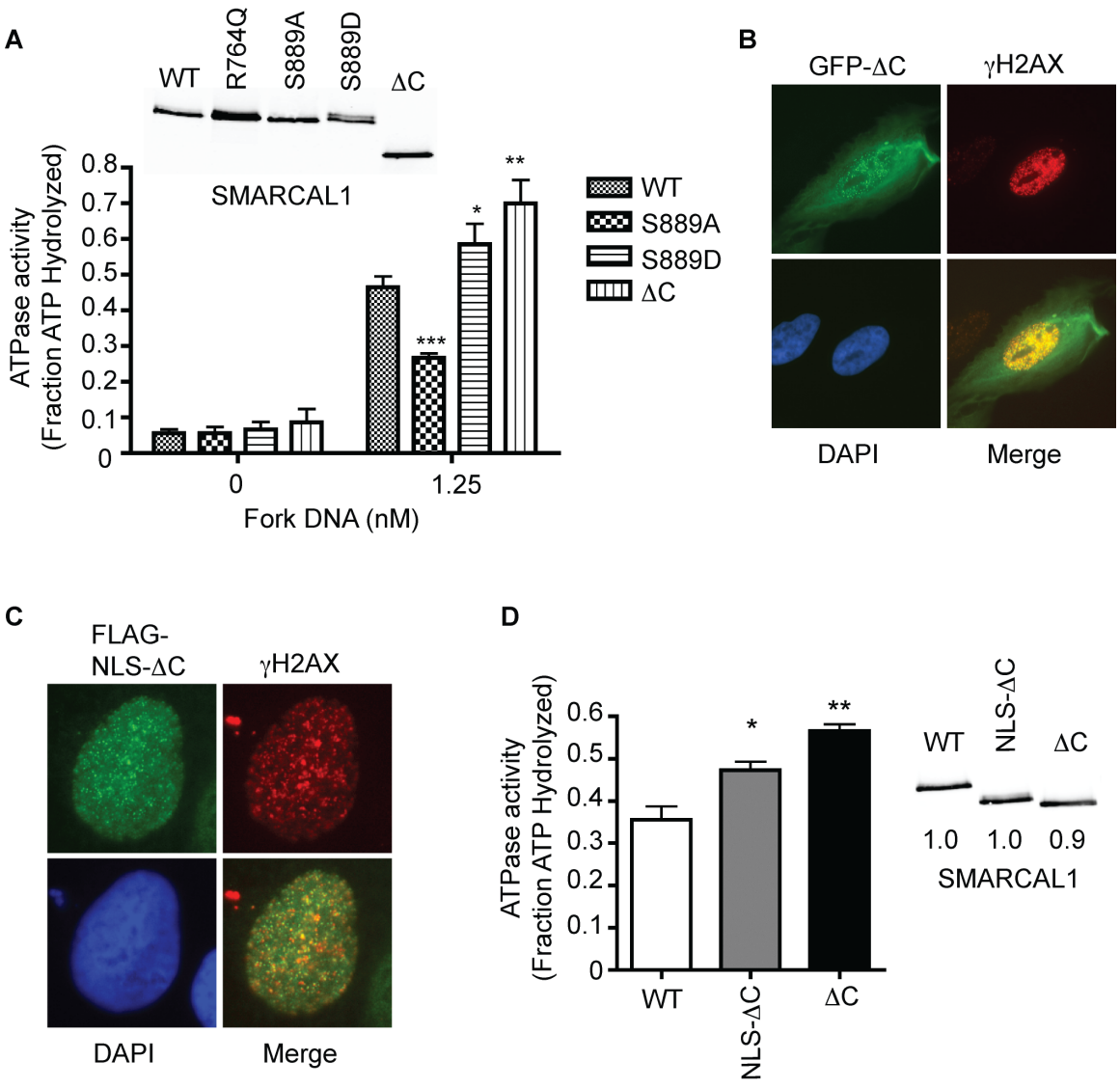

Figure 6.5. The C-terminus of SMARCAL1 contains an auto-inhibitory domain. (A,D) Protein was purified from HEK293T cells that were transiently overexpressing FLAG-SMARCAL1. ATPase activity was measured in the presence of 1.25nM fork DNA. (A) R764Q was not included in all replicates so data is not shown. (B,C) U2OS cells were transfected with lower expressing pLL5.0-GFP SMARCAL1 vectors. The day after transfection cells were seeded onto coverslips. The next day cells were treated with HU for 5 h, then fixed and stained with γH2AX antibodies and mounted on slides with ProLongGold +DAPI Error bars are standard deviation (n=3). Quantification of FLAG-SMARCAL1 by WB represents an average of protein input across replicates. For (A) (***)*P*=0.0005, (**)*P*=0.0049, (*)*P*=0.031. For (D), (*) p=0.043; (**) p=0.0031. All two-tailed unpaired t-test comparing WT and mutants.

an auto-inhibitory domain that must be relieved for SMARCAL1 to function efficiently.

S889 is phosphorylated in undamaged cells

We mapped phosphopeptides of the S889A mutant to identify the pS889 peptide and determine whether it changed after HU. We were able to detect the loss of phosphopeptide-*f* in S889A maps (Figure 6.6). Phosphopeptide-*f* consistently migrated underneath an unknown phosphopeptide (marked with an asterisk) and above the S173 phosphopeptide-*b*. Because of the HU-induced streak pattern of phosphopeptide-*b* we were unable to determine conclusively if the pS889 peptide changed in response to HU.

In order to characterize the regulation of S889 phosphorylation, we generated a phospho-specific antibody to a pS889 peptide. We determined that this antibody recognized WT protein purified from untreated cells (Figure 6.7A). The antibody had some affinity for S889D, but did not recognize the S889A mutant. The detection of WT protein was also phosphatase sensitive (Figure 6.7B). We conclude that this antibody specifically recognizes phosphorylated S889 in SMARCAL1. The pS889 antibody was then used to determine whether S889 phosphorylation was induced following replication stress. Although this evidence is still preliminary, we found that the pS889 antibody signal is reduced when FLAG-SMARCAL1 is purified from cells treated with HU (Figure 6.7C). These data suggest that SMARCAL1 is phosphorylated at S889 during normal conditions and then removed during times of replication stress.

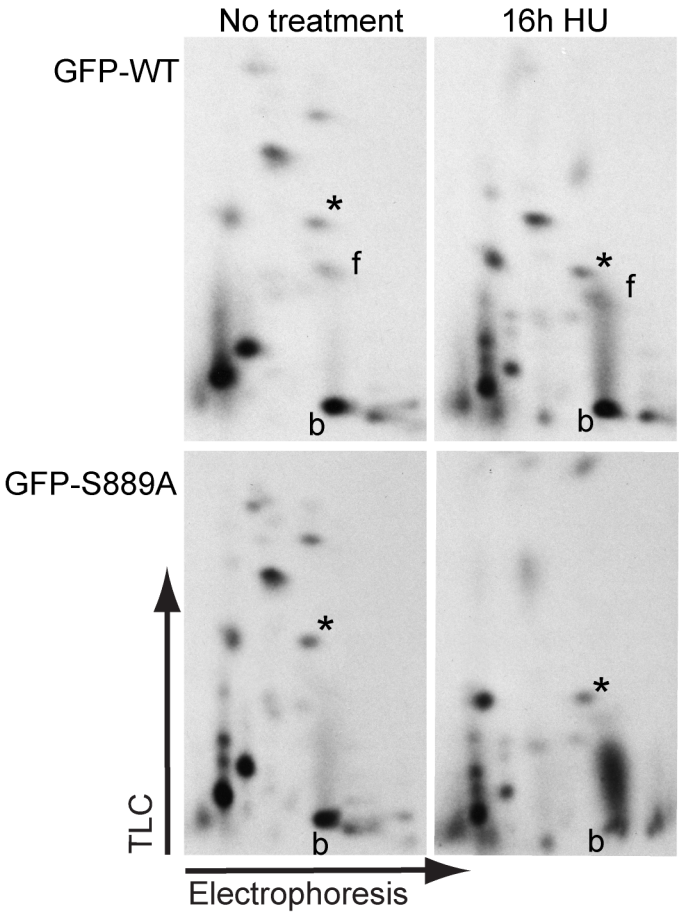

Figure 6.6. Phosphopeptide mapping of the S889A mutant identifies a pS889 containing peptide. HEK293T cells were transfected with expression vectors (pLL5.0-GFP-GW backbone for low expression) encoding GFP-tagged WT and S889A SMARCAL1. The day after transfection cells were passaged and HU was added for 16hrs where indicated. Cells were metabolically labeled and peptides processed as in (Figure 5.4). Phosphopeptide-*f* reproducibly migrates above the S173 phosphopeptide-*b* and below the unknown phosphopeptide marked with an asterisk.

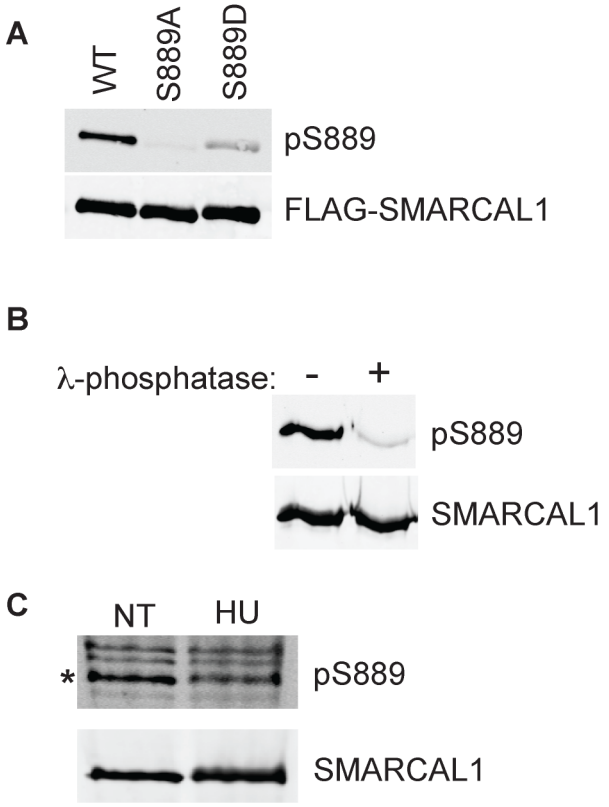

Figure 6.7. S889 is phosphorylated in undamaged cells. (A,B) HEK293T cells were transfected with FLAG-SMARCAL1 vectors and two days later protein was purified from cell lysates. After FLAG-purification, protein was eluted from FLAG antibodies with FLAG peptide (300μg/mL). (B) Eluted protein was incubated with l-phosphatase buffer (-/+) enzyme for 45min at 30°C. Samples were separated by SDS-PAGE, transferred to nitrocellulose and immunoblotted with pS889. Where indicated (FLAG-SMARCAL1), FLAG antibodies were used to detect SMARCAL1. In other cases, pS889 blots were stripped and reprobed with SMARCAL1 antibodies. (C) Cells were synchronized by thymidine for 16hrs and released for 1.5h before harvesting (NT) or adding HU for 4 hours. (C) Cells were transfected as in (A). (*) Denotes SMARCAL1 band. The non-specific bands in (C) appear when samples are processed on FLAG beads rather than eluting (as in A and B)

Discussion

Deregulated SMARCAL1 activity at replication forks is associated with genome instability indicating that fine-tuning the amount of SMARCAL1 activity at replication forks is critical for genome maintenance. We have determined that the **C**-terminus of SMARCAL1 contains an auto-**I**nhibitory **D**omain, or CID, that is regulated by S889 phosphorylation. The phenotypes of S889 phospho-mutants reflect defects in CID regulation. A mutant that cannot be phosphorylated is constitutively inhibited, while a phospho-mimetic mutant or ΔCID mutant is constitutively uninhibited. Our data suggest that phospho-regulation of CID function is important for regulating SMARCAL1 activation at stalled replication forks.

Regulation of S889 phosphorylation

Our results indicate that SMARCAL1 is maintained in a CID-resistant state by constitutive phosphorylation of S889. The kinase that phosphorylates S889 is unknown. The sequence surrounding S889 does not fit any described kinase consensus. We did identify a potential protein interaction with the protein kinase homeodomain-interacting protein kinase 3 isoform 2 (HIPK3) in a yeast-2-hybrid screen for SMARCAL1 interacting proteins. HIPK3 is implicated in regulating RNA processing pathways. The pS889 antibody can be used to screen available kinase inhibitors and narrow the list of possible kinases that phosphorylate S889.

Importantly, preliminary data indicates that S889 phosphorylation may decrease in cells following treatment with HU. In Chapter V we determined that an inactivating phosphorylation in the hinge domain of SMARCAL1 is an important function of DDR-dependent S652 phosphorylation. Removal of S889 phosphorylation at stalled replication forks would further promote limiting SMARCAL1 activity. This would relieve inhibition of the CID. The interplay between S652 and S889 phosphorylation warrants further investigations.

S889 phosphorylation primes SMARCAL1 for activation

Unique domains or subdomains are often attached to the amino- or carboxy- terminus of the ATPase motor in SNF2 and the related SF2 helicase proteins. In some cases, these domains have been demonstrated to inhibit the ATPase motor by either blocking initial nucleic acid binding site or influencing ATPase domain orientation (129). The CID of SMARCAL1 could function similarly. For domains that block nucleic acid binding, removal of the inhibitory domain usually removes the nucleic-acid requirement for ATPase activity (130). We do not observe significant increase in fork-DNA independent activation of the ΔC mutant. Therefore, we favor a model in which the CID affects the orientation of the ATPase domain.

For the SF2 helicase DDX19, an N-terminal helix of DDX19 wedges itself between the two ATPase domains and prevents the closed conformation of the ATPase domains (131). We propose a similar mechanism for the CID of SMARCAL1 (Figure 6.8). The CID makes important contacts with the ATPase

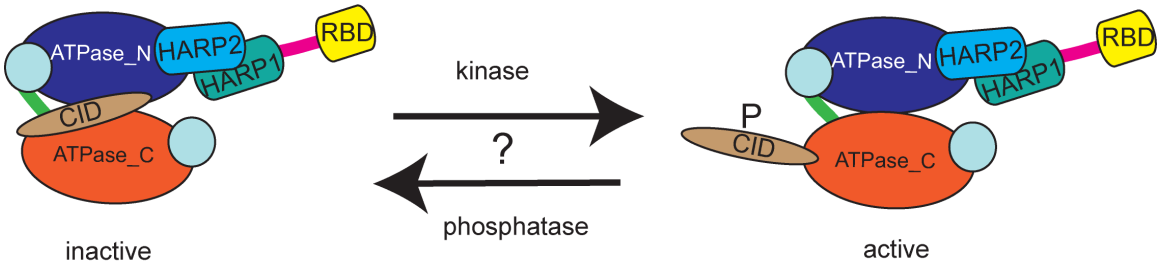

Figure 6.8. Phosphorylation of S889 relieves an auto-inhibitory domain in SMARCAL1. An unknown kinase phosphorylates S889 to relieve the inhibitory conformation of the CID. Following replication stress, phosphorylated S889 is removed potentially by a DDR phosphatase. This would return the CID to the inhibitory conformation.

motor in order to inhibit ATP hydrolysis. Phosphorylation of S889 alters the conformation of the CID and relieves inhibition.

Structural information about how the C-terminus contacts the catalytic core of SMARCAL1 will be necessary to understand how phosphorylation at S889 may regulate this domain. Alternatively, S889 could regulate an interaction with an inhibitory protein. It is possible that this protein is present in our ATPase reactions since we purify SMARCAL1 from human cells. This alternative would suggest that S889 regulates a protein-protein interaction in the CID. Whether or not the CID directly contacts the ATPase motor or binds another protein to facilitate inhibition can be determined, but the predicted functional outcomes are the same.

CHAPTER VII

DISCUSSION AND FUTURE DIRECTIONS

Further characterization of genome maintenance genes

Our genome maintenance screens identified 171 genes with potential functions in genome maintenance pathways. Using published literature and functional annotation programs we placed many of the genes into four major functional groups: the ATM/ATR-related DNA damage response, mitosis, chromatin regulation, and RNA metabolism. However, in most cases direct evidence for gene function in these pathways is lacking. We investigated the biological functions of three proteins identified in our screens, DDB1, CINP and SMARCAL1. As validation for our screen methodology, we found that all three promote genome integrity (58,66,81).

A dataset including over 100 uncharacterized gene products awaits further validation. As functional genomic screens continue to develop as tools for discovery, primary-screening methodologies will grow in complexity and yield more specific information about gene function. Only three years after completing our genome maintenance screen, our methodologies have advanced far beyond the initial limitations of an entirely manual immunofluorescence screen. A post-doctoral fellow in our laboratory, Dr. Gina Kavanaugh, has recently completed a genome-wide RNAi screen, nearly fully-automated, in which she measured four

different indicators of replication stress and DNA damage. I expect that cross-referencing this data set with our previous data set will yield new mechanistic information for how these genes function in genome maintenance pathways.

Developing hypotheses for loss of function phenotypes in the RNAi screen is somewhat easier than interpreting overexpression phenotypes. However, identification of sixteen oncogenes in the overexpression screen emphasizes the capacity of this methodology to identify novel oncogenes. We expect that other genes identified in the screen are oncogenic and drive abnormal cell division cycles that challenge genome integrity and promote cancer progression. To test this hypothesis, we need to determine whether these gene alterations are sufficient to drive cancer *in vivo*.

Overexpression of *Hox* genes has been associated with a subset of human myeloid leukemias (132). Importantly, *Hox gene* overexpression can drive transformation of mouse hemopoietic cells, providing direct evidence for *Hox* involvement in leukemic transformation (133,134). The methodology used in these studies could also be used to validate genes we identified in our overexpression screen. The basic experimental scheme involves (i) cloning cDNAs into high expressing retroviral vectors, (ii) viral transduction of mouse primary bone marrow cells, (iii) reconstituting bone marrow depleted mice with infected bone marrow cells, and (iv) assessing leukemic growth *in vivo.*

Enhancing our understanding of the cellular response to DNA damage will help us better understand cancer development and progression. Our knowledge

of genome maintenance pathways and how cancer cells manipulate these pathways will also help refine how we classify and treat the disease.

SMARCAL1

My thesis work led to the identification of the SNF2 protein SMARCAL1 as a genome maintenance protein. My characterization of SMARCAL1 functions, along with evidence from other investigators (58-61), led to the discovery that SMARCAL1 annealing helicase activity promotes genome integrity specifically at stalled replication forks. Since this initial discovery, SMARCAL1 activities have been further defined. SMARCAL1 travels with at least some replication forks in untreated cells and accumulates at stalled replication forks when cells are treated with replication inhibitors (54). Biochemical evidence suggests that in addition to re-annealing RPA-ssDNA, SMARCAL1 can catalyze fork regression and remodel a variety of fork-like structures. These activities can promote replication restart and prevent replication-associated DNA double strand breaks *in vivo.* Indeed, SMARCAL1 deficiencies cause fork restart defects after replication stress, sensitivity to replication stress agents, and aberrant fork structures that result in DSBs (54,58,59).

In the second part of my thesis, I investigated the functional implications of SMARCAL1 phosphorylation by DDR kinases. I identified a number of DDR-dependent and DDR-independent sites. My characterization of phospho-mutants provides evidence to support a model in which DDR-regulated phosphorylation

limits SMARCAL1 activity at stalled replication forks by modulating conformational changes required for ATP hydrolysis.

SIOD: a genetic model for studying SMARCAL1 replication stress functions

Defects in the activities of many replication fork-remodeling enzymes are associated with human syndromes (45,47,135). The requirement for fork remodeling *in vivo* has been elucidated from studies investigating the cellular defects of patients. Mutations in *SMARCAL1* cause SIOD (65). In our characterization of SMARCAL1 genome maintenance functions, we established the first link between SIOD and a defect in the replication stress response. We showed that an R764Q mutant associated with the severest form of SIOD fails to complement the replication-associated DNA damage caused by silencing SMARCAL1 (58).

If the phenotypes of SIOD are, in part, due to defects in SMARCAL1 function at stalled replication forks, we would expect to observe an increase in DNA damage in SIOD patient cells. Indeed, fibroblasts derived from an SIOD patient exhibit elevated levels of γH2AX when compared to cells that have been complemented with functional SMARCAL1 (Figure 7.1A). An RPA binding-mutant (ΔN) and the R764Q SMARCAL1 mutant have no affect on the amount of γH2AX, indicating that the localization of annealing helicase activity is required for reducing DNA damage accumulation. Additionally, we find that SIOD cells exhibit increased RPA phosphorylation following HU treatment (Figure 7.1B). These data support that SIOD cells are prone to accumulating DNA damage as

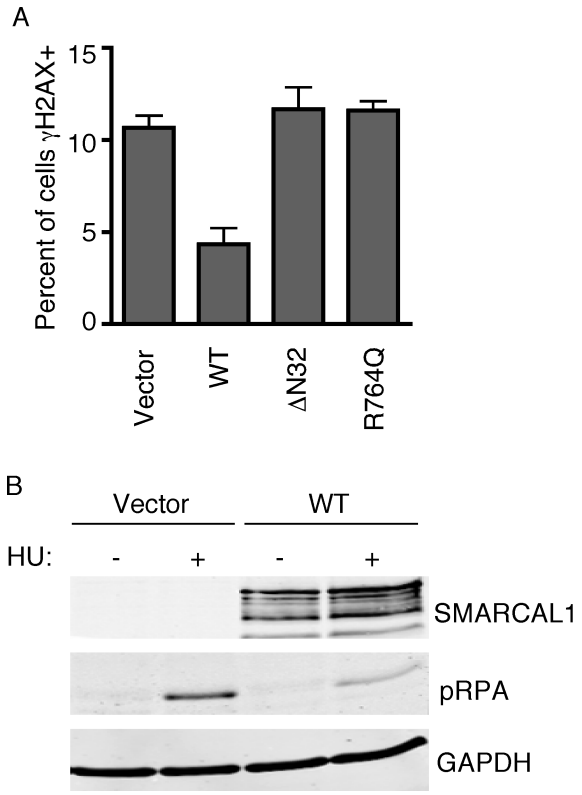

Figure 7.1. SIOD patient fibroblasts exhibit increased DNA damage signaling. (A) SIOD patient-derived human fibroblasts stably expressing vector only, WT, ΔN, or R764Q SMARCAL1 were cultured on coverslips and stained with antibodies to γH2AX and appropriate secondary antibodies. The percentage of cells with γH2AX staining was scored; error bars are standard deviation 300 cells were counted with each cell type analyzed in triplicate. (B) Cells with an integrated empty vector (Vector) or vector expressing wild-type SMARCAL1 (WT) were left untreated (NT) or treated with 2 mM hydroxyurea (HU) for 5 h then released into normal growth media for 0 or 8 h as indicated. Total cell lysates were prepared and immunoblotted with antibodies to SMARCAL1, phosphorylated RPA2, or total RPA2.

they cycle and accumulate more phospho-RPA-ssDNA during replication stress. Furthermore, they suggest SMARCAL1 activities during replication stress are important for preventing SIOD. Complementation of these defects in SIOD fibroblasts will be a useful tool for determining the functional requirements for phosphorylation and CID regulation of SMARCAL1 at stalled replication forks.

SIOD-associated mutations in the hinge domain

Missense mutations associated with SIOD are particularly informative for determining important functional domains in SMARCAL1. For example, R764 in SMARCAL1 is a conserved arginine that is predicted to make important contacts with ATP in the active site cleft of SNF2 and SF2 ATPases (121). The R764Q mutation associated with severe SIOD results in an ATPase-dead SMARCAL1, supporting the structural observation that this arginine is involved in ATP hydrolysis. Some of the patient mutants that I think would be particularly interesting to characterize I have mapped on a model of SMARCAL1 domain structure for reference (Figure 7.2). Characterizing the functional defects of these mutants will lead to a better understanding of how SMARCAL1 is regulated to support genome integrity.

A number of SIOD associated mutations (R644W, R645C, R645H, K647Q, and K647T) cluster in the hinge domain of SMARCAL1 suggesting this is a critical domain for SMARCAL1 function as well (65,136). The hinge domain likely provides flexibility required for a critical conformational change that positions active site residues in the ATPase_N and ATPase_C domains for ATP

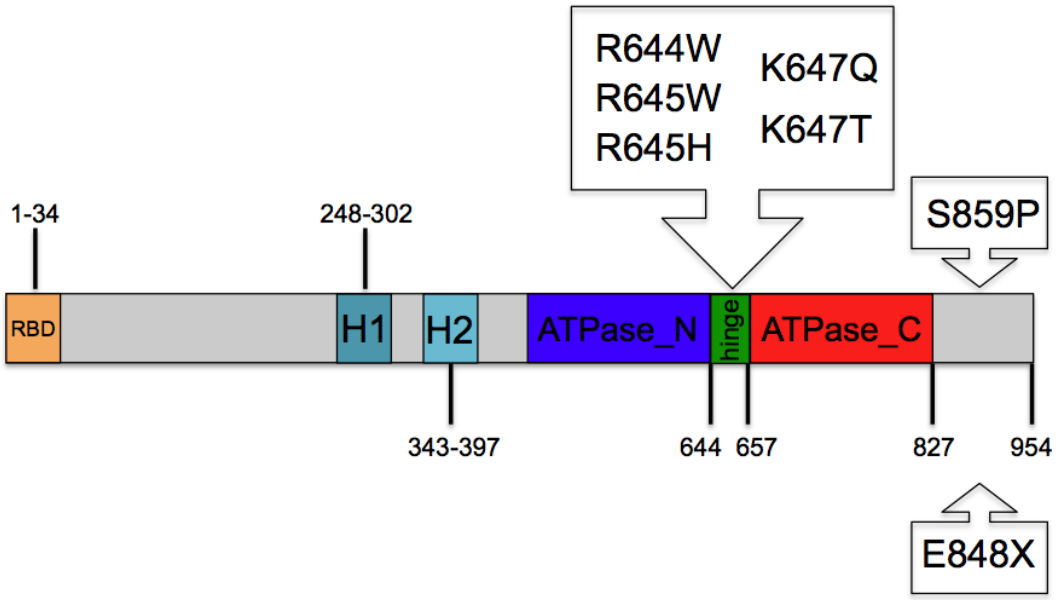

Figure 7.2. SIOD patient mutations map to potential regulatory domains of SMARCAL1. Missense mutations in *SMARCAL1* generally lead to SMARCAL1 mutants that are expressed but lack necessary SMARCAL1 functions. These mutations may indicate critical regulatory domains for SMARCAL1 activities that maintain genome integrity and prevent SIOD. The mutations mapped here do not represent all SIOD patient mutations, only those that have implications for the potential regulatory domains I have identified in my thesis. (reference http://bioinf.uta.fi/SMARCAL1base/ for a complete list)

hydrolysis. As discussed below, I identified a S652D mutant that is less active likely due to inhibition of hinge domain function.

I expect that the SIOD-associated mutations in the hinge domain also inhibit hinge flexibility and therefore inhibit SMARCAL1 enzymatic activity. This is likely the cause of SIOD in these patients. It is possible that mutating R644, R645, K647, and K647 to alanines may preserve flexibility and ATPase function. In support of this hypothesis, the R645H mutation, which preserves the basic residue at 645 is associated with a mild form of SIOD. Suggesting it may retain partial function. Demonstrating that the functional defect of these mutations can be rescued by mutating residues to more permissive amino acids will be important to validate the essential role of hinge domain flexibility in the activities of SMARCAL1 and other SNF2 enzymes.

SIOD-associated mutation E848X is a ΔC mutant

The E848X mutation has been identified in at least one allele of seven patients with SIOD (65,137). Five of the seven patients have severe SIOD. My results indicate that this truncated form would not localize to the nucleus efficiently. Although I was able to detect a portion of the ΔC protein localized to stalled replication forks, it is not clear whether this amount would be sufficient to support SMARCAL1 genome maintenance activities. Based on in vitro activity for this mutant, the amount that does localize would be hyperactive. Most likely the E848X mutant causes SIOD because this mutant lacks an NLS. However, another more intriguing alternative may occur. The E848X mutant may also be

hyperactive and the amount that is nuclear can disrupt normal replication fork processing activities. If this were true, it would be the first example where hyperactive SMARCAL1 can cause SIOD.

Finally, a new SIOD-associated SMARCAL1 mutation has recently been identified in a patient seen here at Vanderbilt, S859P. Because of the geometry of its side chain, proline can cause kinks in alpha helices. This mutation could induce a conformational change in the C-terminus of SMARCAL1. The effects of this mutation on S889 phosphorylation or C-terminal auto-inhibition should be investigated. Characterization of the S859P mutant may lead to a better understanding of how the C-terminus of SMARCAL1 regulates SMARCAL1 activities.

Defining SMARCAL1 interacting proteins

SNF2 family members differ significantly in the protein sequence flanking the SNF2 domain (138). Protein domains outside the SNF2 domain often dictate the functional specificity of the protein. For example, protein-protein interactions in the N-terminus of the SNF2 protein INO80 direct localization of INO80 to sites of DNA DSBs where it functions to remodel nucleosomes and promote the recruitment of DNA processing factors (139). I determined that SMARCAL1 interacts with RPA through its N-terminus. The genome maintenance activities of SMARCAL1 in cells depend upon this interaction (58,59,61).

Identifying additional SMARCAL1 interacting partners will be important to further elucidate SMARCAL1 functions at replication forks and potentially

uncover other cellular functions of SMARCAL1. We have investigated SMARCAL1 complexes purified from both HU-treated and untreated cells by mass spectrometry. Thus far, we observe no damage-inducible changes in SMARCAL1 complexes. Additionally, many of the interactions seem to be dependent on RPA. Indirect interactions through RPA do not exclude the possibility that forming these complexes in vivo is critical for SMARCAL1 function. It does however complicate the biochemistry and interpretation of results. Disrupting the SMARCAL1-RPA interaction abolishes SMARCAL1 replication stress function in cells.

The conditions used for purifying complexes may have not been ideal for identifying more transient SMARCAL1 interactions. Important interactions may be temporally regulated in the response to DNA damage. Therefore, an analysis of SMARCAL1 complexes purified from synchronous cell populations exposed for different times in HU may yield new interacting partners.

We also initiated a two-hybrid screen performed by Hybrigenics using their ULTImate Y2H methodology. Using full-length SMARCAL1 as bait, we identified 33-interacting clones, representing seven different proteins (FAM178A, HIPK3, KCTD13, PIAS2, RPA2, UBE2I, ZNF92). The identification of RPA2 subunits in this screen validates this methodology for identifying SMARCAL1 interacting proteins. Interestingly, UBE2I and PIAS2 function together to sumoylate proteins. UBE2I is the only SUMO-E2 conjugating enzyme in mammals, while many different PIAS proteins direct substrate specificity for SUMO conjugation

(140,141). The identification of both in the Y2H screen suggests they may be biologically relevant for SMARCAL1 function.

The interaction between SMARCAL1 and UBE2I or PIAS2 has not been validated, nor have I been able to detect SMARCAL1 sumoylation. However, SUMO modifications are very labile and are difficult to detect. SMARCAL1 contains a sumoylation consensus site (ΨKxE) at K647 in the hinge region. Interestingly, S652 is positioned in a phospho-sumoyl switch motif (ΨKxExxS) that couples sequential phosphorylation and sumoylation (142). A SUMO regulatory network may be involved in the DDR-dependent regulation of SMARCAL1 at replication forks (Figure 7.3). Presumably, addition of SUMO in the hinge region would enhance the inhibitory effect already induced by S652 phosphorylation. If the interaction with UBE2I and PIAS2 can be confirmed, this would be sufficient rationale to investigate whether SMARCAL1 is a substrate for sumoylation in vitro. These studies would further define the mechanisms by which DDR kinases limit SMARCAL1 activity at stalled replication forks.

Phospho-regulation of SMARCAL1

We identified several CDK consensus sites in the N-terminus of SMARCAL1. CDK activity is important for regulating the cell cycle specific activities of other SNF2 proteins. CDK phosphorylation of the SNF2 protein, PICH (Plk1 interacting checkpoint helicase) is required to expose a Plk1 binding site in PICH and localize it to the centromere where it functions in the spindle checkpoint (143). The human SWI/SNF complex composed of Brg1 and Brm

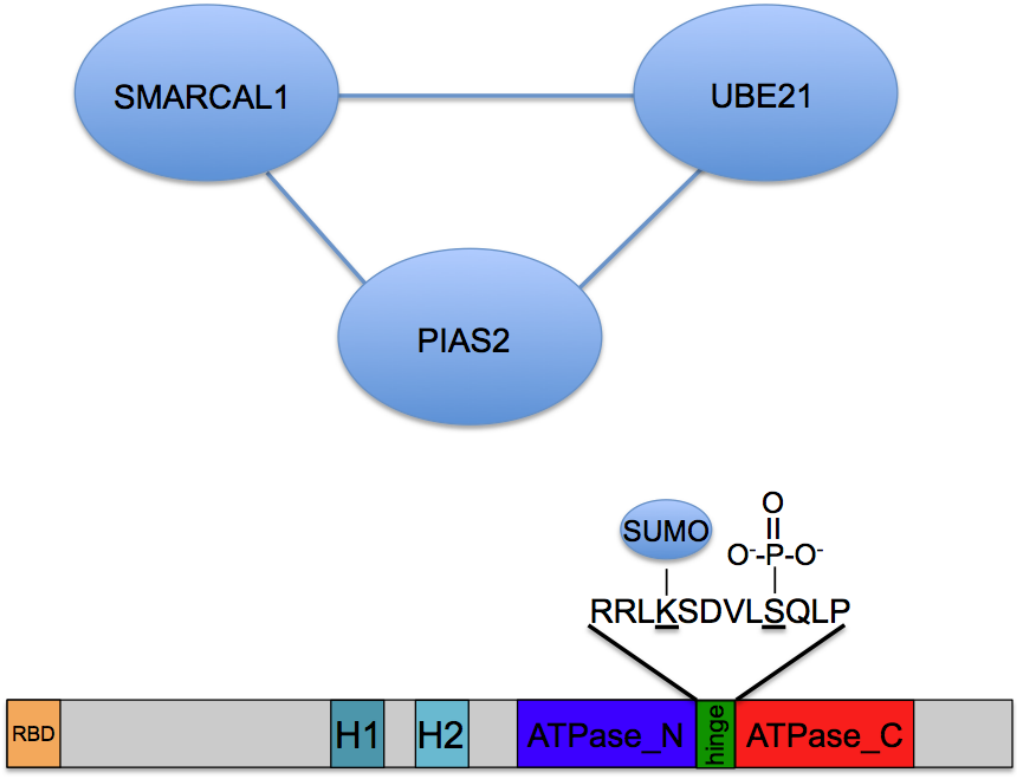

Figure 7.3. Potential SUMO-regulatory network identified by yeast-2-hybrid screen. A yeast-2-hybrid screen with full-length SMARCAL1 as bait identified fragments of UBE2I and PIAS2 as interacting clones. UBE2I and PIAS2 function together to conjugate SUMO to lysine on target proteins. Sumoylation is preferred at lysines within a certain consensus (ΨKxE). Sumoylation can also be directed by phosphorylation downstream of the lysine (ΨKxExxS) at a serine. SMARCAL1 contains this consensus in the hinge domain and S652 is located in the phospho-consensus site. Sumoylation may be another regulatory step in inactivating the hinge domain.

alters nucleosome structure in vitro and may regulate transcription via chromatin remodeling in cells (144). BRG and BRM undergo phosphorylation when cells enter mitosis. When purified from cells at different stages in the cell cycle, the phosphorylated complex from mitotic cells lacks nucleosome disrupting activity in vitro (145). Reversibly inactivating the SWI/SNF complex as cells traverse mitosis may be required for forming the tightly packed chromatin structure important for completing mitosis.

The phosphorylation of S112, S123, S129, and S198 in the N-terminus of SMARCAL1 indicate the potential for cell-cycle dependent regulation of SMARCAL1. In fact, S198 phosphorylation was identified in a large-scale proteomic analysis of mitotic phosphorylation (146). Additionally, a peptide containing phosphorylated S172 and S181 was identified among 494 unique phosphopeptides that increased in samples from cells exposed to low doses of irradiation (119). Phosphopeptide mapping of SP-phospho-mutants will reveal to what extent these sites contribute to SMARCAL1 phospho-regulation. Maps of SMARCAL1 purified from synchronous populations of cells will indicate whether phosphorylation depends on cell cycle.

We also uncovered at least two major sites of phosphorylation located on the same peptide as S919 that are phosphorylated in SMARCAL1 purified from normal growing conditions. This suggests that DDR-independent phospho-regulation of the C-terminus is also important for SMARCAL1 functions. There are 14 potential sites in the S919 peptide and identifying which sites are

phosphorylated will be important to understand the functional requirements for DDR-independent and DDR-dependent phosphorylation in this region.

A C-terminal auto-inhibitory domain is regulated by S889 phosphorylation

SMARCAL1 is phosphorylated in the absence of DNA damage. S889 phosphorylation was identified in our mass spectrometry analysis and I determined using p-S889 antibodies that this site is phosphorylated independent of DNA damage. Functional characterization of S889 mutants in vitro and in cells suggests that S889 phosphorylation may regulate a C-terminal auto-inhibitory domain of SMARCAL1.

The ΔC mutant in these studies truncates SMARCAL1 following aa860 (Figure 7.4A). Further defining this domain is important for studying its inhibitory function. New mutants generated from scanning alanine mutagenesis or truncating various lengths from the C-terminus (Figure 7.4B) can be rapidly screened in the ATPase assay to determine the minimal domain required to inhibit SMARCAL1 activity. Also, our analysis indicates that the 860-end truncation also removes an NLS. Defining the minimal inhibitory domain may circumvent this issue. If the NLS is within the defined domain, then it will be important to use expression vectors that add a NLS. An NLS vector was sufficient to localize ΔC to the nucleus.

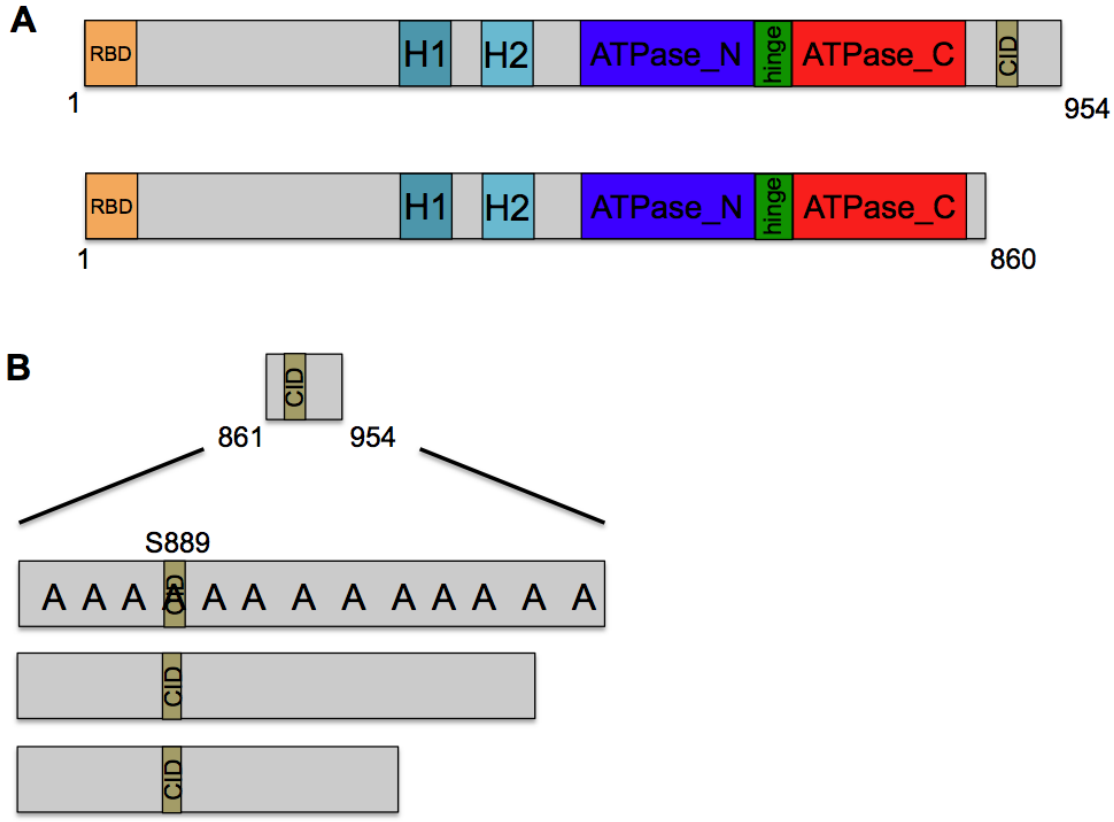

Figure 7.4. Strategy for determining the minimal CID in the C-terminus of SMARCAL1. (A) The ΔC mutant truncates the C-terminus 861-954. (B) This domain contains a CID and also a NLS. A scanning alanine mutagenesis strategy can identify important residues in this region that contribute to CID function. Alternatively, smaller truncations can be made from the C-terminus..

Phosphorylation by DDR kinases

I identified three damage-inducible phosphorylation sites in SMARCAL1: S173, S652, and S919. As determined by phosphopeptide mapping, phosphorylation of S173 and S919 following HU is dependent on DDR kinases. Likewise, detection of S652 phosphorylation by p-S652 antibodies is reduced after DDR kinase inhibition. ATR, ATM, and DNA-PK directly phosphorylate SMARCAL1 in vitro. However, I did not determine whether S173, S652, or S919 are directly targeted by these kinases. Determining whether ATM, ATR, and DNA-PK directly phosphorylate these sites in vitro and whether there is any kinase specificity for each residue may shed light on the functional significance of these sites in cells.

ATR is activated early at stalled replication forks while ATM and DNA-PK activation depend on fork collapse, or DSB formation. If there is any selectivity for S173, S652, or S919, this may suggest DDR kinases regulate SMARCAL1 activity at different stages in replication fork processing. ATM/ATR selectivity has been observed for phosphorylation sites in the replication stress response helicase, WRN (126). The model I propose for DDR-regulated phosphorylation assumes that phosphorylation at S173, S652, and S919 contribute to the same effect: limiting SMARCAL1 activity at stalled replication forks. I have inferred most of my model based on the observation that SMARCAL1 is less active when purified from cells treated with HU for 16hrs. My findings could ultimately reflect the culmination of many steps of regulation at different times in the response to

HU-stalled forks that ultimately lead to less active SMARCAL1. However, the potential for positive regulatory effects of DDR phosphorylation is still possible.

There is at least one additional DDR-dependent phosphorylation site in SMARCAL1 that does not fit the S/TQ consensus. Phosphopeptide-*a* is reduced following DDR kinase inhibition but is not one of the ten S/TQs present in SMARCAL1. This site may be a substrate for another kinase activated by ATM, ATR, or DNA-PK. Alternatively, phosphopeptide-*a* could be one of the SP sites identified by mass spectrometry and increases with 16hrs in HU because this essentially synchronizes cells in S phase.

Hypotheses for S173 and S919 phospho-regulation

Whether or not S173 and S919 phosphorylation represent functionally important regulatory events is unclear. I determined that S173 and S919 do not affect SMARCAL1 ATPase activity in vitro. However, this interpretation is largely based on the assumption that the S173D and S919D mutants sufficiently mimic phosphorylation at these sites. Our evidence that the C-terminus contains a CID that is relieved by S889 phosphorylation suggests that the C-terminus undergoes conformational changes relevant to SMARCAL1 enzymatic activities. It is possible that DDR-dependent phosphorylation of S919 can also contribute to changes in the C-terminal domain structure.

Alternatively, S919 phosphorylation could create a binding site for a regulatory protein. We observe a decrease in pS889 following HU-treatment. S919 phosphorylation may promote recruitment of a phosphatase that removes

S889 phosphorylation. Removing pS889 would support the DDR-regulated inactivation of SMARCAL1, restoring the CID to an inhibitory conformation (Figure 7.5). The S919D mutant does not affect the ATPase of SMARCAL1 in vitro suggesting that this mutant does not affect pS889 status. Although the protein is purified from human cells, we overexpress FLAG-SMARCAL1 at very high levels and do not expect that the protein is regulated efficiently. It is not likely that a phosphatase would be present in high enough levels to remove all of pS889. The relationship between S919 and S889 phosphorylation should be investigated in conditions that express low levels of SMARCAL1.

It is less obvious how phosphorylation of SMARCAL1 in the N-terminus would regulate SMARCAL1 activities. At least at the primary amino acid sequence level, S173 is closest to the first HARP domain in SMARCAL1. We have no evidence that it affects HARP function. S173 phosphorylation presumably occurs in a region that is phosphorylated at many other sites. Maybe S173 phosphorylation promotes regulation of these sites following DNA damage. It is just as likely that S173 controls protein-protein interactions required for SMARCAL1 regulation. We synthesized biotin-conjugated S173 peptides and phosphopeptides that can be used for streptavidin purification of proteins that specifically bind the S173 phosphopeptide in nuclear extracts. The same should be done for S919 peptides. This methodology bypasses the need for upstream regulatory events and will hopefully reveal some aspect of S173 and S919 function.

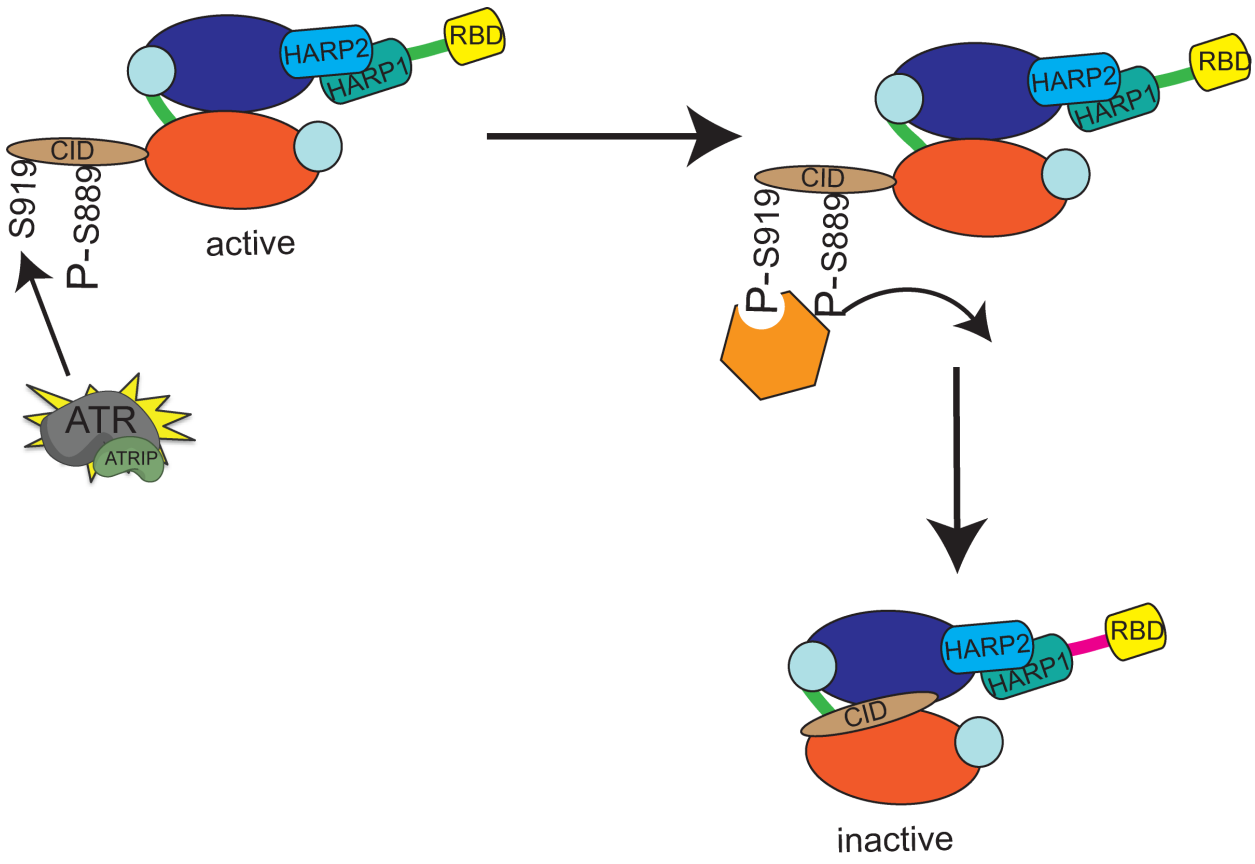

Figure 7.5. Model for S919-dependent regulation of the CID. Phosphorylation of S919 in the C-terminus could regulate CID function. Activation of ATR at stalled replication forks leads to the phosphorylation of S919. Phospho-S919 recruits a phosphatase (orange polygon) to S889 to remove phosphorylation. Following de-phosphorylation of S889 the CID returns to an inhibitory conformation leading to inactive SMARCAL1.

S652 phosphorylation regulates the SMARCAL1 hinge domain

The hinge domain is a flexible linker that enables the ATPase_N and ATPase_C domains of SNF2 proteins to rotate into a closed-active conformation. The rotation of these domains is necessary for ATP hydrolysis and efficient DNA translocation. Thus, changes in the flexibility of this hinge can affect the activities of SNF2 enzymes. The importance of this domain in SMARCAL1 is emphasized by the fact that a number of SIOD-associated mutations occur at hinge domain residues. Introduction of a phospho-mimetic amino acid at S652 disrupts the ATPase activity of SMARCAL1 in vitro. Since an alanine at this position has no affect on ATPase activity, we conclude that the serine itself is not required for activity; rather introducing a negatively charged side group or phosphate at this site in vivo most likely disrupts critical hinge domain functions. We observe damage-induced phosphorylation of S652 indicating that it is regulated in cells and may disrupt hinge function. However, evidence that S652 is the inactivating phosphorylation that contributes to the decreased activity of SMARCAL1 purified from HU-treated cells remains circumstantial. Further studies are necessary to provide a direct link between DDR-regulated phosphorylation at S652 and the observed decrease in SMARCAL1 from HU treated cells.

A comprehensive model for SMARCAL1 phospho-regulation at stalled replication forks

SMARCAL1 acts as an annealing helicase at forks where the polymerase and helicase activities have been uncoupled due to DNA lesions or other perturbations. This activity is important to prevent degeneration of the fork into a

MUS81-dependent double-strand break. The absence of this activity contributes to the human disease SIOD. Furthermore, too much SMARCAL1 activity is equally challenging to genome maintenance. I have identified a number of phosphorylation events that would be excellent candidates for regulating SMARCAL1 activities at replication forks. I have summarized my interpretations and predictions for SMARCAL1 regulation during the replication stress response in a comprehensive model (Figure 7.6). Although many steps are quite speculative, the general theme is that phosphorylation provides an 'on-off' switch for SMARCAL1 activities at stalled replication forks.

The SMARCAL1-RPA interaction localizes annealing helicase activity to stalled replication forks. This step in SMARCAL1 regulation is critical for its genome maintenance function. SMARCAL1 is likely recruited to forks in its 'on' conformation with pS889 relieving the CID. Once localized, I favor the model that SMARCAL1 binds the parental fork DNA and re-anneals the parental strands to promote the first steps of fork reversal. After SMARCAL1 binds DNA and ATP, the rotation of the ATPase_N and ATPase_C domains drives ATP hydrolysis and subsequent translocation. The HARP domains of SMARCAL1 are important for converting this force into strand annealing and RPA displacement.

ATR is activated at stalled replication forks. During the translocation step, the hinge domain of SMARCAL1 is exposed to ATR kinase activity as a result of the conformational change upon DNA and ATP binding. Phosphorylation of S652 reduces the flexibility of the hinge domain, inhibiting another round of ATP hydrolysis for this molecule of SMARCAL1. The disruption of the hinge domain

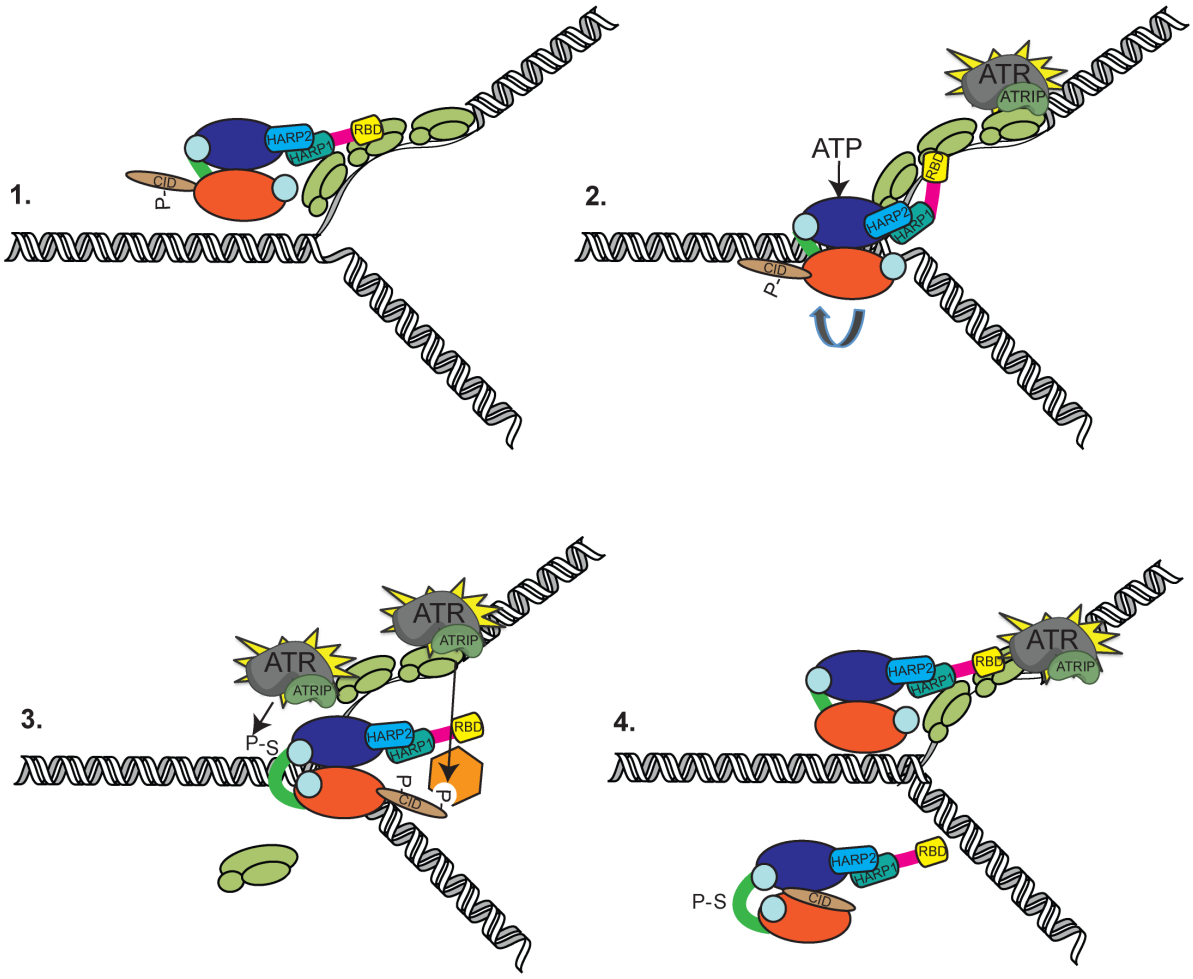

Figure 7.6. A comprehensive model for SMARCAL1 phospho-regulation at stalled replication forks. 1. In normal conditions S889 phosphorylation is maintained, relieving the CID inhibitory conformation. In the presence of replication stress, SMARCAL1 accumulates at stalled replication forks in its active form via its N-terminal RPA binding domain (RBD). (2) Localization to forks concentrates SMARCAL1 near its DNA substrates and subsequently SMARCAL1 binds the parental fork DNA. (3) DNA binding promotes ATP binding and transition from the open-inactive state to the closed-active state. ATP hydrolysis drives SMARCAL1 translocation on the DNA to re-anneal the parental strands while removing RPA. The conformational change exposes S652 to DDR kinases like ATR, which is activated at stalled replication forks. Phosphorylation of S652 in the hinge domain hinders ATPase domain flexibility thereby inhibiting the next enzymatic cycle of ATP binding and hydrolysis. Additionally, S919 is phosphorylated by ATR and recruits a phosphatase to remove S889 phosphorylation. This reactivates the CID and further inhibits SMARCAL1 activity at the fork. (4) The inactivated molecule dissociates and a new molecule of SMARCAL1 can bind. This mechanism provides the opportunity for SMARCAL1 to catalyze necessary fork remodeling activities, but with limited processivity.

may also affect DNA binding and reduce the maintenance time of SMARCAL1 on its substrate.

ATR most likely phosphorylates S919 in the C-terminal domain as well. Our evidence suggests that significant damage-induced phosphorylation does not occur prior to DNA binding. It is unclear whether S919 would depend on a DNA-dependent conformational change to be exposed to ATR kinase activity, considering the C-terminal inhibitory domain may already be in an open conformation. The order of events is unclear, but if S919 did regulate the status of S889 phosphorylation I would expect a mechanism to be in place to prevent S919 phosphorylation until after SMARCAL1 translocates. If ATR phosphorylates S919 upon SMARCAL1 localization, SMARCAL1 would be turned 'off' before even binding DNA. S919 may also be masked until after DNA binding. Following translocation S919 is phosphorylated and activates the CID. The CID could further disrupt DNA binding and promote removal of SMARCAL1 from the fork DNA.

Testing this model is anything but simple. A major challenge of projects related to phosphorylation is the difficulty in interpreting the phenotypes of phospho-mutants. Phospho-mimetic mutants may or may not mimic phosphorylation. For example, S173D and S919D may be active in vitro because they are not phospho-mimetic. Although we can make predictions based on the location of these sites in the primary sequence, we have little understanding of how the accessory domains of SMARCAL1 interact with each other or the ATPase motor. SMARCAL1 regulation is obviously complex and the accessory

domains seem to influence the ATPase domain more than expected. Further characterization of SMARCAL1 interacting proteins and even phospho-peptide interacting proteins will surely clarify more of the mechanisms of S173 and S919 phosphorylation. As for S652, evidence strongly supports that phosphorylation in the hinge domain is inhibitory. Structural information is necessary to validate this mechanism.

My thesis work has generated a number of testable models for SMARCAL1 regulation in normal and replication-stressed conditions. The overexpression phenotypes of SMARCAL1 demonstrate that SMARCAL1 activities must be limited at replication forks. However, SMARCAL1 deficiencies indicate that SMARCAL1 activity is required to promote fork stability. The activities of SMARCAL1 must be unique enough to warrant recruitment of an enzyme that could potentially threaten genome stability if not regulated.

APPENDIX A

RAW DATA FOR OVEREXPRESSION SCREEN

This appendix lists the 97 genes identified in the cDNA overexpression screen in the Appendix Table. The table includes the quantitation of γH2AX foci and calculated *p* values.

Table S4. Mean γH2AX foci in GFP-positive U2OS cells following co-transfection of pCMV-SPORT6 cDNA and GFP expression plasmids. The *p* value was calculated in comparison to an empty vector control using an unpaired, two-tailed t-test and was adjusted for a false discover rate of 0.05 using the Benjamini and Hochberg procedure (n≥3).

Gene ID	Gene Name	Symbol	Mean	Std Dev	adj. *p* value
N/A	Control (empty vector)	N/A	0.87	0.92	N/A
3005	H1 histone family, member 0	H1F0	2.33	1.51	0.016
7730	zinc finger protein 177	ZNF177	2.33	0.58	0.023
27237	Rho guanine exchange factor (GEF) 16	ARHGEF16	2.33	1.53	0.043
56033	BARX homeobox 1	BARX1	2.33	1.53	0.043
63979	fidgetin-like 1	FIGNL1	2.33	1.53	0.042
10399	guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1	GNB2L1	2.43	1.72	0.014
4292	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	MLH1	2.67	1.03	0.001
10514	MYB binding protein (P160) 1a	MYBBP1A	2.67	0.58	0.007
843	caspase 10, apoptosis-related cysteine peptidase	CASP10	2.67	1.53	0.016
94	activin A receptor type II-like 1	ACVRL1	2.67	2.08	0.028
5709	proteasome (prosome, macropain) 26S subunit, non-ATPase, 3	PSMD3	2.75	1.50	0.007
373156	glutathione S-transferase kappa 1	GSTK1	2.75	2.22	0.020
26260	F-box protein 25	FBXO25	3.00	1.00	0.003
27166	PRELI domain containing 1	PRELID1	3.00	1.83	0.005
7913	DEK oncogene (DNA binding)	DEK	3.00	3.00	0.030
6137	ribosomal protein L13	RPL13	3.29	3.50	0.023
3783	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4	KCNN4	3.33	0.58	7.43E-04
5307	paired-like homeodomain 1	PITX1	3.33	0.58	7.32E-04
64288	zinc finger protein 323	ZNF323	3.33	1.53	0.002
6128	ribosomal protein L6	RPL6	3.67	2.31	0.003
63924	cell death-inducing DFFA-like effector c	CIDEC	3.67	2.52	0.003
10473	high mobility group nucleosomal binding domain 4	HMGN4	3.75	0.96	7.23E-05
4286	microphthalmia-associated transcription factor	MITF	3.75	3.50	0.010
7026	nuclear receptor subfamily 2, group F, member 2	NR2F2	4.00	1.00	1.27E-04
7737	ring finger protein 113A	RNF113A	4.33	1.53	1.14E-04
255252	leucine rich repeat containing 57	LRRC57	4.33	2.89	0.001
90441	zinc finger protein 622	ZNF622	4.33	3.21	0.002
5925	retinoblastoma 1 (including osteosarcoma)	RB1	4.67	2.08	1.32E-04
128866	chromatin modifying protein 4B	CHMP4B	4.67	2.89	6.40E-04
220202	atonal homolog 7 (Drosophila)	ATOH1	4.67	3.21	0.001
2150	coagulation factor II (thrombin) receptor-like 1	F2RL1	4.75	3.40	0.001
7022	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma)	TFAP2C	5.00	2.65	1.76E-04
8341	histone cluster 1, H2bn	HIST1H2BN	5.22	2.22	2.07E-06
55957	lin-37 homolog (C. elegans)	LIN37	5.33	1.15	3.20E-06
148327	cAMP responsive element binding protein 3-like 4	CREB3L4	5.33	1.15	3.27E-06
7704	zinc finger and BTB domain containing 16	ZBTB16	5.33	2.08	2.49E-05
9040	ubiquitin-conjugating enzyme E2M (UBC12 homolog, yeast)	UBE2M	5.67	0.58	5.55E-07
58492	zinc finger protein 77	ZNF77	5.67	1.53	3.06E-06
6208	ribosomal protein S14	RPS14	5.67	4.37	8.47E-04
3217	homeobox B7	HOXB7	5.67	4.55	0.001
3020	H3 histone, family 3A	H3F3A	5.83	3.06	0.002
65117	arginine/serine-rich coiled-coil 2	RSRC2	6.00	5.74	0.003
3428	interferon, gamma-inducible protein 16	IFI16	6.25	4.92	8.48E-04
8738	CASP2 and RIPK1 domain containing adaptor with death domain	CRADD	6.25	5.44	0.002
210	aminolevulinate, delta-, dehydratase	ALAD	6.33	0.58	1.04E-07
4200	malic enzyme 2, NAD(+)-dependent, mitochondrial	ME2	6.33	3.51	6.48E-05
79576	NFKB activating protein	NKAP	6.67	4.04	9.68E-05
2115	ets variant gene 1	ETV1	7.33	4.16	3.91E-05
3183	heterogeneous nuclear ribonucleoprotein C (C1/C2	HNRPC	7.33	4.93	1.44E-04
3280	hairy and enhancer of split 1, (Drosophila)	HES1	7.33	4.93	1.46E-04
1747	distal-less homeobox 3	DLX3	8.33	3.06	5.92E-07
3202	homeobox A5	HOXA5	8.67	3.06	3.45E-07
23764	v-maf musculoaponeurotic fibrosarcoma oncogene homolog F (avian)	MAFF	8.67	4.62	1.15E-05
83444	zinc finger, HIT type 4	ZNHIT4	8.67	7.37	8.40E-04
121599	Spi-C transcription factor (Spi-1/PU.1 related)	SPIC	9.67	1.53	1.30E-09
971	CD72 molecule	CD72	10.00	2.94	1.47E-08
5684	proteasome (prosome, macropain) subunit, alpha type, 3	PSMA3	10.00	3.61	1.80E-07
1746	distal-less homeobox 2	DLX2	10.33	0.58	7.35E-11
4332	myeloid cell nuclear differentiation antigen	MNDA	10.50	5.80	9.28E-06
5347	polo-like kinase 1 (Drosophila)	PLK1	10.67	2.31	2.54E-09
339488	transcription factor AP-2 epsilon (activating enhancer binding protein 2 epsilon)	TFAP2E	10.67	2.52	4.12E-09
51564	histone deacetylase 7	HDAC7A	11.00	1.73	3.53E-10
126789	pseudouridylate synthase-like 1	PUSL1	11.00	4.97	8.84E-07
4998	origin recognition complex, subunit 1-like (yeast)	ORC1L	12.33	0.58	5.39E-12
2114	v-ets erythroblastosis virus E26 oncogene homolog 2 (avian)	ETS2	12.33	2.52	5.51E-10
27345	potassium large conductance calcium-activated channel, subfamily M, beta member 4	KCNMB4	13.00	3.00	9.49E-10
55506	H2A histone family, member Y2	H2AFY2	13.33	2.08	1.88E-11
9324	high mobility group nucleosomal binding domain 3	HMGN3	13.67	1.53	7.92E-12
4152	methyl-CpG binding domain protein 1	MBD1	14.00	4.00	4.20E-09
55621	TRM1 tRNA methyltransferase 1 homolog (S. cerevisiae)	TRMT1	15.00	4.36	3.81E-09
6625	small nuclear ribonucleoprotein 70kDa polypeptide (RNP antigen)	SNRP70	16.33	3.79	3.15E-10
2119	ets variant gene 5 (ets-related molecule)	ETV5	18.00	7.00	5.47E-08
5494	protein phosphatase 1A (formerly 2C), magnesium-dependent, alpha isoform	PPM1A	19.33	2.89	2.23E-12
50485	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a-like 1	SMARCAL1	19.56	5.20	2.08E-11
3178	heterogeneous nuclear ribonucleoprotein A1	HNRPA1	20.00	1.73	4.31E-14
2774	guanine nucleotide binding protein (G protein), alpha activating activity polypeptide, olfactory type	GNAL	20.00	6.56	5.72E-09
10949	heterogeneous nuclear ribonucleoprotein A0	HNRPA0	21.17	2.99	2.48E-14
27440	cat eye syndrome chromosome region, candidate 5	CECR5	21.33	5.51	3.39E-10
80712	ESX homeobox 1	ESX1	22.33	11.06	6.61E-07

2118	ets variant gene 4 (E1A enhancer binding protein, E1AF)	ETV4	23.67	2.08	2.47E-14
1994	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 (Hu antigen R)	ELAVL1	24.67	7.77	2.49E-09
5883	RAD9 homolog A (S. pombe)	RAD9A	25.33	11.24	1.57E-07
55249	YY1 associated protein 1	YY1AP1	25.67	11.02	1.05E-07
81620	chromatin licensing and DNA replication factor 1	CDT1	27.00	4.24	2.04E-13
3191	heterogeneous nuclear ribonucleoprotein L	HNRPL	27.00	6.00	3.29E-11
126272	EP300 interacting inhibitor of differentiation 2B	EID-3	30.33	8.74	6.12E-10
54828	breast carcinoma amplified sequence 3	BCAS3	31.67	3.21	1.24E-14
22827	poly-U binding splicing factor 60KDa	PUF60	32.00	7.79	5.10E-11
284004	hexosaminidase (glycosyl hydrolase family 20, catalytic domain) containing	HEXDC	33.67	4.04	2.20E-14
8139	giant axonal neuropathy (gigaxonin)	GAN	35.00	4.36	2.39E-14
51125	golgi autoantigen, golgin subfamily a, 7	GOLGA7	36.50	5.13	7.65E-14
9937	DNA cross-link repair 1A (PSO2 homolog, S. cerevisiae)	DCLRE1A	38.17	14.43	1.04E-08
92335	protein kinase LYK5	LYK5	39.86	6.84	2.26E-14
6434	splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila)	SFRS10	41.33	5.51	2.37E-14
2078	v-ets erythroblastosis virus E26 oncogene homolog (avian)	ERG	50.67	5.51	7.15E-15
6432	splicing factor, arginine/serine-rich 7, 35kDa	SFRS7	52.00	22.52	7.46E-08

APPENDIX B

DEFINING GENOME MAINTENANCE PATHWAYS USING FUNCTIONAL GENOMIC APPROACHES[3]

Introduction

Our knowledge of the genome maintenance activities that happen in human cells comes from the study of all kingdoms of life since most of these activities are highly evolutionarily conserved. Genetic and biochemical studies of bacteria provide insights into replication and DNA repair proteins that also operate in eukaryotic systems. *Archaebacteria* provide a rich resource for structural biologists since it is often easier to purify and crystallize proteins from the thermophilic species. The genetics of simple eukaryotes like *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* that can divide as either haploid or diploid cells provided some of our first mechanistic understanding of the cell division cycle (148,149).

Higher eukaryotes including the invertebrates *Caenorhabditis elegans* and *Drosophila*, as well as the vertebrates *Xenopus* and zebrafish have unique strengths as research models. For example, Xenopus eggs contain highconcentrations of genome maintenance proteins and have been used to

[3] The majority of this chapter has previously been published in reference 147.
Bansbach, C.E. and Cortez, D. (2011) Defining genome maintenance pathways using functional genomic approaches. *Crit Rev Biochem Mol Biol*, **46**, 327-341.

understand the mechanisms of DNA damage checkpoints using complex biochemical strategies. Finally, mammalian systems have been historically important and are increasingly providing resources for new discovery.

Some genome maintenance discoveries came directly from human systems through bedside to bench research. This is especially true in the DNA damage response field since many human diseases are caused by mutations in DDR genes. For example, clinicians studying patients with the cancer predisposition disease ataxia telangiectasia described how at the cellular level this disease is characterized by an inability to properly arrest the cell cycle in response to ionizing radiation (150) (151). These were some of the first descriptions of cell cycle checkpoints. Positional cloning of the gene responsible for ataxia telangiectasia (*ATM*), identified one of the most important kinases that signals in response to DNA DSBs (152).

Other notable examples include the identification of DSB repair and signaling proteins like breast and ovarian cancer susceptibility protein 1 (BRCA1) and Nijmegen breakage syndrome protein (NBS1) (153-156); the identification of the DNA crosslink repair proteins defective in patients with Fanconi anemia (FANC proteins) (157); and the description of proteins involved in the repair of ultraviolet (UV)-radiation damage via nucleotide excision repair pathways from people with xeroderma pigmentosum (XP proteins) (158). The primary limitations of this approach for understanding genome maintenance are that it is time and resource-intensive and restricted by the size of the human population.

Yeast systems historically provided the simplest methods to identify genes involved in genome maintenance pathways. Classic mutagenesis screens including the Hartwell and Nurse cell division cycle screens (148,149,159), DNA damaging agent sensitivity screens (160-162), and mitotic arrest deficient screens (163,164) generated lists of yeast strains that were converted to lists of genes as the mutant alleles were identified. As DNA sequencing provided a complete sequence of the yeast genome, these mutagenesis screens gave way to more systematic functional genomic approaches based on libraries of yeast strains with targeted knockouts in every open reading frame (ORF) or tags integrated at every ORF to allow high-throughput protein-protein interaction and localization studies (165-167).

Recent technological advances including high-density microarrays, mass spectrometry, and especially RNA interference (RNAi) have made functional genomic approaches feasible in diploid organisms including human cells. In many cases, these approaches have been copied from the work done in yeast. For example, cell division cycle (168-171), mitotic arrest (172), and DNA damage sensitivity screens (173-176) have all been completed with RNAi approaches in cell culture (see Table 3.1 for a list of screens discussed in this review).

Some of the experimental advantages of human cell culture such as the ease of imaging are now being exploited to complete novel high content microscopy screens (66,177-181). Furthermore, other screens like the proteomic identification of hundreds of ATM substrates have been pioneered in human cell culture and only later translated into yeast (88,115,116,182). Overall, these high-

throughput approaches have identified hundreds of new proteins in genome maintenance pathways and rapidly accelerated genome maintenance research.

In this Appendix, I will review functional genomic approaches to identify and characterize genome maintenance pathways. A brief introduction of each approach will be followed by specific examples of their successful use in discovery. I will focus primarily on approaches in diploid, multi-cellular eukaryotic system and direct readers interested in functional genomics using yeast systems to other excellent reviews (183,184).

Gene expression profiling

One of the first functional genomic approaches used in higher eukaryotic organisms was gene expression profiling (185). The general idea is to track transcriptional changes in response to a genome maintenance challenge. This is most commonly done using microarrays to measure mRNA abundance. Many investigators analyzed relative mRNA expression levels as a function of cell cycle position, time after exposure to a DNA- or spindle-damaging agent, or genetic background (i.e. p53 overexpression or ATM-deficiency) (186-188).

There are two rationales for gene expression profiling in the interrogation of genome maintenance pathways. First, genes with activities in a specific cell cycle phase or in response to DNA damage may be more highly expressed in cells under these conditions. Second, the outcome of many genome maintenance pathways such as the DDR depends on a change in transcription,

so tracking transcriptional changes can be used as an assay to monitor signaling through these pathways.

A well-studied example is the G1 arrest or apoptosis induced by p53 in response to DNA damage, which is mediated by p53-dependent transcriptional changes in multiple genes (189,190). DNA damage leads to activation of p53 through post-translational modifications that promote its stabilization. By identifying p53-target genes, investigators have learned how cells regulate the G1-S transition and apoptosis. These studies have been complemented by the identification of p53 response elements throughout the genome using chromatin immunoprecipitation experiments. The reader is directed to several recent reviews to learn more about these methods to study p53 function (190,191).

The limitations of this type of approach are that regulation of genome maintenance pathways often proceeds through transcriptional-independent mechanisms and that many transcriptional changes may be only indirectly related to the process being studied. Proteins like ATM that initiate signaling in response to DNA damage are largely unregulated at the transcriptional level. Also, many transcriptional changes occur as a general reaction to any type of environmental stress so they may not be particularly useful in dissecting a specific pathway (192). Finally, expression profiling provides little information on the function of the identified genes.

High throughput proteomics

Proteomic approaches to studying genome maintenance pathways include techniques that measure protein abundance, protein-protein interactions, post-translational modifications, and/or subcellular localization. Assays include protein microarrays, yeast two-hybrid, mass spectrometry, and fluorescent imaging. These are inherently more technically difficult, expensive, and lower throughput than expression profiling. However, they offer the significant advantage of easier functional interpretation.

Protein microarrays

Protein microarrays or protein chips are arrays of proteins or peptides deposited on a glass or silicon surface (193). They are useful to measure protein abundance, protein-protein interactions, and post-translational modifications (Figure B.1). This technology has seen only limited use largely due to issues of cost and robustness of the assays. However, a recent publication on the use of protein microarrays to identify ubiquitination substrates of the anaphase promoting complex (APC) highlights the potential of this emerging technology (194).

The APC is an E3 ubiquitin ligase that polyubiquitinates proteins to regulate the cell cycle (195). For example, APC-dependent ubiquitination of cyclin B is responsible for the cycling of its abundance and required for exit from mitosis (196-198). The Kirschner group used what they termed extract-based

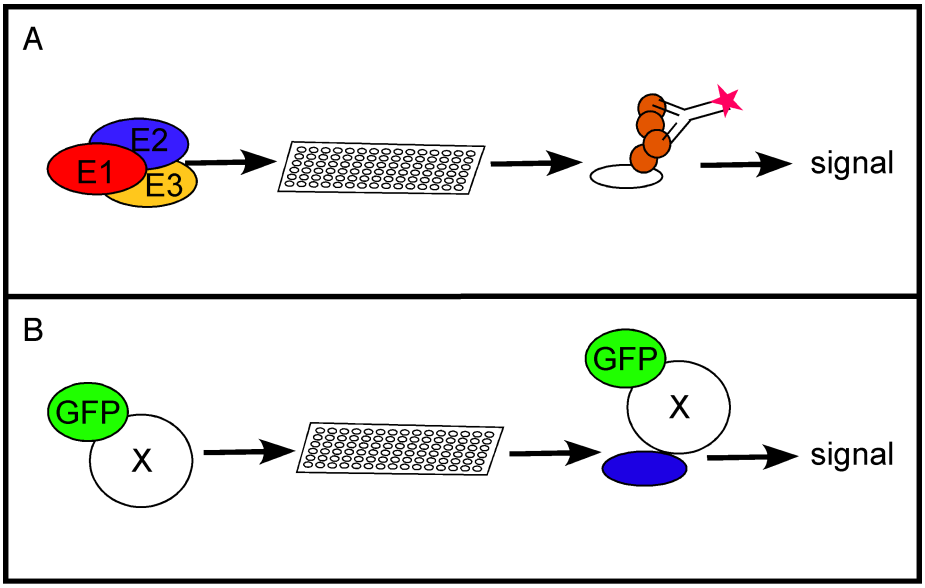

Figure B.1. Protein microarrays. (A) Protein microarrays can be used to identify post-translational modifications. For example, Merbl et al., identified substrates of the APC ubiquitin conjugating enzyme using protein microarrays (194). **(B)** Protein microarrays can also be used to identify protein-protein interactions using fluorescently tagged proteins as the probe.

functional assays combined with protein microarrays to define APC substrates (194). They isolated extracts from HeLa cells that were arrested at the mitotic checkpoint. These extracts catalyzed the ubiquitination of protein spotted on a solid surface that was detectable using an anti-polyubiquitin antibody. They defined APC substrates by comparing the activities of these extracts on the protein microarray to extracts in which the APC was inactivated. Of the 8,000 proteins arrayed on the chips, they found that 132 were differentially modified.

To validate the methodology for the identification of APC substrates, they selected six mitotic proteins from their list that were previously unknown targets of the APC complex (Nek9, Calm2, p27, RPS6KA4, cyclin G2, and DDA3) and demonstrated through more standard assays that they were indeed APC substrates. In addition, four proteins not previously linked to mitosis (Zap-70, MAP3K11, RPL30, and Dyrk3) were chosen for validation. Strikingly, two of four (RPL30 and Dyrk3) were also shown to be APC substrates. RPL30 is a component of the 60S ribosome. Dyrk3 is a dual specificity protein kinase that has functions in regulating cell survival in response to energy stresses and in modulating erythropoiesis (199,200). It is unclear why these proteins would be regulated by the APC, but their identification in this screening effort points to either an unappreciated function for them in the cell cycle or a broader function for the APC.

This example provides several lessons about the use of protein microarrays. First, the power of working directly with proteins is that results immediately lead to some mechanistic understanding of the pathway. The

example also demonstrates that complex cell extracts can be used to recapitulate a genome maintenance activity. This provides several advantages especially when a genome maintenance activity has yet to be fully reconstituted with purified proteins. However, it also illustrates the current limitations of this technology. The most obvious limitation is the current protein chips, which contain only a small fraction of the proteome. It is also difficult to know the percentage of active, folded, and sterically unhindered proteins retained on the microarray.

Methods for identifying protein-protein interactions

Many genome maintenance pathways depend on assemblies of multi-subunit protein machines. Thus, the identification of protein-protein interactions is a primary method for discovering new genome maintenance proteins and understanding their regulation. While there are many methodologies to accomplish this goal, the standard methods rely on interrogating proteins or complexes individually in a one protein per project fashion. More recently, investigators have worked to convert these approaches into higher throughput formats and to create a full “interactome” description of protein-protein interactions that occur within a cell (201).

The two most common interactome methodologies are yeast two-hybrid (Y2H) and affinity purification combined with mass spectrometry. Efforts began in yeast systems (202) and expanded to other organisms. One impediment to using these methods in higher eukaryotes is the need for full-length ORFs cloned

into appropriate vectors. Fortunately, orfeome cloning projects using recombination-based vectors are rapidly removing this obstacle. Y2H is a sensitive method to detect binary interactions (Figure B.2) while affinity purification followed by mass spectrometry (Figure B.3) is better suited to identify stable protein complexes that occur in the native context.

Vidal and colleagues used interactome mapping to generate a protein-protein interaction map for proteins involved in the DDR in *Caenorhabditis elegans* (203). They focused on 75 worm orthologs of known DDR proteins, and initially used Y2H to generate pairwise protein interaction maps among these proteins. From this data they concluded that they had approximately a 50% success rate in identifying predicted interactions. More importantly they used 67 of the genes to perform unbiased proteome-wide Y2H screens. A total of 165 interacting sequences were recovered, many encoding novel proteins not previously linked to the DDR.

This example illustrates both the power of interactome mapping and the current deficiencies in the technologies. The major limitation of the technology is that it is not easily scaled to a high-throughput approach. Thus, we only see a very limited portion of the interactome landscape in any single study. A second problem is the issue of false-positives. Each of the “hits” from a Y2H or mass spectrometry experiment can only be considered a potential protein-protein interaction. Performing these experiments with large numbers of baits helps eliminate the false-positives through the identification of common contaminants. However, validation using an independent method is essential. Finally, false-

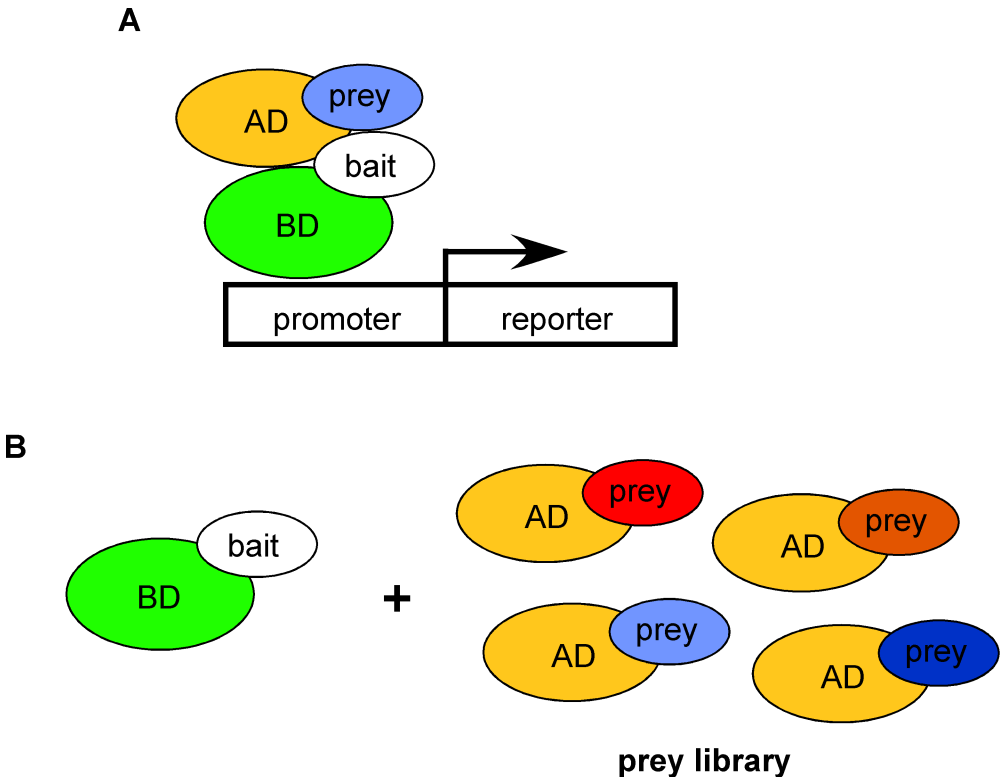

Figure B.2. Yeast two-hybrid (Y2H) methods for generating interactomes. Yeast 2 hybrid screens detect interactions between bait proteins and prey proteins. Although useful to interrogate pairs of proteins (B1), these screens are easily easily expanded to screen cDNA libraries to generate interactome maps (B2)

negatives will be difficult to avoid using any single method so combining results from multiple methods will be important.

Mass spectrometry methods for identifying post-translational modifications

In addition to identifying protein complexes, mass spectrometry can be used to detect post-translational modifications such as phosphorylation sites (Figure B.3). Unbiased phosphoproteomic screens have yielded large databases of thousands of phosphorylation sites (see (204) and references therein). Typically, these studies employ a phospho-protein enrichment strategy using technologies such as metal affinity chromatography. More recently, they have been combined with quantitative mass spectrometry methodologies to compare cells or tissues from two or more conditions.

This methodology was applied to identify cell cycle-related phosphorylation (205). In this study, the authors combined stable isotope labeling in cell culture (SILAC) with affinity purification of kinases using resins attached to five kinase inhibitors. They differentially labeled cells synchronized in S or M phase with isotopes, performed the kinase enrichment strategy, and after trypsinization, enriched phosphopeptides prior to identification by mass spectrometry. This technique yielded 1182 phosphopeptides about 50% of which are upregulated in mitotic cells and 10% upregulated in replicating cells. Many new sites on known mitotic regulatory proteins were described including sites on CDC2, PLK1, and AurB. Further experiments will be needed to understand their functional significance. The large number of changes in phosphopeptides

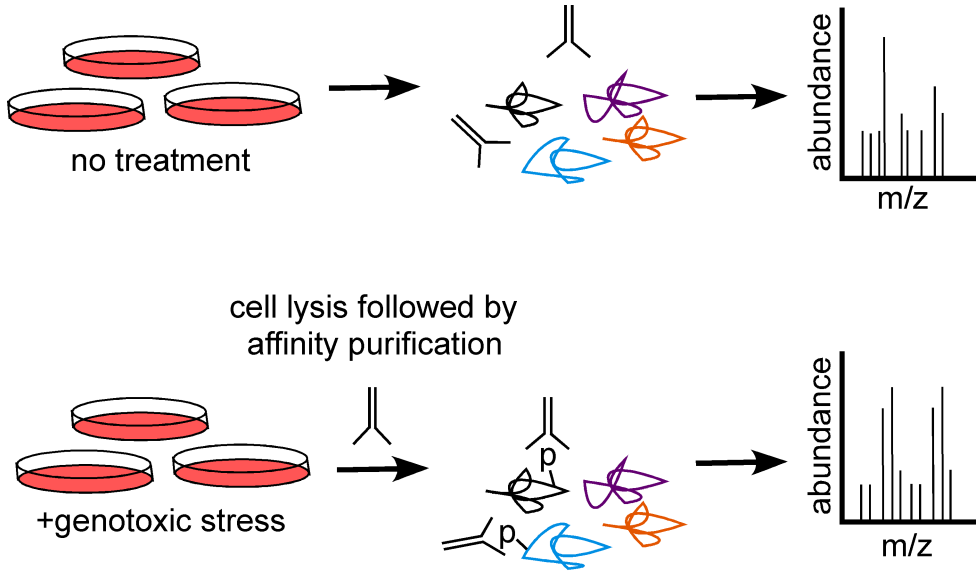

Figure B.3. Affinity purification coupled to mass spectrometry. This proteomic approach is useful to interrogate protein complexes purified from their native cellular context or as depicted, to identify differential post-translational modifications following genotoxic stress.

identified indicates that progression through mitosis involves large changes in the phospho-proteome.

A second example comes from the study of the kinases that operate in the DNA damage response. The DDR kinases ATM and ATR have a strong preference for phosphorylating serines or threonines that are followed by glutamines (206). This bias can be used to pinpoint putative phosphorylation sites within known substrates. More importantly, it can be exploited to identify novel ATM and ATR substrates. The first example where this strategy was used successfully was our identification of MCM2 and MCM3 as substrates of ATM and ATR (207). In that proof-of-concept study, we raised a phosphopeptide specific antibody to random phosphopeptides that contained a phosphorylated SQ motif. This antibody was then used to immunopurify proteins from damaged cells followed by mass spectrometry to identify the phosphorylated proteins. Another group followed a similar approach to identify MCM3 phosphorylation sites (208). Importantly, these proof of concept studies have now been extended to identify hundreds of potential ATM- or ATR-catalyzed phosphorylation sites.

Elledge and colleagues performed the largest study of this kind (88). They assembled a large panel of phosphopeptide specific antibodies that recognized phospho-SQ or -TQ epitopes. Reasoning that these antibodies likely cross-react with other phospho-SQ/TQ epitopes, they used them to immunopurify tryptic peptides derived from undamaged and ionizing-radiation treated cells. By performing quantitiative mass spectrometry, they defined over 900 regulated phosphorylation sites in a little over 700 proteins. This study greatly expanded

the number of phosphorylated residues identified on known components of the DDR. For example, they identified 27 new sites on the BRCA1, 53BP1, and TOPBP1 proteins. The sites identified on MDC1 have since been defined as ATM-dependent phosphorylation sites and shown to be required for MDC1 to promote 53BP1 foci formation at DNA DSBs (178).

In addition to identifying new sites on known DDR proteins, the method also identified damage-dependent phosphorylation events on proteins not previously implicated in the cellular response to DNA damage. These include proteins linked to RNA processing, gene expression, chromatin remodeling, and developmental processes. Additional studies by other laboratories using similar approaches have further increased the total of number of potential ATM and ATR substrates (115,182).

Although further experimentation is necessary to validate that any one of these phosphorylation sites is really catalyzed by ATM/ATR and determine whether these phosphorylations are functionally important, these lists provide a rich resource for investigators using other genomic or reductionist approaches to understanding DDR mechanisms. Finally, this general strategy can be used for other post-translational modifications like methylation, ubiquitylation, sumoylation, or acetylation provided that efficient affinity purifications can be achieved. See Boisvert et al., for an example in which several arginine methylated DNA repair proteins were identified (209).

The primary drawback of the mass spectrometry approach is that it is biased to identify the most abundant post-translationally modified peptides. Low

abundant proteins that are only modified transiently and sub-stoichiometrically are likely to be poorly represented. The second issue is that there may be significant amounts of post-translational modifications that have no or redundant functional consequences. Thus, the catalogue of modifications is only a starting point for reductionist methods of studying function.

Screens examining protein subcellular localization

The subcellular localization of a protein provides a powerful indicator of whether it functions in a genome maintenance pathway. For example, proteins localized to the kinetochore likely regulate chromosome segregation during mitosis. Proteins localized to telomeres may be important in chromosome end replication or protection. Proteins localized to replisomes likely function in DNA replication.

There are two general strategies for using subcellular localization to identify genome maintenance proteins (Figure B.4). The first is to make a library of tagged cDNAs and examine localization by fluorescence microscopy (Figure B.4A). The second is to purify cellular structures like telomeres and identify the associated proteins by mass spectrometry (Figure B.4B). The first approach has seen only limited application primarily as a secondary screening strategy for other functional genomic approaches. The second approach, however, has seen widespread application with great success.

Dejardin and Kingston developed a technique they called proteomics of isolated chromatin segments (PICh) to selectively purify proteins associated with

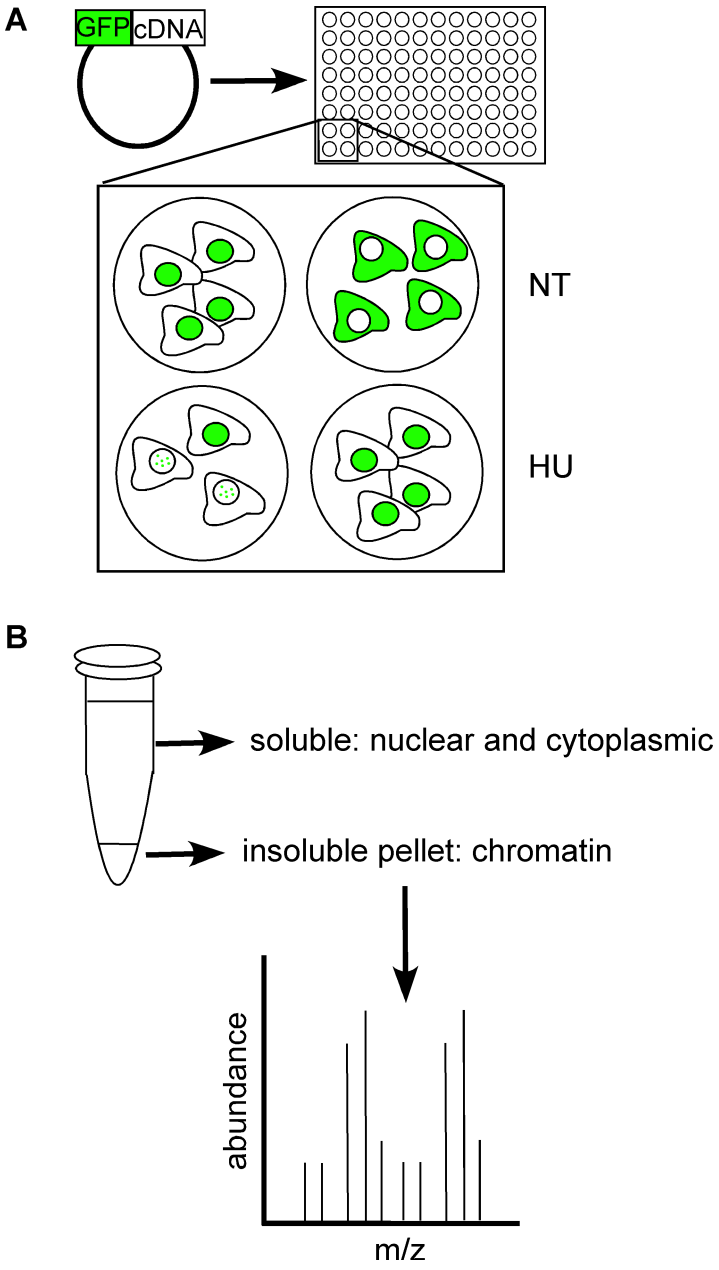

Figure B.4. Analysis of protein localization. Determining the subcellular localization of a protein may indicate whether it is likely to function in a genome maintenance pathway. (A) Fluorescence microscopy to examine subcellular localization can identify proteins that change localization in response to genotoxic drugs like hydroxyurea (HU). (B) Biochemical purification of cellular components coupled to mass spectrometry provides an alternative method of defining proteins that reside in cellular structures relevant to genome maintenance activities.

specific DNA loci (210). The steps of PICh include formaldehyde-crosslinking cells to crosslink DNA-protein and protein-protein complexes, solubilizing chromatin, denaturing the DNA to allow for DNA probe hybridization, purification of hybrids, elution, and finally analysis of eluted proteins by mass spectrometry. Since DNA hybridization for capture is not sensitive to high concentrations of ionic detergent, PICh allows for stringent capture and purification conditions that reduce non-specific binding of proteins. They applied this technique to identify proteins associated with mammalian telomeres. Their experiments successfully identified 85% of known telomere binding proteins and over 100 new proteins. Several of these are orphan nuclear receptors including COUP-TF2 and TR4. These proteins actually ranked higher on their abundance list than two known shelterin components. Interestingly, these proteins were found in telomere preparations from cells using the ALT method of telomere elongation specifically suggesting they have some function at ALT telomeres.

A few technical challenges must be overcome to use PICh for the classification of chromatin composition relevant to other genome maintenance activities. One of the primary concerns with this methodology is DNA-probe design. In many cases such as damage caused by a genotoxic agent, site-specificity is not available. The other major hurdle is that it may not be easily adapted to identifying proteins at DNA sequences that are found at one or few copies in the genome.

A second localization-based screen identified proteins whose localization to chromatin increased after cells were exposed to ionizing radiation (211). Both

DSB repair and DDR signaling proteins accumulate at the site of the DSB. This accumulation is often sufficient to observe a significant redistribution of the protein to the chromatin. Biochemical cell fractionation into soluble and chromatin compartments can be used to measure this change in localization. In this example, the authors determined that several subunits of the NuRD (nucleosome remodeling and histone deacetylation) chromatin-remodeling complex (CHD4, CHD3, MTA1 and MTA2) were recruited to chromatin after IR. These proteins have been best studied in the context of transcription where they work to remodel nucleosomes and thereby change gene expression (212). DSB repair, especially homology-directed repair, requires access of the DSB repair proteins to the DNA surrounding the break (213). Chromatin remodeling proteins likely work to regulate this access. Chromatin remodeling proteins are frequently identified in other functional genomic screens for genome maintenance proteins (see below) emphasizing that genome maintenance also requires chromatin maintenance.

RNAi screening

Genetics are difficult in diploid organisms because two copies of a gene must be manipulated to study loss of function phenotypes. Since RNAi-mediated silencing of genes occurs independently of gene copy number, it provides an opportunity to perform loss of function genetic screens in diploid cells. Thus, the genetic screens that were previously limited to organisms like *S. cerevisiae* and

S. pombe can now be done in human, mouse, and insect cell culture or even whole organisms like *C. elegans*.

There are currently three primary genome-wide RNAi platforms. Small interfering RNAs (siRNAs) are chemically synthesized silencing reagents that can be arrayed in multi-well plates. Formats include individual siRNAs per well or pools of multiple siRNAs targeting a gene in each well. Enzymatically prepared siRNAs (esiRNAs) are produced through enzymatic methodologies instead of chemical synthesis and create arrayed pools of siRNAs targeting each gene. Finally, short-hairpin RNA (shRNA) vectors are DNA vectors encoding RNAs that are processed to yield the silencing RNA. These are often produced in viral vector systems allowing stable integration into the genome of infected cells. Both arrayed and large pool formats of shRNA libraries are available.

Drug hypersensitivity screens

Designing an RNAi screen begins with an assay suitable for high-throughput screening. The simplest screen designs relevant to genome maintenance pathways are drug hypersensitivity screens. The rationale is that by treating cells with a drug that places a high burden on a specific genome maintenance pathway, researchers can identify genes whose loss of function causes hypersensitivity to the drug. These types of screens have been used in many genetically tractable organisms like yeast for decades, are simple to execute, and are compatible with all RNAi formats.

Since drug hypersensitivity screens rely on growth inhibition or cell death as the primary assay, they are compatible with pooled libraries of shRNAs (Figure B.5). In this format, a population of cells is infected at a low multiplicity of infection such that each cell harbors only a single shRNA. Since the shRNA should produce a stable knockdown, the cell population can be split in half, treated with a drug, and then examined for the relative abundance of each shRNA in the two populations. If a shRNA silences a gene required for cells to survive or proliferate in the presence of the drug, then that shRNA will be lost from the treated cell population but retained in the control population. The abundance of each shRNA can be determined by sequencing the shRNAs recovered from the two populations. Alternatively, microarrays can be used to detect barcodes incorporated into the shRNA library.

Using a pooled hypersensitivity screen, Elledge and colleagues sought to identify novel components of the DDR using cellular sensitivity to ionizing radiation as the assay (173,176). The screen conducted in the human U2OS osteosarcoma cell line used a library of approximately 75,000 shRNA vectors pooled into six sub-libraries. 850 shRNAs targeting 813 genes were identified that caused significant hypersensitivity to ionizing radiation. These genes were grouped into functional categories using the PANTHER (214). Unclassified genes accounted for 30% of the gene set. Multiple genes known to function in DDR networks were identified, and two new components of the DSB repair and DNA damage checkpoint signaling pathways were examined at the mechanistic

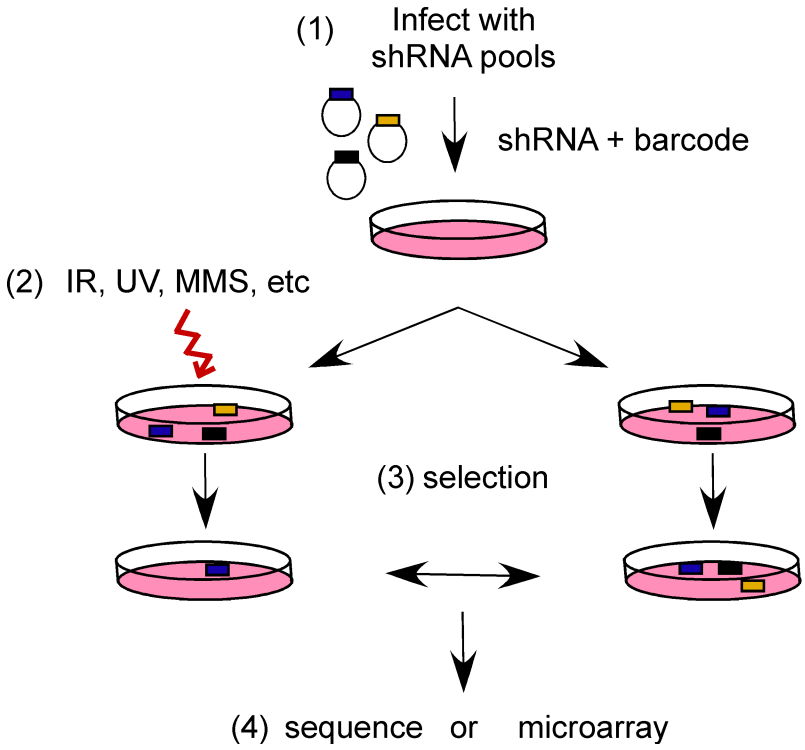

Figure B.5. Pooled RNAi screening. Pooled format screens facilitate drug hyper-sensitivity screening. (1) Barcoded shRNA vector libraries are introduced into a population of cells, (2) split into treated and untreated groups, (3) grown for several generations to allow selection against a subset of the shRNA expressing cells; and (4) analyzed by sequencing or microarray technology to identify differences in shRNA abundance.

level, NBA1 and the Triple T complex. NBA1 localizes to sites of DNA damage, associates with BRCA1 and is required for the recruitment of BRCA1 to damage sites (173). The Triple T complex consists of three proteins TTI1, TTI2, and TEL2 that regulates the stability of the phosphoinositide-3-kinase-related protein kinases ATM and ATR (176).

Drug hypersensitivity screens can also be performed using libraries of arrayed shRNA, siRNAs or esiRNAs. Arrayed hypersensitivity screens measure the effects of single gene knockdown on cells in each well of a multi-well plate (Figure B.6A). Individual or pools of siRNAs targeting a single gene are typically introduced into cells by reverse transfection using lipid-based methods. Reagents that change color or fluorescence relative to the number of proliferating cells in a well can be used as a measure of cell viability. Microplate spectrophotometers enable multi-well processing of this assay endpoint.

Whitehurst and colleagues used an arrayed RNAi hyper-sensitivity screen to identify gene functions that drive the cellular response to paclitaxel in lung cancer cells (215). Paclitaxel is a chemotherapeutic drug that targets microtubules to inhibit mitosis and is commonly used to treat patients with non-small lung cancers. Four unique siRNAs targeting each of 21,127 genes were transfected into a non-small lung cancer cell line in a one gene per well format with each well containing a pool of four siRNAs. A concentration below the half maximal inhibitory concentration for paclitaxel was selected to enrich for genes that when silenced rendered cells hypersensitive to low doses of drug. The

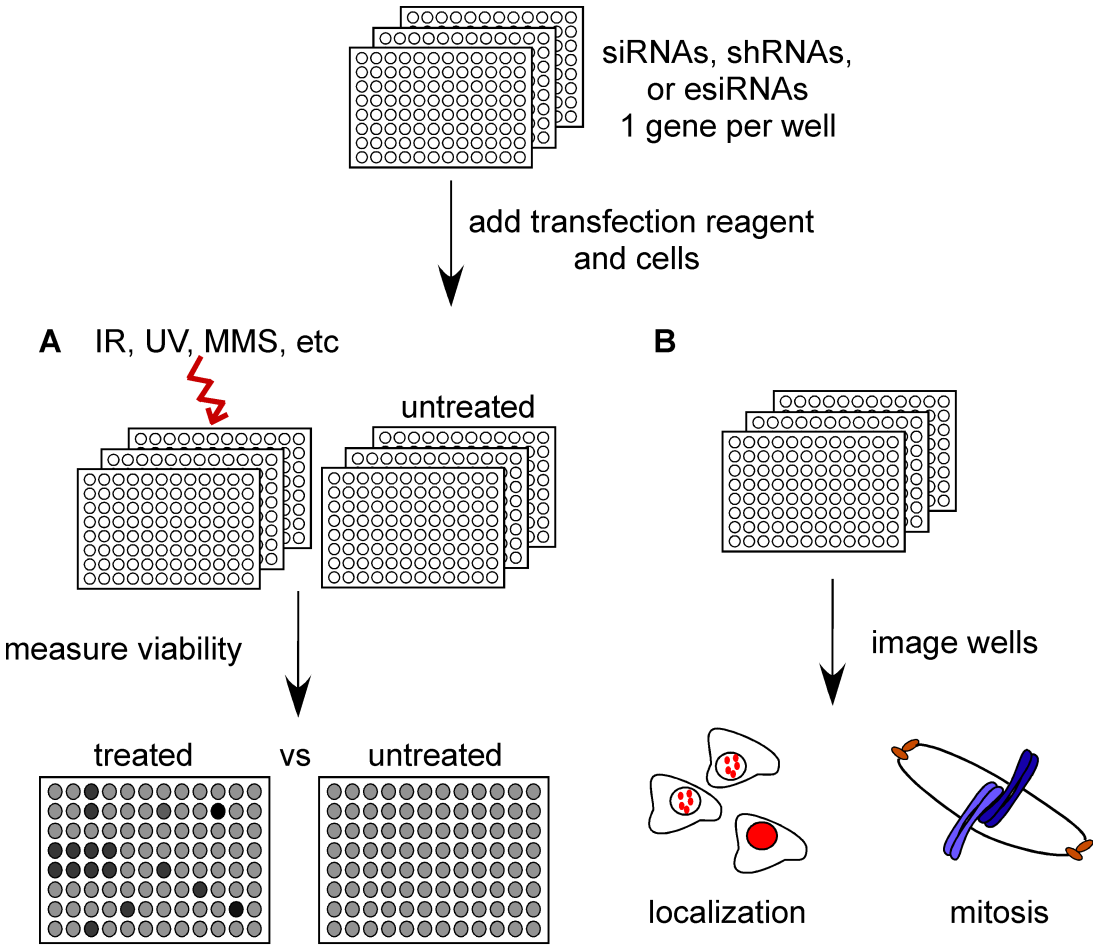

Figure B.6. Arrayed RNAi screening. Arrayed format RNAi libraries can be used both for (A) hyper-sensitivity or (B) high-content screens.

screen was completed in triplicate and normalizations to internal plate controls were made to eliminate plate-to-plate technical variations.

To identify significant hits, a statistical algorithm was applied to calculated viability ratios yielding an estimated false discovery rate below 5%. As validation for their analysis, siRNAs targeting core proteasome components were identified as hits and proteasome inhibitors have already been demonstrated to enhance the sensitivity of paclitaxel in cancer cells (216). They identified 87 genes that fit their criteria and reconfirmed that silencing caused sensitivity to paclitaxel with four additional siRNAs in pools and individually. Importantly, they confirmed these affects across multiple cell lines and defined genes that when silenced specifically sensitized cancer cells to low doses of paclitaxel compared to normal cell lines. Thus, the hits from this screen are potential drug targets for lung cancer since they selectively sensitize cancer cells and not normal cells (217,218).

These screens offer several lessons about RNAi-based functional genomics approaches. It is striking that in the shRNA pooled screen for IR sensitivity the vast majority of genes found in the primary screen were only identified by a single shRNA despite multiple shRNAs per gene being present in the library (173,176). This suggests that the shRNA targeting was inefficient, or a large number of the hits were due to off-target effects. The pooled screens require that a single shRNA vector inserted into the cellular genome produces sufficient shRNA to silence the gene effectively. Even with optimal vector designs, this requirement may not often be fulfilled. This issue will increase the false-

negatives in the screen. Off-target effects leading to false-positives are also a concern necessitating that further follow-up experimentation be performed to rule them out. Finally, the large pooled format cannot ensure that every shRNA vector is equally represented or expressed in the library.

The arrayed screens likely suffer less from the issue of insufficient targeting. Algorithms for designing siRNAs are sufficiently robust to yield high percentages of highly effective siRNAs. Furthermore, they are transfected into the cells at very high copy number. However, off-target effects are also a concern in this format. In the paclitaxel arrayed screen on-target effects were confirmed by validating hits from the pooled siRNA primary screen with four individual siRNAs (215). Alternatively, if the library is arrayed such that several individual siRNAs are used independently, then the results can be quickly scanned for genes that have at least two siRNAs yielding the same hypersensitive phenotype. This raises the largest obstacle to the arrayed screens – cost. Typically, whole genome screens are done in formats where small pools of siRNAs targeting a single gene are combined into a single well. This saves money in performing the screen but necessitates de-convolution of the pools to eliminate off-target effects. The arrayed screens also require specialized liquid handling systems that may not be available to a typical laboratory.

Synthetic genetic interaction screens

Synthetic genetic interactions arise when the phenotype from the loss of function of two genes are compared to the phenotypes of the individual gene effects. Often these interactions take the form of synthetic lethality. Synthetic genetic interaction screens in yeast are currently being done using synthetic genetic array (SGA) technologies and are very useful for dissecting complex biological pathways (72,219,220). SGA provides a format for the high-throughput construction of double mutants at the genome-wide level. Similar methodologies are not available in higher eukaryotes. Instead, investigators are turning to three approaches to perform synthetic genetic interaction screens: drug sensitivity screens, screens combining RNAi molecules targeting two different genes, and screens using matched pairs of cell lines in which a specific gene is mutated or wild-type.

The first involves using selective drugs in combination with RNAi to target proteins in the cell and is particularly useful in identifying drug targets or genetic backgrounds where a drug might be useful in treating cancer. For example, several synthetic lethal screens have been performed with recently developed inhibitors of the poly-ADP ribose polymerase (PARP) enzymes (221-223). These inhibitors selectively kill cells deficient in homologous recombination (HR) based repair mechanisms. Thus, they are being developed for treating BRCA-deficient tumors since BRCA1 and BRCA2 are required for HR repair. Investigators reasoned that loss of function mutations in other genes are likely to yield defects in HR and performed synthetic lethal screens with PARP inhibitors to identify

these genes. These screens not only define novel HR genes but also may point to additional clinical settings where PARP inhibitors can be useful.

The alternative two approaches to perform synthetic genetic interaction screens do not use selective drugs. The first is to combine two RNAi molecules targeting different genes. The second is to use matched pairs of cell lines in which a specific gene is mutated or wild-type. The first approach may be more quickly applied to any gene but suffers from issues of whether the simultaneous knockdown of two genes is as efficient as the single knockdown. It also suffers from the potential for synthetic off-target effects, which may be very difficult to control. Thus, the superior results that can be obtained from the second approach using matched pairs of cell lines with a defined loss of function mutation at the gene level may be worth the added time and cost of generating the cell lines.

High content RNAi screens

Hypersensitivity and synthetic lethal screens have relatively low information content because cell proliferation or viability is not a phenotype specific to genome maintenance pathways. Thus, large numbers of genes are identified in these types of screens. Even in yeast, these screens often identify large percentages of the genome as being required to survive or proliferate in response to a drug. Thus, higher-content assays are increasingly being employed. Essentially, anything that can be imaged can be tested using high content microscopy (HCM). High-resolution images can be generated rapidly in

microtiter plates containing up to 1536 wells. Image processing software can yield information as varied as DNA content, cell size, and protein localization. Thus high content screening can be used to interrogate specific genome maintenance pathways at high resolution, and the screens typically yield smaller numbers of “hits”. These types of screens are best suited to arrayed RNAi library formats (Figure B.6B).

One example of a high content screen relevant to genome maintenance is a screen to define cell division cycle genes in human cells (170). In this study, video microscopy was used to measure changes in the timing of cell division in HeLa cells in real time. In addition, fixed cells were stained for tubulin and DNA to visualize physical aberrations in mitotic structures. Screening a set of esiRNAs representing 5,305 genes led to the classification of genes with functional roles in spindle formation, mitotic arrest, and cytokinesis, including 7 genes with no previous assigned function and 23 with annotated functions other than cell division. The functional readouts for this study were specific and facilitated direct classification of gene groups.

Additional examples of high content screens include several recent screens that studied the DNA damage response by microscopy. Two screens were reported that used the detection of phosphorylated ATM/ATR substrates as the primary assay (66,181). The largest of these screens used an siRNA library targeting the entire genome and visualized phosphorylation of the histone variant H2AX. H2AX is dispersed throughout the genome in approximately 10% of nucleosomes (224). Phosphorylation of H2AX at serine 139 (γH2AX) by ATM

and ATR is a well-defined early event at or near the site of DDR activation (225-227). γH2AX spreads kilobases of DNA beyond the break site, which allows visualization of bright γH2AX nuclear foci by microscopy (228). γH2AX has classically been used as a marker for DNA DSBs but is also phosphorylated at stalled replication forks and perhaps at other lesions where ATR is active.

The Cimprich group monitored γH2AX phosphorylation following gene silencing in the absence of any added genotoxic agent with the rationale that genes involved in genome maintenance pathways like DNA repair or replication prevent DNA damage from accumulating in cells (181). Silencing these genes would increase the DNA damage burden, thereby increasing the activation of ATM or ATR and phosphorylation of H2AX. The authors grouped their results into confidence levels defined by the strength of the phenotype, reproducibility, and numbers of individual siRNAs that yielded the same results. They defined 581 genes as high confidence hits. In addition to genes that function in the known genome maintenance pathways, they identified genes predicted to function in many other cellular metabolism activities including a large group of RNA processing proteins. After further investigation, they determined that at least a subset of these genes are needed to prevent RNA-DNA hybrids from forming during transcription which presumably lead to RNA polymerase stalling followed by collisions with the replication machinery in S-phase. Importantly, RNA processing proteins were also enriched in ATM/ATR substrate identification screens (88) suggesting that RNA processing is a critical component of genome maintenance.

DNA damage also causes the recruitment of many DDR proteins to intranuclear foci that are easily visualized by immunofluoresence, including p53 binding protein (53BP1) (229,230). 53BP1 localizes to sites of DNA DSBs and functions in DSB repair via end joining (231-236). The recruitment of 53BP1 to DNA damage foci depends on γH2AX and MDC1 (237). The underlying mechanisms of focal recruitment and retention of these DNA repair factors at DNA breaks has been a major question in the field. RNAi screens based on imaging 53BP1 localization were designed to address this question and led to the discovery of RNF8 and RNF168, which are ubiquitn conjugating enzymes that regulate the chromatin microenvironment important for DNA DSB repair (178-180).

Durocher and colleagues identified RNF8 in a genome-wide siRNA screen for the abundance of 53BP1 foci in U2OS cells 24hrs after IR (178). They reasoned that gene functions required upstream of 53BP1 foci formation at DSBs would result in significant changes in abundance of 53BP1 foci when silenced. Pools of four siRNAs per gene were arrayed in 384 well formats and transfected into cells in quadruplicate. Using the high throughput acquisition and analysis capabilities of the Opera imaging system, they identified 500 siRNAs resulting in significant reduction in 53BP1 foci compared to controls. RNF8 was the top scoring hit following known upstream regulators of 53BP1 like γH2AX and MDC1. In a follow up study, 59 candidate genes from the original screen were targeted with additional esiRNAs to further validate 53BP1 loss. RNF168 was the top hit

in this secondary validation and was subsequently defined to act downstream of RNF8, to assemble K63-linked ubiquitin chains at sites of DNA damage (180).

Lukas and colleagues also identified RNF168 in a screen based on similar rationale and functional readout but using distinct screening methodologies (179). They applied a protocol that was developed for screening siRNAs spotted on chambered slides (238). This technique represents another method for siRNA arrayed screens, siRNA microarrays. In this particular screen, U2OS were seeded in chambers containing 384 individual siRNA spots that were embedded in optimized transfection mixtures. After three days, the central area of each spot, containing approximately 150 cells, was analyzed for number of 53BP1 foci. A total of 21541 human genes were targeted with multiple siRNAs. When the screen was published, 15000 genes had been analyzed and 42 had at least one siRNA that resulted in reduced 53BP1 foci. Only three of 42 hits had two siRNAs score positive, MDC1, RNF8, and the previously uncharacterized RING finger protein, RNF168. These results may reflect a current disadvantage of this methodology, a high false negative rate. This problem may be due to inefficient siRNA knockdown or the small numbers of cells that were analyzed per data point. Increasing replicates or cells analyzed may suffice for reducing false-negatives.

RNAi screens in multi-cellular organisms

The benefits of RNAi genetic screens are not limited to cell culture. RNAi screens in the *C. elegans* worm provide a method for analyzing loss of function

phenotypes in an intact animal. Delivery of RNAi into worms is performed either by injection of dsRNA directly into head or tail, soaking the worms in a solution of RNAi molecules, or by feeding animals bacteria that express dsRNA targeting a specific gene (239). The effects of RNAi in *C.elegans* can be analyzed by whole mount staining or other phenotypic assays.

A genome-wide RNAi screen in *C. elegans* identified genes that protect animal cells against IR (240). The screen was completed with worm larvae cultured in the presence of bacteria expressing RNAi molecules. 19000 *C. elegans* genes were screened in an arrayed, multi-well format. Worm cultures were exposed to IR and subsequently imaged to measure for IR-induced cell cycle arrest and apoptosis in germ cells. After undergoing stringent filtering using dose response analyses, 45 genes emerged as IR protective genes. 43 of the 45 have clear human homologs emphasizing the extent to which the DDR pathway is conserved across evolution. Genes were cataloged based on their previously known functions, requirement for resistance to other DNA damaging agents, and homology across human, fly, and yeast.

Importantly, the human homologs of seven genes (*ATM*, *ITGA6*, *NIPBL*, *CAND1/TIP120*, *NOB1*, *WWP2*, and *TOPBP1*) were targeted by siRNA and measured for IR sensitivity, and all seven were confirmed as genes required for IR resistance in human cells. Several of these genes like *ATM* and *TOPBP1* are not surprising since they were already known to function in DNA damage response pathways. However, less clear was why *NOB1* (a protein implicated in ribosome assembly) or *WWP2* (a ubiquitin ligase) were identified. Interestingly,

WWP2 has since been shown to function as an E3 ligase for PTEN (241). This may provide a link to the DNA damage response since PTEN function may modulate DNA repair activities (242).

Another *C. elegans* screen used a *gfp-LacZ* reporter of frameshifts and small insertions/deletions to screen for DNA instability in somatic cells (243). This method enabled visual screening of cells harboring mutations that reverted the out of frame *LacZ* gene to the correct reading frame. The authors found 61 out of 16,606 genes screened resulted in a mutator phenotype. The orthologs of the DNA mismatch repair genes were among this list as would be expected. In addition to DNA repair genes, they found genes such as the histone deacetylases *HDAC1* and *HDAC2* which regulate chromatin organization and genes like *cdc5 and cdc14* that regulate cell cycle checkpoints. 29 of the genes were not described previously but several have interesting protein domains such as exonuclease, histone acetyltransferase, and SET domains suggesting they could be involved in DNA metabolism.

General considerations about RNAi screens

RNAi screens provide an increasingly important source of discovery for genome maintenance researchers (Table B.1). A question like how do cells repair a double-strand break can be approached at multiple levels with RNAi screens. The simplest but least specific may be the ionizing radiation sensitivity screens. The most specific, but also most technologically difficult will be screens

Table B.1. Examples of RNAi screens for genome maintenance discussed in text.

Format	Genotoxic Challenge	Cell Type	RNAi type	Library size (RNAi; genes targeted)	Hits	Genome Maintenance Activity	Ref.
Pooled-hypersensitivity	IR	U2OS	barcoded shRNA	74,905; 32,293	813	DNA damage response to DSB	(98)
Arrayed-hypersensitivity	paclitaxel	NCI-H1-155	siRNAs	84,508; 21,127	87	mitosis; spindle assembly checkpoint	(140)
Arrayed-synthetic lethal	PARPi	CAL51	siRNAs	3,116; 779	6	homologous recombination; DSB repair	(149)
Arrayed-synthetic lethal	PARPi	CAL51	siRNAs	460; 230	5	homologous recombination; DSB repair	(147)
Arrayed-synthetic lethal	PARPi	U2OS	siRNAs	~250; 73	29	homologous recombination; DSB repair	(148)
Cell cycle image screen	none	HeLa	esiRNAs	5,305 genes	37	cell cycle; cytokinesis; spindle assembly	(92)
H2AX image screen	none	HeLa	siRNAs	~84,000; ~21,000	581	spontaneous DSB prevention	(103)
H2AX image screen	none	HeLa, U2OS	shRNAs, siRNAs	6386; 2287	93	spontaneous DSB prevention	(65)
53BP1 image screen	IR	U2OS	siRNAs	genome-wide	500	DSB-induced chromatin microenvironment	(100)
53BP1 image screen	none	U2OS	siRNAs microarrays	~50,000; 21,541	42	DSB-induced chromatin microenvironment	(101)
Organism	IR	*C.elegans*	RNAi food	19,000 genes	45	IR-induced cell cycle arrest; apoptosis	(166)
Organism	none	*C.elegans*	RNAi food	16,757; 16,606	61	spontaneous mutation prevention	(169)

based on direct measurements of repair. In between are the high content imaging screens scoring for markers like γH2AX or 53BP1 foci. The choice of screen assay in part determines the optimal screening platform. In practice, the most useful information may come from combining the results of multiple screens. Certainly the confidence that any previously unstudied gene functions in a DSB repair pathway is significantly increased by the independent identification of that gene in more than one screen.

Eliminating false-positives is critical for the results of the screen to be useful. The most common source of false-positives is off-target effects of the RNAi. Since silencing does not require perfect base pair complementarity between the siRNA and the mRNA, each RNAi molecule has the potential to silence multiple genes. The most common method to eliminate off-target effects is to require that more than one RNAi molecule targeting the same gene yield the same phenotypic consequence. If antibodies to the targeted protein are available then correlating protein knockdown efficiency with strength of phenotype adds further confidence. Finally, complementation of the knockdown phenotype with a resistant cDNA is the best method of demonstrating an on-target effect. However, this approach is not feasible in a high-throughput format.

RNAi screens likely suffer from high rates of false-negatives. Even when two investigators use similar methodologies, the overlap in gene hits is often low. False-negatives come from inefficient knockdown, cell type specific effects, and deficiencies in the genome-wide libraries. The cell-type specific effects are likely to be large since most screens use tumor-derived cells that have large numbers

of genetic abnormalities already. These differences can be exploited especially in synthetic lethal type screens where the important question is what makes two or more cell types behave differently.

All of these RNAi screens depend on a statistical determination of what is a “hit”. Defining the hits requires a measurement of difference from the control, variability, and elimination of off-target effects. Many different statistical methods have been developed but determining biological significance requires more in-depth analyses. Functional annotation programs such as PANTHER can help but these issues point out the need for secondary screens to validate hits and provide further guidance on future directions. Ideally, these secondary screens should also be amenable to high-throughput assays. Ultimately, more traditional reductionist approaches are needed to fully validate and understand the results of any RNAi screen.

Conclusion

Expression profiling, proteomics, and RNAi screening methodologies have proven successful in the investigation of genome maintenance activities. When considering the power of these methodologies for discovery, it is important to realize that in many cases they were made with first generation technologies. In the next ten years we can expect the development of arrayed full-genome epitope-tagged cDNA libraries for faster genome-wide proteomics; protein microarray chips will expand to include a much higher percentage of the

proteome; and RNAi libraries will continue to improve.

The amount of data being generated is quickly outpacing the ability of a single investigator to assimilate. Thus, bioinformatic analyses must be developed alongside of improving high-throughput screening technologies. Better databases to store, report, search, and integrate data sets from RNAi screens are also needed.

The rapidly expanding data sets from expression profiling, proteomics, and RNAi screening methods are generating lists of new genome maintenance proteins. These lists contain genes that function in pathways like chromatin regulation that have obvious links to the DNA metabolic processes that maintain the genome. But they also contain evidence that many other intracellular pathways impact genome integrity. Our challenge is to complete the catalogue and explain at a mechanistic level how these activities are integrated within the cell to protect genome stability and prevent disease.

REFERENCES

1. Ikner, A. and Shiozaki, K. (2005) Yeast signaling pathways in the oxidative stress response. *Mutat Res*, 569, 13-27.

2. Friedberg, E.C. (2003) DNA damage and repair. *Nature*, 421, 436-440.

3. Bell, S.P. and Dutta, A. (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem*, 71, 333-374.

4. Nasmyth, K. (2001) Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu Rev Genet*, 35, 673-745.

5. Campos, E.I. and Reinberg, D. (2009) Histones: annotating chromatin. *Annu Rev Genet*, 43, 559-599.

6. Harper, J.W. and Elledge, S.J. (2007) The DNA damage response: ten years after. *Mol Cell*, 28, 739-745.

7. Rouse, J. and Jackson, S.P. (2002) Interfaces between the detection, signaling, and repair of DNA damage. *Science*, 297, 547-551.

8. Harrison, J.C. and Haber, J.E. (2006) Surviving the breakup: the DNA damage checkpoint. *Annu Rev Genet*, 40, 209-235.

9. Bonilla, C.Y., Melo, J.A. and Toczyski, D.P. (2008) Colocalization of sensors is sufficient to activate the DNA damage checkpoint in the absence of damage. *Mol Cell*, 30, 267-276.

10. Lieber, M.R., Lu, H., Gu, J. and Schwarz, K. (2008) Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. *Cell Res*, 18, 125-133.

11. Gapud, E.J., Dorsett, Y., Yin, B., Callen, E., Bredemeyer, A., Mahowald, G.K., Omi, K.Q., Walker, L.M., Bednarski, J.J., McKinnon, P.J. *et al.* (2011) Ataxia telangiectasia mutated (Atm) and DNA-PKcs kinases have overlapping activities during chromosomal signal joint formation. *Proc Natl Acad Sci U S A*, 108, 2022-2027.

12. Durocher, D. and Jackson, S.P. (2001) DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? *Curr Opin Cell Biol*, 13, 225-231.

13. Fanning, E., Klimovich, V. and Nager, A.R. (2006) A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res*, 34, 4126-4137.

14. Cortez, D., Guntuku, S., Qin, J. and Elledge, S.J. (2001) ATR and ATRIP: partners in checkpoint signaling. *Science*, 294, 1713-1716.

15. Brown, E.J. and Baltimore, D. (2000) ATR disruption leads to chromosomal fragmentation and early embryonic lethality. *Genes Dev*, 14, 397-402.

16. de Klein, A., Muijtjens, M., van Os, R., Verhoeven, Y., Smit, B., Carr, A.M., Lehmann, A.R. and Hoeijmakers, J.H. (2000) Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice. *Curr Biol*, 10, 479-482.

17. Brown, E.J. and Baltimore, D. (2003) Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance. *Genes Dev*, 17, 615-628.

18. Krude, T., Jackman, M., Pines, J. and Laskey, R.A. (1997) Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. *Cell*, 88, 109-119.

19. Zegerman, P. and Diffley, J.F. (2007) Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature*, 445, 281-285.

20. Boos, D., Sanchez-Pulido, L., Rappas, M., Pearl, L.H., Oliver, A.W., Ponting, C.P. and Diffley, J.F. (2011) Regulation of DNA replication

through Sld3-Dpb11 interaction is conserved from yeast to humans. *Curr Biol*, 21, 1152-1157.

21. Pacek, M. and Walter, J.C. (2004) A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *Embo J*, 23, 3667-3676.

22. Stukenberg, P.T., Turner, J. and O'Donnell, M. (1994) An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. *Cell*, 78, 877-887.

23. Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D. and Labib, K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol*, 8, 358-366.

24. Liu, Y., Kao, H.I. and Bambara, R.A. (2004) Flap endonuclease 1: a central component of DNA metabolism. *Annu Rev Biochem*, 73, 589-615.

25. Bell, S.P. and Dutta, A. (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem*, 71, 333-374.

26. Van, C., Yan, S., Michael, W.M., Waga, S. and Cimprich, K.A. (2010) Continued primer synthesis at stalled replication forks contributes to checkpoint activation. *J Cell Biol*, 189, 233-246.

27. Chang, D.J. and Cimprich, K.A. (2009) DNA damage tolerance: when it's OK to make mistakes. *Nat Chem Biol*, 5, 82-90.

28. Raschle, M., Knipscheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J.D., Ellenberger, T.E., Scharer, O.D. and Walter, J.C. (2008) Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell*, 134, 969-980.

29. Sogo, J.M., Lopes, M. and Foiani, M. (2002) Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. *Science*, 297, 599-602.

30. Osman, F. and Whitby, M.C. (2007) Exploring the roles of Mus81-Eme1/Mms4 at perturbed replication forks. *DNA Repair (Amst)*, 6, 1004-1017.

31. Petermann, E., Orta, M.L., Issaeva, N., Schultz, N. and Helleday, T. (2010) Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol Cell*, 37, 492-502.

32. Seigneur, M., Bidnenko, V., Ehrlich, S.D. and Michel, B. (1998) RuvAB acts at arrested replication forks. *Cell*, 95, 419-430.

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

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

35. Ralf, C., Hickson, I.D. and Wu, L. (2006) The Bloom's syndrome helicase can promote the regression of a model replication fork. *J Biol Chem*, 281, 22839-22846.

36. Franchitto, A., Pirzio, L.M., Prosperi, E., Sapora, O., Bignami, M. and Pichierri, P. (2008) Replication fork stalling in WRN-deficient cells is overcome by prompt activation of a MUS81-dependent pathway. *J Cell Biol*, 183, 241-252.

37. Davies, S.L., North, P.S. and Hickson, I.D. (2007) Role for BLM in replication-fork restart and suppression of origin firing after replicative stress. *Nat Struct Mol Biol*, 14, 677-679.

38. Mazina, O.M., Rossi, M.J., Deakyne, J.S., Huang, F. and Mazin, A.V. (2012) Polarity and Bypass of DNA Heterology during Branch Migration of Holliday Junctions by Human RAD54, BLM, and RECQ1 Proteins. *J Biol Chem*, 287, 11820-11832.

39. Michel, B., Flores, M.J., Viguera, E., Grompone, G., Seigneur, M. and Bidnenko, V. (2001) Rescue of arrested replication forks by homologous recombination. *Proc Natl Acad Sci U S A*, 98, 8181-8188.

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

41. Letessier, A., Millot, G.A., Koundrioukoff, S., Lachages, A.M., Vogt, N., Hansen, R.S., Malfoy, B., Brison, O. and Debatisse, M. (2011) Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature*, 470, 120-123.

42. Lundin, C., Erixon, K., Arnaudeau, C., Schultz, N., Jenssen, D., Meuth, M. and Helleday, T. (2002) Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells. *Mol Cell Biol*, 22, 5869-5878.

43. Saintigny, Y., Delacote, F., Vares, G., Petitot, F., Lambert, S., Averbeck, D. and Lopez, B.S. (2001) Characterization of homologous recombination induced by replication inhibition in mammalian cells. *Embo J*, 20, 3861-3870.

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

45. Pichierri, P., Ammazzalorso, F., Bignami, M. and Franchitto, A. (2011) The Werner syndrome protein: linking the replication checkpoint response to genome stability. *Aging (Albany NY)*, 3, 311-318.

46. Whitby, M.C. (2010) The FANCM family of DNA helicases/translocases. *DNA Repair (Amst)*, 9, 224-236.

47. Brosh, R.M., Jr. and Bohr, V.A. (2007) Human premature aging, DNA repair and RecQ helicases. *Nucleic Acids Res*, 35, 7527-7544.

48. Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S. and Wigley, D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, 97, 75-84.

49. Shaw, G., Gan, J., Zhou, Y.N., Zhi, H., Subburaman, P., Zhang, R., Joachimiak, A., Jin, D.J. and Ji, X. (2008) Structure of RapA, a Swi2/Snf2 protein that recycles RNA polymerase during transcription. *Structure*, 16, 1417-1427.

50. Blastyak, A., Pinter, L., Unk, I., Prakash, L., Prakash, S. and Haracska, L. (2007) Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. *Mol Cell*, 28, 167-175.

51. Achar, Y.J., Balogh, D. and Haracska, L. (2011) Coordinated protein and DNA remodeling by human HLTF on stalled replication fork. *Proc Natl Acad Sci U S A*, 108, 14073-14078.

52. Yusufzai, T. and Kadonaga, J.T. (2008) HARP is an ATP-driven annealing helicase. *Science*, 322, 748-750.

53. Yusufzai, T. and Kadonaga, J.T. (2010) Annealing helicase 2 (AH2), a DNA-rewinding motor with an HNH motif. *Proc Natl Acad Sci U S A*, 107, 20970-20973.

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

55. Ratnakumar, K., Duarte, L.F., LeRoy, G., Hasson, D., Smeets, D., Vardabasso, C., Bonisch, C., Zeng, T., Xiang, B., Zhang, D.Y. *et al.* (2012) ATRX-mediated chromatin association of histone variant macroH2A1 regulates alpha-globin expression. *Genes Dev*, 26, 433-438.

56. Hanawalt, P.C. and Spivak, G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol*, 9, 958-970.

57. Sandhu, S., Wu, X., Nabi, Z., Rastegar, M., Kung, S., Mai, S. and Ding, H. (2012) Loss of HLTF function promotes intestinal carcinogenesis. *Mol Cancer*, 11, 18.

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

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

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

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

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

63. Coleman, M.A., Eisen, J.A. and Mohrenweiser, H.W. (2000) Cloning and characterization of HARP/SMARCAL1: a prokaryotic HepA-related SNF2 helicase protein from human and mouse. *Genomics*, 65, 274-282.

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

65. Boerkoel, C.F., Takashima, H., John, J., Yan, J., Stankiewicz, P., Rosenbarker, L., Andre, J.L., Bogdanovic, R., Burguet, A., Cockfield, S. *et al.* (2002) Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia. *Nat Genet*, 30, 215-220.

66. Lovejoy, C.A., Xu, X., Bansbach, C.E., Glick, G.G., Zhao, R., Ye, F., Sirbu, B.M., Titus, L.C., Shyr, Y. and Cortez, D. (2009) Functional genomic screens identify CINP as a genome maintenance protein. *Proc Natl Acad Sci USA*, 106, 19304-19309.

67. Mendez, J. and Stillman, B. (2000) Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins

during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol Cell Biol*, 20, 8602-8612.

68. van der Geer, P. and Hunter, T. (1994) Phosphopeptide mapping and phosphoamino acid analysis by electrophoresis and chromatography on thin-layer cellulose plates. *Electrophoresis*, 15, 544-554.

69. Nakatani, Y. and Ogryzko, V. (2003) Immunoaffinity purification of mammalian protein complexes. *Methods Enzymol*, 370, 430-444.

70. Halazonetis, T.D., Gorgoulis, V.G. and Bartek, J. (2008) An oncogene-induced DNA damage model for cancer development. *Science*, 319, 1352-1355.

71. Bartkova, J., Horejsi, Z., Koed, K., Kramer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C. *et al.* (2005) DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature*, 434, 864-870.

72. Bartkova, J., Rezaei, N., Liontos, M., Karakaidos, P., Kletsas, D., Issaeva, N., Vassiliou, L.V., Kolettas, E., Niforou, K., Zoumpourlis, V.C. *et al.* (2006) Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature*, 444, 633-637.

73. Gorgoulis, V.G., Vassiliou, L.V., Karakaidos, P., Zacharatos, P., Kotsinas, A., Liloglou, T., Venere, M., Ditullio, R.A., Jr., Kastrinakis, N.G., Levy, B. *et al.* (2005) Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature*, 434, 907-913.

74. Di Micco, R., Fumagalli, M., Cicalese, A., Piccinin, S., Gasparini, P., Luise, C., Schurra, C., Garre, M., Nuciforo, P.G., Bensimon, A. *et al.* (2006) Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature*, 444, 638-642.

75. Mallette, F.A., Gaumont-Leclerc, M.F. and Ferbeyre, G. (2007) The DNA damage signaling pathway is a critical mediator of oncogene-induced senescence. *Genes Dev*, 21, 43-48.

76. Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D.C., Lukas, J., Bekker-Jensen, S., Bartek, J. and Shiloh, Y. (2006) Chromatin

relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nat Cell Biol*, 8, 870-876.

77. Kumar-Sinha, C., Tomlins, S.A. and Chinnaiyan, A.M. (2008) Recurrent gene fusions in prostate cancer. *Nature reviews. Cancer*, 8, 497-511.

78. Trenz, K., Errico, A. and Costanzo, V. (2008) Plx1 is required for chromosomal DNA replication under stressful conditions. *Embo J*, 27, 876-885.

79. Yim, H. and Erikson, R.L. (2009) Polo-like kinase 1 depletion induces DNA damage in early S prior to caspase activation. *Mol Cell Biol*, 29, 2609-2621.

80. Jin, J., Arias, E.E., Chen, J., Harper, J.W. and Walter, J.C. (2006) A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol Cell*, 23, 709-721.

81. Lovejoy, C.A., Lock, K., Yenamandra, A. and Cortez, D. (2006) DDB1 maintains genome integrity through regulation of Cdt1. *Mol Cell Biol*, 26, 7977-7990.

82. Bassermann, F., Frescas, D., Guardavaccaro, D., Busino, L., Peschiaroli, A. and Pagano, M. (2008) The Cdc14B-Cdh1-Plk1 axis controls the G2 DNA-damage-response checkpoint. *Cell*, 134, 256-267.

83. Yamamoto, Y., Matsuyama, H., Kawauchi, S., Matsumoto, H., Nagao, K., Ohmi, C., Sakano, S., Furuya, T., Oga, A., Naito, K. *et al.* (2006) Overexpression of polo-like kinase 1 (PLK1) and chromosomal instability in bladder cancer. *Oncology*, 70, 231-237.

84. von Lindern, M., Fornerod, M., van Baal, S., Jaegle, M., de Wit, T., Buijs, A. and Grosveld, G. (1992) The translocation (6;9), associated with a specific subtype of acute myeloid leukemia, results in the fusion of two genes, dek and can, and the expression of a chimeric, leukemia-specific dek-can mRNA. *Mol Cell Biol*, 12, 1687-1697.

85. Meyn, M.S., Lu-Kuo, J.M. and Herzing, L.B. (1993) Expression cloning of multiple human cDNAs that complement the phenotypic defects of ataxia-

telangiectasia group D fibroblasts. *American journal of human genetics*, 53, 1206-1216.

86. Rubin, E., Wu, X., Zhu, T., Cheung, J.C., Chen, H., Lorincz, A., Pandita, R.K., Sharma, G.G., Ha, H.C., Gasson, J. *et al.* (2007) A role for the HOXB7 homeodomain protein in DNA repair. *Cancer research*, 67, 1527-1535.

87. Alvaro, D., Lisby, M. and Rothstein, R. (2007) Genome-wide analysis of Rad52 foci reveals diverse mechanisms impacting recombination. *PLoS genetics*, 3, e228.

88. Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y. *et al.* (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, 316, 1160-1166.

89. Hartwell, L.H. and Kastan, M.B. (1994) Cell cycle control and cancer. *Science*, 266, 1821-1828.

90. Kastan, M.B. and Bartek, J. (2004) Cell-cycle checkpoints and cancer. *Nature*, 432, 316-323.

91. McKinnon, P.J. and Caldecott, K.W. (2007) DNA strand break repair and human genetic disease. *Annual review of genomics and human genetics*, 8, 37-55.

92. Barzilai, A., Biton, S. and Shiloh, Y. (2008) The role of the DNA damage response in neuronal development, organization and maintenance. *DNA Repair (Amst)*, 7, 1010-1027.

93. Abraham, R.T. (2004) PI 3-kinase related kinases: 'big' players in stress-induced signaling pathways. *DNA Repair (Amst)*, 3, 883-887.

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

95. Byun, T.S., Pacek, M., Yee, M.C., Walter, J.C. and Cimprich, K.A. (2005) Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev*, 19, 1040-1052.

96. Binz, S.K., Sheehan, A.M. and Wold, M.S. (2004) Replication protein A phosphorylation and the cellular response to DNA damage. *DNA Repair (Amst)*, 3, 1015-1024.

97. Fillingham, J., Keogh, M.C. and Krogan, N.J. (2006) GammaH2AX and its role in DNA double-strand break repair. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 84, 568-577.

98. Tabancay, A.P., Jr. and Forsburg, S.L. (2006) Eukaryotic DNA replication in a chromatin context. *Current topics in developmental biology*, 76, 129-184.

99. Flanagan, J.F. and Peterson, C.L. (1999) A role for the yeast SWI/SNF complex in DNA replication. *Nucleic Acids Res*, 27, 2022-2028.

100. Papamichos-Chronakis, M. and Peterson, C.L. (2008) The Ino80 chromatin-remodeling enzyme regulates replisome function and stability. *Nat Struct Mol Biol*, 15, 338-345.

101. Vincent, J.A., Kwong, T.J. and Tsukiyama, T. (2008) ATP-dependent chromatin remodeling shapes the DNA replication landscape. *Nat Struct Mol Biol*, 15, 477-484.

102. Shimada, K., Oma, Y., Schleker, T., Kugou, K., Ohta, K., Harata, M. and Gasser, S.M. (2008) Ino80 chromatin remodeling complex promotes recovery of stalled replication forks. *Curr Biol*, 18, 566-575.

103. Muthuswami, R., Truman, P.A., Mesner, L.D. and Hockensmith, J.W. (2000) A eukaryotic SWI2/SNF2 domain, an exquisite detector of double-stranded to single-stranded DNA transition elements. *J Biol Chem*, 275, 7648-7655.

104. Fernandez-Capetillo, O., Lee, A., Nussenzweig, M. and Nussenzweig, A. (2004) H2AX: the histone guardian of the genome. *DNA Repair (Amst)*, 3, 959-967.

105. Stucki, M. and Jackson, S.P. (2006) gammaH2AX and MDC1: anchoring the DNA-damage-response machinery to broken chromosomes. *DNA Repair (Amst)*, 5, 534-543.

106. Mer, G., Bochkarev, A., Gupta, R., Bochkareva, E., Frappier, L., Ingles, C.J., Edwards, A.M. and Chazin, W.J. (2000) Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell*, 103, 449-456.

107. Unsal-Kacmaz, K., Chastain, P.D., Qu, P.P., Minoo, P., Cordeiro-Stone, M., Sancar, A. and Kaufmann, W.K. (2007) The human Tim/Tipin complex coordinates an Intra-S checkpoint response to UV that slows replication fork displacement. *Mol Cell Biol*, 27, 3131-3142.

108. Hickson, I., Zhao, Y., Richardson, C.J., Green, S.J., Martin, N.M., Orr, A.I., Reaper, P.M., Jackson, S.P., Curtin, N.J. and Smith, G.C. (2004) Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. *Cancer research*, 64, 9152-9159.

109. Leahy, J.J., Golding, B.T., Griffin, R.J., Hardcastle, I.R., Richardson, C., Rigoreau, L. and Smith, G.C. (2004) Identification of a highly potent and selective DNA-dependent protein kinase (DNA-PK) inhibitor (NU7441) by screening of chromenone libraries. *Bioorganic & medicinal chemistry letters*, 14, 6083-6087.

110. Kumagai, A., Lee, J., Yoo, H.Y. and Dunphy, W.G. (2006) TopBP1 activates the ATR-ATRIP complex. *Cell*, 124, 943-955.

111. Ball, H.L., Ehrhardt, M.R., Mordes, D.A., Glick, G.G., Chazin, W.J. and Cortez, D. (2007) Function of a conserved checkpoint recruitment domain in ATRIP proteins. *Mol Cell Biol*, 27, 3367-3377.

112. Michel, B., Grompone, G., Flores, M.J. and Bidnenko, V. (2004) Multiple pathways process stalled replication forks. *Proc Natl Acad Sci U S A*, 101, 12783-12788.

113. Lambert, S., Froget, B. and Carr, A.M. (2007) Arrested replication fork processing: interplay between checkpoints and recombination. *DNA Repair (Amst)*, 6, 1042-1061.

114. Postow, L., Crisona, N.J., Peter, B.J., Hardy, C.D. and Cozzarelli, N.R. (2001) Topological challenges to DNA replication: conformations at the fork. *Proc Natl Acad Sci U S A*, 98, 8219-8226.

115. Stokes, M.P., Rush, J., Macneill, J., Ren, J.M., Sprott, K., Nardone, J., Yang, V., Beausoleil, S.A., Gygi, S.P., Livingstone, M. *et al.* (2007) Profiling of UV-induced ATM/ATR signaling pathways. *Proc Natl Acad Sci USA*, 104, 19855-19860.

116. Smolka, M.B., Albuquerque, C.P., Chen, S.-h. and Zhou, H. (2007) Proteome-wide identification of in vivo targets of DNA damage checkpoint kinases. *Proc Natl Acad Sci USA*, 104, 10364-10369.

117. Shechter, D., Costanzo, V. and Gautier, J. (2004) Regulation of DNA replication by ATR: signaling in response to DNA intermediates. *DNA Repair (Amst)*, 3, 901-908.

118. Postow, L., Woo, E.M., Chait, B.T. and Funabiki, H. (2009) Identification of SMARCAL1 as a component of the DNA damage response. *J Biol Chem*, 284, 35951-35961.

119. Yang, F., Waters, K.M., Miller, J.H., Gritsenko, M.A., Zhao, R., Du, X., Livesay, E.A., Purvine, S.O., Monroe, M.E., Wang, Y. *et al.* (2010) Phosphoproteomics profiling of human skin fibroblast cells reveals pathways and proteins affected by low doses of ionizing radiation. *PLoS One*, 5, e14152.

120. Boyle, W.J., van der Geer, P. and Hunter, T. (1991) Phosphopeptide mapping and phosphoamino acid analysis by two-dimensional separation on thin-layer cellulose plates. *Methods Enzymol*, 201, 110-149.

121. Durr, H., Korner, C., Muller, M., Hickmann, V. and Hopfner, K.P. (2005) X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. *Cell*, 121, 363-373.

122. Sprouse, R.O., Brenowitz, M. and Auble, D.T. (2006) Snf2/Swi2-related ATPase Mot1 drives displacement of TATA-binding protein by gripping DNA. *Embo J*, 25, 1492-1504.

123. Lewis, R., Durr, H., Hopfner, K.P. and Michaelis, J. (2008) Conformational changes of a Swi2/Snf2 ATPase during its mechano-chemical cycle. *Nucleic Acids Res*, 36, 1881-1890.

124. Fitzgerald, D.J., DeLuca, C., Berger, I., Gaillard, H., Sigrist, R., Schimmele, K. and Richmond, T.J. (2004) Reaction cycle of the yeast Isw2 chromatin remodeling complex. *Embo J*, 23, 3836-3843.

125. Tripathi, V., Kaur, S. and Sengupta, S. (2008) Phosphorylation-dependent interactions of BLM and 53BP1 are required for their anti-recombinogenic roles during homologous recombination. *Carcinogenesis*, 29, 52-61.

126. Ammazzalorso, F., Pirzio, L.M., Bignami, M., Franchitto, A. and Pichierri, P. (2010) ATR and ATM differently regulate WRN to prevent DSBs at stalled replication forks and promote replication fork recovery. *Embo J*, 29, 3156-3169.

127. Karmakar, P., Piotrowski, J., Brosh, R.M., Jr., Sommers, J.A., Miller, S.P., Cheng, W.H., Snowden, C.M., Ramsden, D.A. and Bohr, V.A. (2002) Werner protein is a target of DNA-dependent protein kinase in vivo and in vitro, and its catalytic activities are regulated by phosphorylation. *J Biol Chem*, 277, 18291-18302.

128. Tyson, J.J. and Novak, B. (2008) Temporal organization of the cell cycle. *Curr Biol*, 18, R759-R768.

129. Hauk, G. and Bowman, G.D. (2011) Structural insights into regulation and action of SWI2/SNF2 ATPases. *Curr Opin Struct Biol*, 21, 719-727.

130. Hauk, G., McKnight, J.N., Nodelman, I.M. and Bowman, G.D. (2010) The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol Cell*, 39, 711-723.

131. Collins, R., Karlberg, T., Lehtio, L., Schutz, P., van den Berg, S., Dahlgren, L.G., Hammarstrom, M., Weigelt, J. and Schuler, H. (2009) The DEXD/H-box RNA helicase DDX19 is regulated by an {alpha}-helical switch. *J Biol Chem*, 284, 10296-10300.

132. Borrow, J., Shearman, A.M., Stanton, V.P., Jr., Becher, R., Collins, T., Williams, A.J., Dube, I., Katz, F., Kwong, Y.L., Morris, C. *et al.* (1996) The

t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. *Nat Genet*, 12, 159-167.

133. Perkins, A., Kongsuwan, K., Visvader, J., Adams, J.M. and Cory, S. (1990) Homeobox gene expression plus autocrine growth factor production elicits myeloid leukemia. *Proc Natl Acad Sci U S A*, 87, 8398-8402.

134. Kroon, E., Krosl, J., Thorsteinsdottir, U., Baban, S., Buchberg, A.M. and Sauvageau, G. (1998) Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *Embo J*, 17, 3714-3725.

135. Moldovan, G.L. and D'Andrea, A.D. (2009) How the fanconi anemia pathway guards the genome. *Annu Rev Genet*, 43, 223-249.

136. Zivicnjak, M., Franke, D., Zenker, M., Hoyer, J., Lucke, T., Pape, L. and Ehrich, J.H. (2009) SMARCAL1 mutations: a cause of prepubertal idiopathic steroid-resistant nephrotic syndrome. *Pediatr Res*, 65, 564-568.

137. Lucke, T., Billing, H., Sloan, E.A., Boerkoel, C.F., Franke, D., Zimmering, M., Ehrich, J.H. and Das, A.M. (2005) Schimke-immuno-osseous dysplasia: new mutation with weak genotype-phenotype correlation in siblings. *Am J Med Genet A*, 135, 202-205.

138. Eisen, J.A., Sweder, K.S. and Hanawalt, P.C. (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Res*, 23, 2715-2723.

139. Conaway, R.C. and Conaway, J.W. (2009) The INO80 chromatin remodeling complex in transcription, replication and repair. *Trends Biochem Sci*, 34, 71-77.

140. Palvimo, J.J. (2007) PIAS proteins as regulators of small ubiquitin-related modifier (SUMO) modifications and transcription. *Biochem Soc Trans*, 35, 1405-1408.

141. Muller, S., Hoege, C., Pyrowolakis, G. and Jentsch, S. (2001) SUMO, ubiquitin's mysterious cousin. *Nat Rev Mol Cell Biol*, 2, 202-210.

142. Yang, X.J. and Gregoire, S. (2006) A recurrent phospho-sumoyl switch in transcriptional repression and beyond. *Mol Cell*, 23, 779-786.

143. Baumann, C., Korner, R., Hofmann, K. and Nigg, E.A. (2007) PICH, a centromere-associated SNF2 family ATPase, is regulated by Plk1 and required for the spindle checkpoint. *Cell*, 128, 101-114.

144. Vignali, M., Hassan, A.H., Neely, K.E. and Workman, J.L. (2000) ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol*, 20, 1899-1910.

145. Sif, S., Stukenberg, P.T., Kirschner, M.W. and Kingston, R.E. (1998) Mitotic inactivation of a human SWI/SNF chromatin remodeling complex. *Genes Dev*, 12, 2842-2851.

146. Dephoure, N., Zhou, C., Villen, J., Beausoleil, S.A., Bakalarski, C.E., Elledge, S.J. and Gygi, S.P. (2008) A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci U S A*, 105, 10762-10767.

147. Bansbach, C.E. and Cortez, D. (2011) Defining genome maintenance pathways using functional genomic approaches. *Crit Rev Biochem Mol Biol*, 46, 327-341.

148. Hartwell, L.H., Culotti, J. and Reid, B. (1970) Genetic control of the cell-division cycle in yeast. I. Detection of mutants. *Proc Natl Acad Sci U S A*, 66, 352-359.

149. Nurse, P. (1975) Genetic control of cell size at cell division in yeast. *Nature*, 256, 547-551.

150. Painter, R.B. and Young, B.R. (1980) Radiosensitivity in ataxia-telangiectasia: a new explanation. *Proc Natl Acad Sci U S A*, 77, 7315-7317.

151. Houldsworth, J. and Lavin, M.F. (1980) Effect of ionizing radiation on DNA synthesis in ataxia telangiectasia cells. *Nucleic Acids Res*, 8, 3709-3720.

152. Savitsky, K., Bar-Shira, A., Gilad, S., Rotman, G., Ziv, Y., Vanagaite, L., Tagle, D.A., Smith, S., Uziel, T., Sfez, S. *et al.* (1995) A single ataxia

telangiectasia gene with a product similar to PI-3 kinase. *Science*, 268, 1749-1753.

153. Futreal, P.A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L.M., Haugen-Strano, A., Swensen, J., Miki, Y. *et al.* (1994) BRCA1 mutations in primary breast and ovarian carcinomas. *Science*, 266, 120-122.

154. Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W. *et al.* (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 266, 66-71.

155. Carney, J.P., Maser, R.S., Olivares, H., Davis, E.M., Le Beau, M., Yates, J.R., 3rd, Hays, L., Morgan, W.F. and Petrini, J.H. (1998) The hMre11/hRad50 protein complex and Nijmegen breakage syndrome: linkage of double-strand break repair to the cellular DNA damage response. *Cell*, 93, 477-486.

156. Matsuura, S., Tauchi, H., Nakamura, A., Kondo, N., Sakamoto, S., Endo, S., Smeets, D., Solder, B., Belohradsky, B.H., Der Kaloustian, V.M. *et al.* (1998) Positional cloning of the gene for Nijmegen breakage syndrome. *Nat Genet*, 19, 179-181.

157. D'Andrea, A.D. (2010) Susceptibility pathways in Fanconi's anemia and breast cancer. *N Engl J Med*, 362, 1909-1919.

158. Cleaver, J.E., Lam, E.T. and Revet, I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat Rev Genet*, 10, 756-768.

159. Nurse, P., Thuriaux, P. and Nasmyth, K. (1976) Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe. *Mol Gen Genet*, 146, 167-178.

160. Birrell, G.W., Giaever, G., Chu, A.M., Davis, R.W. and Brown, J.M. (2001) A genome-wide screen in Saccharomyces cerevisiae for genes affecting UV radiation sensitivity. *Proc Natl Acad Sci U S A*, 98, 12608-12613.

161. Bennett, C.B., Lewis, L.K., Karthikeyan, G., Lobachev, K.S., Jin, Y.H., Sterling, J.F., Snipe, J.R. and Resnick, M.A. (2001) Genes required for ionizing radiation resistance in yeast. *Nat Genet*, 29, 426-434.

162. Cejka, P. and Jiricny, J. (2008) Interplay of DNA repair pathways controls methylation damage toxicity in Saccharomyces cerevisiae. *Genetics*, 179, 1835-1844.

163. Li, R. and Murray, A.W. (1991) Feedback control of mitosis in budding yeast. *Cell*, 66, 519-531.

164. Hoyt, M.A., Totis, L. and Roberts, B.T. (1991) S. cerevisiae genes required for cell cycle arrest in response to loss of microtubule function. *Cell*, 66, 507-517.

165. Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B. *et al.* (2002) Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, 418, 387-391.

166. Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H. *et al.* (1999) Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science*, 285, 901-906.

167. Gelperin, D.M., White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H. *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev*, 19, 2816-2826.

168. Björklund, M., Taipale, M., Varjosalo, M., Saharinen, J., Lahdenperä, J. and Taipale, J. (2006) Identification of pathways regulating cell size and cell-cycle progression by RNAi. *Nature*, 439, 1009-1013.

169. Kittler, R., Pelletier, L., Heninger, A.-K., Slabicki, M., Theis, M., Miroslaw, L., Poser, I., Lawo, S., Grabner, H., Kozak, K. *et al.* (2007) Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat Cell Biol*, 9, 1401-1412.

170. Kittler, R., Putz, G., Pelletier, L., Poser, I., Heninger, A.-K., Drechsel, D., Fischer, S., Konstantinova, I., Habermann, B., Grabner, H. *et al.* (2004) An

endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature*, 432, 1036-1040.

171. Mukherji, M., Bell, R., Supekova, L., Wang, Y., Orth, A.P., Batalov, S., Miraglia, L., Huesken, D., Lange, J., Martin, C. *et al.* (2006) Genome-wide functional analysis of human cell-cycle regulators. *Proc Natl Acad Sci USA*, 103, 14819-14824.

172. Draviam, V.M., Stegmeier, F., Nalepa, G., Sowa, M.E., Chen, J., Liang, A., Hannon, G.J., Sorger, P.K., Harper, J.W. and Elledge, S.J. (2007) A functional genomic screen identifies a role for TAO1 kinase in spindle-checkpoint signalling. *Nat Cell Biol*, 9, 556-564.

173. Wang, B., Hurov, K., Hofmann, K. and Elledge, S.J. (2009) NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. *Genes Dev*, 23, 729-739.

174. Smogorzewska, A., Desetty, R., Saito, T.T., Schlabach, M., Lach, F.P., Sowa, M.E., Clark, A.B., Kunkel, T.A., Harper, J.W. and Colaiácovo, M.P. (2010) A Genetic Screen Identifies FAN1, a Fanconi Anemia-Associated Nuclease Necessary for DNA Interstrand Crosslink Repair. *Mol Cell*, 39, 36-47.

175. O'Connell, B.C., Adamson, B., Lydeard, J.R., Sowa, M.E., Ciccia, A., Bredemeyer, A.L., Schlabach, M., Gygi, S.P., Elledge, S.J. and Harper, J.W. (2010) A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Mol Cell*, 40, 645-657.

176. Hurov, K.E., Cotta-Ramusino, C. and Elledge, S.J. (2010) A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. *Genes Dev*, 24, 1939-1950.

177. O'Donnell, L., Panier, S., Wildenhain, J., Tkach, J.M., Al-Hakim, A., Landry, M.-C., Escribano-Diaz, C., Szilard, R.K., Young, J.T.F., Munro, M. *et al.* (2010) The MMS22L-TONSL Complex Mediates Recovery from Replication Stress and Homologous Recombination. *Mol Cell*.

178. Kolas, N.K., Chapman, J.R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F.D., Panier, S., Mendez, M., Wildenhain, J., Thomson, T.M. *et*

al. (2007) Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science*, 318, 1637-1640.

179. Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D.H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J. *et al.* (2009) RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell*, 136, 435-446.

180. Stewart, G.S., Panier, S., Townsend, K., Al-Hakim, A.K., Kolas, N.K., Miller, E.S., Nakada, S., Ylanko, J., Olivarius, S., Mendez, M. *et al.* (2009) The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell*, 136, 420-434.

181. Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F. and Solow-Cordero, D.E. (2009) A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Mol Cell*, 35, 228-239.

182. Mu, J.-J., Wang, Y., Luo, H., Leng, M., Zhang, J., Yang, T., Besusso, D., Jung, S.Y. and Qin, J. (2007) A Proteomic Analysis of Ataxia Telangiectasia-mutated (ATM)/ATM-Rad3-related (ATR) Substrates Identifies the Ubiquitin-Proteasome System as a Regulator for DNA Damage Checkpoints. *Journal of Biological Chemistry*, 282, 17330-17334.

183. Cagney, G., Alvaro, D., Reid, R.J.D., Thorpe, P.H., Rothstein, R. and Krogan, N.J. (2006) Functional genomics of the yeast DNA-damage response. *Genome Biol*, 7, 233.

184. Davidson, M.B. and Brown, G.W. (2009) Dissecting the DNA damage response using functional genomics approaches in S. cerevisiae. *DNA Repair (Amst)*, 8, 1110-1117.

185. Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270, 467-470.

186. Chaudhry, M.A., Chodosh, L.A., McKenna, W.G. and Muschel, R.J. (2002) Gene expression profiling of HeLa cells in G1 or G2 phases. *Oncogene*, 21, 1934-1942.

187. Amundson, S.A., Bittner, M. and Fornace, A.J., Jr. (2003) Functional genomics as a window on radiation stress signaling. *Oncogene*, 22, 5828-5833.

188. Rashi-Elkeles, S., Elkon, R., Weizman, N., Linhart, C., Amariglio, N., Sternberg, G., Rechavi, G., Barzilai, A., Shamir, R. and Shiloh, Y. (2006) Parallel induction of ATM-dependent pro- and antiapoptotic signals in response to ionizing radiation in murine lymphoid tissue. *Oncogene*, 25, 1584-1592.

189. Levine, A.J. and Oren, M. (2009) The first 30 years of p53: growing ever more complex. *Nat Rev Cancer*, 9, 749-758.

190. Menendez, D., Inga, A. and Resnick, M.A. (2009) The expanding universe of p53 targets. *Nat Rev Cancer*, 9, 724-737.

191. Beckerman, R. and Prives, C. (2010) Transcriptional regulation by p53. *Cold Spring Harb Perspect Biol*, 2, a000935.

192. Birrell, G.W., Brown, J.A., Wu, H.I., Giaever, G., Chu, A.M., Davis, R.W. and Brown, J.M. (2002) Transcriptional response of Saccharomyces cerevisiae to DNA-damaging agents does not identify the genes that protect against these agents. *Proc Natl Acad Sci U S A*, 99, 8778-8783.

193. MacBeath, G. and Schreiber, S.L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science*, 289, 1760-1763.

194. Merbl, Y. and Kirschner, M.W. (2009) Large-scale detection of ubiquitination substrates using cell extracts and protein microarrays. *Proc Natl Acad Sci U S A*, 106, 2543-2548.

195. Peters, J.M. (2006) The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nat Rev Mol Cell Biol*, 7, 644-656.

196. King, R.W., Deshaies, R.J., Peters, J.M. and Kirschner, M.W. (1996) How proteolysis drives the cell cycle. *Science*, 274, 1652-1659.

197. King, R.W., Peters, J.M., Tugendreich, S., Rolfe, M., Hieter, P. and Kirschner, M.W. (1995) A 20S complex containing CDC27 and CDC16

catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell*, 81, 279-288.

198. Sudakin, V., Ganoth, D., Dahan, A., Heller, H., Hershko, J., Luca, F.C., Ruderman, J.V. and Hershko, A. (1995) The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol Biol Cell*, 6, 185-197.

199. Guo, X., Williams, J.G., Schug, T.T. and Li, X. (2010) DYRK1A and DYRK3 promote cell survival through phosphorylation and activation of SIRT1. *J Biol Chem*, 285, 13223-13232.

200. Li, K., Zhao, S., Karur, V. and Wojchowski, D.M. (2002) DYRK3 activation, engagement of protein kinase A/cAMP response element-binding protein, and modulation of progenitor cell survival. *J Biol Chem*, 277, 47052-47060.

201. Cusick, M.E., Klitgord, N., Vidal, M. and Hill, D.E. (2005) Interactome: gateway into systems biology. *Hum Mol Genet*, 14 Spec No. 2, R171-181.

202. Fromont-Racine, M., Rain, J.C. and Legrain, P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet*, 16, 277-282.

203. Boulton, S.J. (2002) Combined Functional Genomic Maps of the C. elegans DNA Damage Response. *Science*, 295, 127-131.

204. Yates, J.R., Ruse, C.I. and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng*, 11, 49-79.

205. Daub, H., Olsen, J.V., Bairlein, M., Gnad, F., Oppermann, F.S., Korner, R., Greff, Z., Keri, G., Stemmann, O. and Mann, M. (2008) Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. *Mol Cell*, 31, 438-448.

206. Kim, S.T., Lim, D.S., Canman, C.E. and Kastan, M.B. (1999) Substrate specificities and identification of putative substrates of ATM kinase family members. *J Biol Chem*, 274, 37538-37543.

207. Cortez, D., Glick, G. and Elledge, S.J. (2004) Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A*, 101, 10078-10083.

208. Shi, Y., Dodson, G.E., Mukhopadhyay, P.S., Shanware, N.P., Trinh, A.T. and Tibbetts, R.S. (2007) Identification of carboxyl-terminal MCM3 phosphorylation sites using polyreactive phosphospecific antibodies. *J Biol Chem*, 282, 9236-9243.

209. Boisvert, F.M., Cote, J., Boulanger, M.C. and Richard, S. (2003) A proteomic analysis of arginine-methylated protein complexes. *Mol Cell Proteomics*, 2, 1319-1330.

210. Dejardin, J. and Kingston, R.E. (2009) Purification of proteins associated with specific genomic Loci. *Cell*, 136, 175-186.

211. Larsen, D.H., Poinsignon, C., Gudjonsson, T., Dinant, C., Payne, M.R., Hari, F.J., Danielsen, J.M.R., Menard, P., Sand, J.C., Stucki, M. *et al.* (2010) The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. *J Cell Biol*, 190, 731-740.

212. Knoepfler, P.S. and Eisenman, R.N. (1999) Sin meets NuRD and other tails of repression. *Cell*, 99, 447-450.

213. Xu, Y. and Price, B.D. (2011) Chromatin dynamics and the repair of DNA double strand breaks. *Cell Cycle*, 10, 261-267.

214. Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A. and Narechania, A. (2003) PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res*, 13, 2129-2141.

215. Whitehurst, A.W., Bodemann, B.O., Cardenas, J., Ferguson, D., Girard, L., Peyton, M., Minna, J.D., Michnoff, C., Hao, W., Roth, M.G. *et al.* (2007) Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature*, 446, 815-819.

216. Lynch, T., Jr. and Kim, E. (2005) Optimizing chemotherapy and targeted agent combinations in NSCLC. *Lung Cancer*, 50 Suppl 2, S25-32.

217. Nezi, L. and Musacchio, A. (2009) Sister chromatid tension and the spindle assembly checkpoint. *Curr Opin Cell Biol*, 21, 785-795.

218. Jordan, M.A. and Wilson, L. (2004) Microtubules as a target for anticancer drugs. *Nat Rev Cancer*, 4, 253-265.

219. Baryshnikova, A., Costanzo, M., Dixon, S., Vizeacoumar, F.J., Myers, C.L., Andrews, B. and Boone, C. (2010) Synthetic genetic array (SGA) analysis in Saccharomyces cerevisiae and Schizosaccharomyces pombe. *Methods Enzymol*, 470, 145-179.

220. Pan, X., Ye, P., Yuan, D.S., Wang, X., Bader, J.S. and Boeke, J.D. (2006) A DNA integrity network in the yeast Saccharomyces cerevisiae. *Cell*, 124, 1069-1081.

221. Lord, C.J., McDonald, S., Swift, S., Turner, N.C. and Ashworth, A. (2008) A high-throughput RNA interference screen for DNA repair determinants of PARP inhibitor sensitivity. *DNA Repair (Amst)*, 7, 2010-2019.

222. Wiltshire, T.D., Lovejoy, C.A., Wang, T., Xia, F., O'connor, M.J. and Cortez, D. (2010) Sensitivity to Poly(ADP-ribose) Polymerase (PARP) Inhibition Identifies Ubiquitin-specific Peptidase 11 (USP11) as a Regulator of DNA Double-strand Break Repair. *Journal of Biological Chemistry*, 285, 14565-14571.

223. Turner, N.C., Lord, C.J., Iorns, E., Brough, R., Swift, S., Elliott, R., Rayter, S., Tutt, A.N. and Ashworth, A. (2008) A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor. *Embo J*, 27, 1368-1377.

224. Bonner, W.M., Redon, C.E., Dickey, J.S., Nakamura, A.J., Sedelnikova, O.A., Solier, S. and Pommier, Y. (2008) GammaH2AX and cancer. *Nat Rev Cancer*, 8, 957-967.

225. Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S. and Bonner, W.M. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem*, 273, 5858-5868.

226. Burma, S., Chen, B.P., Murphy, M., Kurimasa, A. and Chen, D.J. (2001) ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J Biol Chem*, 276, 42462-42467.

227. Ward, I.M. and Chen, J. (2001) Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. *J Biol Chem*, 276, 47759-47762.

228. Rogakou, E.P., Boon, C., Redon, C. and Bonner, W.M. (1999) Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J Cell Biol*, 146, 905-916.

229. Schultz, L.B., Chehab, N.H., Malikzay, A. and Halazonetis, T.D. (2000) p53 binding protein 1 (53BP1) is an early participant in the cellular response to DNA double-strand breaks. *J Cell Biol*, 151, 1381-1390.

230. Rappold, I., Iwabuchi, K., Date, T. and Chen, J. (2001) Tumor suppressor p53 binding protein 1 (53BP1) is involved in DNA damage-signaling pathways. *J Cell Biol*, 153, 613-620.

231. Iwabuchi, K., Basu, B.P., Kysela, B., Kurihara, T., Shibata, M., Guan, D., Cao, Y., Hamada, T., Imamura, K., Jeggo, P.A. *et al.* (2003) Potential role for 53BP1 in DNA end-joining repair through direct interaction with DNA. *J Biol Chem*, 278, 36487-36495.

232. Ward, I.M., Reina-San-Martin, B., Olaru, A., Minn, K., Tamada, K., Lau, J.S., Cascalho, M., Chen, L., Nussenzweig, A., Livak, F. *et al.* (2004) 53BP1 is required for class switch recombination. *J Cell Biol*, 165, 459-464.

233. Nakamura, K., Sakai, W., Kawamoto, T., Bree, R.T., Lowndes, N.F., Takeda, S. and Taniguchi, Y. (2006) Genetic dissection of vertebrate 53BP1: a major role in non-homologous end joining of DNA double strand breaks. *DNA Repair (Amst)*, 5, 741-749.

234. Manis, J.P., Morales, J.C., Xia, Z., Kutok, J.L., Alt, F.W. and Carpenter, P.B. (2004) 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. *Nat Immunol*, 5, 481-487.

235. Difilippantonio, S., Gapud, E., Wong, N., Huang, C.Y., Mahowald, G., Chen, H.T., Kruhlak, M.J., Callen, E., Livak, F., Nussenzweig, M.C. *et al.* (2008) 53BP1 facilitates long-range DNA end-joining during V(D)J recombination. *Nature*, 456, 529-533.

236. Xie, A., Hartlerode, A., Stucki, M., Odate, S., Puget, N., Kwok, A., Nagaraju, G., Yan, C., Alt, F.W., Chen, J. *et al.* (2007) Distinct roles of chromatin-associated proteins MDC1 and 53BP1 in mammalian double-strand break repair. *Mol Cell*, 28, 1045-1057.

237. FitzGerald, J.E., Grenon, M. and Lowndes, N.F. (2009) 53BP1: function and mechanisms of focal recruitment. *Biochem Soc Trans*, 37, 897-904.

238. Erfle, H., Neumann, B., Liebel, U., Rogers, P., Held, M., Walter, T., Ellenberg, J. and Pepperkok, R. (2007) Reverse transfection on cell arrays for high content screening microscopy. *Nat Protoc*, 2, 392-399.

239. Kamath, R.S. and Ahringer, J. (2003) Genome-wide RNAi screening in Caenorhabditis elegans. *Methods*, 30, 313-321.

240. van Haaften, G., Romeijn, R., Pothof, J., Koole, W., Mullenders, L.H.F., Pastink, A., Plasterk, R.H.A. and Tijsterman, M. (2006) Identification of conserved pathways of DNA-damage response and radiation protection by genome-wide RNAi. *Current Biology*, 16, 1344-1350.

241. Maddika, S., Kavela, S., Rani, N., Palicharla, V.R., Pokorny, J.L., Sarkaria, J.N. and Chen, J. (2011) WWP2 is an E3 ubiquitin ligase for PTEN. *Nat Cell Biol*.

242. Chalmers, A.J., Lakshman, M., Chan, N. and Bristow, R.G. (2010) Poly(ADP-ribose) polymerase inhibition as a model for synthetic lethality in developing radiation oncology targets. *Semin Radiat Oncol*, 20, 274-281.

243. Pothof, J., van Haaften, G., Thijssen, K., Kamath, R.S., Fraser, A.G., Ahringer, J., Plasterk, R.H. and Tijsterman, M. (2003) Identification of genes that protect the C. elegans genome against mutations by genome-wide RNAi. *Genes Dev*, 17, 443-448.