Award Number: W81XWH-F€ÆÆ€GGî

TITLE: Ø˘}&Œ]}Á¡~ÆZØŒ]ÞÖH¡Å§Á§@¡ÄÖÞŒÄÖæ¡æ�‡^ÄÜ^•][}•^

PRINCIPAL INVESTIGATOR: Óæœ}&æÄÙ¡ã¡à˘

CONTRACTING ORGANIZATION: V@˘ÁXæ}å^¦àã¡ÓÄ\}ã̧^¦•ã˘
Þæ˘@¸ã̧¦^ÃÄVÞÁÄ¦Í¡GÍ€Á

REPORT DATE: June 20FG

TYPE OF REPORT: Annual Ù˘{{æˆ^

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
01-06-2012	Annual Summary	1 JUN 2011- 31 MAY 2012

4. TITLE AND SUBTITLE

Function of ZFAND3 in the DNA Damage Response

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-10-1-0226

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Bianca Sirbu

E-Mail: bianca.m.sirbu@anderbilt.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Vanderbilt University
Nashville, TN 37240

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT The DNA-damage response (DDR) pathway maintains genomic integrity and prevents tumorigenesis. Human precancerous lesions exhibit high levels of DNA damage particularly during DNA replication. This observation has led to the idea that unresolved problems accrued during DNA replication can contribute to tumorigenesis. Mechanistically, such replication stress can arise when replication forks arrest at DNA lesions. Activation of the DDR leads to repair and restart of damaged replication forks using tumor-suppressor proteins such as BRCA1 (breast cancer 1). I hypothesize that multiple pathways are important in breast cancer to respond to the elevated levels of replication stress known to exist in both pre-cancerous lesions and tumors. I identified one putative new replication stress response protein (Zfand3) and characterized its regulation. We have also developed a new method termed iPOND (isolation of proteins on nascent DNA) to examine Zfand3 and other DDR protein dynamics and functions at moving and stalled replication forks in human cells. Our studies analyzed the temporal and spatial regulation of damaged fork repair. These findings provide insights into why defects in specific DDR pathways cause genome instability during DNA replication. These studies have implications for understanding breast cancer etiology.

15. SUBJECT TERMS

genome maintenance, DNA damage response, replication stress

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	34	19b. TELEPHONE NUMBER (include area code)

Table of Contents

Introduction

Maintenance of genomic integrity is essential to prevent carcinogenesis (1,2). The genome is under constant attack from both exogenous and endogenous sources of damage, such as oxidative by-products of normal metabolism. To combat the propagation of damaged DNA, cells have evolved a signaling pathway called the DNA damage response (DDR) (3). In response to damage, the DDR halts cell cycle progression, allows time for DNA repair, and initiates programmed cell death of heavily damaged cells (3,4). Mutations in genes that function in the DDR, such as mutations in BRCA1 (breast cancer 1), highlight the significance of DDR genes for cellular survival and prevention of breast cancer. Findings from functional genomic screens from our laboratory have identified ZFAND3 (zinc finger AN1-type domain containing protein 3) as a potential new DDR gene (5). ZFAND3 prevents premature entry into mitosis in the presence of DNA damage and is a putative interacting partner of TopBP1 (topoisomerase II binding protein 1). This proposal tests the hypothesis that ZFAND3 functions in the DNA damage response pathway to promote genome integrity.

Body

As indicated in the statement of work, the aims of this proposal include: characterizing ZFAND3 function in the DDR, identifying ZFAND3 interacting proteins, and examining ZFAND3 regulation after DNA damage.

Task 1: Characterize the function of ZFAND3 in the DNA damage response (partially completed)

To further analyze the function of ZFAND3 in the DNA damage response, we first examined the cell cycle profile of ZFAND3 depleted cells. ZFAND3 silenced cells continue to progress through the cell cycle with a very slight decrease in the percentage of cells in the G1 phase (Fig. 1). This
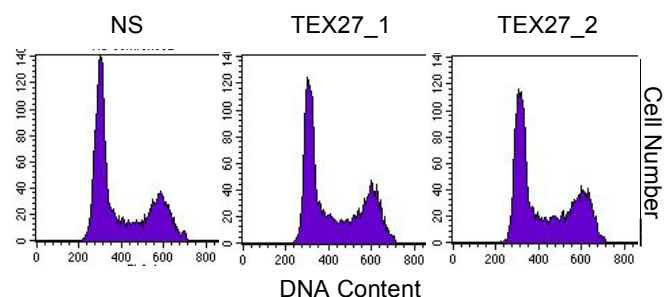
Figure 1. Cell cycle analysis of ZFAND3 depleted cells. ZFAND3 was silenced using two distinctive siRNAs (TEX27_1 and TEX27_2), cells were stained with propidium iodide, and DNA content was examined by flow-cytometry.

indicates that ZFAND3 is not essential for cell cycle progression. Furthermore, it suggests that ZFAND3's role in eliciting a G2/M arrest in response to ionizing radiation (preliminary data) is not an indirect effect of silencing ZFAND3 and arresting cells in the cell cycle.
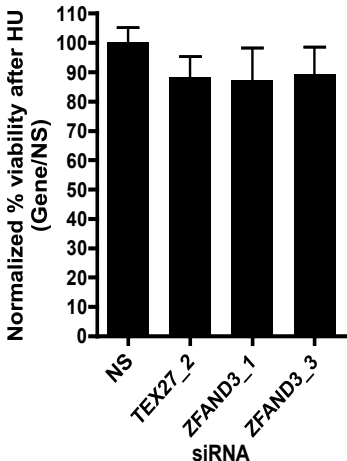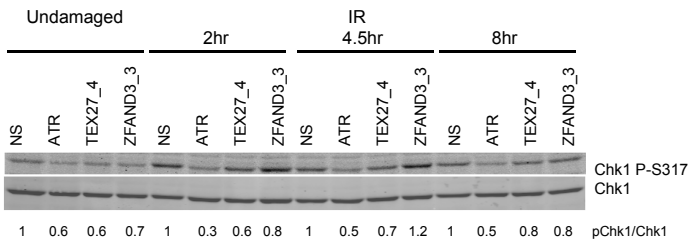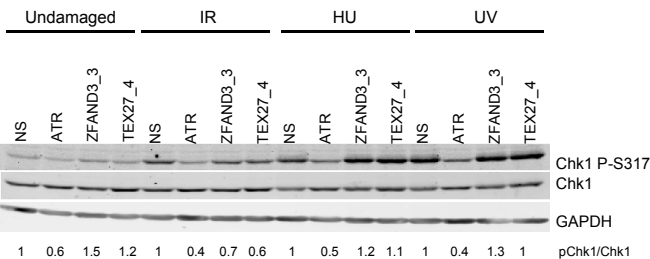
Next, as proposed, we investigated whether ZFAND3 maintains cellular viability after DNA damage. ZFAND3 silencing did not significantly increase cellular sensitivity to the replication stress reagent hydroxyurea (HU). This data suggests that ZFAND3 is not a major determinant of cell survival to damage encountered during DNA replication. It remains to be tested whether ZFAND3 depletion hypersensitizes cells to double-strand break reagents such as ionizing radiation (IR) or mitomycin C.

Figure 2. ZFAND3 silencing does not hypersensitize cells to HU. Sensitivity to HU was determined after depletion of ZFAND3 (using siRNAs labeled TEX27_2, ZFAND3_1, and ZFAND3_3) by measuring the cell viability in a colorimetric proliferation assay. Normalized % cellular viability after HU was calculated as the ratio of treated/untreated for each ZFAND3/non-targeting (NS) siRNA. Error bars represent SEM from three independent repeats. Statistical significance was calculated using a t-test as described (Lovejoy 2009).

Lastly, the involvement of ZFAND3 in DNA damage signaling following treatment with various genotoxic agents was examined. Signaling from the apical DDR kinase ATR (ataxia telangiectasia-mutated (ATM) and Rad3-related) occurs through phosphorylation of the transducing kinase Chk1 (checkpoint kinase 1) (3). While silencing of ATR impaired phosphorylation of Chk1 after HU and UV radiation, ZFAND3 depletion had no noticeable effect on Chk1 phosphorylation (Fig. 3). A minor defect in signaling was observed when cells depleted

of ZFAND3 were treated with IR, particularly at the two-hour time-point (Fig. 3B). Taken together, ZFAND3 contributes minimally to ATR-dependent checkpoint signaling after IR. It remains possible that the knock-down

Figure 3. ZFAND3 regulation of DDR signaling. **(A)** Cells were treated with 5Gy of IR, 2mM of HU, or 50J/M^2 of UV radiation and collected after 1.5hr, 6hr, or 1hr, respectively. Immunoblotting with indicated antibodies against phosphorylated Chk1 or total Chk1 was detected and quantified with an Odyssey scanner. **(B)** Cells treated with 5Gy of IR were collected at the indicated time-points. Immunoblotting was performed as above.

efficiency of ZFAND3 targeting siRNAs is poor. To address this concern, we raised an antibody targeting ZFAND3 (see Task 2 below, Fig. 5).
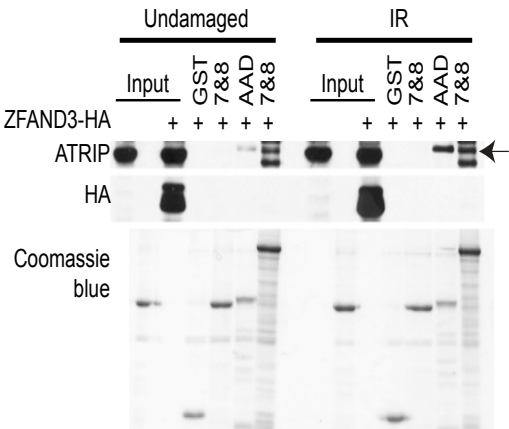
Task 2: Identify ZFAND3 interacting proteins (partially completed)

For this task, we focused on confirming the yeast two-hybrid (Y2H) interaction between TopBP1 and ZFAND3. Several GST-fused fragments of TopBP1 were incubated with cell lysates expressing HA-tagged ZFAND3 (Fig. 4). No interaction between TOPBP1 and ZFAND3 was detectable irrespective of DNA damage. However, the ATR interacting protein ATRIP bound to TopBP1 in this assay, as previously reported (6). Co-immunoprecipitation experiment confirmed these results under various lysis conditions (data not shown). Therefore, the ZFAND3 interaction with TopBP1 could not be confirmed using the proposed methodologies.

Figure 4. ZFAND3 interaction with TopBP1 is undetectable in TopBP1-GST pulldown assays. Nuclear extracts left untreated or damaged with 5Gy IR were prepared from 293T cells transfected with vectors encoding ZFAND3-HA or empty vector. Extracts were incubated with GST-tagged recombinant TopBP1 fragments (BRCT repeats 7&8 and ATR activation domain (AAD)) immobilized on glutathione beads (Mordes 2008). Proteins bound to TopBP1 were eluted, resolved by SDS-PAGE, and immunoblotted with the indicated antibodies. Input is 5% of total cell extract from the pulldown assays. Arrow indicates band corresponding to ATRIP.
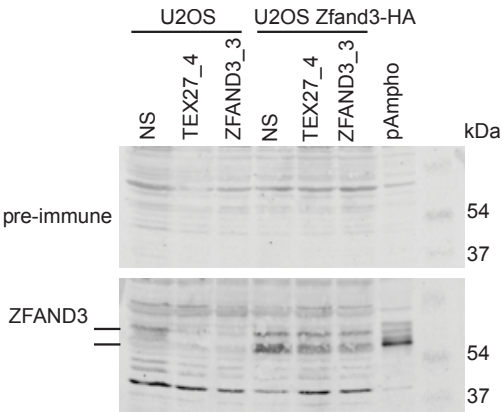
We generated an antibody to detect ZFAND3 (Fig. 5). The antibody recognizes ZFAND3-HA overexpressed in U2OS and Phoenix ampho cells. Silencing ZFAND3 does not appear to significantly alter the levels of bands visible in cells expressing ZFAND3-HA. This antibody may recognize some endogenous ZFAND3 protein (see bands in untagged cell line), but further testing will be required to validate this antibody. The availability of a strong antibody is critical to assess ZFAND3 knockdown in functional studies (Task 1) and for use in protein-protein interactions.

Task 3: Determine ZFAND3 regulation after DNA damage (completed)

The final task of the proposal was to examine the

Figure 5. Antibody raised targeting FAND3. Cell lines expressing HA-tagged Zfand3 or no overexpressed protein were tested using pre-immune and ZFAND3 antibody after depletion of endogenous ZFAND3 using siRNA. pAmpho cells over-express HA-tagged ZFAND3 (above 54kDa). Two lines indicate potential bands for ZFAND3

whether ZFAND3 expression or protein stability is regulated by genotoxic stress. Initial results suggested that when ZFAND3 is expressed from a heterologous promoter its levels increase after ionizing radiation (data not shown). Thus we examined the half-life of HA-tagged ZFAND3 protein (Fig. 6). ZFAND3 protein half-life is not reproducibly altered by ionizing radiation. Furthermore, we failed to detect changes in ZFAND3 mRNA in published data sets examining DNA damage-regulated transcripts. Thus, we do not believe ZFAND3 expression is regulated by DNA damage.
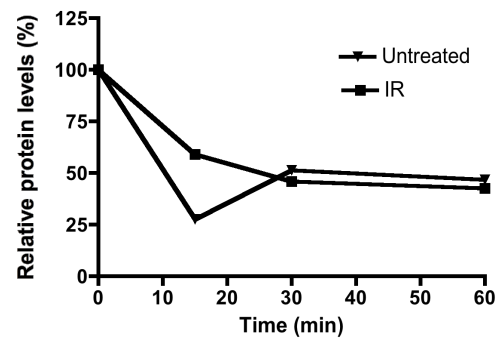
Figure 6. Analysis of ZFAND3 protein half-life. Relative protein levels of ZFAND3-HA were determined over time after cycloheximide treatment in the presence and absence of IR. Protein levels were quantified using an Odyssey system.

New research direction:

As reported last year, I have expanded my studies of genome maintenance during DNA replication using a new technology that I developed called iPOND (7,8). In the past year I have used this methodology in combination with mass spectometry to identify proteins associated with active, stalled, and collapsed replication forks. We are still analyzing the large data-set but ZFAND3 was not identified in this analysis.

Key Research Accomplishments

- Examined ZFAND3's role in the DNA damage response and found that ZFAND3 depletion causes very mild phenotypes in DDR signaling after two hours of IR damage, but not HU or UV radiation.

- ZFAND3 does not appear to play a significant role in cell cycle progression or in protecting cell viability after replication stress.

- Biochemical studies of ZFAND3 protein-protein interactions were unable to confirm an interaction with TopBP1.

- Confirmation of an antibody raised to recognize endogenous ZFAND3 requires further testing.

- ZFAND3 mRNA and protein half-life are not regulated by DNA damage.

Training

My training in molecular cancer biology has continued in the past year. Primary training activities include attendance at the following seminar series: Vanderbilt-Ingram Cancer Center, bi-monthly seminars, Center in Molecular Toxicology weekly seminars, Department of Biochemistry weekly seminars, and periodic Genome Maintenance and Breast Cancer SPORE seminars. I also participated in several research symposia including DOD Breast Cancer Program Era of Hope Conference and the Vanderbilt Institute of Chemical and Physical Biology retreat. I presented a poster on my genome maintenance research at both of these symposia.

I have also continued to have regular meetings with my research advisor, Dr. David Cortez as well as my thesis committee. I presented both written and oral progress reports at these meetings and received feedback on research directions.

Reportable outcomes

1. All proposed coursework indicated in Statement of Work has been completed
2. Attended Vanderbilt Ingram Cancer Center bi-monthly seminars
3. Attended the yearly Genome Maintenance Seminar and Molecular Toxicology seminars relevant to cancer etiology
4. Poster presentations of research findings
    a. Title: Analysis of protein dynamics at active, stalled and collapsed replication forks
    b. Authors: **Sirbu BM**, Couch FB, Feigerle JT, Bhaskara S, Hiebert SW, Cortez D
    c. Meeting: Department of Defense Breast Cancer Program Era of Hope Conference August 2011
5. Poster presentations of research findings
    a. Title: Analysis of protein dynamics at active, stalled and collapsed replication forks
    b. Authors: **Sirbu BM**, Couch FB, Feigerle JT, Bhaskara S, Hiebert SW, Cortez D
    c. Retreat: Vanderbilt Institute of Chemical and Physical Biology August 2011
6. Awarded Vanderbilt Prize Scholar 2011
7. Awarded Vanderbilt Ingram Cancer Center Graduate Student of the Year 2012
8. Received funding from Swim Across America $50,000 to support cancer research 2012
9. Published detailed protocol for the iPOND methodology (see appendix)
    a. Title: Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromating using isolation of proteins on nascent DNA.
    b. Authors: **Sirbu BM**, Couch FB, Cortez D
    c. Published: February 2012 *Nature Protocols.* 25(12):1320-7
10. Co-authored on publication describing the role of the genome maintenance protein SMARCAL1 (see appendix)
    a. Title: SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication
    b. Authors: Betous R, Mason AC, Rambo RP, Bansbach CE, Badu-Nkansah A, **Sirbu BM**, Eichman BF, Cortez D.
    c. Published: January 2012 *Genes and Development* **26**(2):151-62*.*

Conclusions and Future Directions

My research on ZFAND3 has characterized its regulation and function in the replication stress response. Based on my studies, ZFAND3 likely has only a minor role in protecting genome integrity in breast cancer cell lines. While this conclusion is disappointing, I have made significant progress on the larger goal of understanding how breast cancer cells maintain their genomes during DNA replication. That progress includes the development and implementation of the iPOND technology to understand protein dynamics at active, stalled and damaged replication forks. These studies resulted in two high-impact publications (7,8) and several research awards. My training activities are continuing to prepare me for a career as a molecular cancer biologist. My research focus in the next year will be to complete the study of proteins at damaged replication forks and publish these results. A major training goal in the next funding period will be to improve my scientific writing skills under the tutelage of Dr. Cortez. To achieve this goal, I am writing a review article for submission to a special issue of DNA Repair.

References

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

2. Schar, P. (2001). Spontaneous DNA damage, genome instability, and cancer—when DNA replication escapes control. Cell 104:329-32.

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

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

5. Lovejoy CA, Xu X, Bansbach CE, Glick GG, Zhao R, Ye F, Sirbu BM, Titus LC, Shyr Y, Cortez D. (2009). Functional genomic screens identify CINP as a genome maintenance protein. Proc Natl Acad Sci U S A 106(46):19304-9.

6. Mordes DA, Glick GG, Zhao R, Cortez D. (2008) TopBP1 activates ATR through ATRIP and a PIKK regulatory domain. Genes Dev 22(11):1478-9.

7.   Sirbu BM, Couch FB, Feigerle JT, Bhaskara S, Hiebert SW, Cortez D. (2011). Analysis of protein dynamics at active, stalled and collapsed replication forks. Genes Dev 25(12):1320-7.

8. Sirbu BM, Couch FB, Cortez D. (2012). Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromating using isolation of proteins on nascent DNA. Nature Protocols 25(12):1320-7.

Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA

Bianca M Sirbu, Frank B Couch & David Cortez

Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. Correspondence should be addressed to D.C. (david.cortez@vanderbilt.edu).

Understanding the processes of DNA replication, chromatin assembly and maturation, and the replication stress response requires the ability to monitor protein dynamics at active and damaged replication forks. Detecting protein accumulation at replication forks or damaged sites has primarily relied on immunofluorescence imaging, which is limited in resolution and antibody sensitivity. Here we describe a procedure to isolate proteins on nascent DNA (iPOND) that permits a high-resolution spatiotemporal analysis of proteins at replication forks or on chromatin following DNA replication in cultured cells. iPOND relies on labeling of nascent DNA with the nucleoside analog 5-ethynyl-2′-deoxyuridine (EdU). Biotin conjugation to EdU-labeled DNA using click chemistry facilitates a single-step streptavidin purification of proteins bound to the nascent DNA. iPOND permits an interrogation of any cellular process linked to DNA synthesis using a 3- to 4-d protocol.

INTRODUCTION

During S-phase, DNA replication and chromatin assembly are coordinated at the replication fork to duplicate the genome and epigenome rapidly and accurately. DNA template damage and other forms of replication stress challenge genetic stability and activate a DNA damage response[1]. This signaling pathway protects and repairs damaged replication forks to promote successful completion of chromosome replication and prevent diseases such as cancer[2].

Immunofluorescence imaging is a useful method to detect proteins in active replisomes or proteins recruited to damaged forks. However, immunofluorescence imaging suffers from low resolution, poor sensitivity and a requirement for highly specific antibodies[3]. Other methods such as chromatin immunoprecipitation (ChIP) have limited applicability to mammalian cell replication because of difficulties in obtaining synchronous cultures and the lack of highly efficient, sequence-specified origins of replication[4]. Purification of replisome protein complexes through protein-protein interactions is useful to identify potential components, but it provides limited spatial information about protein localization.

To overcome these technical challenges, we developed iPOND[5]. In addition to its use for monitoring replisome dynamics, iPOND provides a method for examining protein recruitment and modification at damaged replication forks and for analyzing chromatin deposition and maturation.

Overview of iPOND

The iPOND methodology enables the purification of proteins bound directly or indirectly to the nascent DNA at replication forks. The method relies on labeling short fragments of nascent DNA with EdU, a nucleoside analog of thymidine[6]. EdU contains an alkyne functional group that permits copper-catalyzed cycloaddition (click chemistry)[7] to a biotin azide to yield a stable covalent linkage (**Fig. 1**). This facilitates a single-step purification of DNA-protein complexes based on the high-affinity biotin-streptavidin interaction.

The iPOND procedure (**Fig. 2**) begins by incubating cells with EdU for a short period of time (typically 2–15 min). The cells are

then fixed with formaldehyde, which serves to both stop DNA replication and cross-link protein-DNA complexes. A click reaction in the presence of copper to conjugate biotin to EdU is completed in detergent-permeabilized cells. Some DNA fragmentation occurs during this step because of copper-catalyzed hydrolysis of the DNA. Cells are then lysed in denaturing conditions and sonication completes the DNA fragmentation producing solubilized DNA-protein complexes. Streptavidin-coated beads purify the nascent, EdU-labeled DNA-protein complexes. Finally, the proteins are eluted from the complexes. Standard immunoblotting or mass spectrometry (MS) methodologies can be used to detect the purified proteins and post-translational modifications.

The spatial and temporal resolution achieved with iPOND depends on EdU incubation time, the rate of DNA synthesis and chromatin fragment size. Experimentally, EdU incubation time and replication rate are the major determinants of iPOND resolution, as the protocol consistently yields chromatin fragments of 100–300 bp. The shortest EdU incubation time we have used to purify replisome components is 2.5 min (ref. 5). As forks move between 750 and 2,000 bp min^{-1} (ref. 8), as much as 5,000 bp could contain EdU during a 2.5-min incubation, yielding a resolution of 5 kb. This is likely to be a substantial underestimation of the resolution because EdU must enter the cell and be phosphorylated before incorporation. Our analysis indicates that a 2.5-min incubation with EdU is sufficient to capture replisome proteins and that longer incubations with EdU are required to isolate newly deposited chromatin[5].

Applications

Thus far, we have used iPOND in three major applications. First, iPOND is useful for identifying proteins associated with active replisomes. This application requires combining iPOND within a pulse-chase experimental framework (**Figs. 3** and **4**). Cells are labeled with EdU for a short time (the pulse), and then EdU is replaced with thymidine for increasing periods of time (the chase). Samples are collected at the end of the pulse and chase periods. A true replisome protein should be detected only in the pulse sample
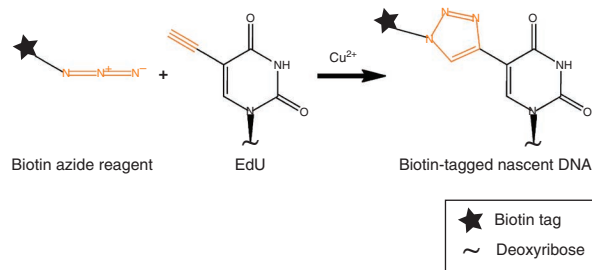
Figure 1 | Click chemistry addition of biotin tags to nascent DNA. EdU incorporated into nascent DNA is covalently tagged with biotin in the copper-catalyzed click reaction. Orange color represents the functional groups involved in the click chemistry reaction.

and not the chase sample. In contrast, other chromatin-bound proteins such as histones may be detected in both samples.

The pulse-chase experimental design is also the method of choice for the second major iPOND application—monitoring changes in chromatin located at various distances from the replication fork. Chromatin reassembly after passage of the replication fork occurs as a function of time and hence distance from the elongating fork[9]. The use of iPOND to purify histones on a segment of EdU-labeled DNA after various times of thymidine chase permits an analysis of how chromatin architecture is restored behind the elongating fork. For example, we used iPOND to document the timing of the deacetylation of newly synthesized histone H4 after deposition[5].

Finally, iPOND can be used to detect protein recruitment or post-translational modifications of proteins at damaged forks. The procedure in this case is to pulse for a short time with EdU, then to add a replication stress agent such as hydroxyurea (HU) or camptothecin (**Fig. 3b**). HU is particularly useful as high concentrations largely stop fork movement, facilitating an analysis of transiently or persistently stalled forks. Combining the DNA damaging protocol with the pulse-chase procedure also enables an examination of DNA damage–dependent events at different distances from the damaged fork. For example, we used this procedure to demonstrate spreading of H2AX phosphorylation from an HU-stalled fork[5]. Thus, the high spatial resolution of iPOND is derived from the capacity to measure the position of protein changes in relation to the replication fork. Theoretically, this system can also be used to monitor long-term changes in chromatin structure after DNA damage or replication stress by simply extending the time frame of the chase.

These three major applications are quite powerful, especially when combined with genetic or small molecule–mediated inactivation of specific pathways that regulate DNA replication, chromatin deposition and maturation, and DNA repair. iPOND is compatible with all proliferating cell types. We have used it successfully in HEK293T, HCT116, NIH3T3 and mouse embryonic fibroblasts (B.M.S. and D.C., unpublished observations). Thus, cell lines engineered to have mutations in specific pathways can be used directly with iPOND without any major modifications to the protocol. iPOND can also be extended for use beyond mammalian cell culture. Any cell type that can incorporate EdU during DNA synthesis (or be engineered to use EdU) can be used. In fact, we have used iPOND to purify DNA-protein complexes from the yeast *Saccharomyces cerevisiae*, although substantial optimization will be required to improve purification efficiency (F.B.C. and D.C., unpublished observations).

In addition to these three documented applications, iPOND can be used to study other processes that involve DNA or even RNA synthesis. An example would be DNA repair synthesis outside of S-phase. Synchronized or terminally differentiated cell cultures could be exposed to DNA damaging agents in the presence of EdU. The late steps in repair of that damage or the re-establishment of chromatin following repair synthesis can be monitored with iPOND. Synchronized cell cultures could also be used to examine the differences in DNA replication, chromatin deposition or DNA repair that occur in early versus late S-phase cells. This approach was recently used by Kliszczak *et al.*[10] to describe a methodology similar to iPOND. Another application could be to monitor DNA synthesis outside of the nucleus such as in mitochondrial DNA if iPOND is combined with a purification step that isolates this organelle. iPOND could theoretically be adapted to analyze even proteins on nascent RNA, as click chemistry has been used to label newly synthesized RNA with the uridine analog 5-ethynyluridine[11].

Finally, combining iPOND with quantitative MS should be a powerful methodology for identifying new replisome and DNA damage response proteins, as well as for monitoring the substantial numbers of post-translational modifications at damaged forks.
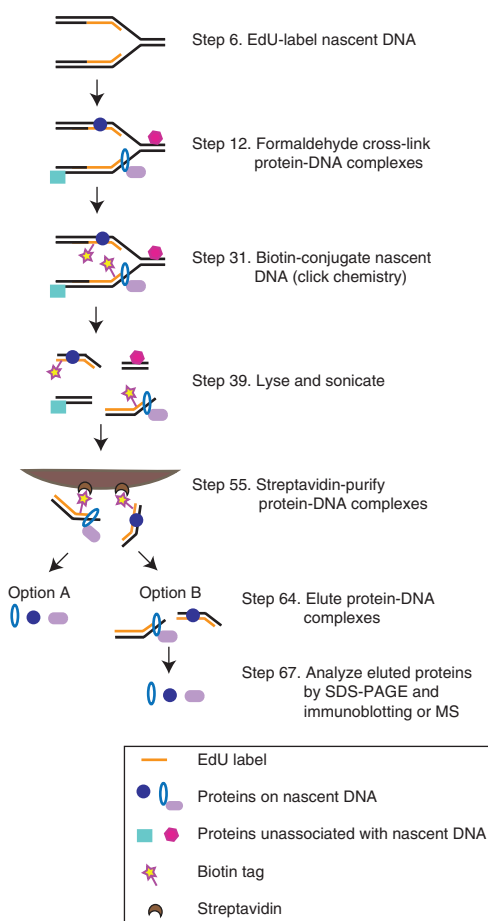
Figure 2 | Schematic overview of the iPOND procedure. The iPOND procedure consists of pulsing cells with EdU to label nascent DNA *in vivo*, formaldehyde cross-linking protein-DNA complexes, covalently tagging EdU-labeled DNA with biotin by using click chemistry, lysing and sonicating cells, purifying the solubilized protein-DNA complexes and eluting bound proteins for analysis by SDS-PAGE, and immunoblotting or MS.
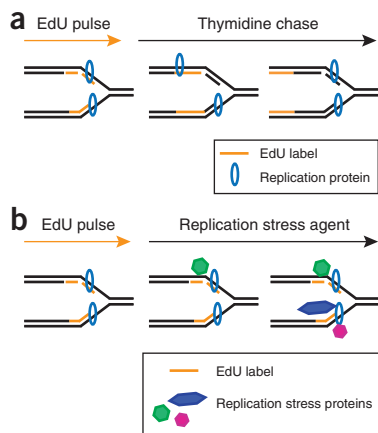
Figure 3 | Schematic of the experimental procedures used to identify replisome or DNA damage proteins and modifications at the replication fork. (**a**) To identify replisome proteins, a pulse-chase variation of the iPOND protocol uses a thymidine chase to move the nascent, EdU-labeled DNA segment away from the replication fork. The chase sample provides a control to distinguish replisome components from general chromatin-binding factors. (**b**) To study proteins and modifications associated with damaged replication forks, an agent that stalls replication forks, such as HU, is added after the EdU-labeling period.

Comparison with other methods

Compared with conventional indirect immunofluorescence, iPOND has an improved sensitivity of detection because even low-abundance replisome proteins such as polymerases are isolated[5]. It also provides improved spatial and temporal resolution. An improved imaging technique permits single-molecule detection of replisome proteins in bacteria[12]; however, unlike imaging, iPOND is compatible with unbiased approaches for protein identification such as MS.

ChIP is a powerful substitute for several iPOND capabilities in organisms such as *S. cerevisiae* that have highly efficient, sequence-defined origins of replication and cell cycle synchronization is easily achieved. ChIP has the advantage of being more sensitive than iPOND as it detects DNA sequences after PCR amplification. However, ChIP requires highly specific, often unavailable antibodies and is not compatible with unbiased approaches such as MS. Moreover, although ChIP has been used in mammalian systems to examine protein recruitment to origins of replication[13], it is generally not useful for studying the dynamic processes associated with fork elongation and chromatin maturation. Finally, adapting ChIP to studying damaged replication forks in mammalian cell culture awaits the development of ways to engineer site-specific DNA lesions that stall forks with high efficiency as has been done using *Xenopus* egg extracts to study interstrand cross-link repair[14,15].

The most comparable technology to iPOND is the immunoprecipitation of nascent DNA-protein complexes with antibodies to halogenated nucleoside analogs, which was used to examine the recruitment of the homologous recombination factor RAD51 to sites of replication fork stalling[16]. However, the relatively low affinity of this antibody-epitope interaction and the requirement for DNA denaturation for antibody access necessitated a very long chlorodeoxyuridine (CldU)-labeling period (40 min), providing little advantage over biochemical fractionation of chromatin. In principle, biotin-dUTP could be used directly to label the nascent DNA, thus avoiding the need to perform the click chemistry reaction. However, biotin-dUTP

is not cell permeable, thus necessitating some cellular manipulation to introduce it into cells, and the large biotin tag may interfere with DNA structure and protein associations with DNA.

Experimental design

Several parameters can be varied within the iPOND protocol depending on the specific experimental purpose. As outlined above, the EdU pulse and chase combinations yield different types of information. In addition, it may be useful to omit the formaldehyde cross-linking step. In particular, formaldehyde cross-linking may complicate analysis of proteins by MS if the cross-links are not fully reversed. Chromatin can be captured with iPOND without cross-linking, provided that Igepal or another nondenaturing detergent is used in the lysis step and that the salt concentration in the wash step is reduced (**Box 1**).

A second experimental design option is to change the elution methodology. For most applications we found that boiling in SDS sample buffer is sufficient to reverse cross-links and solubilize proteins after purification (**Fig. 2**, elution option A). However, this method also releases any proteins that bind to the bead matrix nonspecifically and does not release the DNA from the beads. The use of a cleavable biotin azide in the click reaction facilitates elution in milder conditions to improve specificity and recovery of the DNA (**Fig. 2**, elution option B). Several cleavable biotin azides have been described[17]. We successfully use a UV-photocleavable biotin-azide synthesized by Ned Porter's group at Vanderbilt[18]. This elution option may also be useful in experimental systems where biotinylation of endogenous proteins is a concern.

Controls

Control samples are essential for interpreting the results. Most importantly, a control for the specificity of the purification is needed. This control is analogous to the preimmune control for coimmunoprecipitation experiments. We typically use a sample that omits the biotin azide during the click reaction (Steps 29 and 30). Alternatively, a sample in which the cells were not incubated with EdU can be used as the control. No DNA-protein complexes should be purified in this control sample. If any protein is detected, it is likely to come from nonspecific interactions with the streptavidin matrix or precipitation of protein during the manipulations.
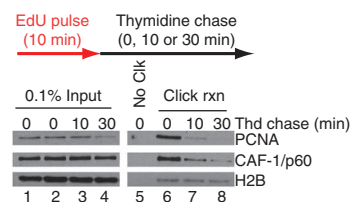
Figure 4 | Example of results obtainable with iPOND. Cells were pulsed with EdU for 10 min and then incubated with thymidine (Thd) for 0, 10 or 30 min as indicated. iPOND was performed as described in the protocol. Eluted proteins were analyzed by SDS-PAGE followed by immunoblotting for the replication proteins PCNA, chromatin assembly factor 1 (CAF-1/p60) and histone H2B. As expected, proteins are detectable in every sample of the input (lanes 1–4). In the absence of click chemistry (No Clk, lane 5, negative control), no proteins are isolated from nascent DNA. PCNA and CAF-1 are enriched specifically at the replication fork (Click rxn, lane 6), but not on nascent DNA that is thymidine chased away from the replication fork (Click rxn, lanes 7 and 8). In contrast, a chromatin-bound protein such as H2B is detectable both at the replication fork (Click rxn, lane 6) and in thymidine-chased samples (Click rxn, lanes 7 and 8).

Box 1 | Native iPOND

iPOND performed without formaldehyde cross-linking (native iPOND) may simplify mass spectrometry analyses of purified histones.

1. Culture 5×10^7 cells in one 150-mm dish per sample.
2. Label the samples with 10 μM EdU for 60 min.
3. Collect the cells by scraping on ice.
4. Collect the pellets by centrifuging at 100g for 5 min at 4 °C.
5. Discard the supernatant and wash the cells with 5 ml of ice-cold PBS. Collect the cells by centrifuging at 100g for 5 min at 4 °C.
6. Discard the supernatant and lyse the cells by resuspension in ice-cold cell lysis buffer with Igepal CA-630 at 1×10^7 cells per ml.
7. Vortex five times for 5 s with 5 s between pulses.
8. Collect nuclei by centrifugation at 100g for 5 min at 4 °C.
9. Discard the supernatant and wash twice in 5 ml cell lysis buffer without Igepal CA-630.
10. Collect nuclei by centrifugation at 100g for 5 min at 4 °C.
11. Discard the supernatant and resuspend the cells in ice-cold nuclei buffer at 2.5×10^7 cells per ml.
12. Set up click reactions using the formula in **Table 1**.
13. Incubate for 1 h on a shaker at 4 °C and protect from light.
14. Collect the nuclei by centrifugation at 100g for 5 min at 4 °C.
15. Discard the supernatant and resuspend in ice-cold nuclei buffer at 2×10^7 cells per ml.
16. Add EDTA to a final concentration of 1 mM and $CaCl_2$ to 2 mM.
17. Warm to 37 °C in a water bath and add micrococcal nuclease to 20 Kurntz units per 1×10^7 cells.
18. Incubate the cells at 37 °C for 3.5 min.
19. Add EDTA to a final concentration of 2 mM to quench the reactions. Collect nuclei by centrifugation at 100g for 5 min at 4 °C.
20. Extract chromatin by discarding the supernatant and resuspending the nuclei in ice-cold extraction buffer at 5×10^7 cells per 3 ml. Rotate for 2 h to overnight at 4 °C, protected from light.
21. Centrifuge at 16,100g for 5 min at 4 °C to remove all insoluble material. Transfer the supernatant to a fresh tube and discard the pellet.
22. Remove 0.5% of the total volume and save it as the 'input' sample. To the remaining lysate, add 20 μl of streptavidin-agarose beads per 1×10^7 cells. Rotate for 1.5 h to overnight at 4 °C, protected from light.
23. Collect the beads by centrifugation at 1,800g for 1 min. Let the beads stand for another minute to settle completely. Aspirate and discard the supernatant.
24. Transfer the beads to a 1.5-ml centrifuge tube.
25. Wash the beads twice with 1 ml of extraction buffer for 5 min at 4 °C.
26. Add an equal volume of 2× SB and heat to 95 °C for 10 min.
27. Separate the recovered proteins with SDS-PAGE and analyze them by immunoblotting or mass spectrometry.

A second control is needed to ensure that a purified protein is actually enriched at replication forks, as opposed to simply being an abundant chromatin-associated protein. This control is a sample in which the EdU is removed and cells are incubated with thymidine for several minutes before collecting (a chase sample). Proteins that travel with the replication fork will only be detected before this thymidine chase. If a protein is detected in the chase sample, this indicates that it is a chromatin-bound protein but not specifically part of the replisome.

Finally, control immunoblots to examine known replisome components, such as proliferating cell nuclear antigen (PCNA), should be performed within each experiment to ensure that the procedure worked as expected.

iPOND limitations and other considerations
Currently, the major limitation of iPOND is the large amount of starting material needed. Each sample requires approximately 1×10^8 cells for efficient iPOND capture of replisome proteins with a 10-min EdU incubation. The large number of cells needed for the procedure is dictated by the sensitivity of the immunoblotting and MS detection methods. This cell number is based on unsynchronized cultures of 293T cells in which about 50% of the cells are in S-phase at the time of the experiment. Synchronizing cells such that 100% are in S-phase would reduce the cells needed, whereas

the use of cell types with fewer replicating cells would increase it. Although these cell numbers are large, they are obtainable by using standard cell culture methods.

iPOND is an ensemble methodology, meaning that the data comes from hundreds of replication forks in millions of cells. It provides a picture of an average replication fork and cannot distinguish the significant heterogeneity between cells in the population or between forks within different genomic regions. Thus, identification of two proteins by iPOND does not mean that those two proteins are necessarily recruited to the same nascent DNA segment. Furthermore, distinguishing the relative distribution of proteins within the chromosomal space at the replication fork is currently not possible with iPOND. Such high-resolution mapping has been achieved with *in vitro* replication systems by using T4 DNA polymerase and primer template DNA that contains a position-specific cross-linkable aryl azide[19]. This elegant study provided topographical information about the location of binding of accessory proteins respective to polymerase interaction with and movement along the DNA template. Finally, iPOND resolution may be improved in a system in which EdU exists as the sole nucleoside to pair with adenosine. This could be achieved in a cellular system such as *Xenopus*, in which dNTPs are added in a controlled manner for incorporation into nascent DNA.

MATERIALS

REAGENTS
- EdU (Invitrogen, cat. no. E10187)
- Thymidine (Sigma, cat. no. T1895)
- Formaldehyde solution (37% (wt/vol); Sigma, cat. no. F1635) **! CAUTION** Formaldehyde is very toxic if inhaled, ingested or absorbed through skin.
- PBS, pH 7.2 (10×; Gibco, cat. no. 70013)
- Glycine (Fisher, cat. no. BP 381)
- Cell lifter (Corning, cat. no. 3008)
- Triton X-100 (Sigma, cat. no. T8787) **! CAUTION** Hazardous in case of eye contact, ingestion or inhalation.
- BSA (Sigma, cat. no. A7030)
- Dimethyl sulfoxide (DMSO; Fisher, cat. no. A4034) **! CAUTION** It readily permeates skin, is a combustible liquid and vapor, and is hygroscopic.
- Copper (II) sulfate pentahydrate ($CuSO_4 \cdot 5H_2O$; Fisher, cat. no. C489) **! CAUTION** It is toxic if swallowed and causes skin and eye irritation.
- (+) Sodium L-ascorbate (Sigma, cat. no. A4034)
- Biotin azide (Invitrogen, cat. no. B10184)
- SDS (Sigma, cat. no. L4390) **! CAUTION** SDS is toxic on contact with skin, harmful if swallowed, and causes skin, eye and respiratory irritation.
- Tris, pH 8.0 and 6.7
- Sodium chloride (NaCl)
- RNase A solution (Sigma, cat. no. R6148)
- Proteinase K (Sigma, cat. no. P5568)
- Glycerol
- Bromophenol blue
- EDTA
- Agarose (Bio-Rad Laboratories, cat. no. 161-3101)
- Dithioerythritol (DTT; Sigma, cat. no. D-8255)
- Aprotinin (Sigma, cat. no. A6279)
- Leupeptin (Sigma, cat. no. L2884)
- Streptavidin agarose (Novagen, cat. no. 69203-3) **▲ CRITICAL** Different bead surfaces and binding capacities will alter iPOND efficiency.
- Trichloroacetic acid (TCA) **! CAUTION** TCA is corrosive.
- Acetone **! CAUTION** It is highly flammable.
- Western Lightning Plus enhanced chemilluminescence substrate (PerkinElmer, cat. no. NEL103001EA)
- Igepal CA-630
- Odyssey infrared imaging system (Li-Cor Biosciences)
- Click reaction stock solutions (biotin azide, $CuSO_4$ and sodium L-ascorbate; see REAGENT SETUP)

EQUIPMENT
- Nylon mesh (90 μm; Small Parts, cat. no. B000FN0PGQ)
- Glass vial screw thread with cap attached for UV photocleavage (Fisher, cat. no. 03-338AA) **▲ CRITICAL** Other vial surfaces may perturb penetration of UV light, necessitating different elution times or conditions.
- Magnetic micro-stirring bar (2 mm diameter × 7 mm length; Fisher, cat. no. 1451363)
- Microtip sonicator for cell lysis and chromatin fragmentation (Misonix 4000 or Fisher Scientific Sonic Dismembrator, model 500)
- Rotating platform for biotin captures
- UV lamp (UVP, cat. no. UVLMS-38 EL Series 3UV lamp, 365/302/254 nm UV 8 Watt)
- Magnetic stir plate
- Microcentrifuge for 1.5-ml microcentrifuge tubes
- Tabletop centrifuge for 15-ml and 50-ml conical tubes
- Cell culture incubator
- Biological safety cabinet

REAGENT SETUP
EdU Dissolve EdU in DMSO to obtain a final concentration of 10 mM. Protect from light. Store in aliquots at −20 °C for up to 1 year. Before use, thaw at 37 °C. To EdU-label cells, pipette 1:1,000 of EdU directly into medium for a final concentration of 10 μM.

Thymidine Dissolve in water to a final concentration of 10 mM. Store in aliquots at −20 °C for up to 1 year. Thaw the solution before use. Use at a final concentration of 10 μM.

PBS, 1× Prepare 1× PBS from 10× PBS stock by diluting 1:10 with water; store at room temperature (RT, 25 °C) for up to 1 year.

Formaldehyde/PBS, 1% (wt/vol) Dilute 37% (wt/vol) formaldehyde 1:37 with PBS. Freshly prepare this reagent and keep it at RT until cell fixation (PROCEDURE, Step 12).

Glycine, 1.25 M Prepare 1.25 M glycine stock in water and store at RT for up to 1 year. Use at 1:10 dilution for a final concentration of 0.125 M glycine.

Permeabilization buffer Prepare a 20% (vol/vol) stock of Triton X-100 in water and keep it at RT. Dilute to 0.25% (vol/vol) Triton X-100 in PBS. Store at 4 °C for several months.

BSA in PBS wash buffer, 0.5% (wt/vol) Prepare 0.5% (wt/vol) BSA in PBS. Filter-sterilize the solution and store it at 4 °C for a couple of weeks.

Biotin azide (1 mM) Dissolve biotin azide in DMSO to a final concentration of 1 mM. Aliquot and store at −20 °C for up to 1 year.

$CuSO_4$ (100 mM) Prepare a stock of 100 mM $CuSO_4$ in H_2O; store at RT for several months.

Sodium L-ascorbate Freshly prepare 20 mg ml^{-1} of (+) sodium L-ascorbate (reducing agent) in H_2O; limit exposure to air and store on ice until needed.

Click reaction mixes To prepare click reaction cocktails, please see **Table 1** for details. Cocktails are freshly prepared for each experiment before the click reaction (PROCEDURE, Step 28).

Lysis buffer Prepare 1% (wt/vol) SDS in 50 mM Tris (pH 8.0). Store at RT for several months. Before use, add protease inhibitors aprotinin and leupeptin to a final concentration of 1 μg ml^{-1}.

Salt wash Prepare 5 M NaCl in water. Dilute to 1 M NaCl with water before use. Store at RT for 1 year.

SDS Laemmli sample buffer (2× SB) Mix 0.4 g of SDS, 2 ml of 100% glycerol, 1.25 ml of 1 M Tris (pH 6.8), and 0.01 g of bromophenol blue in 8 ml of H_2O. Store at −20 °C for up to 1 year. Before use, add 1 M DTT to a final concentration of 0.2 M.

Cross-link reversal solution Mix 2 μl of 0.5 M EDTA, 4 μl of 1 M Tris (pH 6.7) and 1 μl of Proteinase K. Freshly prepare this solution.

TABLE 1 | Click reaction cocktails for a sample with 1×10^8 cells.

Reagent	Stock	Final	Control reaction volume (ml)	Experimental reaction volume (ml)
PBS, 1×			4.35	4.35
DMSO			0.05	
Biotin azide	1 mM	10 μM		0.05
Sodium ascorbate	100 mM	10 mM	0.5	0.5
$CuSO_4$	100 mM	2 mM	0.1	0.1
Total volume			5.0	5.0

Adjust volumes proportionally for actual cell numbers.

Box 2 | Cross-link reversal and DNA analysis

To examine DNA fragmentation size, cross-links are reversed from lysates collected before and after DNA sonication, bound proteins are digested, DNA fragments are separated on an agarose gel and analyzed under UV light.

1. Before sonication (PROCEDURE, Step 39), remove 5 µl of lysate and place it on ice. This is the presonication sample.
2. After sonication and sample filtration (Step 44), remove 5 µl of lysate and place it on ice. This represents the postsonication sample.
3. To all samples, add 90 µl of H_2O and 4 µl of 5 M NaCl.
4. Incubate the samples at 65 °C for 4–16 h.
5. Add 1 µl of RNase A (20 mg ml^{-1}) to each sample.
6. Incubate the samples in a 37 °C water bath for 30 min.
7. Prepare the cross-link reversal solution (see REAGENT SETUP).
8. Add 7 µl of cross-link reversal solution to each sample.
9. Incubate the samples at 45 °C for 1–2 h.
10. During the incubation time, pour a 1.5% (wt/vol) agarose/TAE gel without ethidium bromide.
11. Add DNA loading dye to 20 µl of sample and load it on a 1.5% (wt/vol) agarose gel.
12. Perform electrophoresis at 75 V for 3 h in 1× TAE buffer to resolve DNA fragments.
13. Stain the gel with ethidium bromide.
14. Visualize DNA fragments under UV light.

Prepare sufficient cross-link reversal solution mix to add 7 µl to each sample in step 8 of **Box 2**.
Cell lysis buffer Mix 10 mM Tris (pH 8.0), 2 mM $MgCl_2$ and 1% (vol/vol) Igepal CA-630. Before use, add protease inhibitors aprotinin and leupeptin to a final concentration of 1 µg ml^{-1}. Freshly prepare this buffer.

Nuclei buffer Mix 15 mM Tris (pH 8.0), 0.125 M sucrose, 15 mM NaCl, 40 mM KCl, 0.5 mM spermidine and 0.15 M spermine. Freshly prepare this buffer.
Extraction buffer Mix 1× PBS with 350 mM NaCl, 2 mM EDTA and 0.1% (vol/vol) Triton X-100. Freshly prepare this buffer.

PROCEDURE
Cell culture preparation ● TIMING 1–7 d
1| Calculate the number of dishes of cells needed for the experiment. Each sample requires at least 1.0×10^8 cells at the time of the EdU pulse. We typically use three 150 mm dishes of HEK293T cells per sample. The number of cells may need to be increased depending on the application and cell type.

2| Expand cell cultures 1 d before EdU incubation (Step 3) to ensure that the cells are growing optimally. Include one extra dish of cells for counting the cell number in Step 3.
▲ CRITICAL STEP For HEK293T cells, the experiment works best when cell confluence is between 4 and 6×10^7 cells per dish on the day of the EdU pulse. Cells must be in log phase of growth and should not be overgrown. Monitor proper incubator temperature and CO_2 content. EdU incorporation is not maximal unless these crucial parameters are met. If you are performing chases, equilibrate the medium to 37 °C and the proper CO_2 content overnight.

EdU labeling of nascent DNA ● TIMING 10 min–8 h
3| Determine the cell number in the extra dish of cells from Step 2. This cell number will be used to calculate the amount of the reagents used for each sample in Step 29.

4| Plan out times to pulse, chase, fix, quench, collect and wash the samples.
▲ CRITICAL STEP Stagger the samples to ensure that each is treated equally throughout the processing steps.

5| To pulse cells with EdU, remove the dishes from the incubator and place them in a biological safety cabinet.

6| Add 23 µl of the 10 mM EdU stock into 23 ml of cell culture medium in each dish to achieve a final EdU concentration of 10 µM. Return the dishes to the incubator for the desired pulse time (e.g., 10 min).

7| If thymidine chases or drug treatments are not being performed, skip to Step 11.

8| To perform thymidine chase or addition of drug, remove the dishes from the incubator and decant the medium.

9| Carefully wash the cells with 5 ml of chase medium and decant. The chase medium should have been pre-equilibrated to 37 °C and the proper CO_2 content.

10| Add 20 ml of chase medium containing 10 μM thymidine or the desired concentration of DNA damaging drug. Return the dishes to the incubator for the desired length of time.
▲ CRITICAL STEP It is important to perform Steps 5–10 as quickly as possible to prevent pH and temperature changes in the medium, which can affect replication rates.

Formaldehyde cross-linking and collection of cells ● TIMING 1 h
11| After EdU pulse and/or chase, decant the medium.

12| Immediately fix the cells on a dish by adding 10 ml of 1% (wt/vol) formaldehyde in PBS and incubating for 20 min at RT.

13| Quench cross-linking by adding 1 ml of 1.25 M glycine.

14| Collect the sample by scraping with a cell lifter and transfer it to a 50-ml conical tube. Note the volume. This is the same volume that should be used for PBS washes in Step 17.

15| Centrifuge for 5 min at 900g, 4 °C.

16| Decant the supernatant.

17| Wash pellets three times with 1× PBS and centrifuge for 5 min at 900g, 4 °C. PBS wash volume is same as fixation volume noted in Step 14. Vortex to resuspend pellets in PBS.

18| After the last wash, decant PBS.
■ PAUSE POINT The samples can be flash-frozen and stored at −80 °C for several weeks.

Cell permeabilization ● TIMING 1 h
19| Resuspend the cells in permeabilization buffer at a concentration of 1×10^7 cells per ml.

20| Incubate the cells at RT for 30 min. During incubation, thaw and prepare the reagents necessary for the click reaction cocktail (see Steps 28 and 29).

21| Spin down for 5 min at 900g, 4 °C.

22| Carefully decant the supernatant.

23| Wash the cells once with cold 0.5% (wt/vol) BSA in PBS, using the same volume as used for permeabilization in Step 19.
▲ CRITICAL STEP BSA prevents the cell pellet from detaching from the wall of a 50-ml conical flask. A loose pellet will lead to the loss of cells in this step.

24| Centrifuge the cells for 5 min at 900g, 4 °C, and then decant the supernatant.
? TROUBLESHOOTING

25| Wash the cells once with PBS using the same volume as used for permeabilization in Step 19.

26| Spin down for 5 min at 900g, 4 °C.

27| Decant the supernatant and place the pellets on ice while completing the preparation for the click reaction cocktail.

Click reaction ● TIMING 2 h
28| Thaw an aliquot of stock biotin azide by placing it on a 37 °C heat block.
▲ CRITICAL STEP If you are using photocleavable biotin azide, keep the reagent protected from light and prepare the click reaction cocktail in the dark.

29| To calculate click reaction cocktail volumes, **Table 1** lists the amounts of each reagent needed per reaction with an example sample size of 1×10^8 cells. The actual volumes should be adjusted on the basis of the cell number measured per sample (Step 3). Note that two click reaction cocktails need to be prepared: one for the control, which contains DMSO, and one for the experimental samples, which contains the biotin azide.

30| Combine the click reaction cocktail reagents on ice in the order listed in **Table 1**.

31| Resuspend the cell pellets from Step 27 in the click reaction cocktail from Step 30 by vortexing.

32| Rotate the reactions at RT for 1–2 h.

33| Centrifuge the samples for 5 min at 900*g*, 4 °C, and decant the supernatants.

34| Wash the cells once with cold 0.5% (wt/vol) BSA in PBS, using the same volume as used in click reaction for one sample.

35| Centrifuge for 5 min at 900*g*, 4 °C and decant supernatant.

36| Wash the cells once with PBS, using the same volume as used in click reaction for one sample.

37| Decant the PBS and invert the tubes on a paper towel to remove all PBS.
■ **PAUSE POINT** The samples can be flash-frozen and stored at −80 °C for a few days.

Cell lysis and sonication ● **TIMING** 1 h
38| Prepare the lysis buffer by adding aprotinin and leupeptin before use (see REAGENT SETUP) and place on ice.

39| Resuspend the samples from Step 37 at a concentration of 1.5×10^7 cells per 100 µl of lysis buffer and transfer them to 1.5-ml centrifuge tubes on ice. To examine DNA fragment size at this step, see **Box 2**.

40| Sonicate the cells by using a microtip sonicator and the following settings: pulse: 20 s constant pulse, 40 s pause; power: 13–16 Watts; repeat pulse 1× for every 200 µl of cell lysate; total pulse time: 4–5 min per sample.
▲ **CRITICAL STEP** Lysates should appear translucent after sonication and not cloudy. Cloudiness is an indicator of an improper ratio of SDS to protein in the lysate or of insufficient sonication. Keep the samples on an ice slurry during sonication to prevent overheating.
? TROUBLESHOOTING

41| Centrifuge the samples for 10 min at 16,100*g*, RT in a tabletop centrifuge.
▲ **CRITICAL STEP** Lysate should appear clear after centrifugation. The presence of a white precipitate or a white film on top of the lysate is indicative of insufficient clearing of the lysate.

42| Filter the supernatant through a 90-µm nylon mesh into a new tube. Place the tube on ice.

43| Note the lysate volume.

44| To examine DNA fragment size at this step, see **Box 2** for cross-link reversal and DNA analysis.

45| Dilute the lysate 1:1 (vol/vol) with cold PBS containing 1 µg ml⁻¹ of aprotinin and leupeptin.
▲ **CRITICAL STEP** Samples have been diluted to contain 0.5% (wt/vol) SDS and 25 mM Tris because less efficient biotin capture is observed in lysates containing 1% (wt/vol) SDS.

46| Note the final capture volume.

47| Remove 15 µl of the lysate to save as the input sample for use in Step 64 and place it on ice. Immediately add 15 µl of 2× SB to this input sample and store at −80 °C. The remaining lysate is used for the streptavidin capture, which is described below.

Streptavidin capture of biotin-tagged nascent DNA and associated proteins ● TIMING 16–20 h

48| To capture biotin-tagged nascent DNA, each sample from Step 47 is incubated with streptavidin-agarose beads at a concentration of 100 µl of bead slurry (50 µl packed volume) per 1×10^8 cells. First, wash sufficient beads for all samples together by centrifuging the bead slurry at 1,800g for 1 min at RT.

49| Slowly and carefully aspirate the storage buffer from the beads.

50| Wash the beads twice with 1:1 (vol/vol) lysis buffer containing aprotinin and leupeptin.

51| Carefully and slowly aspirate the supernatant after each wash in Step 50.

52| Wash the beads once with 1:1 (vol/vol) PBS containing aprotinin and leupeptin; carefully aspirate the supernatant.

53| Resuspend the beads in 1:1 (vol/vol) PBS containing protease inhibitors.

54| Add an equal volume of beads to each sample from Step 47 with a pipette tip that is cut at the end.

55| Rotate the biotin captures in a cold room for 16–20 h (in the dark if photocleavable biotin azide is used).

56| Centrifuge the streptavidin-agarose beads with the captured DNA and associated proteins for 3 min at 1,800g, RT.

57| Very slowly and carefully aspirate most of the supernatant.
▲ CRITICAL STEP The supernatant should be light blue/clear with no precipitate.
? TROUBLESHOOTING

58| Add 1 ml of cold lysis buffer (no additives needed) to wash the beads.

59| Rotate at RT for 5 min.

60| Centrifuge for 1 min at 1,800g at RT and carefully aspirate and discard the supernatant.

61| Wash the beads once with 1 ml of 1 M NaCl.

62| Rotate and pellet the beads by repeating Steps 59 and 60.

63| Repeat the lysis buffer washes (Steps 58–60) two more times.

Elution of proteins bound to nascent DNA ● TIMING 1–4 h

64| Protein elution can be performed using option A (boiling in 2× SB) or option B (UV photocleavage), depending on the amount of background observed in the negative control. Option B is best suited for proteins that show substantial background and require larger amounts of starting material for detection.
(A) Boiling in 2× SB
  (i) After the last wash in Step 63, aspirate all of the supernatant. Protein-DNA complexes isolated on the beads are called the capture sample.
 (ii) To elute proteins bound to nascent DNA, add 2× SB to packed beads from Step 64A(i) (1:1, vol/vol of packed beads; e.g., 100 µl 2× SB/100 µl packed beads).
(iii) Incubate the capture sample from Step 64A(ii) and the input sample from Step 47 for 25 min at 95 °C to reverse cross-links.
    ▲ CRITICAL STEP Typically, both the input and iPOND-purified capture samples should be examined concurrently.
(iv) Centrifuge the boiled samples for 1 min at 1,800g, RT. The supernatant is the '2× eluted capture' sample and is ready to use in standard SDS-PAGE and immunoblotting procedures (see Step 65).
(B) UV photocleavage, TCA concentration and boiling in 2× SB
  (i) After the last wash in Step 63, wash one additional time with 1× PBS containing leupeptin and aprotinin as in Steps 59 and 60.

(ii) Centrifuge for 1 min at 1,800g, RT, and carefully aspirate the supernatant.

(iii) Add 1:1 (vol/vol) of 1× PBS containing protease inhibitors to the packed beads and resuspend by pipetting.

(iv) Transfer the resuspended beads into a glass vial with a mini magnetic stir bar.

(v) Place the glass vial containing the sample on a magnetic stir plate and adjust to stir on the lowest possible speed.

(vi) Position a UV lamp as close to the glass vial as possible. UV-photoelute at 365 nm for 1–2 h at RT.

(vii) Transfer the bead slurry from the glass vial into a 1.5-ml centrifuge tube.

(viii) Centrifuge the tube for 1 min at 1,800g, RT to pellet the beads.

(ix) Carefully remove the supernatant into a fresh tube. This is the 'UV-photoeluted capture' sample in PBS.

(x) Optionally, to concentrate the sample using TCA precipitation, proceed to the next step. Otherwise, add 1:1 (vol/vol) of 2× SB to the UV-photoeluted capture sample, boil at 95 °C for 25 min to reverse cross-links, and then proceed to analysis of proteins (Step 65).

(xi) Add ice-cold 100% TCA to the UV photoeluted capture sample from Step 64B(ix) to achieve a final concentration of 15% (vol/vol) TCA.

(xii) Incubate the sample on ice for 30 min.

(xiii) Centrifuge at 16,100g for 30 min in a cold room.

(xiv) Carefully remove the supernatant and save it for troubleshooting.

(xv) Wash the pellet with 1 ml of ice-cold acetone.

(xvi) Centrifuge for 10 min at 16,100g in cold room.

(xvii) Carefully remove the supernatant and save it for troubleshooting.

(xviii) Air-dry the pellet for 2–3 min until the smell of acetone is undetectable.

▲ CRITICAL STEP If the pellet is not visible at this step, spin down the supernatant saved from Step 64B(xiv), and then repeat Step 64B(xv–xviii). If no pellet is observed, spin down the supernatant previously saved from Step 64B(xvii) and repeat Step 64B(xviii).

? TROUBLESHOOTING

(xix) Add 30 µl of 2× SB to the protein pellet to resuspend the sample.

(xx) Incubate the capture sample (from Step 64B(x) if it is not TCA precipitated or from Step 64B(xix) if it is TCA precipitated) and the input sample (from Step 47) for 25 min at 95 °C. The samples are ready for use in standard SDS-PAGE and immunoblotting procedures.

Analysis of eluted proteins using western blotting ● TIMING 2–3 d

65| Prepare a standard SDS-PAGE gel[20]. To examine purification of positive controls concurrently (a replication protein and a histone, e.g., PCNA and H3, respectively), it is useful to prepare a 15% (wt/vol) gel.

66| To detect purified proteins from input and capture samples (from Step 64A(iv) or Step 64B(xx)), load the equivalent of 3 to 6 × 10^7 cells per well from the total protein capture (e.g., 3 to 6 × 10^7 of 1 × 10^8). This means that each sample of 1 × 10^8 cells yields sufficient sample for analysis of 2–3 immunoblots. For input samples, load the equivalent of 0.1% (vol/vol) input per well.

▲ CRITICAL STEP Depending on antibody quality, different proteins may require more cells for detection than others. This will require empirical determination.

67| Perform electrophoresis to resolve proteins on the basis of molecular weight, and then proceed with standard immunoblotting with desired antibodies according to supplier instructions or with MS analysis[21].

68| Proteins can be detected by using chemiluminescence (e.g., Western Lightning Plus) or quantitative immunoblotting with the Odyssey infrared imaging system.

? TROUBLESHOOTING

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 2**.

TABLE 2 | Troubleshooting table.

Step	Problem	Reason	Solution
24	Poor cell recovery	The cells were not pelleted sufficiently during the centrifugation	Increase the time or speed of the centrifugation. Be sure that the wash solution contains BSA
40	Cell lysate is cloudy after sonication	Sonication did not completely lyse cells or SDS-protein complexes precipitate from solution	Increase sonication times and be sure to avoid foaming of samples. Ensure that the proper volume of lysis buffer was used in Step 39
57	White precipitate layer is observed above beads after centrifugation of biotin captures	Lipids from cell membranes were not properly pelleted after sonication	Make certain that lysate is clear after sonication and centrifugation. If a white layer is observed on top of the cell lysate, remove the lysate and clear again by centrifugation
64B(xviii)	No pellet is observed after air-drying the TCA-concentrated iPOND eluate	Sample was lost during TCA precipitation	Centrifuge the supernatant saved in Step 64B(xiv). Proceed with Step 64B(xv–xviii). If no pellet is observed, centrifuge supernatant previously saved in Step 64B(xvii). Continue with Step 64B(xviii)
68	High background signal in the control sample	Protein binds to streptavidin beads nonspecifically	Use elution option B; increase the number of washes in Steps 62 and 63
	Poor signal for control proteins such as PCNA in the experimental sample	Poor EdU incorporation	Increase the number of cells used in each sample and ensure that the cells are growing well prior to experiment
	Poor detection of the protein of interest in the input samples	Poor antibody or formaldehyde cross-linking interferes with epitope detection	Optimize immunoblotting conditions or change antibody. Consider increasing the boiling time in Step 65A(iii) or Step 65B(xx) to completely reverse the formaldehyde cross-links

● **TIMING**
Steps 1 and 2, preparation of cell cultures: 1–7 d
Steps 3–10, EdU labeling and thymidine/HU chase: variable, typically 10 min–8 h
Steps 11–18, cell fixation/collection: 1 h
Steps 19–27, cell permeabilization: 1 h
Steps 28–37, click reaction: 2 h
Steps 38–47, cell lysis and sonication: 1 h
Steps 48–55, biotin capture: 16–20 h
Steps 56–64, washes and protein elution: variable, typically 1–4 h
Steps 65–68, analysis of eluted proteins: variable, typically 2–3 d

ANTICIPATED RESULTS
Typically, 1×10^8 cells are EdU labeled and processed by using iPOND to yield sufficient material for immunoblotting with 2–3 antibodies. A protein is interpreted to be enriched at the replication fork if the following conditions are met: the protein is detected in a click reaction sample that has been EdU labeled (**Fig. 4**, lane 6); the protein is not detected in a sample that omits the click reaction (**Fig. 4**, lane 5); and the protein level is progressively decreased in the thymidine chase samples (**Fig. 4**, lanes 7–8). Chromatin-bound proteins will appear to be enriched specifically after the click reaction, but they will also be detected in the thymidine chase sample (**Fig. 4**, e.g., histone H2B). Replication stress proteins recruited to damaged forks will be detected only after a chase into a replication stress reagent (**Fig. 3b**).

AUTHOR CONTRIBUTIONS B.M.S. developed the protocol and F.B.C. made modifications to omit the formaldehyde cross-linking step. D.C. conceived and supervised the project. B.M.S. and D.C. wrote the manuscript.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

Published online at http://www.natureprotocols.com/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Cimprich, K.A. & Cortez, D. ATR: an essential regulator of genome integrity. *Nat. Rev. Mol. Cell Biol.* **9**, 616–627 (2008).
2. Hoeijmakers, J.H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366–374 (2001).
3. Berkovich, E., Monnat, R.J. Jr. & Kastan, M.B. Assessment of protein dynamics and DNA repair following generation of DNA double-strand breaks at defined genomic sites. *Nat. Protoc.* **3**, 915–922 (2008).
4. Bell, S.P. & Dutta, A. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* **71**, 333–374 (2002).
5. Sirbu, B.M. *et al.* Analysis of protein dynamics at active, stalled, and collapsed replication forks. *Genes Dev.* **25**, 1320–1327 (2011).
6. Salic, A. & Mitchison, T.J. A chemical method for fast and sensitive detection of DNA synthesis *in vivo*. *Proc. Natl. Acad. Sci. USA* **105**, 2415–2420 (2008).
7. Moses, J.E. & Moorhouse, A.D. The growing applications of click chemistry. *Chem. Soc. Rev.* **36**, 1249–1262 (2007).
8. Herrick, J. & Bensimon, A. Global regulation of genome duplication in eukaryotes: an overview from the epifluorescence microscope. *Chromosoma* **117**, 243–260 (2008).
9. Probst, A.V., Dunleavy, E. & Almouzni, G. Epigenetic inheritance during the cell cycle. *Nat. Rev. Mol. Cell Biol.* **10**, 192–206 (2009).
10. Kliszczak, A., Rainey, M., Harhen, B., Boisvert, F. & Santaocanale, C. DNA-mediated chromatin pull-down for the study of chromatin replication. *Scientific Rep.* **1**, 1–7 (2011).
11. Jao, C.Y. & Salic, A. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc. Natl. Acad. Sci. USA* **105**, 15779–15784 (2008).
12. Reyes-Lamothe, R., Sherratt, D.J. & Leake, M.C. Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* **328**, 498–501 (2010).
13. Aladjem, M.I. Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat. Rev. Genet.* **8**, 588–600 (2007).
14. Raschle, M. *et al.* Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* **134**, 969–980 (2008).
15. Ben-Yehoyada, M. *et al.* Checkpoint signaling from a single DNA interstrand crosslink. *Mol. Cell* **35**, 704–715 (2009).
16. Petermann, E., Orta, M.L., Issaeva, N., Schultz, N. & Helleday, T. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol. Cell* **37**, 492–502 (2010).
17. Szychowski, J. *et al.* Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J. Am, Chem. Soc.* **132**, 18351–18360 (2010).
18. Kim, H.Y., Tallman, K.A., Liebler, D.C. & Porter, N.A. An azido-biotin reagent for use in the isolation of protein adducts of lipid-derived electrophiles by streptavidin catch and photorelease. *Mol. Cell Proteomics* **8**, 2080–2089 (2009).
19. Capson, T.L., Benkovic, S.J. & Nossal, N.G. Protein-DNA cross-linking demonstrates stepwise ATP-dependent assembly of T4 DNA polymerase and its accessory proteins on the primer-template. *Cell* **65**, 249–258 (1991).
20. Gallagher, S.R. One-dimensional SDS gel electrophoresis of proteins. *Curr. Prot. Mol. Biol.* 10.2A.1–10.2A.34 (1999).
21. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).

SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication

Rémy Bétous,[1] Aaron C. Mason,[2] Robert P. Rambo,[3] Carol E. Bansbach,[1] Akosua Badu-Nkansah,[1] Bianca M. Sirbu,[1] Brandt F. Eichman,[1,2] and David Cortez[1,4]

[1]Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA; [2]Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37240, USA; [3]Life Sciences Division, Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

SMARCAL1 (SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily A-like1) maintains genome integrity during DNA replication. Here we investigated its mechanism of action. We found that SMARCAL1 travels with elongating replication forks, and its absence leads to MUS81-dependent double-strand break formation. Binding to specific nucleic acid substrates activates SMARCAL1 activity in a reaction that requires its HARP2 (Hep-A-related protein 2) domain. Homology modeling indicates that the HARP domain is similar in structure to the DNA-binding domain of the PUR proteins. Limited proteolysis, small-angle X-ray scattering, and functional assays indicate that the core enzymatic unit consists of the HARP2 and ATPase domains that fold into a stable structure. Surprisingly, SMARCAL1 is capable of binding three-way and four-way Holliday junctions and model replication forks that lack a designed ssDNA region. Furthermore, SMARCAL1 remodels these DNA substrates by promoting branch migration and fork regression. *SMARCAL1* mutations that cause Schimke immunoosseous dysplasia or that inactivate the HARP2 domain abrogate these activities. These results suggest that SMARCAL1 continuously surveys replication forks for damage. If damage is present, it remodels the fork to promote repair and restart. Failures in the process lead to activation of an alternative repair mechanism that depends on MUS81-catalyzed cleavage of the damaged fork.

SMARCAL1 (SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily A-like1), also known as DNA-dependent ATPase A and HARP (Hep-A-related protein), is a member of the SNF2 family of ATPases (Flaus et al. 2006). Many of these proteins use the energy of ATP hydrolysis to translocate along DNA and thereby remodel DNA structures or DNA–protein interactions. They function in many cellular processes, including transcription, DNA replication, and DNA repair.

Biallelic mutations in *SMARCAL1* cause the human disease Schimke immunoosseous dysplasia (SIOD) (Boerkoel et al. 2002). SIOD symptoms commonly include skeletal dysplasia, T-cell immunodeficiency, and kidney failure (Boerkoel et al. 2000). At the cellular level, SMARCAL1 deficiency causes increased DNA replication-associated damage (Bansbach et al. 2009, 2010; Postow et al. 2009; Yuan et al. 2009) and sensitizes cells to DNA-damaging agents that inhibit DNA replication (Bansbach et al. 2009; Ciccia et al. 2009; Yuan et al. 2009). SMARCAL1 localizes to damaged replication factories via an interaction with the ssDNA-binding protein replication protein A (RPA) (Bansbach et al. 2009; Ciccia et al. 2009; Yuan et al. 2009; Yusufzai et al. 2009), and this interaction is essential for its genome maintenance function (Bansbach et al. 2009; Yuan et al. 2009). SMARCAL1 is phosphorylated by checkpoint kinases in response to DNA damage (Bansbach et al. 2009; Postow et al. 2009). *SMARCAL1* mutants derived from SIOD patients fail to rescue the genome maintenance defects caused by *SMARCAL1* deficiency (Bansbach et al. 2009, 2010; Yuan et al. 2009). Thus, SMARCAL1 acts at damaged replication forks to maintain genome stability, and defects in this activity may underlie at least some of the phenotypes associated with SIOD (Bansbach et al. 2010).

[4]Corresponding author.
E-mail david.cortez@vanderbilt.edu.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.178459.111.

The mechanism of how SMARCAL1 acts to repair damaged forks remains largely unknown. Biochemically, SMARCAL1 can bind to DNA that contains single- and double-stranded regions such as forks and DNA hairpins (Muthuswami et al. 2000; Yusufzai and Kadonaga 2008). DNA binding activates its ATPase activity, and this activity promotes DNA single-strand annealing even in the presence of RPA (Yusufzai and Kadonaga 2008). The N-terminal RPA-binding domain of SMARCAL1 is not necessary for this DNA strand-annealing activity (Bansbach et al. 2009; Yusufzai et al. 2009), but patient-derived mutants lack this function. The molecular basis for this activity may not be simply translocation along dsDNA, since the related protein RAD54 cannot perform this function despite being a robust translocase (Yusufzai and Kadonaga 2008).

SMARCAL1 is a multidomain protein. The ATPase domain, which lies in the C-terminal half of the protein, is split into two regions of primary amino acid sequence by a 115-amino-acid linker sequence. The N-terminal half of the protein contains a highly sequence conserved RPA-binding domain (Bansbach et al. 2009; Ciccia et al. 2009; Postow et al. 2009; Yuan et al. 2009; Yusufzai et al. 2009), a 200-amino-acid region of low sequence conservation without predicted domain structure, and two HARP domains. The HARP domains are 55 amino acids in length with high sequence similarity but unknown function and structure. They are separated by 40 amino acids, and the second HARP domain is linked to the ATPase domain by an additional 47 amino acids.

Fusing the HARP domains to the ATPase domain of the SNF2 proteins BRG1 or HELLS is sufficient to reconstitute DNA-dependent ATPase and annealing helicase activities, suggesting that the HARP domains are important determinants of the SMARCAL1 enzyme specificity (Ghosal et al. 2011). Paradoxically, the closest homolog of SMARCAL1 in humans, annealing helicase 2 (AH2, also known as ZRANB3), also has annealing helicase activity despite a different domain structure and no unambiguous HARP domains (Yusufzai and Kadonaga 2010).

In this study, we took genetic, biochemical, and biophysical approaches to understand how SMARCAL1 functions to maintain genome integrity. We found that SMARCAL1 travels with at least some elongating replication forks, and the MUS81 structure-specific endonuclease cleaves damaged forks in SMARCAL1-deficient cells. The HARP2 domain is essential for DNA binding, and both biochemical and small-angle X-ray scattering (SAXS) data indicate that the HARP2+SNF2 domains provide the minimal enzymatic unit. The HARP domain resembles the DNA-binding domain of the PUR-α protein and has limited ability to bind DNA on its own. Surprisingly, we found that SMARCAL1 can bind three-way and four-way DNA structures and model replication forks. Furthermore, SMARCAL1 branch-migrates the four-way junction and catalyzes extensive fork regression of model replication forks. These data provide mechanistic insight into how SMARCAL1 functions and suggest that it remodels stalled replication forks through fork regression and branch migration to promote replication fork restart

and prevent replication-associated DNA double-strand breaks.

Results

SMARCAL1 is present at DNA replication forks during an unperturbed S phase and prevents MUS81-dependent double-strand breaks

Previous analyses indicated that SMARCAL1 localizes to nuclear foci that colocalize with replisomes in response to agents that induce replication stress (Bansbach et al. 2009; Ciccia et al. 2009; Postow et al. 2009; Yuan et al. 2009; Yusufzai et al. 2009). This localization is dependent on an interaction with the replisome protein RPA. Silencing SMARCAL1 using RNAi causes elevated levels of γH2AX in replicating cells (Bansbach et al. 2009; Postow et al. 2009; Yuan et al. 2009). To determine whether SMARCAL1 actually is a component of active replisomes, we used the iPOND procedure (Sirbu et al. 2011) to purify active and stalled replication forks. SMARCAL1 is purified with nascent, 5-ethynyl-2′-deoxyuridin (EdU)-labeled DNA at elongating replication forks even when replication is not perturbed (Fig. 1A). It is not purified with
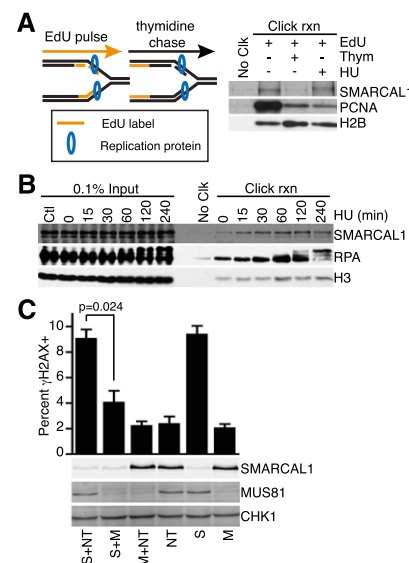
Figure 1. SMARCAL1 acts at replication forks to prevent MUS81-catalyzed double-strand breaks. (*A*) Cells were labeled with EdU for 10 min, the EdU was removed, and thymidine was added for 20 min or HU was added for 3 h prior to purifying the nascent DNA–protein complexes using the iPOND procedure. (*B*) EdU-labeled cells were treated with 2 mM HU for the indicated lengths of time prior to performing iPOND. The "No Clk" controls in *A* and *B* are samples treated with EdU only, but no biotin-azide was added during the click reaction. (*C*) U2OS cells were transfected with siRNAs targeting SMARCAL1 (S), MUS81 (M), or nontargeting (NT) as indicated. Three days after transfection, the cells were either stained with antibodies to γH2AX or harvested for immunoblotting with the indicated antibodies. The percentage of cells staining positive for γH2AX was determined by immunofluorescent imaging from three independent experiments. Cells with >10 foci were counted as positive. Error bars are the standard deviation (SD; *n* = 3).

the EdU-labeled DNA once the labeled DNA segment is no longer adjacent to the fork (after a chase in medium lacking EdU), indicating that it travels with at least some moving replisomes. As expected, SMARCAL1 is also found at forks stalled with hydroxyurea (HU) (Fig. 1A,B), and its mobility on SDS-PAGE gels is altered in these circumstances due to phosphorylation by checkpoint kinases (Bansbach et al. 2009).

The MUS81 endonuclease cleaves some blocked and damaged replication forks, generating a double-strand break and initiating recombination-based repair mechanisms (Osman and Whitby 2007). Therefore, we hypothesized that the high level of γH2AX found in *SMARCAL1* silenced cells could be due to double-strand breaks catalyzed by MUS81. To address this question, we measured the proportion of cells containing γH2AX after SMARCAL1 and/or MUS81 depletion. As expected, silencing *SMARCAL1* caused an induction of γH2AX, while silencing *MUS81* had no effect (Fig. 1C). *MUS81* silencing prevented γH2AX induction in *SMARCAL1* silenced cells without significantly altering the efficiency of *SMARCAL1* silencing (Fig. 1C). Thus, γH2AX induction after SMARCAL1 depletion is MUS81-dependent.

SMARCAL1 binds a wide variety of DNA substrates that combine ssDNA and dsDNA

Our results suggest that SMARCAL1 either processes or prevents the formation of MUS81 substrates. Little is known about SMARCAL1 substrate specificity other than it prefers to bind DNA with both single- and double-stranded characteristics rather than ssDNA or dsDNA (Supplemental Fig. 1A; Muthuswami et al. 2000; Yusufzai

and Kadonaga 2008), and its ATPase activity is activated upon DNA binding. To clarify the DNA determinants that mediate SMARCAL1 DNA binding and activation, we investigated a broad range of possible DNA substrates. We first evaluated how long the ssDNA arms of a fork need to be and found that significant SMARCAL1 binding is observable even with a fork length of only 5 nucleotides (nt) per arm (Supplemental Fig. 1B,C). Increasing the arm lengths beyond 5 nt increases the binding affinity. We also observe a second DNA–protein complex forming when the ssDNA region is lengthened to 20 nt or more. The second, higher-molecular-weight complex may contain more than one SMARCAL1 molecule.

We next varied the length of one of the ssDNA arms while keeping the other constant and found that the length of the second arm did not influence binding affinity (Supplemental Fig. 1D,E). In fact, SMARCAL1 bound equivalently to a fork and an ssDNA overhang substrate. Both DNA substrates stimulated SMARCAL1 ATPase activity as well (Supplemental Fig. 1F). Furthermore, DNA substrates with either a 5′ or 3′ recessed end bind and stimulate SMARCAL1 ATPase activity equivalently (Fig. 2A–C).

At a replication fork, the free 5′ end of the nascent nucleic acid on the lagging stand template would consist of a short RNA primer rather than DNA. To test whether SMARCAL1 can bind and be activated on the lagging strand, we examined a nucleic acid substrate that mimics this chimeric nucleic acid structure. A RNA–DNA primer substrate bound and stimulated SMARCAL1 equivalently to the DNA–DNA substrate (Supplemental Fig. 1G–I).

Next, we assessed how the length of ssDNA alters SMARCAL1-binding affinity. Five nucleotides are suffi-
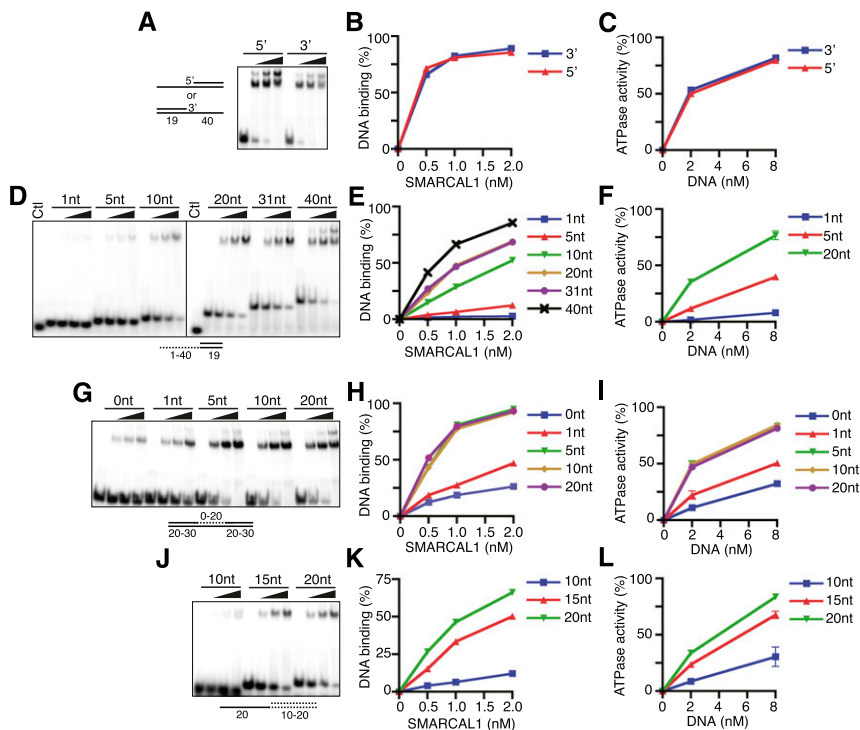
Figure 2. Characterization of the DNA-binding and DNA-stimulated ATPase activities of full-length SMARCAL1. (*A,D,G,J*) Increasing amounts of SMARCAL1 were incubated with the indicated oligonucleotide substrates prior to polyacrylamide gel electrophoresis. The control (Ctl) in *D* is an unhybridized single-stranded oligonucleotide. (*B,E,H,K*) Quantitation of a representative DNA-binding experiment. (*C,F,I,L*) Increasing amounts of DNA substrate were added to SMARCAL1, and ATPase activity was measured as the percentage of ATP hydrolyzed. Error bars represent the mean ± SD from three independent experiments. In cases in which no error bars are visible, the SD is smaller than the symbol size. The sequences of the oligonucleotides are listed in Supplemental Table 1, and a description of which oligonucleotides were used in each experiment is presented in Supplemental Table 2.

cient to allow some binding and elicit significant ATPase activity (Fig. 2D–F). SMARCAL1 binding and ATPase activity increase as the length of the ssDNA increases. SMARCAL1 also binds and is stimulated efficiently by a gapped DNA substrate. Maximum binding and activation require only a five single-stranded nucleotide gap, and even a nick can elicit some activity (Fig. 2G–I; Supplemental Fig. 1J,K). When a bubble replaces the gap, increasing the length of the mismatched nucleotides to 16 significantly increases affinity (Supplemental Fig. 1L,M). Thus, the length of ssDNA needed for optimal binding and activation of SMARCAL1 is shortest when it is presented in the context of a gap.

We also investigated how the length of dsDNA affects binding and ATPase stimulation of SMARCAL1. Optimal SMARCAL1 binding and ATPase activation requires 20 dsDNA nucleotides (Fig. 2J–L). Greater dsDNA lengths yield no further improvement in SMARCAL1 affinity (data not shown). Fifteen nucleotides of dsDNA on either side of a 5-nt gap are sufficient to elicit maximal SMARCAL1 binding (Supplemental Fig. 1N,O).

Taken together, these results show that SMARCAL1 binds and is activated by any nucleic acid structure that contains both single- and double-stranded regions, including an RNA–primer template. The optimal length of ssDNA that elicits binding depends on the structural context of the DNA, with 5 nt being sufficient for a gap and longer lengths promoting better binding to a forked or single-stranded overhang substrate. The optimal length of dsDNA is ~15 nt. Finally, the dsDNA and ssDNA must be within the same molecule, since adding these separately to SMARCAL1 does not elicit any binding (Supplemental Fig. 1A).

HARP2 but not HARP1 is required for SMARCAL1 DNA-binding, ATPase, and annealing helicase activities

To understand how SMARCAL1 binds DNA, we examined the affinity of a series of truncated SMARCAL1 proteins for a forked DNA substrate (Fig. 3A). While deletion of the first 198 and the last 84 amino acids had no effect on SMARCAL1 DNA binding, deletion of the first 424 amino acids containing the HARP domains severely compromises the DNA-binding and ATPase activities of SMARCAL1 (Fig. 3B–E).

These results led us to hypothesize that the HARP domains may be essential for SMARCAL1 DNA binding. To test this hypothesis, we assessed the behavior of a series of HARP domain mutants (Fig. 4A). SMARCAL1 lacking the first HARP domain (ΔHARP1) binds to and is activated by a forked DNA substrate, although with slightly reduced affinity compared with wild-type SMARCAL1 (Fig. 4B,C; Supplemental Fig. 2A). In contrast, deleting the second HARP domain alone or in combination with the first HARP domain (SMARCAL1-ΔHARP2 and ΔHARP1+2) severely attenuated both DNA binding and ATPase activation. The effects of the deletions were even more severe when assayed with a 5-nt, single-stranded gap substrate (Fig. 4D,E).
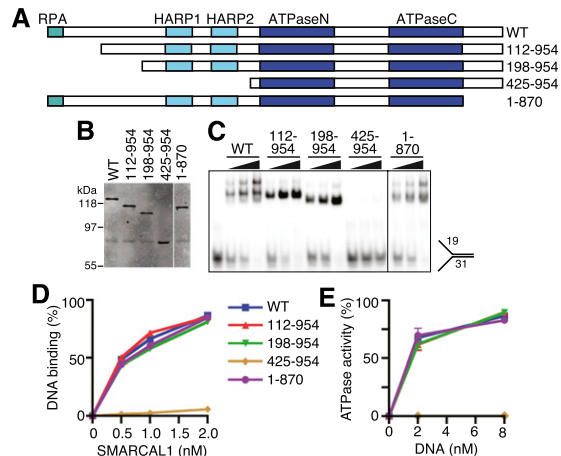
Figure 3. The SMARCAL1 N terminus containing the HARP domains is necessary for DNA-binding and ATPase activity. (*A*) Diagram of the SMARCAL1 proteins used to identify domains required for function. Wild type (WT) in all figures is full-length SMARCAL1. (*B*) Overexpressed SMARCAL1 proteins were purified from HEK-293T cells and examined on an SDS-PAGE gel by immunoblotting. (*C,D*) Increasing amounts of purified SMARCAL1 proteins were incubated with the forked DNA substrate to measure DNA binding. (*E*) Increasing amounts of forked DNA were added to the SMARCAL1 fragments to measure DNA-stimulated ATPase activity. Error bars represent the mean ± SD from three independent experiments. In cases in which no error bars are visible, the SD is smaller than the symbol size. The DNA substrates corresponding to each symbol and line color are the same in *D* and *E*.

To confirm the deletion results, we generated point mutants in HARP1 and HARP2. The HARP domains are evolutionarily conserved (Fig. 4F). We mutated two of the invariant residues within each domain to alanine (HARP1 W277A/F279A and HARP2 W372A/F379A). These mutants exhibit DNA-binding and ATPase activity similar to the corresponding complete deletion of the domain (Fig. 4G–I). Interestingly, we found that the decreased DNA-binding and ATPase activity of the HARP1-WF mutant yielded only a slight impairment of SMARCAL1 annealing helicase activity, while mutation of the HARP2 domain completely abolished the ability of SMARCAL1 to anneal an RPA-coated plasmid substrate (Fig. 4J). The complete deletion of HARP1 also had no effect on the SMARCAL1 annealing helicase activity (Supplemental Fig. 2B). These results suggest that HARP2 is critical for the DNA-binding, ATPase, and annealing helicase activities of SMARCAL1. HARP1 may have a supporting role in facilitating SMARCAL1 function.

Finally, we asked whether the HARP domains themselves have any DNA-binding activity. We found that a SMARCAL1 fragment encompassing both HARP domains (amino acids 198–425) is sufficient to bind forked DNA, albeit with much lower affinity than the full-length protein (Supplemental Fig. 3A–C). The HARP domain–DNA complex did not migrate as a discrete band in the electrophoretic mobility shift assay; however, we were able to supershift the DNA–protein
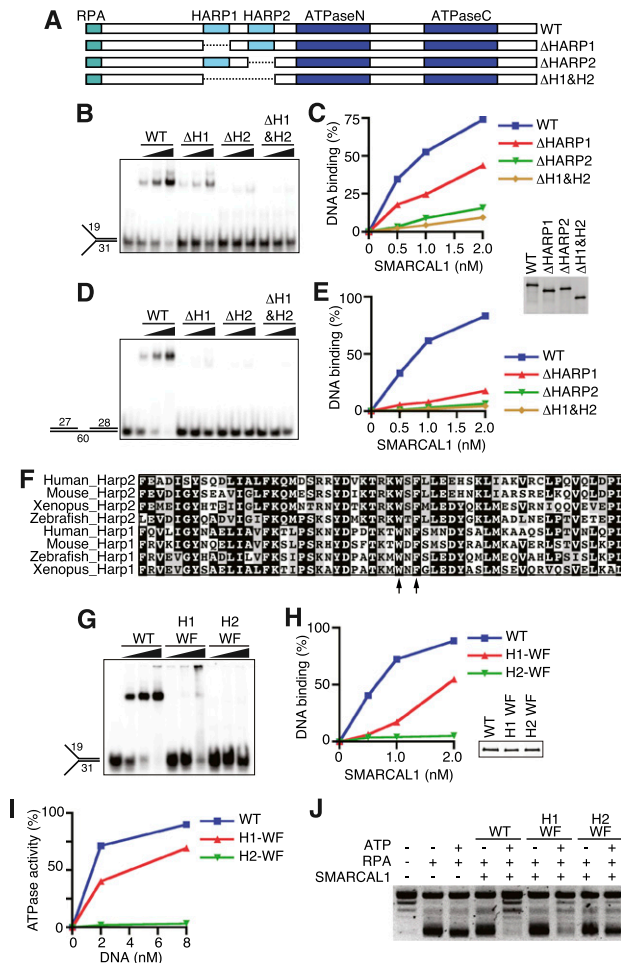
Figure 4. The HARP2 domain of SMARCAL1 is required for annealing helicase activity. (*A*) Diagram of the SMARCAL1 HARP domain deletion mutants purified after overexpression in HEK-293T cells. DNA binding was measured with increasing concentrations of a forked DNA substrate (*B,C*) or 5-nt gap DNA substrate (*D,E*). (*F*) Sequence alignment of the HARP1 and HARP2 domains of human, mouse, *Xenopus laevis*, and zebrafish SMARCAL1. The arrows point to the two residues mutated in the WF mutants used in *G–J*. (*G,H*) Forked DNA binding of the wild type and SMARCAL1 HARP-WF mutants purified from baculovirus-infected insect cells. Note that the HARP1-WF mutant reproducibly shifted much of the DNA substrate into the well of the gel at higher concentrations of protein. (*I*) Increasing amounts of forked DNA were added to the SMARCAL1 mutants to measure DNA-stimulated ATPase activity. Error bars represent the mean ± SD from three independent experiments. In cases in which no error bars are visible, the SD is smaller than the symbol size. (*J*) Annealing helicase activities of SMARCAL1 wild-type and mutant proteins. The concentration of the SMARCAL1 proteins in this assay is 15 nM. The insets in *C* and *H* are immunoblots confirming that equal concentrations of SMARCAL1 proteins were used.

complex with an antibody that recognizes the recombinant HARP1+2 protein fragment, confirming the complex was not due to a contaminant in the protein purification.
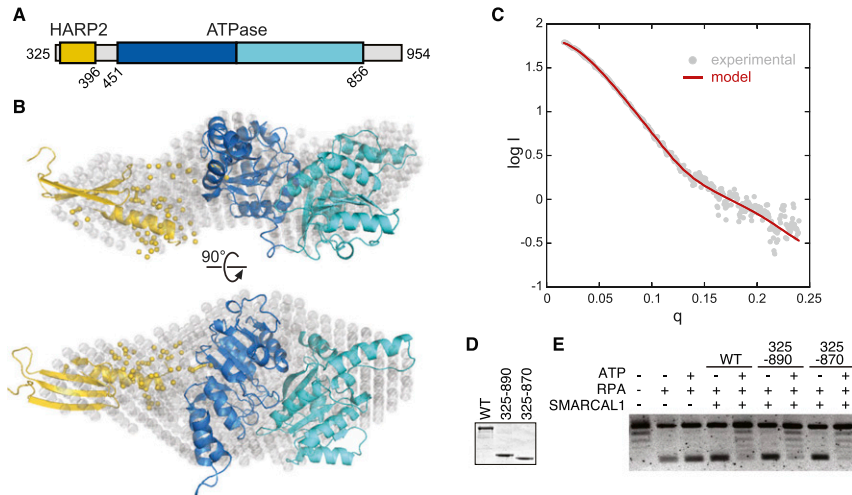
The HARP2-ATPase constitutes a structural core motor domain

Our biochemical results demonstrate the importance of the HARP2 domain in SMARCAL1 function. To gain mechanistic insight into how SMARCAL1 might use this novel domain, SAXS experiments were performed to determine the spatial arrangement of the HARP2 and ATPase domains in solution (Fig. 5A). Limited proteolysis of the full-length protein purified from insect cells revealed a proteolytically resistant fragment consisting of the HARP2-ATPase regions (Supplemental Fig. 4A). Kratky analysis of SMARCAL1(325–954) revealed parabolic features, suggesting that the protein is globular with distinct domains (Supplemental Fig. 4B,C). The radius of gyration (r_g) obtained from the Guinier region was 33.0 ± 0.3 Å (Supplemental Fig. 4D), indicating that the 75-kDa protein is elongated when compared with glucose isomerase, a spherical protein at 173 kDa with a similar R_g of 32 Å.

SAXS data provide complete structural information and can be used to distinguish between different conformations of a high-resolution model or build a complete atomistic model from known domains (Rambo and Tainer 2010). Therefore, we used the SAXS data of SMARCAL1 and homology models of both the HARP2 and ATPase domains to determine the solution state of the protein. To date, there are no known structural homologs of the HARP domain. However, we discovered by sequence–structure comparison (Shi et al. 2001) that there is good agreement between the predicted secondary structural elements of the HARP domains with tandem PUR repeats observed in the structure of the purine-rich element-binding protein PUR-α (Supplemental Fig. 5; Graebsch et al. 2009). PUR repeats are ~140-residue motifs consisting of anti-parallel β-β-β-β-α topology that bind ssDNA and dsDNA and thus provide a reasonable structural model for the HARP domains. A model of the core ATPase domain was also created based on the crystal structure of *Sulfolobus sulfotaricus* (Sso) Rad54, which shares 23% sequence identity and 58% overall similarity with SMARCAL1 (Supplemental Fig. 6).

Preliminary normal mode analysis (Suhre and Sanejouand 2004) was performed on the core ATPase domain to produce a family of alternative conformations. Each conformation was then combined with the HARP2 model for partial ab initio modeling using a simulated annealing search algorithm. The models converged into an elongated structure that was independently validated by the close resemblance to the three-dimensional (3D) molecular envelope generated from the SAXS data using GASBOR (Fig. 5B), and the remarkable agreement between the experimental scattering curve and the theoretical curve calculated from the docking model (Fig. 5C). The resulting HARP2-ATPase model revealed that the HARP2 and ATPase motifs form one continuous domain in the absence of DNA, suggesting that their association constitutes a structural and functional core domain necessary to drive translocation. To test this idea, we assayed whether the HARP2-ATPase protein is sufficient to

Figure 5. HARP2-ATPase constitutes an active structural core domain. (*A*) Construct used for SAXS measurements. (*B*) The SAXS model constructed from HARP2 (residues 325–396, gold) and ATPase (residues 451–856, blue) homology models superimposed on the ab initio molecular envelope determined by GASBOR (gray spheres). The yellow spheres represent region 397–450 modeled in BUNCH. (*C*) The theoretical scattering curve (red) from the model shown in *B* is superimposed on the experimental SAXS data (gray circles) with a goodness of fit $\chi = 1.5$. Coomassie-stained gel of wild-type or truncated SMARCAL1 proteins (*D*) used in an annealing helicase assay (*E*).

catalyze strand annealing. Indeed, SMARCAL1(325–870) and SMARCAL1(325–890) are both efficient ATP-dependent annealing helicases (Fig. 5D,E).

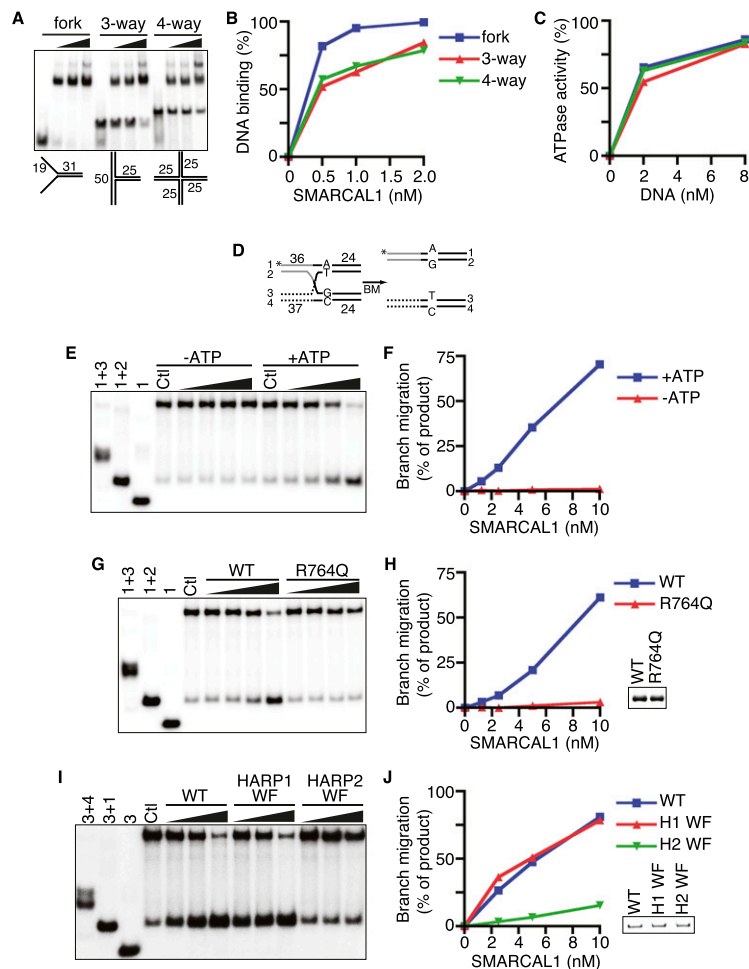### SMARCAL1 can bind and branch-migrate a four-way junction

The DNA-binding activities of SMARCAL1 characterized thus far suggest that SMARCAL1 may have dsDNA- and ssDNA-binding surfaces. Combined with the energy of ATP hydrolysis, SMARCAL1 may translocate along the DNA in a way that leads to single-strand annealing. To determine whether these properties could yield any other enzymatic consequences, we expanded our search for SMARCAL1 substrates to more complex DNA structures, including three-way and four-way Holliday junctions. Surprisingly, despite lacking any designed ssDNA regions, SMARCAL1 could bind these DNA substrates with only slightly reduced affinity compared with a fork substrate (Fig. 6A,B). Furthermore, both DNA substrates activated the SMARCAL1 ATPase (Fig. 6C).

Given that these structures bind SMARCAL1 and stimulate its ATPase activity, we asked whether SMARCAL1 could also induce branch migration like Rad54 (Bugreev et al. 2006). We prepared a synthetic Holliday junction consisting of two homologous and two heterologous arms, similar to those used in previous branch migration studies (Fig. 6D; Gari et al. 2008b). Indeed, SMARCAL1 catalyzed branch migration in an ATP-dependent manner (Fig. 6E,F). As expected, the SIOD patient-derived ATPase-defective mutant (R764Q) failed to promote branch migration despite having the ability to bind DNA (Fig. 6G,H; Yusufzai and Kadonaga 2008). To test the importance of the HARP domains in this process, we examined the branch migration properties of HARP1-WF and HARP2-WF SMARCAL1 mutants. While the HARP1 mutant was able to branch-migrate the Holliday junction as efficiently as the wild-type protein, the HARP2 mutant had severely attenuated activity (Fig. 6I,J).

SMARCAL1 can bind and branch-migrate a replication fork

Previous studies indicate that SMARCAL1 acts at stalled replication forks but may not have an essential function in homology-directed double-strand break repair. Double-strand breaks are only thought to form in normal cells at persistently stalled forks (Petermann et al. 2010; Sirbu et al. 2011). Thus, we investigated whether SMARCAL1 could bind and process other branched structures that might exist at a transiently stalled fork. Specifically, we compared SMARCAL1 affinity to model forks with no nascent DNA strands, a leading strand, a lagging strand, or both. Strikingly, we found that SMARCAL1 binds to and is activated by each of these structures (Fig. 7A–C). To determine whether SMARCAL1 can catalyze remodeling of these replication fork structures, we prepared a substrate to monitor fork regression (Fig. 7D; Gari et al. 2008b). SMARCAL1 catalyzed displacement of the two "nascent" DNA strands and annealing of the parental strands (Fig. 7E,F). Again, the SIOD patient-derived R764Q mutation eliminated this activity.

SMARCAL1 does not possess any helicase activity (Yusufzai and Kadonaga 2008), so it is unlikely that it could unwind the nascent strands before annealing both parental and both nascent strands together. To confirm that the SMARCAL1 fork reversal activity is coordinated without the formation of ssDNA intermediates, we labeled the model nascent leading strand of the synthetic replication fork and performed a time-course assay. We found that only a double-stranded product consisting of the two nascent strands is formed without the appearance of any ssDNA intermediates (Supplemental Fig. 7A–C). We conclude that SMARCAL1 processes replication fork structures by coupling unwinding and annealing in a concerted manner to yield fork regression. As expected, the SMARCAL1 HARP1-WF mutant is able to regress the replication fork as efficiently as the wild-type protein, whereas mutations in the SMARCAL1 HARP2 domain eliminate fork regression activity (Supplemental Fig. 7D–F). Thus, HARP2 but not HARP1 is critical for SMARCAL1 fork regression activity.

Figure 6. SMARCAL1 binds and branch-migrates Holliday junctions. The ability of SMARCAL1 to bind (*A*,*B*) and be activated (*C*) by forked, three-way, and four-way Holiday junctions was compared. The DNA substrates corresponding to each symbol and line color are the same in *B* and *C*. (*D*) Four-way branch migration substrate used in *E–J*. The ^{32}P-labeled DNA strand (#1) for the experiments shown in *E–H* is indicated with an asterisk. Strand #3 was labeled for the experiment shown in *I* and *J*. (*E*) Increasing amounts of SMARCAL1 were incubated with the four-way branch migration substrate in the absence or presence of ATP as indicated. (*G*,*I*) Increasing amounts of wild-type (WT), R764Q, HARP1-WF, or HARP2-WF SMARCAL1 proteins were incubated with the DNA substrate in the presence of ATP. The first three lanes in *E*, *G*, and *I* are size standards generated by annealing the indicated oligonucleotides. The control (Ctl) samples are the annealed branch migration substrate in the absence of recombinant protein. (*F*,*H*,*J*) Quantitation of the reactions from *E*, *G*, and *I*, respectively. The amount of product in the control reactions (from spontaneous branch migration) was set at zero in each experiment, and all other samples are measured relative to the control sample. All reactions in *E–J* were performed for 20 min prior to termination and gel electrophoresis to characterize the products. The *insets* in *H* and *J* are Coomassie-stained gels confirming that equal concentrations of SMARCAL1 proteins were used.

Finally, we tested whether SMARCAL1 can catalyze fork regression and sustained migration on a plasmid-sized substrate that more closely models a stalled replication fork. We created a joint molecule by annealing gapped plasmids (Fig. 7G). This substrate mimics a stalled fork in which the lagging strand is 14 nt longer than the leading strand (Ralf et al. 2006; Blastyak et al. 2007). The extent of fork regression of this substrate was detected by restriction enzyme digestion to liberate a linear 5′-labeled lagging strand. SMARCAL1 efficiently catalyzed remodeling of this substrate, yielding substantial amounts of a regressed fork corresponding to movement of at least 836 base pairs (bp) (Fig. 7H). This reaction is dependent on the amount of SMARCAL1 added to the reaction and requires ATP hydrolysis, since ATPγS completely blocked remodeling of the substrate.

Discussion

Previous studies by our group and others defined SMARCAL1 as a replication stress response protein that acts to preserve genome integrity during DNA replication (Bansbach et al. 2009; Ciccia et al. 2009; Driscoll and Cimprich 2009; Postow et al. 2009; Yusufzai et al. 2009). Immunofluorescent imaging demonstrated that

SMARCAL1 accumulates at damaged replication forks due to its interaction with RPA. We now report that SMARCAL1 associates with active, elongating replisomes, and its absence causes MUS81-dependent DNA damage. Significantly, we found that SMARCAL1 exhibits a much broader range of enzymatic activities than previously recognized, including an ability to promote branch migration of Holliday junctions and fork reversal of model replication forks. Concerted fork regression and branch migration coupled to DNA polymerization provides one mechanism to allow DNA damage bypass and replication restart (Petermann and Helleday 2010). SMARCAL1 depletion does not significantly slow the overall rate of DNA replication but is required for efficient DNA replication restart of stalled or collapsed replication forks (Ciccia et al. 2009). Thus, SMARCAL1 may continuously survey replisomes and promote efficient restart of stalled forks through its fork remodeling activity. In the absence of SMARCAL1, slowed or damaged forks are cleaved by MUS81, perhaps as an alternative mechanism of fork repair.

In addition, our results indicate that all SMARCAL1 activities require the HARP2 and SNF2-like ATPase domains. The HARP2 domain is required for DNA binding, and the HARP2-ATPase domains together form the
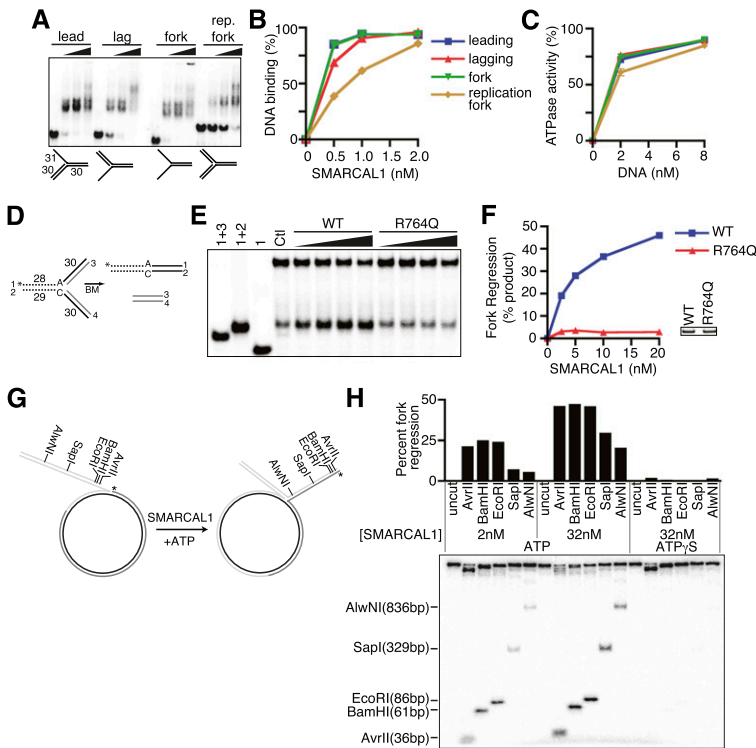
Figure 7. SMARCAL1 catalyzes fork regression of model replication forks. (*A,B*) Increasing amounts of SMARCAL1 were incubated with the indicated substrates to measure DNA binding. (*C*) ATPase activity of SMARCAL1 was measured in the presence of increasing amounts of leading, lagging, fork, and replication fork substrate. Symbols and line colors correspond to the same substrates as in *B*. Error bars represent the mean ± SD from three independent experiments. In cases in which no error bars are visible, the SD is smaller than the symbol size. (*D*) Diagram of the model replication fork substrates used to measure fork regression activity in *E* and *F*. A single mismatch is present at the fork junction to prevent spontaneous fork migration. The labeled strand (#1) is indicated by an asterisk. (*E,F*) Increasing amounts of SMARCAL1 (wild type [WT]) or R764Q SMARCAL1 were incubated with the annealed substrate for 20 min, the reaction was terminated, and products were separated by gel electrophoresis for analysis. The first three lanes in *E* are size standards generated by annealing the indicated oligonucleotides. The control (Ctl) sample is the annealed fork regression substrate in the absence of recombinant protein. The amount of product in the control reaction (from spontaneous regression of the model replication fork substrate) was set at zero in each experiment, and all other samples are measured relative to the control sample. The *inset* in *F* is a Coomassie-stained gel confirming that equal concentrations of SMARCAL1 proteins were used. (*G*) Diagram of the annealed gapped plasmid substrate used to measure SMARCAL1-catalyzed fork regression in *H*. The ^{32}P-labeled DNA end is indicated with an asterisk. (*H*) Restriction digests with the indicated enzymes were completed following incubation of the plasmid substrate with the indicated concentrations of SMARCAL1 in the presence of ATP or ATPγS. The liberated, ^{32}P-labeled DNA fragment was visualized on a polyacrylamide gel. The extent of fork regression was calculated as the amount of liberated fragment compared with the total radioactivity in the reaction. A representative experiment is shown.

functional enzymatic unit of SMARCAL1. Significant sequence similarity between the HARP domain and the DNA-binding domain of the PUR proteins combined with our SAXS data allowed us to derive a model of the solution state structure of the SMARCAL1 core enzyme. The HARP2 and ATPase motifs dock together and constitute a structural and functional core necessary to drive ATP-dependent translocation.

The HARP2 domain likely provides specificity to the action of the ATPase motor domain, thereby converting the energy of ATP hydrolysis into functional strand annealing, branch migration, and fork reversal. This type of activity could be facilitated by insertion of the HARP domain as a kind of wedge at the branch point within these structures (Fig. 8). One possibility supported by our data is that the compact HARP2-ATPase core enzyme contains both dsDNA- and ssDNA-binding surfaces encoded in the ATPase and HARP2 domains, respectively. DNA binding induces a conformational change, promoting ATP hydrolysis and protein translocation. A model for how this could function to promote fork regression is provided by the bacterial RecG protein, which shares some enzymatic activities with SMARCAL1 (Atkinson and McGlynn 2009). Further structural data, including high-resolution structures of SMARCAL1 with a bound DNA substrate, will be required to fully test this hypothesis.

In contrast to the HARP2 domain, the HARP1 domain makes a modest contribution to the DNA-binding and ATPase activities of SMARCAL1 and is largely dispensable for its annealing, branch migration, and fork regression functions. While vertebrate SMARCAL1 proteins contain two HARP domains, invertebrate SMARCAL1 proteins contain only a single HARP domain adjacent to the ATPase domain, suggesting that only a single HARP domain is essential for its evolutionarily conserved functions.

Our conclusions about the important function of the HARP2 domain are generally consistent with a recent report that found that the HARP domains are important for the annealing helicase activity of SMARCAL1 (Ghosal et al. 2011). However, the Chen group (Ghosal et al. 2011) reported that deleting both HARP1 and HARP2 together did not impair either DNA-binding or ATPase activity despite eliminating the annealing helicase activity. In contrast, our data with both deletion and point mutants clearly point to a requirement for the HARP2 domain for all SMARCAL1 enzymatic functions. We tested multiple proteins purified from both insect and human cells using several different DNA substrates and always found that the HARP2 domain was critical for DNA binding, ATPase activity, strand annealing, and branch migration. We do not have an explanation for this discrepancy.

The ability of SMARCAL1 to efficiently bind to Holliday junctions and model replication forks that lack ssDNA
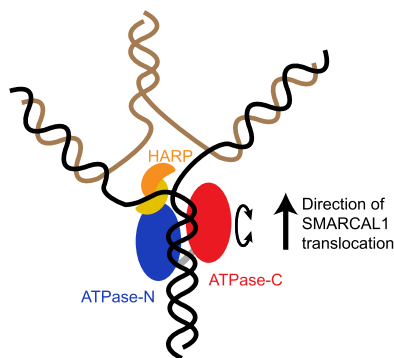
Figure 8. Model for how the translocase activity of the SMARCAL1 HARP2-ATPase core catalyzes fork regression. Existing structures of SNF2 translocases demonstrate that ATPase-N and ATPase-C lobes are capable of adopting different relative conformations and suggest that such conformational changes (depicted as a circular arrow) in response to the ATPase-binding and hydrolysis cycle may drive translocation along dsDNA (Durr et al. 2005; Thoma et al. 2005; Lewis et al. 2008). The SAXS model shows that the HARP2 domain in SMARCAL1 is physically associated with the ATPase-N lobe and may aid in the specialized annealing activity through ssDNA or junction binding. Translocation displaces the nascent DNA strands, induces fork regression, and promotes junction migration.

regions was unexpected, since SMARCAL1 has very little affinity to dsDNA compared with the optimal substrates containing at least 15 nt of dsDNA and 5 nt of ssDNA. One possibility is that SMARCAL1 captures a small amount of these structures as the dsDNA regions near the junction or fork breath to expose ssDNA. Our data indicate that only a small amount of ssDNA or even just a nick is necessary for SMARCAL1 DNA binding when it is in the context of a gap. Likewise, only a small amount may be needed in the context of these more complicated structures. We also observed no significant specificity of human SMARCAL1 for 3′ or 5′ recessed junctions. Furthermore, a 5′ recessed junction containing a model RNA–DNA primer, as would be found during lagging strand replication, efficiently binds and activates SMARCAL1. This contrasts with a previous report that found a preference for a 3′-hydroxyl recessed end (Muthuswami et al. 2000). The origin of this difference may be because the previous report used a fragment of bovine SMARCAL1, whereas we used full-length human SMARCAL1 in our studies.

The ATP-dependent activity of SMARCAL1 to remodel Holliday junctions and replication forks and prevent DNA damage during S phase is reminiscent of the activities of other proteins, including FANCM, WRN, RAD5, BLM, and HLTF (Constantinou et al. 2000; Ralf et al. 2006; Blastyak et al. 2007; Franchitto et al. 2008; Gari et al. 2008a,b; Opresko et al. 2009; Achar et al. 2011). All of these proteins are thought to be recruited to damaged replication forks, but it is unclear whether any travel with active forks like SMARCAL1. In contrast to SMARCAL1, none of these proteins contain a HARP domain or exhibit annealing helicase activity. Instead, most are DNA helicases. Thus, the enzymatic mecha-

nisms by which they remodel replication fork structures are likely to be different. Why there are so many different enzymes that can catalyze similar reactions on DNA is unclear. It is possible that some of these enzymes work coordinately at the same damaged fork. In this regard, it is interesting that the loss of WRN, like SMARCAL1, also causes MUS81-dependent fork cleavage (Franchitto et al. 2008), and we and others have found WRN in SMARCAL1 purifications, suggesting a possible physical interaction (Ciccia et al. 2009; data not shown). Coordination of their enzymatic activities might help remodel damaged forks in cells where many other replisome and repair proteins may be present. However, these proteins must also have distinct functions, since inactivating mutations cause different human diseases.

In summary, our data suggest that SMARCAL1 surveys DNA replication forks. When it detects a problem, it uses its DNA-stimulated ATPase motor to remodel the fork by catalyzing strand annealing, branch migration, and fork reversal to promote efficient fork repair. These activities are encoded within the HARP2-SNF2 ATPase domains, which form a functional enzyme flanked by regulatory sequences. Absence of SMARCAL1 forces the use of alternative fork repair mechanisms that involve MUS81-dependent DNA double-strand breaks.

Materials and methods

Cell culture

HEK-293T and U2OS cells were cultured in DMEM supplemented with 7.5% FBS. Sf9 cells were cultured in Insect XPRESS medium with 7.5% FBS at 27°C.

Antibodies

The antibodies used were as follows: Flag-M2 (Sigma), γH2AX and GAPDH (Millipore), RPA (Bethyl Laboratories), H3 (Abcam), and MUS81 (Novus). The SMARCAL1 antibody was described previously (Bansbach et al. 2009).

Detection of γH2AX

γH2AX foci were detected by indirect immunofluorescent imaging of fixed U2OS cells 72 h after transfection with siRNA as previously described (Lovejoy et al. 2009).

iPOND

The iPOND technique was performed as described previously (Sirbu et al. 2011). Briefly, cells were labeled for 10 min with EdU, then treated with 2 mM HU for increasing amounts of time. Alternatively, after the EdU labeling period, 10 μM thymidine was added to the growth medium for 20 min as a "chase" sample. This concentration of thymidine does not block replication but is sufficient to ensure that no additional EdU is incorporated. After cross-linking with formaldehyde and a click reaction to conjugate biotin to the EdU-labeled nascent DNA, protein–DNA complexes were isolated with streptavidin beads, cross-links were reversed, and the eluted proteins were analyzed by immunoblotting. The "no click" control omitted the biotin-azide during the click reaction.

Protein purification

Flag-SMARCAL1, His-SMARCAL1(325-954), HARP1-WF, and HARP2-WF were purified from baculovirus-infected cells essentially as described previously (Yusufzai and Kadonaga 2008) except that cells were lysed in TNT buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 0.2 mM PMSF, 1 mg/mL leupeptin, 1 mg/mL aprotinin, and 0.1% Triton X-100. Proteins for structural studies were purified by Ni-NTA affinity, ion exchange, and gel filtration chromatography. To purify SMARCAL1 proteins from human cells, HEK-293T cells were transfected with pLPCX-Flag-HA-SMARCAL1 plasmids using Lipofectamine 2000 (Invitrogen). Seventy-two hours after transfection, the cells were lysed in TNT buffer for 30 min on ice. After high-speed centrifugation, the cleared lysates were incubated with Flag-M2 beads (Sigma) for 3 h at 4°C. The beads were washed three times in wash buffer (TNT buffer containing 0.3 M LiCl) and twice in SMARCAL1 buffer (20 mM HEPES at pH 7.6, 20% glycerol, 0.1 M KCl, 1.5 mM MgCl$_2$, 0.2 mM EDTA, 1 mM DTT, 0.2 mM PMSF, 0.01% IGEPAL CA-630). The bound proteins were eluted in SMARCAL1 buffer containing 0.25 mg/mL Flag peptide on ice, flash-frozen, and stored at −80°C.

DNA-binding, annealing helicase, and ATPase assays

The gel mobility shift assays for DNA-binding, annealing helicase, and SMARCAL1 ATPase assays were performed as described previously (Yusufzai and Kadonaga 2008) with the following modifications. For the gel mobility shift assay, increasing concentrations of purified SMARCAL1 (0, 0.5, 1, 2 nM final concentrations) were combined with radiolabeled oligonucleotide probe (1 nM final concentration) in binding buffer supplemented with 0.2% IGEPAL CA-630. The samples were loaded into a 5% polyacrylamide 0.5× TBE gel (82 × 28.5 cm, 1 mm thick), and subjected to electrophoresis in 0.5× TBE for 2 h and 30 min at 50 V at 4°C. The gels were dried and quantified using a Molecular Imager FX (Bio-Rad). DNA-binding reactions were performed at least twice, and a representative experiment is shown with quantitation. For the annealing helicase assay, the topoisomerase I was purchased (Invitrogen), and pBluescript was used as the plasmid substrate. For the ATPase assay, increasing concentrations of oligonucleotides (0, 2, or 8 nM final concentration) were incubated with purified SMARCAL1 (8 nM final concentration) in a final volume of 10 μL, and the reactions were incubated for 30 min at 30°C. The results are presented as the percent of ATP hydrolyzed to ADP during the reaction. ATPase assays were performed a minimum of three times each, and graphs depict means and standard deviation error bars. All oligonucleotide sequences are described in Supplemental Table 1, and all DNA substrates are described in Supplemental Table 2. All figures show a representative experiment from at least two replicates.

Homology modeling

The HARP repeats were identified as evolutionary structural homologs to PUR-α repeats using the FUGUE sequence–structure homology recognition server (Shi et al. 2001). The HARP2 (amino acids 325–396) homology model was constructed using the crystal structure of PUR-α (Protein Data Bank [PDB] ID 3K44) (Graebsch et al. 2009) residues 41–185 as a template. The ATPase model (SMARCAL1 residues 451–856) was generated from residues 455–891 of the SsoRad54 crystal structure (PDB ID 1Z63) (Durr et al. 2005). In both cases, the SMARCAL1 sequences were threaded onto the crystal structure using Swiss PDB Viewer, and the model was optimized using Swiss Model (http://swissmodel.expasy.org).

SAXS data collection and model building

SAXS data were collected at the SIBYLS beamline at the Advanced Light Source and prepared as described (Hura et al. 2009). Specifically, SAXS data were collected on SMARCAL1(325–954) in buffer containing 20 mM HEPES (pH 7.6), 200 mM NaCl, 2 mM MgCl$_2$, 0.5 mM TCEP, 5% glycerol, and 1% sucrose. The protein sample was prepared for SAXS as described (Kazantsev et al. 2011) using a Shodex KW402.5 size exclusion column. The peak fraction was analyzed for SAXS as a 2/3 dilution series starting from 3 mg/mL. Three exposure times (0.5, 1, and 6 sec) were taken at 25°C and 12 keV. Guinier and Kratky analysis was performed as described (Putnam et al. 2007; Rambo and Tainer 2011). Linearity of the Guinier region for each exposure demonstrated a lack of radiation damage and aggregation (Supplemental Fig. 4D). SAXS profiles were overlaid, inspected for concentration-dependent scattering, and merged (Hura et al. 2009). For modeling, the composite scattering curve was generated from data from 1-sec exposures of 2 and 3 mg/mL samples. The maximum dimension (116 Å) was determined using GNOM (Svergun 1992). Atomistic-based modeling of the SAXS data was achieved with the program BUNCH (Petoukhov and Svergun 2005) using HARP and ATPase homology models. The models were treated as independent domains in a simulated annealing algorithm to determine their relative spatial arrangements. Missing residues between the HARP and ATPase domains (397–450) were modeled as dummy residues as described (Petoukhov and Svergun 2005). Ab initio modeling was performed with GASBOR using 630 dummy residues. Ten independent modeling runs were performed and averaged (Volkov and Svergun 2003) to produce a final macromolecular envelope. The final model targeted residues 325–856, consistent with a Porod volume of 91,148 Å^3 calculated from the SAXS data. The missing C-terminal 99 residues were not included in the modeling based on proteolytic sensitivity of the C terminus (Supplemental Fig. 4A).

Branch migration and fork regression assays

Oligonucleotide #48 was end-labeled with [γ-^{32}]ATP and T4 polynucleotide kinase (New England Biolabs) and purified through a G25 column (GE Healthcare). To prepare tailed or forked intermediates, 250 nM complementary ssDNA oligonucleotides (#48/#54 and #55/#56 for the branch migration, and #48/#50 and #53/#54 for the fork regression) were annealed in 20 μL of SSC buffer (15 mM NaCitrate at pH 7.0, 150 mM NaCl) in a PCR machine. To prepare the branch migration and the fork regression substrate, 32 nM ^{32}P-labeled and 48 nM nonlabeled DNA intermediates were incubated in reaction buffer (40 mM Tris at pH 7.5, 20 mM KCl, 5 mM MgCl$_2$, 100 μg/mL BSA, 2 mM ATP, 2 mM DTT) for 30 min at 37°C. The DNA substrates were diluted threefold in reaction buffer and mixed with increasing amounts of SMARCAL1 in a 20-μL reaction volume. The reaction was completed for 20 min at 37°C and terminated by the addition of 3× stop buffer (0.9% SDS, 50 mM EDTA, 40% glycerol, 0.1% bromophenol blue, 0.1% xylene cyanol). Samples were loaded into 8% polyacrylamide 1× TBE gels (82 × 28.5 cm, 1 mm thick) and subjected to electrophoresis in 1× TBE for 90 min at 80 V at room temperature. The gels were dried and quantified using a Molecular Imager FX (Bio-Rad).

The plasmid-sized replication fork model substrate was generated and purified as described (Blastyak et al. 2007). Recombinant SMARCAL1 purified from insect cells was incubated with 0.5 nM substrate for 20 min at 37°C in reaction buffer (20 mM Tris at pH 7.5, 50 mM KCl, 5 mM MgCl$_2$, 100 μg/mL BSA, 2 mM ATP or ATPγS, 1 mM DTT). The reaction was quenched by the addition of 10 mM ATPγS and 10 mM MgCl$_2$. One microliter

(2–20 U, depending on the enzyme) of the indicated restriction enzymes was added to the reaction and further incubated for 30 min at 37°C. The reaction products were then separated on a 6% polyacrylamide gel. The gel was dried and quantified using a Molecular Imager FX (Bio-Rad).

References

Achar YJ, Balogh D, Haracska L. 2011. Coordinated protein and DNA remodeling by human HLTF on stalled replication fork. *Proc Natl Acad Sci* **108:** 14073–14078.

Atkinson J, McGlynn P. 2009. Replication fork reversal and the maintenance of genome stability. *Nucleic Acids Res* **37:** 3475–3492.

Bansbach CE, Betous R, Lovejoy CA, Glick GG, Cortez D. 2009. The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. *Genes Dev* **23:** 2405–2414.

Bansbach CE, Boerkoel CF, Cortez D. 2010. SMARCAL1 and replication stress: An explanation for SIOD? *Nucleus* **1:** 245–248.

Blastyak A, Pinter L, Unk I, Prakash L, Prakash S, 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.

Boerkoel CF, O'Neill S, Andre JL, Benke PJ, Bogdanovic R, Bulla M, Burguet A, Cockfield S, Cordeiro I, Ehrich JH, et al. 2000. Manifestations and treatment of Schimke immuno-osseous dysplasia: 14 new cases and a review of the literature. *Eur J Pediatr* **159:** 1–7.

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

Bugreev DV, Mazina OM, Mazin AV. 2006. Rad54 protein promotes branch migration of Holliday junctions. *Nature* **442:** 590–593.

Ciccia A, Bredemeyer AL, Sowa ME, Terret ME, Jallepalli PV, Harper JW, Elledge SJ. 2009. The SIOD disorder protein SMARCAL1 is an RPA-interacting protein involved in replication fork restart. *Genes Dev* **23:** 2415–2425.

Constantinou A, Tarsounas M, Karow JK, Brosh RM, Bohr VA, Hickson ID, West SC. 2000. Werner's syndrome protein (WRN) migrates Holliday junctions and co-localizes with RPA upon replication arrest. *EMBO Rep* **1:** 80–84.

Driscoll R, Cimprich KA. 2009. HARPing on about the DNA damage response during replication. *Genes Dev* **23:** 2359–2365.

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

Flaus A, Martin DM, Barton GJ, Owen-Hughes T. 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* **34:** 2887–2905.

Franchitto A, Pirzio LM, Prosperi E, Sapora O, Bignami M, 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.

Gari K, Decaillet C, Delannoy M, Wu L, Constantinou A. 2008a. Remodeling of DNA replication structures by the branch point translocase FANCM. *Proc Natl Acad Sci* **105:** 16107–16112.

Gari K, Decaillet C, Stasiak AZ, Stasiak A, Constantinou A. 2008b. The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. *Mol Cell* **29:** 141–148.

Ghosal G, Yuan J, Chen J. 2011. The HARP domain dictates the annealing helicase activity of HARP/SMARCAL1. *EMBO Rep* **12:** 574–580.

Graebsch A, Roche S, Niessing D. 2009. X-ray structure of Pur-α reveals a Whirly-like fold and an unusual nucleic-acid binding surface. *Proc Natl Acad Sci* **106:** 18521–18526.

Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL II, Tsutakawa SE, Jenney FE Jr, Classen S, Frankel KA, Hopkins RC, et al. 2009. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat Methods* **6:** 606–612.

Kazantsev AV, Rambo RP, Karimpour S, Santalucia J Jr, Tainer JA, Pace NR. 2011. Solution structure of RNase P RNA. *RNA* **17:** 1159–1171.

Lewis R, Durr H, Hopfner KP, Michaelis J. 2008. Conformational changes of a Swi2/Snf2 ATPase during its mechano-chemical cycle. *Nucleic Acids Res* **36:** 1881–1890.

Lovejoy CA, Xu X, Bansbach CE, Glick GG, Zhao R, Ye F, Sirbu BM, Titus LC, Shyr Y, Cortez D. 2009. Functional genomic screens identify CINP as a genome maintenance protein. *Proc Natl Acad Sci* **106:** 19304–19309.

Muthuswami R, Truman PA, Mesner LD, Hockensmith JW. 2000. A eukaryotic SWI2/SNF2 domain, an exquisite detector of double-stranded to single-stranded DNA transition elements. *J Biol Chem* **275:** 7648–7655.

Opresko PL, Sowd G, Wang H. 2009. The Werner syndrome helicase/exonuclease processes mobile D-loops through branch migration and degradation. *PLoS ONE* **4:** e4825. doi: 10.1371/journal.pone.0004825.

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

Petermann E, Helleday T. 2010. Pathways of mammalian replication fork restart. *Nat Rev Mol Cell Biol* **11:** 683–687.

Petermann E, Orta ML, Issaeva N, Schultz N, 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.

Petoukhov MV, Svergun DI. 2005. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J* **89:** 1237–1250.

Postow L, Woo EM, Chait BT, Funabiki H. 2009. Identification of SMARCAL1 as a component of the DNA damage response. *J Biol Chem* **284:** 35951–35961.

Putnam CD, Hammel M, Hura GL, Tainer JA. 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40:** 191–285.

Ralf C, Hickson ID, Wu L. 2006. The Bloom's syndrome helicase can promote the regression of a model replication fork. *J Biol Chem* **281:** 22839–22846.

Rambo RP, Tainer JA. 2010. Bridging the solution divide: Comprehensive structural analyses of dynamic RNA, DNA, and

protein assemblies by small-angle X-ray scattering. *Curr Opin Struct Biol* **20:** 128–137.

Rambo RP, Tainer JA. 2011. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95:** 559–571.

Shi J, Blundell TL, Mizuguchi K. 2001. FUGUE: Sequence–structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J Mol Biol* **310:** 243–257.

Sirbu BM, Couch FB, Feigerle JT, Bhaskara S, Hiebert SW, Cortez D. 2011. Analysis of protein dynamics at active, stalled, and collapsed replication forks. *Genes Dev* **25:** 1320–1327.

Suhre K, Sanejouand YH. 2004. ElNemo: A normal mode Web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res* **32:** W610–W614. doi: 10.1093/nar/gkh368.

Svergun DI. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **25:** 495–503.

Thoma NH, Czyzewski BK, Alexeev AA, Mazin AV, Kowalczykowski SC, Pavletich NP. 2005. Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *Nat Struct Mol Biol* **12:** 350–356.

Volkov VV, Svergun DI. 2003. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* **36:** 860–864.

Yuan J, Ghosal G, Chen J. 2009. The annealing helicase HARP protects stalled replication forks. *Genes Dev* **23:** 2394–2399.

Yusufzai T, Kadonaga JT. 2008. HARP is an ATP-driven annealing helicase. *Science* **322:** 748–750.

Yusufzai T, Kadonaga JT. 2010. Annealing helicase 2 (AH2), a DNA-rewinding motor with an HNH motif. *Proc Natl Acad Sci* **107:** 20970–20973.

Yusufzai T, Kong X, Yokomori K, Kadonaga JT. 2009. The annealing helicase HARP is recruited to DNA repair sites via an interaction with RPA. *Genes Dev* **23:** 2400–2404.